

Received 30 May 2024, accepted 14 June 2024, date of publication 24 June 2024, date of current version 10 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3417928

RESEARCH ARTICLE

Direct Filter Learning From Iterative Reconstructed Images for High-Quality Analytical CBCT Reconstruction Using FDK-Based Neural Network

HONGLIANG QI¹, CHAO LONG², HANWEI LI¹, SHUANG HUANG³, DEBIN HU¹, YUAN XU⁴, AND HONGWEN CHEN¹

¹Department of Clinical Engineering, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China

²School of Graduate Studies, La Consolacion University Philippines, Malolos 3000, Philippines

³Foshan Shunde Heyou Hospital, Foshan 528399, China

⁴School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

Corresponding authors: Yuan Xu (yuanxu@smu.edu.cn) and Hongwen Chen (chw47922@126.com)


This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFA1204203 and Grant 2023YFC2414601; in part by the Key-Area Research and Development Program of Guangdong under Grant 2019B111103001; and in part by the President Foundation of Nanfang Hospital, Southern Medical University, under Grant 2022B016.

ABSTRACT Purpose: We propose an FDK-based neural network to directly learn the filter from an iterative reconstruction (IR) algorithm and apply the learned filter in the FDK algorithm to obtain a high-quality CBCT reconstruction. **Methods:** The FDK algorithm is transformed into a linear expression of several matrix multipliers and embedded into neural network layers. Then, the FDK-based neural network framework is built including two fundamental modules and four core network layers. This network model can learn a filter directly from the iteratively reconstructed CBCT images by cascading the network layers of cosine weighting, filtering, backprojection, and leaky rectified linear unit and setting filter as the only trainable parameter. Preliminary and simulation studies performed on abdominal CT datasets are conducted to explore the correctness and effectiveness of the learned filter. Then, the head and neck CT data and Catphan phantom are utilized to demonstrate the generalization performance of the learned filter. **Results:** Preliminary study shows that the learned filter is consistent with the target filter, and the mean absolute difference is around 0.001. Compared with conventional FDK, the FDK-based neural network shows a better image quality with the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) increasing by 67% and 6%, respectively. In terms of the line-pair slice in the Catphan phantom, the SSIM and PSNR are improved by 13.75% and 42.78%, respectively. **Conclusions:** The FDK-based neural network can reconstruct high-quality images by directly learning a filter from the label images and provides a new perspective on solving the time-consuming problem of IR methods.

INDEX TERMS CBCT iterative reconstruction algorithm, CBCT FDK algorithm, neural network, learnable filter.

I. INTRODUCTION

Traditional three-dimensional reconstruction algorithms for cone-beam computed tomography (CBCT) are mainly

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

divided into two categories: iterative reconstruction (IR) and analytical reconstruction (AR). IR leverages multiple forward projection and backprojection operations to constantly narrow the difference between calculated and measured projections to generate accurate reconstructed images. Many IR methods have been developed for image regularization,

such as total variation (TV) [1], [2], nonlocal regularization methods [3], [4], and low-rank models [4]. IR is of good image quality, but the reconstruction speed is generally slow. Conversely, AR represented by the Feldkamp-Davis-Kress (FDK) algorithm [5] is well explained and fast. The conventional FDK reconstruction includes operations of cosine weighting, ramp filtering, and distance-weighted backprojection, where the ramp filter is used to eliminate the blur effect of backprojection. The filter plays an important role in the FDK algorithm and determines the image quality of the FDK reconstruction in most cases. Filters with different frequency attributions may lead to distinguishable image quality [6], [7]. Marinus et al. [8] proposed a computationally efficient numerical algorithm for computing a reconstruction filter that provided decent CT reconstruction across a broad range of scenarios. Yuuki et al. [9] designed a modified Shepp-Logan filter with high-frequency enhancement and performed it on the CBCT projections of two skull phantoms. The spatial resolution of the periodontal ligament space was substantially improved. However, the above methods are not flexible to manually optimize the filter weights according to different scenarios.

Deep learning (DL) methods can break through the limitations of traditional algorithms and bring better flexibility via automatic training. Some literature reviews have summarized and sorted out the existing applications of DL in medical CT image reconstruction, and made reasonable speculation and elaboration on the possibility of its future development [10], [11], [12], [13]. Relevant studies can be divided into three categories. The first is the DL-based image post-processing from image domain to image domain. For example, with the good performance of a convolutional neural network on image feature extraction, DL-based methods are applied to transform the limited-angle CT images into full-angle CT images [14], sparse-view images into full-view images [15], [16], and low-resolution CT scans into high-resolution reconstruction [17]. The second is DL-based reconstruction directly from the projection domain to the image domain. Researchers have proposed many novel deep neural networks by introducing state-of-the-art DL methods into CT reconstruction, such as DEER [18], iCT-Net [19], LEARN [20], DRONE [21], and MAGIC [22]. The above two categories are concentrated on improving the image quality or highlighting the region of interest through a well-trained deep neural network at the cost of computation complexity and poor mathematical interpretability.

The last category is to combine the conventional reconstruction algorithms with DL methods. Chen et al. [23] proposed a new CT image reconstruction framework named AirNet by fusing analytical and iterative algorithms with a deep neural network. Zhang et al. [24] proposed a new DL-based model, MetaInv-Net, for sparse-view CT image reconstruction with the backbone network architecture built by unrolling an iterative algorithm. Würfl et al. [25] proposed the FBP neural network for limited-angle CT reconstruction. They mapped the FBP algorithm to the layers of a neural

network, and the classic CT reconstruction of the parallel-beam, fan-beam, and cone-beam was realized at a learnable level. Wang et al. [26] further improved the performance of the FBP neural network by adding the encoding and decoding network called U-Net [27] into the image domain. This new type of DL reconstruction framework has become popular in recent years with its simplicity and interpretability. However, one critical issue needs to be pointed out. The FBP neural network [25] was applied to a limited angle problem, showing an effective restoration of missing data by training the Parker redundant weights [28]. The learning rate related to filter weights is set to zero and unchangeable during the whole training process, which makes the neural network pay no attention to the dominant influence of the filter on the reconstructed image quality. Research on utilizing the neural network to learn a filter to achieve high-quality analytic reconstruction is limited.

Considering the fact that FDK is fast and TV-based iterative methods are slow but with better quality, a neural network was designed to learn the filter in an FDK-based network to improve the image quality, so that image quality can be improved without increasing reconstruction time. Specifically, in this study, an FDK-based neural network framework is built to learn a filter directly from the iteratively reconstructed CBCT images by cascading the operations of cosine weighting, filtering, and backprojection and setting the filter as the only trainable parameter. The high-quality images reconstructed by the TV regularized IR algorithm based on the Alternating Direction Method of Multipliers [29], [30] (ADMM-based TV) are regarded as the label of the network. Our work is mainly carried out around two points: leveraging the power of the neural network to break through the deficiency of manually designed filter weights and providing a new perspective to solve the time-consuming problem of IR methods.

The remainder of this work is organized as follows: Section II first explains how to build the FDK-based neural network by introducing the conventional FDK algorithm used in CBCT imaging and then describes the datasets used in the study. Section III shows the representative reconstructed images and related evaluation metrics in the study. Sections IV and V present the discussion and conclusion, respectively.

