Residual Neural Network for Filter Kernel Design in Filtered Back-Projection Method of CT Image Reconstruction

Jintian Xu[1], Chengjin Sun[1], Yixing Huang[2], Xiaolin Huang[1]

[1]Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China
[2]Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg
xiaolinhuang@sjtu.edu.cn

Abstract. Filtered back-projection (FBP) has been widely applied for computed tomography (CT) image reconstruction as a fundamental algorithm. Most of the filter kernels used in FBP are designed by analytic methods. Recently, the precision learning-based ramp filter (PL-Ramp) has been proposed to formulate FBP to directly learn the reconstruction filter. However, it is difficult to introduce regularization terms in this method, which essentially provides a massive solution space. Therefore, in this paper, we propose a neural network based on residual learning for filter kernel design in FBP, named resFBP. With such a neural network, it is possible for us to limit the solution space by introducing various regularization terms or methods to achieve better reconstruction quality on the test set. The experiment results demonstrate that both quality and reconstruction error of the proposed method has great superiority over FBP and also outperforms PL-Ramp when projection data are polluted by Poisson noise or Gaussian noise.

1 Introduction

Computed tomography (CT) is a technology to obtain internal information of an unknown object and has been extensively used in many areas, such as medical imaging and electron microscopy in materials science. Filtered back-projection (FBP), one of the most popular methods for CT image reconstruction, requires a large number of noise-free projections to yield accurate reconstructions. It is applied widely for its low computational cost and ease of implementation. However, in practice, noisy projections and projection discretization can make the image reconstructed by FBP suffer from severe artifacts. To suppress artifacts, the filter kernel designs for specific cases have been proposed to improve the quality of reconstruction images.

The methods for filter kernel design can be divided into two categories: 1) analytic design, 2) data-driven learning. For analytic design, Ram-Lak filter [1] is mathematically optimized from the ramp filter for the discrete image and

projection. Shepp-Logan filter [2], imposing a different noise assumption, is one common filter for FBP. To derive new filters, Wei et al. [3] propose a filter design methodology in the real space, and apply this methodology to deduce new filters as well as classic filters. New filters have been demonstrated to produce equivalent image quality in comparison to classic filters.

For data-driven methods, a novel scheme is proposed by Shi et al. [4] to design the reconstruction filters to replace the ramp filter. The reconstruction filters are optimized to drive the reconstruction matrix approach δ-matrix as close as possible with constraints. Wang et al. [5] propose an end-to-end network FBP-Net, combining the FBP algorithm with a denoiser neural network, to directly reconstruct positron emission tomography (PET) images from sinograms. The frequency filter is adaptively learned in the FBP part. Syben et al. [6] propose the precision learning-based ramp filter (denoted as PL-Ramp) to optimize discrete filter kernels by back-propagation, which solves the problem of manually designing filters. Experiments have proved that the initialized ramp filter can automatically approximate the hand-crafted Ram-Lak filter by training, and it greatly reduces the burden of manual design and allows the filter to be trained to adapt to more noise conditions.

Although PL-Ramp has been proved to have the ability to learn a discrete optimal reconstruction filter from the ramp filter, its performance on training noisy data is not as good as expected according to our experiments. Moreover, PL-Ramp does not have any regularization term to avoid over-fitting. To solve the problems mentioned above, in this paper, we propose a residual neural network to learn the filter in FBP, named resFBP. In detail, resFBP includes a residual fully connected neural network and a residual convolution neural network. Experiments on Poisson noise and Gaussian noise data demonstrate that our proposed method can better suppress the artifacts and has lower reconstruction root-mean-squared error (RMSE) than PL-Ramp.