II. METHOD

A. WORKFLOW OF THE FDK-BASED NEURAL NETWORK

The framework of our FDK-based neural network including two fundamental modules and four core network layers is illustrated in Fig. 1. The FDK-based neural network cascades cosine weighting layer, filtering layer, distance weighted backprojection layer, and leaky rectified linear unit (Leaky ReLU) [31] activation layer to perform FDK reconstruction in a learnable way. In the training process, the projection after logarithmic transformation is first weighted by the cosine weighting layer to properly eliminate the influence of cone angle formed by the distance and angle between the

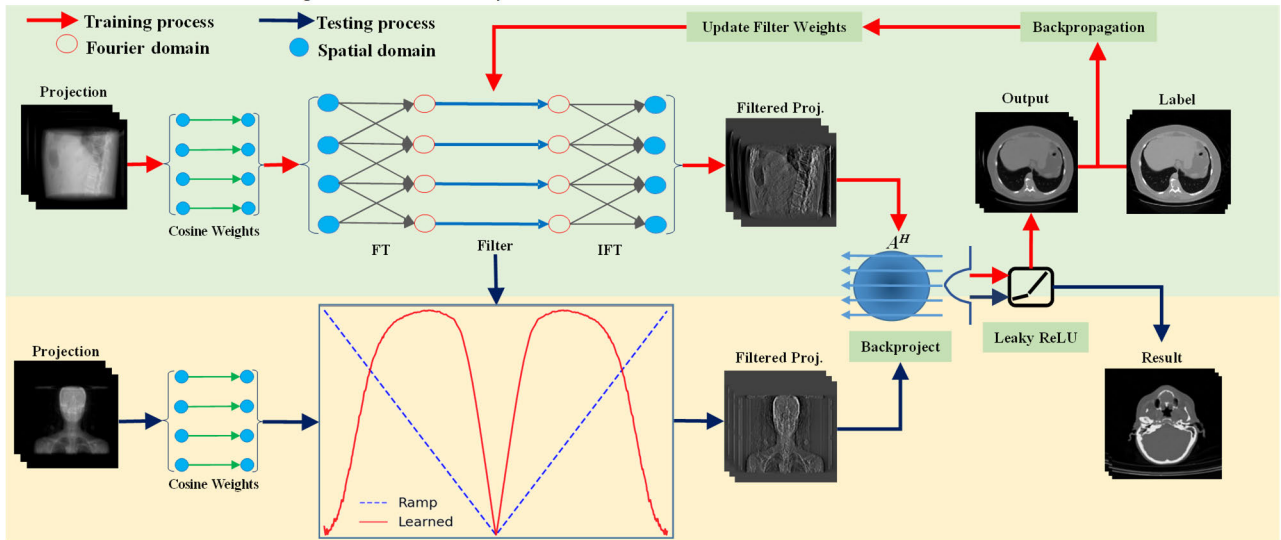


FIGURE 1. FDK-based neural network framework. The reconstruction filter in the training process is initialized as the ramp filter (blue dotted line). When the preset convergence condition is satisfied, the learned filter (red solid line) is saved and applied to the FDK reconstruction.

voxel and the X-ray source. Then, Fourier transform (FT) is applied to the weighted projection, the filtering operation is implemented by multiplying FT weighted projection with a one-dimensional filter in the frequency domain, and the inverse Fourier transform (IFT) is applied to obtain the filtered projection. Next, the filtered projections at different scanning angles are backprojected along the opposite direction of the X-ray to get the reconstructed images. As the output is physically defined as a linear attenuation coefficient, the Leaky ReLU activation layer with a leaky parameter 0.15 is used to constrain the non-negative nature of the reconstruction to some extent. The filter weights are the only trainable parameters, which are initialized with a projection-dependent ramp filter. The other initial weights directly originate from the FDK algorithm and maintain constant values during the training process. The neural network is trained by propagating the analytical gradients of these layers and updating the filter weights when the training process is finished. Finally, the learned filter is saved and applied to replace the ramp filter of the FDK algorithm to achieve the CBCT reconstruction in the testing phase.

B. FDK-BAESD NEURAL NETWORK

The schematic diagram of FDK reconstruction is shown in Fig. 2, where the x , y , and z axes form the world coordinate system. The FDK algorithm is given by:

$$f(r, \varphi, z) = \frac{1}{2} \int_0^{2\pi} \frac{D}{D-s} \int_{-\infty}^{+\infty} \frac{D}{\sqrt{D^2 + l^2 + \hat{z}^2}} g(l, \hat{z}, \beta) h(l' - s) dl d\beta \quad (1)$$

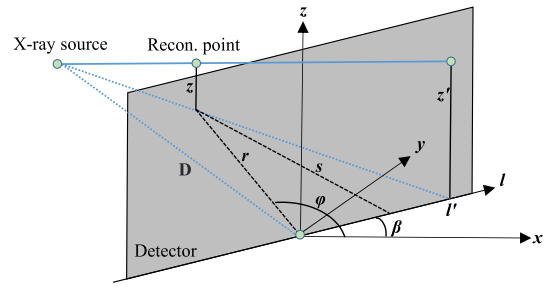


FIGURE 2. Coordinate diagram of FDK reconstruction.

where (r, φ, z) denotes three-dimensional coordinate in the polar system and z refers to the coordinate along the rotation axis. D denotes the Euclidean distance from the X-ray source to the center of the detector, and s is the Euclidean distance from the reconstructed point to the virtual detector. Generally, the virtual detector is located in the center of rotation and parallel to the actual detector, which simplifies the analysis and implementation of the algorithm. $g(l, \hat{z}, \beta)$ is the CBCT projection produced by the object $f(r, \varphi, z)$; $h(l)$ is the convolution kernel of the one-dimensional row-wise ramp filter; $D/\sqrt{D^2 + l^2 + \hat{z}^2}$ and $D/D - s$ are the cosine weighting factor of the incidence angle of cone-beam and the distance weight of cone-beam backprojection, respectively; and l and \hat{z} denote the coordinates in the horizontal and vertical direction of the considered detector pixel, respectively.

To map the FDK algorithm into a neural network framework, the whole reconstruction procedure of the CT imaging system is described as a discrete linear expression, so that the above reconstruction process can be represented by several matrix multiplier operators and embedded into a neural network. The derivation from the acquisition of projection g of

an object x can be described as:

$$Ax = g \quad (2)$$

where A is the system matrix describing the geometry of the imaging system. The reconstruction can be obtained by:

$$x = A^{-1}g \quad (3)$$

where A^{-1} denotes the inverse system matrix. In terms of the CBCT system, we assume that the $x=[x_1, x_2, \dots, x_J]$ and $g=[g_1, g_2, \dots, g_J]$ are I -dimensional and J -dimensional vectors, respectively, the system matrix A is a non-square matrix with the shape of $I \times J$, in which the inverse transformation is unable to be taken directly. Consequently, the Moore–Penrose generalized inverse matrix A^+ is introduced to replace A^{-1} and rewrite Eq. (3) as:

$$x = A^H \left(AA^H \right)^+ g \quad (4)$$

where A^H is the conjugate transpose of system matrix A , also called adjoint projection matrix, which can be implemented as the backprojection operator in the study. According to the analytical filtered backprojection approach, the operator $(AA^H)^+$ corresponds to convolution between the cosine weighted projection with the ramp filter. Then, considering the FDK reconstruction algorithm given by Eq. (1), Eq. (4) can be rewritten as:

$$x = A^H \cdot K * (W_{\cos} \cdot g) \quad (5)$$

where the acquired projection g multiplies the pixel-wise cosine weights W_{\cos} and then makes a convolution operation between weighted cone-beam projections and one-dimensional ramp filter kernel K along the l -direction. Finally, the backprojection via the adjoint projection matrix A^H is used to obtain the reconstruction x . The transformation of the reconstruction process into matrix operators makes FDK-based neural network strongly interpretable.

The FDK-based neural network framework provides two fundamental modules and four core network layers to perform cosine weighting and filtered backprojection to reconstruct images. In the network, a cone-beam geometry module is defined by the detector shape, detector resolution, volume shape, volume resolution, scanning range, number of projections, source-to-detector distance, and source-to-axis distance. Then, a circular trajectory module constructs the scanning trajectory according to the geometric attributes and generates a group of projection matrices to describe the whole imaging process. We construct the network layers by embedding the C++ and CUDA-based FDK algorithm into the TensorFlow platform. By specifying these layers that implement the operators W_{\cos} , K , and A^H , the FDK reconstruction is mapped to a neural network directly. Specifically, the operator W_{\cos} is modeled as a two-dimensional multiplication layer, called the cosine weighting layer, where the input tensors correspond to the projection after logarithmic transformation. This layer uses the cosine function of the incidence angle to weight the pixels in projection, projecting all

ray vectors to the central ray direction. The number of weights is consistent with the total pixel number of the detector and equal to the number of nodes in this layer. The operator K implements the convolution operation within the Fourier frequency domain to avoid the high computation complexity in the spatial domain, where the FT and IFT are the layers straightforwardly provided by the TensorFlow framework. Then, the one-dimensional filter is realized as the weight of an element-wise multiplication layer, where the filter size is equal to the spatial length of the detector along the vertical axis and equal to the number of nodes in this layer. In a word, the filter layer consists of the FT layer, IFT layer, and multiplication layer, which takes the cosine-weighted projections as input tensors. The filtering layer is expected to adjust the frequency distribution of the projection, eliminating the latent blur effect of backprojection.

Compared with the direct conversion from the projection domain to the image domain by the fully connected layer, which leads to a sharp increase in the number of parameters in the whole network model, we choose to directly implement the backprojection operator A^H as a kernel function inspired by Syben et al.'s package in TensorFlow [32]. With the help of the CMake tool, the corresponding C++ code functions are packaged and compiled into a dynamic link library for external Python function calls, so as to cascade with the cosine weighting layer and filtering layer within the TensorFlow platform. This backprojection layer leverages the projection matrices obtained from the circular trajectory module to reconstruct the filtered projection in a voxel-driven manner. Moreover, a texture interpolation mode is provided for memory management of the GPU to interpolate the different measured line integrals.

Finally, it is necessary to limit the negative output in terms of the reconstruction of linear attenuation values. However, when training a neural network, the rectified linear unit (ReLU) function will lead to silent neurons, causing defects in the updated weights. The Leaky ReLU activation layer with the leaky parameter 0.15 is used to address the negative output that should not be set to zero directly, avoiding the “value defects” and “dead neuron” problem in the training process.

By automatically calculating the gradient vectors of the loss function corresponding to all network layers, the training process is continuously well performed to estimate the error in the image domain and update the filter weights in the projection domain.

C. ADMN-BASED TV RECONSTRUCTION

In this study, high-quality reconstructed images of ADMM-based TV are regarded as labels to train an analytical filter. The objective function of the ADMM-based TV reconstruction is formulated as:

$$x = \arg \min_x \frac{\mu}{2} \|Ax - y\|_2^2 + \|x\|_{TV} \quad (6)$$

where $\|\cdot\|_2$ denotes the 2-norm in the projection vector space. The first fidelity term is used to maintain consistency between

the reconstructed image and the observed value, and the vector μ is a balance factor for adjusting the weight between the fidelity term and the second penalty term. Given that TV is able to maintain the feature edge of the reconstructed image and suppress artifacts and noise well, the penalty term is defined as follows:

$$\|x\|_{TV} = \int |\nabla x(i)| di \quad (7)$$

where $\nabla x(i)$ represents the gradient of image x at voxel i . In Eq. (6), this problem is difficult to be solved directly due to the coupling of image gradient and binary norm. Therefore, an auxiliary variable is introduced to decouple them, that is:

$$(x, z) = \arg \min_{x,z} \frac{\mu}{2} \|Ax - y\|_2^2 + \|z\|_1 \text{ s.t. } \nabla x - z = 0 \quad (8)$$

Then, the ADMM method is introduced to solve the convex optimization problem with constraints expressed by:

$$L_A(x, z, w) \triangleq \frac{\mu}{2} \|Ax - y\|_2^2 + \|z\|_1 + w^T (\nabla x - z) + \frac{\lambda}{2} \|\nabla x - z\|_2^2 \quad (9)$$

where w is the dual variable or Lagrange multiplier and μ is a positive weighting parameter corresponding to the penalty term.

The ADMM-based TV relies on the collaboration between the fidelity hyperparameter μ and the penalty hyperparameter λ to reconstruct an ideal image. Generally, when the fidelity hyperparameter remains unchanged, the larger the penalty hyperparameter λ is, the smoother the reconstructed images are. For simplicity, the ADMM-based TV reconstruction with λ at 0.01 and 10 are separately named as standard TV reconstruction and smoothed TV reconstruction, whereas μ at 10 is fixed. A comparison of ADMM-based TV reconstruction with different penalty hyperparameters is shown in Fig. 3.

D. EXPERIMENTAL MATERIALS

To evaluate the performance of our proposed FDK-based neural network, two open source CT datasets are collected in this study [33], [34]. In total, 52 abdomen CT data and another 10 head and neck CT data are selected and resized to a resolution of $110 \times 512 \times 512$ with a voxel size of $0.16 \text{ cm} \times 0.078 \text{ cm} \times 0.078 \text{ cm}$. We perform the cone-beam forward projection on the CT image to generate projection and take the logarithm transformation. For each CT dataset, 360 projection images are generated using a two-dimensional detector of 512×384 pixels with each pixel having a size of $0.12 \text{ cm} \times 0.12 \text{ cm}$. The source-to-detector and source-to-object distances are 150 cm and 100 cm, respectively. The angular increment between each projection is one degree. Finally, the projection and reconstruction counterpart are regarded as the sample and label with respect to the neural network, and the original CT data is the corresponding ground truth.

A preliminary study is conducted to verify whether the filter learned by the FDK-based neural network is consistent

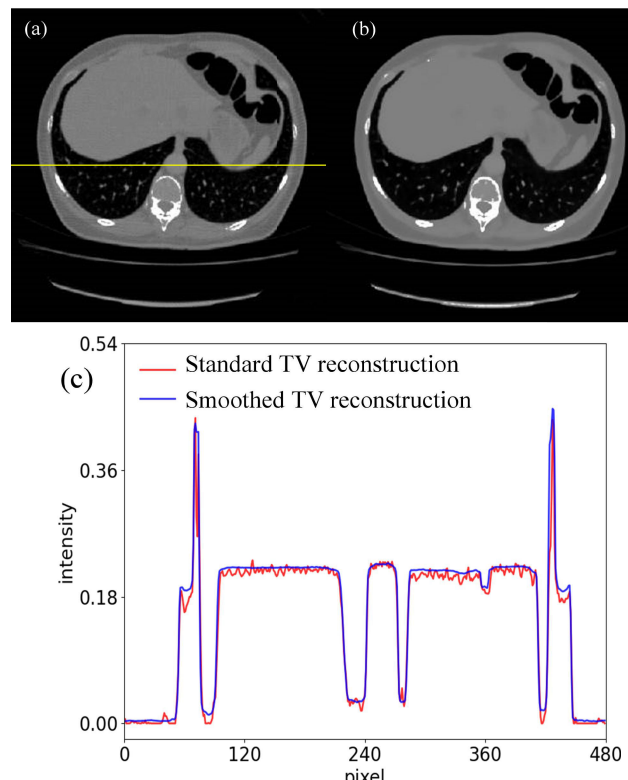


FIGURE 3. ADMM-based TV reconstruction with different hyperparameter λ . (a) Standard TV reconstruction, (b) smoothed TV reconstruction, and (c) corresponding profiles along the yellow line.

with the assumption. The ramp filter is set as the initial filter, and the FDK reconstruction with hamming filter and cosine filter is regarded as the label to train the network. The hamming filter and cosine filter are implemented as:

$$Filter_{Hamming} = Filter_{Ramp} * [\alpha + (1 - \alpha) * \cos(\omega * cutoff)] \quad (10)$$

$$Filter_{Cosine} = Filter_{Ramp} * \cos\left(\frac{\omega}{2} * cutoff\right) \quad (11)$$

where α is a positive constant value less than 1.0 (usually at 0.54) and ω is the range of spatial frequency. In our study, the cutoff is set as a hyperparameter to extend the distribution of the filter curve, where 0.5 is set for the hamming filter and 2.0 for the cosine filter. During the preliminary study, we randomly select 20 abdomen projection-reconstruction datasets to train the FDK-based neural network. These data are divided into 15 for training, 3 for validation, and 2 for testing.