2 Materials and methods

2.1 resFBP

The architecture of the proposed resFBP is illustrated in Fig. 1. The whole procedure can be described as following: the projection data first undergo Fourier

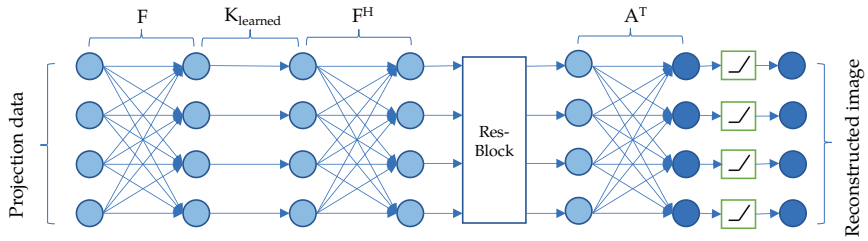


Fig. 1. The general architecture of resFBP.

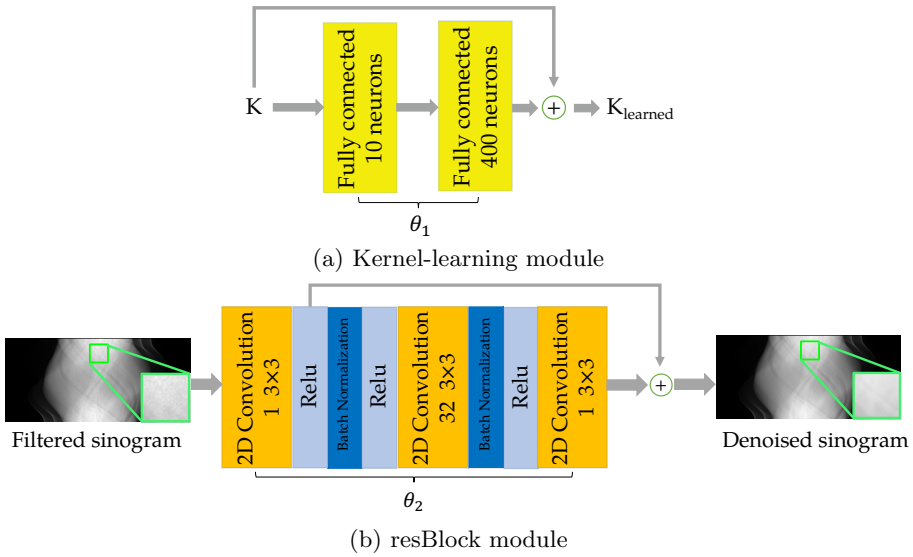
transform F, and then are filtered in the frequency domain with the learned filter K_{learned} before undergoing inverse Fourier transform F^H. Next, the result is refined through the resBlock, and finally is back-projected to get the reconstructed image with rectified linear unit (ReLU) activation. In Fig. 1, the light blue circle and the dark blue circle represent data in the projection domain and in the image domain, respectively.

Mathematically, the objective function of resFBP can be described as the following function:

$$f(\theta_1, \theta_2) = \frac{1}{2} \left\| \boldsymbol{A}^\top \boldsymbol{K_2}(\theta_2) \boldsymbol{F^H} \boldsymbol{K_1}(\theta_1) \boldsymbol{F} \boldsymbol{p} - \boldsymbol{x} \right\|_2^2 + R(\theta_1), \tag{1}$$

in which $\boldsymbol{p}$ denotes the projection data; $\boldsymbol{x}$ denotes the image to be reconstructed; $\boldsymbol{F}$ and $\boldsymbol{F^H}$ denote Fourier transform and inverse Fourier transform respectively; θ_1 denotes the weights and biases of all layers in the kernel-learning module; θ_2 denotes the kernel values of convolution layers and scale variables of batch normalization in the resBlock module; $\boldsymbol{K_1}(\theta_1)$ and $\boldsymbol{K_2}(\theta_2)$ denote effects of kernel-learning module and resBlock module on projection data; $f(\theta_1, \theta_2)$ denotes the loss function; $\boldsymbol{A^T}$ denotes the transpose of the projection operator $\boldsymbol{A}$; $R(\theta_1)$ denotes the l_2-norm regularization term of weights in the kernel-learning module. In the training stage, θ_1 and θ_2 are simultaneously updated by the corresponding gradients of the objective function.