Given that the accuracy of the learned filter is verified in the preliminary study, with the purpose of reconstructing high-quality images based on the analytical method, we conduct a simulation study and re-train the neural network using the standard TV reconstruction to learn a filter directly from the reconstructed images. The ADMM-based TV reconstruction is performed on the 52 abdomen projections to get the reconstructed images. Then, we divide them into 40 for

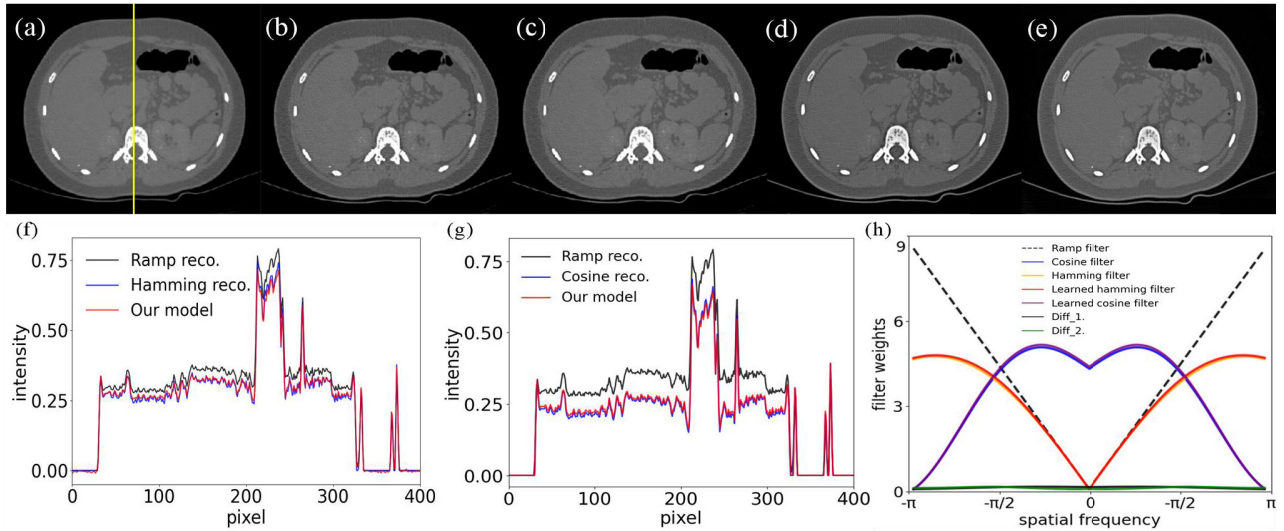


FIGURE 4. Results of the preliminary study reconstructed by (a) ramp filter, (b) hamming filter, (c) learned hamming filter, (d) cosine filter, and (e) learned cosine filter. (f) Profiles along the yellow line of (a), (b), and (c). (g) Profiles along the yellow line of (a), (d), and (e). (h) Comparison of filter shape in the frequency domain. Diff_1 and Diff_2 refer to the absolute difference between the learned cosine filter and the cosine filter, and the learned hamming filter and the hamming filter, respectively.

training, 8 for validation, and 4 for testing. In the generalization study, we replace the ramp filter in the conventional FDK algorithm with the learned filter and perform the generalization study on the 10 collected head and neck CT scans and a Catphan phantom to investigate the effectiveness of the learned filter. Finally, we re-train the neural network by a new group of label data, the smoothed TV reconstruction, to learn a new filter to explore the flexibility of the neural network.

E. IMPLEMENTATION DETAILS

When all the important components of the FDK-based neural network are in place, the loss function L , mathematically expressed as Eq. (12), is chosen to train the neural network.

$$L = (A \cdot e^{-(1)((x^2+y^2)/2\sigma^2)}) * (\frac{1}{N} \sum_{i \in I} |y'_i - y_i|) \quad (12)$$

where $*$ denotes the convolution operation. Essentially, the left-hand side of the $*$ is a Gaussian function where the items A and σ^2 are the pre-setting amplitude and variance, respectively, and the right-hand counterpart denotes the mean absolute error, in which y' and y are the prediction and reference, respectively. N refers to the total pixel numbers of the image. Additionally, the Adam optimizer with learning rate of 5×10^{-4} is utilized for updating the filter weights. By leveraging cross-validation and early-stopping strategies, the model is converged after 1085 epochs with a batch size of 1.

All experiments were implemented on the TensorFlow platform with Python programming language, and training was performed on a computer equipped with an NVIDIA

Titan Xp with 12 gigabytes of VRAM and an Intel Core i7-7700 @ 3.60 GHz with 55 gigabytes of RAM.

F. EVALUATION METRICE

Two common metrics, structural similarity (SSIM) [35] and peak signal-to-noise ratio (PSNR) [36], are employed for evaluation. The computation of SSIM is given by:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (13)$$

where μ_x and μ_y are the mean intensities of two images x and y , respectively. σ_x and σ_y separately denote the standard deviation value, σ_{xy} is the covariance of x and y , and c_1 and c_2 are two constants to maintain stability.

The PSNR between reference and testing images with the same size is defined by:

$$PSNR = 10 \times \log_{10}(\frac{f_{max}^2}{MSE}) \quad (14)$$

where MSE is the mean squared error between the two images and f_{max} denotes the possible maximum gray value of the reference image.

III. RESULTS

A. PRELIMINARY STUDY

Fig. 4 shows the results of the preliminary study. Fig. 4(a) is the FDK reconstruction with original ramp filter. Figs. 4(b) and (c) refer to the FDK reconstruction with hamming filter and corresponding learned filter, respectively. The comparison of profiles along the yellow line is shown in Fig. 4(f). Figs. 4(b) and (c) are consistent with each other, as well as the corresponding profile distribution. The same conclusion can be drawn from the network trained by the

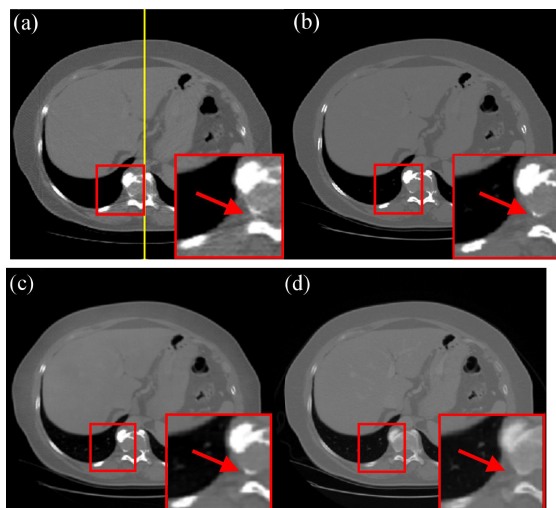


FIGURE 5. Abdominal CT images reconstructed by (a) conventional FDK, (b) standard TV reconstruction, and (c) FDK-based neural network. (d) Ground truth.

FDK reconstruction with cosine filter, which is shown in Figs. 4(a), (d), and (e).

The shapes of the five filters involved in the preliminary study are shown in Fig. 4(h). In terms of two different filter label cases, consistent filter shape and distribution can be learned by the FDK-based neural network after sufficient training. Moreover, the mean value of the absolute difference between the learned filter and corresponding label filter is around 0.001. The good consistency between the learned filter and the label filter in the preliminary study proves the accuracy of the FDK-based neural network and demonstrates the feasibility of applying the learned filter in the FDK algorithm.

B. PRELIMINARY STUDY

In the simulation study, the FDK-based neural network is re-trained by the high-quality reconstructed images of the standard TV reconstruction. The results of an abdominal image reconstructed by different methods are shown in Fig. 5. The corresponding slice of the CT dataset is regarded as the ground truth for comparison. To better visualize the performance of different methods, we magnify the region indicated by the red rectangle in Fig. 5. These enlarged parts are presented in the bottom right corner. As can be seen, the conventional FDK reconstruction is not ideal for the restoration of soft tissues and bone structures, suffering from visible artifacts. By contrast, the trained FDK-based neural network is consistent with the ADMM-based TV reconstruction, generating a noticeable improvement compared with conventional FDK reconstruction. The FDK-based neural network shows better consistency with the ground truth than the conventional FDK with respect to the profiles plotted in Fig. 6(a). Moreover, it is comparable to the ADMM-based TV. The shape of the learned filter is shown in Fig. 6(b).

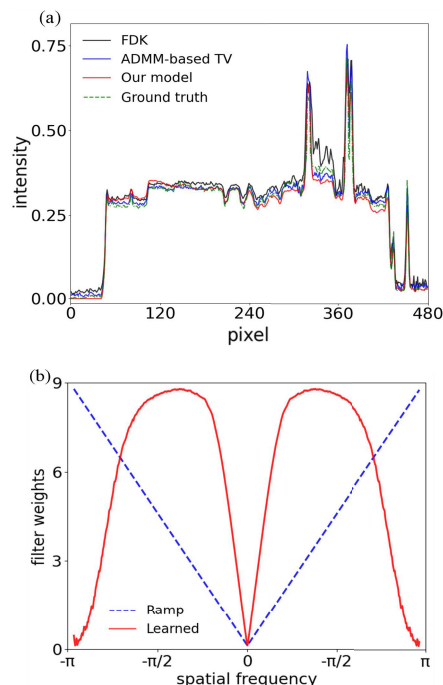


FIGURE 6. (a) Profiles along the yellow line of Fig. 5. (b) Shape of the learned filter in the frequency domain compared with the initial ramp filter.

TABLE 1. Comparison of PSNR, SSIM, STANDARD DEVIATION, and runtime in Fig. 5.

Method	Metrics	PSNR (dB)	SSIM	standard deviation	Runtime(s)
FDK		19.38	0.89	106.68	1.96
Proposed		32.29	0.94	69.84	2.25
ADMM-based TV		33.53	0.96	49.77	94.60

Table 1 lists the PSNR, SSIM, and runtime (refers to the time required for complete 3D reconstruction) of Fig. 5. The FDK-based neural network gains improvements by 67% for PSNR and 5% for SSIM compared with the conventional FDK. Besides, it is remarkably similar to the ADMM-based TV reconstruction. Additionally, the reconstruction speed of the FDK-based neural network is nearly 42 times higher than that of the ADMM-based TV reconstruction.

C. GENERALIZATION STUDY

In this study, we first replace the ramp filter of the conventional FDK algorithm with the learned filter and perform the generalization study on the collected head and neck CT data and a Catphan phantom to investigate the potential of the learned filter. Then, we re-train the neural network with the smoothed TV reconstruction and observe whether the learned smooth filter can be effectively applied to the FDK algorithm to reconstruct the image consistent with the new label, so as to explore the flexibility of the neural network.

TABLE 2. Quantitative assessments of different methods in the generalization study.

	SSIM			PSNR (dB)			standard deviation		
	FDK	FDK with learned filter	ADMM-based TV	FDK	FDK with learned filter	ADMM-based TV	FDK	FDK with learned filter	ADMM-based TV
Head	0.85	0.89	0.91	18.65	25.34	28.47	63.62	45.78	40.98
Neck	0.86	0.91	0.95	20.23	24.82	28.55	82.59	33.63	23.76
Catphan: line-pair slice	0.80	0.91	0.93	16.55	23.63	28.82	82.81	52.09	17.13
Catphan: circular spots slice	0.85	0.92	0.96	22.45	26.82	30.64	61.68	27.84	18.74

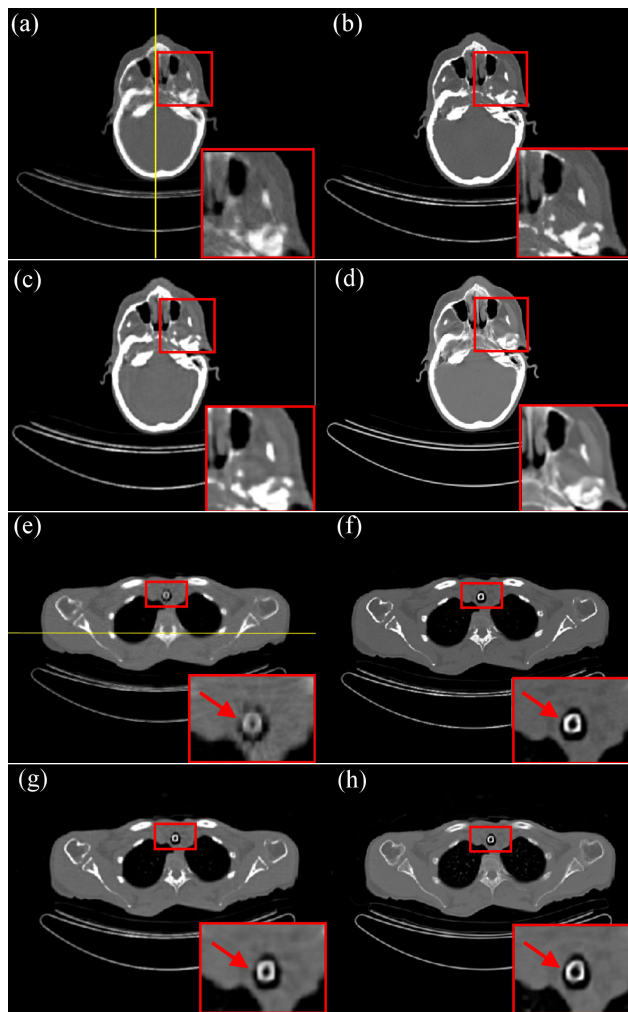


FIGURE 7. Generalization study on the Catphan phantom. (a), (e) Conventional FDK; (b), (f) standard TV reconstruction; (c), (g) FDK with learned filter; (d), (h) Ground truth.

The relevant reconstructions of the head and neck CT data are shown in Fig. 7. FDK with the learned filter shows good consistency with the ADMM-based TV and yields a smaller difference from the reference compared with the conventional

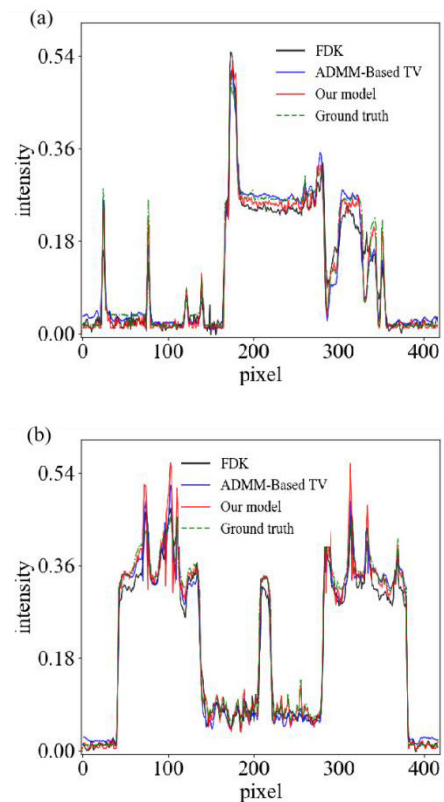


FIGURE 8. (a) and (b) are the profiles along the yellow line shown in the first and second rows of Fig. 9, respectively.

FDK reconstruction. The conventional FDK suffers from severe boundary ambiguity and distortion of the intubation at the patient’s trachea marked by the red arrow in Fig. 7(e). Better performance in recovering the shape and size of the intubation can be observed in Fig. 7(g). The FDK with the learned filter also reaches a higher fitting with the ground truth than conventional FDK reconstruction in terms of the corresponding profiles depicted in Figs. 8(a) and (b).

The results of the generalization study on the Catphan phantom are shown in Figs. 9 and 10. The first row in Fig. 9 is the comparison of the line-pair image reconstructed

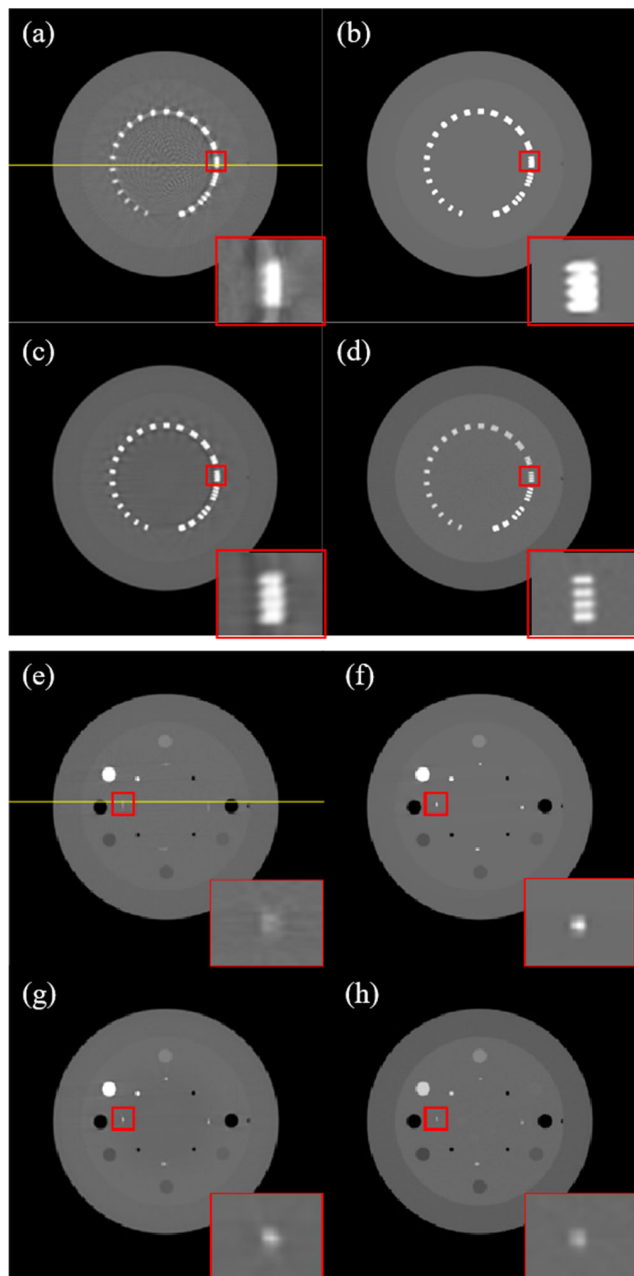


FIGURE 9. Generalization study on the head (the first row) and neck (the second row) CT data. (a), (e) Conventional FDK; (b), (f) standard TV reconstruction; (c), (g) FDK with learned filter; (d), (h) Ground truth.

by different methods. The conventional FDK suffers from severe noise and artifacts as shown in Fig. 9(a), and the structure of line-pairs is blurred. Although the FDK with learned filter is still slightly fuzzy, it can restore the structure basically and maintain good consistency with the standard TV reconstruction. A comparison of the circular spot slice is shown in Figs. 9(e)-(h). The conventional FDK cannot reconstruct these small spots correctly, and some distortion occurs. However, the FDK with the learned filter reconstructs almost all details and avoids the distortion of fine structure,

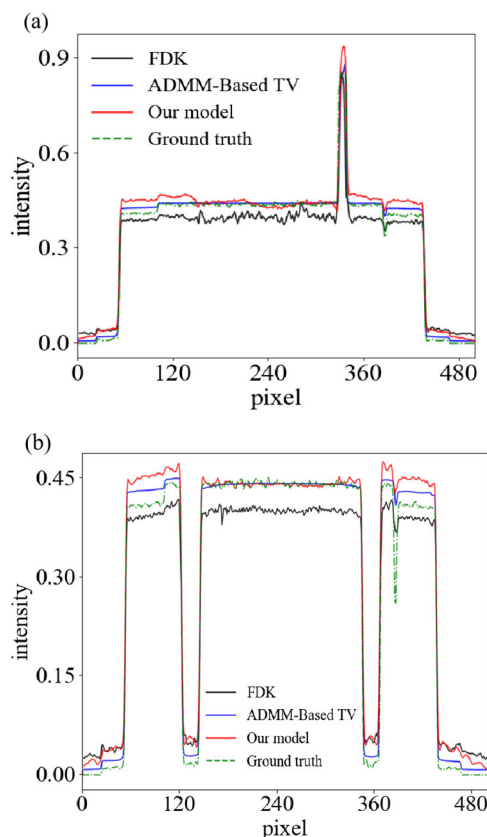


FIGURE 10. (a) and (b) are the profiles along the yellow line shown in the first and second rows of Fig. 7, respectively.

thus obtaining a better image quality and resolution than the conventional FDK reconstruction.

The evaluation metrics in the generalization study are listed in Table 2. FDK with learned filter produces better image quality than the conventional FDK, yielding smaller difference from the ground truth and eliminating more noise and artifacts. In the Catphan phantom study, SSIM and PSNR are improved by 13.75% and 42.78%, respectively. Additionally, the quantitative results of FDK with learned filter is much closer to the ADMM-based TV counterpart, which demonstrates the feasibility of the learned filter method.

For further observation, the penalty hyperparameter λ is reset at 10 to obtain the smoothed TV reconstruction to re-train the FDK-based neural network. The abdominal images reconstructed by different methods and corresponding comparison of profiles are shown in Figs. 11 and 12(a), respectively. Similar smoothing effect on the texture can be observed in the FDK with the learned smooth filter and ADMM-based TV. The corresponding profiles also show a flatter trend compared with the conventional FDK reconstruction and the ground truth. In addition, the FDK with the learned smooth filter demonstrates better performance on recovering the shape and boundary of the structure (magnified region in the red rectangle) than the conventional FDK. Fig. 12(b) shows the shape of the learned smooth

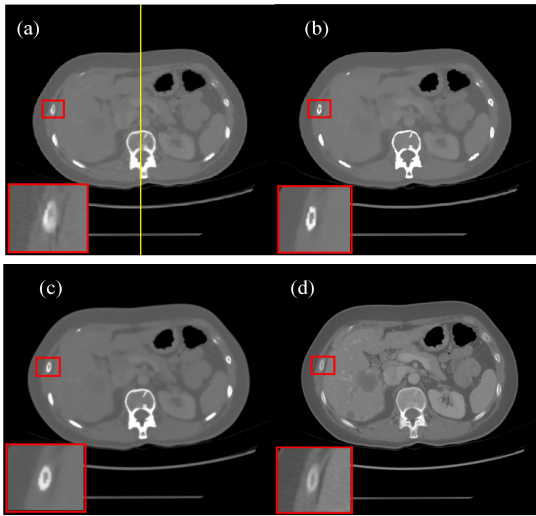


FIGURE 11. Reconstruction of abdominal image by different methods. (a) Conventional FDK, (b) smoothed TV reconstruction, (c) FDK with learned smooth filter, and (d) ground truth.

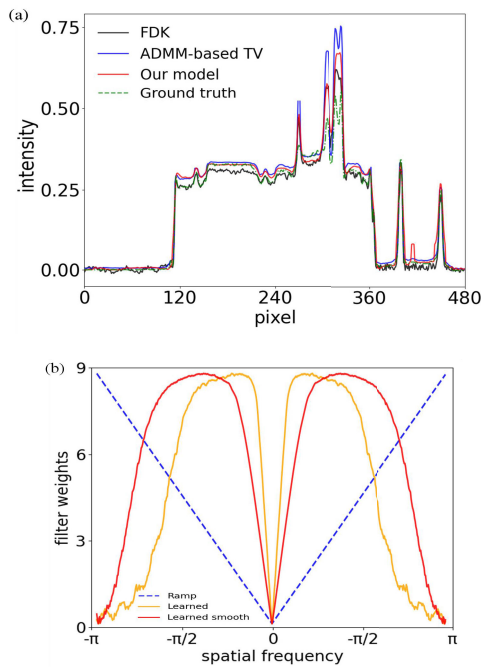


FIGURE 12. (a) Profiles along the yellow line in Fig. 11. (b) Shape of the learned smooth filter in the frequency domain compared with the initial ramp filter.

filter. Compared with the filter learned from the standard TV reconstruction, the learned smooth filter presents a narrower curve distribution.

Table 3 compares the SSIM and PSNR of images in Fig. 11. The smoothed TV reconstruction and the FDK with the learned smooth filter get higher PSNR and lower SSIM than conventional FDK. One reasonable explanation is that the ADMM-based TV with a large penalty parameter generates a striking difference from the ground truth, and the structural

TABLE 3. Compares the SSIM and PSNR of images in Fig. 11.

Method	PSNR (dB)	SSIM	standard deviation
FDK	20.02	0.87	69.23
Proposed	22.71	0.74	109.50
ADMM-based TV	24.63	0.77	122.52

details are less visible, which leads to a lower evaluation performance than the conventional FDK reconstruction in terms of SSIM. However, the noise level is reduced by the smoothing effect. The smoothed TV reconstruction and the FDK with the learned smooth filter still perform better than the conventional FDK on the PSNR metric.

IV. DISCUSSION

In this study, we build an FDK-based neural network by transforming the conventional FDK reconstruction algorithm into the matrix multiplier operators and embedding them into neural network layers. With the purpose to learn a filter directly from the high-quality iterative algorithm, filter within the FDK reconstruction is set as the only trainable parameter in the neural network. A preliminary study is conducted to verify the accuracy and feasibility of the learned filter by applying hamming filter and cosine filter to FDK reconstruction to train the FDK-based neural network. The preliminary study shows good fitting between the learned filter and the target filter, which is consistent with the comparison of image slice and profiles. Then, simulation and generalization studies are performed by training the FDK-based neural network utilizing the ADMM-based TV reconstruction. The FDK with learned filter is demonstrated to reconstruct higher-quality images than conventional FDK, suppressing the distribution of noise and artifacts effectively. Moreover, the conventional FDK is not ideal for line-pair and circular spots. Some distortion and deficiency can be observed in the reconstructed images, whereas the FDK with learned filter improves the accuracy of the reconstruction, advancing the spatial resolution and density resolution to some extent.

The strong advantages of the FDK-based neural network are its simplicity and interpretability, which are significant to clinical practice in medical CT imaging. This reconstruction framework can achieve better performance based on the DL strategies to train the reconstruction filter. Two representative related works, the FBPCNet [37] and CNN-based BPF [38], utilize a deep convolution neural network to post-process the images reconstructed by the FBP method or direct backprojection, achieving good results based on a previous reconstruction. The FDK-based neural network is of lower computational complexity and capable of learning a specific filter directly from the selected label data.

Two other applications of the FDK-based neural network are discussed as follows. First, it is meaningful to further

develop the potential of the reconstruction filter by applying the FDK-based neural network in sparse-view or limited-angle scenarios, which are not applicable for conventional FDK reconstruction. We also look forward to verifying its strong practicability by seeking more high-quality medical CT images available for network training. Second, as a heuristic attempt to replace IR method with an analytical algorithm, this study provides a new perspective to solve the time-consuming problem with respect to IR. From the generalization study shown in Section III, compared with the conventional FDK reconstruction, applying the learned filter in the FDK algorithm obtain better results indeed, but there are still relative differences compared with the corresponding ADMM-based TV reconstruction. Moreover, when we take the smoothed TV reconstruction as the label data to train the FDK-based neural network, the FDK reconstruction with the learned filter appears slightly fuzzy in addition to smoothness. This problem can be solved by introducing additional deep neural networks, such as the three-dimensional U-Net [39] and generative adversarial network [40], which gives credit to the strong scalability of the FDK-based neural network.

V. CONCLUSION

This study enriched the application of the neural network in CT image reconstruction. We put forward a new solution for optimizing the AR filter by directly learning a filter from the high-quality images, which can maintain the same level of computational complexity as the conventional FDK reconstruction, effectively improving the image quality of reconstruction.

REFERENCES

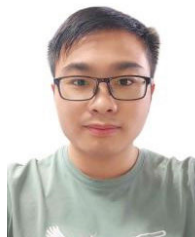
- [1] E. Y. Sidky, C. M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-Ray Sci. Technol.*, vol. 14, pp. 119–139, Jan. 2009.
- [2] G. H. Chen, T. Jie, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection sets," *Proc.-Soc. Photo-Opt. Instrum. Eng.*, vol. 6856, Jun. 2008, Art. no. 685618.
- [3] J. Liu, H. Ding, S. Molloi, X. Zhang, and H. Gao, "TICMR: Total image constrained material reconstruction via nonlocal total variation regularization for spectral CT," *IEEE Trans. Med. Imag.*, vol. 35, no. 12, pp. 2578–2586, Dec. 2016.
- [4] H. Gao, Y. Zhang, L. Ren, and F. Yin, "Principal component reconstruction (PCR) for cine CBCT with motion learning from 2D fluoroscopy," *Med. Phys.*, vol. 45, no. 1, pp. 167–177, Jan. 2018.
- [5] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 1, no. 6, p. 612, 1984.
- [6] J. Farhat, A. Alatoom, and S. Dhoui, "Quantitative evaluation of 3D reconstruction using filtered back-projection on XCAT phantom," in *Proc. Int. Conf. Commun., Signal Process., their Appl. (ICCSIPA)*, Mar. 2019, pp. 1–6.
- [7] K. Srinivasan, M. Mohammadi, and J. Shepherd, "Investigation of effect of reconstruction filters on cone-beam computed tomography image quality," *Australas. Phys. Eng. Sci. Med.*, vol. 37, no. 3, pp. 607–614, Sep. 2014.
- [8] M. J. Lagerwerf, D. M. Pelt, W. J. Palenstijn, and K. J. Batenburg, "A computationally efficient reconstruction algorithm for circular cone-beam computed tomography using shallow neural networks," *J. Imag.*, vol. 6, no. 12, p. 135, Dec. 2020.
- [9] Y. Houno, T. Hishikawa, K.-I. Gotoh, M. Naitoh, A. Mitani, T. Noguchi, E. Arijii, and Y. Kodera, "Optimizing the reconstruction filter in cone-beam CT to improve periodontal ligament space visualization: An in vitro study," *Imag. Sci. Dentistry*, vol. 47, no. 3, p. 199, 2017.
- [10] G. Wang, J. C. Ye, and B. De Man, "Deep learning for tomographic image reconstruction," *Nature Mach. Intell.*, vol. 2, no. 12, pp. 737–748, Dec. 2020.
- [11] L. Shi, B. Liu, H. Yu, C. Wei, L. Wei, L. Zeng, and G. Wang, "Review of CT image reconstruction open source toolkits," *J. X-Ray Sci. Technol.*, vol. 28, no. 4, pp. 619–639, Aug. 2020.
- [12] C. Arndt, F. Güttler, A. Heinrich, F. Bürckenmeyer, I. Diamantis, and U. Teichgräber, "Deep learning CT image reconstruction in clinical practice," *RöFo-Fortschritte Gebiet Röntgenstrahlen Bildgebenden Verfahren*, vol. 193, no. 3, pp. 252–261, Mar. 2021.
- [13] 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.
- [14] B. Zhou, S. K. Zhou, J. S. Duncan, and C. Liu, "Limited view tomographic reconstruction using a cascaded residual dense spatial-channel attention network with projection data fidelity layer," *IEEE Trans. Med. Imag.*, vol. 40, no. 7, pp. 1792–1804, Jul. 2021.
- [15] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1407–1417, Jun. 2018.
- [16] 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.
- [17] I. Guha, S. A. Nadeem, C. You, X. Zhang, S. M. Levy, G. Wang, J. C. Torner, and P. K. Saha, "Deep learning based high-resolution reconstruction of trabecular bone microstructures from low-resolution CT scans using GAN-CIRCLE," *Proc. SPIE, Int. Soc. Opt. Eng.*, vol. 11317, Feb. 2020, Art. no. 113170U, doi: 10.1117/12.2549318.
- [18] H. Xie, H. Shan, W. Cong, C. Liu, X. Zhang, S. Liu, R. Ning, and G. Wang, "Deep efficient end-to-end reconstruction (DEER) network for few-view breast CT image reconstruction," *IEEE Access*, vol. 8, pp. 196633–196646, 2020.
- [19] M. Arjovsky, S. Chintala, L. Xe, and O. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 214–223.
- [20] 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.
- [21] W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabhuti, and G. Wang, "DRONE: Dual-domain residual-based optimization network for sparse-view CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3002–3014, Nov. 2021.
- [22] W. Xia, Z. Lu, Y. Huang, Z. Shi, Y. Liu, H. Chen, Y. Chen, J. Zhou, and Y. Zhang, "MAGIC: Manifold and graph integrative convolutional network for low-dose CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3459–3472, Dec. 2021.
- [23] G. Chen, X. Hong, Q. Ding, Y. Zhang, H. Chen, S. Fu, Y. Zhao, X. Zhang, H. Ji, G. Wang, Q. Huang, and H. Gao, "AirNet: Fused analytical and iterative reconstruction with deep neural network regularization for sparse-data CT," *Med. Phys.*, vol. 47, no. 7, pp. 2916–2930, Jul. 2020.
- [24] H. Zhang, B. Liu, H. Yu, and B. Dong, "MetaInv-Net: Meta inversion network for sparse view CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 2, pp. 621–634, Feb. 2021.
- [25] T. Würfl, 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.
- [26] Y. Wang, T. Yang, and W. Huang, "Limited-angle computed tomography reconstruction using combined FDK-based neural network and U-Net," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Montreal, QC, Canada, Jul. 2020, pp. 1572–1575.
- [27] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.
- [28] D. L. Parker, "Optimal short scan convolution reconstruction for fan beam CT," *Med. Phys.*, vol. 9, no. 2, pp. 254–257, Mar. 1982.

- [29] R. Glowinski and T. Oden, *Numerical Methods for Nonlinear Variational Problems*. New York, NY, USA: Springer, 1985.
- [30] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Comput. Math. with Appl.*, vol. 2, no. 1, pp. 17–40, 1976.
- [31] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines Vinod Nair," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 807–814.
- [32] C. Syben, M. Michen, B. Stimpel, S. Seitz, S. Ploner, and A. K. Maier, "Technical note: PYRO-NN: Python reconstruction operators in neural networks," *Med. Phys.*, vol. 46, no. 11, pp. 5110–5115, Nov. 2019.
- [33] J. Leuschner, M. Schmidt, D. Otero Bague, and P. Maaß, "The LoDoPaB-CT dataset: A benchmark dataset for low-dose CT reconstruction methods," 2019, *arXiv:1910.01113*.
- [34] J. Luo, J. Yang, and Y. Wang, "Research on improvement of filtering process in filtered backprojection reconstruction algorithm," in *Proc. IEEE 2nd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Mar. 2017, pp. 539–543.
- [35] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [36] W. M. Morrow, R. B. Paranjape, R. M. Rangayyan, and J. E. L. Desautels, "Region-based contrast enhancement of mammograms," *IEEE Trans. Med. Imag.*, vol. 11, no. 3, pp. 392–406, Feb. 1992.
- [37] 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.
- [38] Y. Ge, Q. Zhang, Z. Hu, J. Chen, W. Shi, H. Zheng, and D. Liang, "Deconvolution-based backproject-filter (BPF) computed tomography image reconstruction method using deep learning technique," 2018, *arXiv:1807.01833*.
- [39] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense, volumetric segmentation from sparse annotation," in *Proc. 19th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Athens, Greece, 2016, pp. 424–432.
- [40] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014, *arXiv:1406.2661*.



HANWEI LI was born in Guangzhou, China, in July 1989. He received the master's degree in biomedical engineering from Southern Medical University, Guangzhou, Guangdong, in 2012.

He is currently an Engineer with the Department of Clinical Engineering, Nanfang Hospital, Southern Medical University, China. His research interests include medical image calibration, reconstruction, processing, and registration.



SHUANG HUANG was born in Zunyi, China, in August 1998. He received the B.E. and master's degrees in biomedical engineering from Southern Medical University, Guangzhou, Guangdong, China, in 2020 and 2023, respectively.

His major field of study was machine learning-based CBCT image reconstruction. He is currently a Junior Medical Physicist with the Department of Medical Physics, Heyou International Hospital, Foshan, China. His research interests include machine learning and deep learning-based medical image processing.



DEBIN HU was born in Ji'an, China, in May 1991. He received the master's degree in biomedical engineering from Southern Medical University, Guangzhou, Guangdong, China, in 2012.

He is currently an Engineer with the Department of Clinical Engineering, Nanfang Hospital, Southern Medical University. His research interests include medical image reconstruction, processing, and registration.



HONGLIANG QI was born in Shenyang, China, in October 1987. He received the B.E. degree in biomedical engineering from Southern Medical University, Guangzhou, Guangdong, China, in 2013, and the master's and Ph.D. degrees in biomedical engineering from Southern Medical University, Guangzhou, Guangdong, in 2013 and 2017, respectively.

His major field of study was CT image reconstruction artifacts correction. He is currently a Senior Engineer with the Department of Clinical Engineering, Nanfang Hospital, Southern Medical University. His research interests include reliability analysis of medical devices and medical imaging.



YUAN XU received the B.E. degree in electronic and information engineering and the master's and Ph.D. degrees in biomedical engineering from Southern Medical University, China. He is currently a Lecturer with the School of Biomedical Engineering, Southern Medical University. He has published a number of research articles in image guided radiation therapy. His research interests include Monte Carlo (MC) simulation, X-ray imaging, and image guided radiation therapy.



CHAO LONG was born in Zhanjiang, China, in September 1984. He received the master's degree in software engineering from Yunnan University, in 2014. He is currently pursuing the Doctor degree in information technology with La Consolacion University Philippines, Philippines.

He is currently an Experimentalist in computer with the School of Biomedical Engineering, Guangdong Medical University, China. His research interests include software development, image technology analysis, and medical image processing.



HONGWEN CHEN was born in Huaiyin, China, in November 1968. She received the master's degree in software engineering from Wuhan University, China.

Currently, she is a Senior Engineer and the Director of the Department of Clinical Engineering, Nanfang Hospital, Southern Medical University, China. Her research interests include full life cycle management and reliability analysis of medical equipment.

...