The biggest difference from the existing work[6] lies in two modules, the kernel-learning module and the resBlock module, which are described in detail in the following. Fig. 2 (a) illustrates the kernel-learning module, in which K_{learned} is the summation of the Ramp filter kernel K and its residual part that goes through a two-layer fully connected neural network. The first layer has only



(a) Kernel-learning module

(b) resBlock module

Fig. 2. Two modules designed for filter kernel learning.

10 neurons and the second layer has 400 neurons, which equals the number of detector pixels. The width (from 10 neurons to 400 neurons) and depth (from 1 layer to 3 layers) of the fully connected network are optimized by minimizing reconstruction error on the validation set. This configuration achieves state-of-the-art performance while requiring the least training parameters. This residual module brings two advantages. On the one hand, it can be adapted to noise projection data while maintaining a similar performance as the Ramp filter in noise-free cases. On the other hand, it is possible to introduce a regularization term, such as a l_2-norm on weights, in training to avoid over-fitting.

The resBlock module is shown in Fig. 2 (b), where a typical ResNetV2 [7] structure is used. This structure has been proved to make information propagation smooth. The resBlock can further denoise the data in the projection domain, improving the quality of the reconstructed image. With batch normalization, it need not add any regularization term. The number (ranging from 1 to 3) of ResNetV2 and the number (ranging from 8 to 64) of convolution kernels of the resBlock module are optimized by grid search on the validation set.

2.2 Experimental setup

In this paper, the resFBP is implemented with PY-RONN [8], which is an open-source platform for applying deep learning to medical image reconstruction. Experiments are conducted on medical images, provided by American Association Physicists Medicine (AAPM, [9]). In the dataset, AAPM provides 5936 CT slice images from 10 patients for 4 kinds of scanning settings. The training set, validation set and test set contain 4795, 823, and 318 CT slice images respectively. In this paper, we use the CT slice images reconstructed by the full dose projection and 1.0 mm image pixel and reshape them to 256×256. The projection data are simulated from the AAPM volumes by parallel-beam scanning, which has 400 1.0 mm detector pixels collecting projection data from 180 uniformly distributed positions at a range of 180°.

The value of the pixel in the original image is a positive integer, which represents the attenuation rate in the Hounsfield unit (HU) with an offset of 1000 HU. In this paper, the value is normalized to [0,1] for training. To simulate noise in the process of reconstruction, Gaussian noise and Poisson noise are added to the projection data. The mean and variance of Gaussian noise are 0 and 0.5, while Poisson noise is simulated by assuming each detector pixel is exposed to 10^5 photons before attenuation. The RMSE between ground truths and reconstructed CT images is employed as the evaluation metric.

The parameters of the fully connected neural network in the kernel-learning module and convolution layer in the resBlock module are initialized with truncated normal distribution with small-variance and zero-mean. The loss function consists of a data fidelity term and a regularization term of weights in the kernel-learning module. The training batch size is 50 and the model is trained for 1000 epochs with the Adam optimizer using a 0.01 learning rate.

To validate the advantages of resFBP, we train PL-Ramp and resFBP using two kinds of training datasets, which are polluted by Poisson noise or Gaussian

Table 1. The reconstruction RMSE (HU) comparison of PL-Ramp and resFBP in Poisson noise and Gaussian noise under two kinds of training data.

Training dataset	Test on Poisson noise			Test on Gaussian noise		
	FBP	PL-Ramp	resFBP	FBP	PL-Ramp	resFBP
None	133.9±1153.2	-	-	90.3±0.689	-	-
Poisson noise	-	90.7±12.7	80.0±8.8	-	88.6±20.9	77.4±14.6
Gaussian noise	-	96.1±16.5	84.2±15.6	-	94.5±23.6	83.1±21.6

noise. Then, we compare the reconstruction RMSE of FBP and the trained models with projection data polluted by Poisson noise and Gaussian noise.

3 Results

Table 1 displays the reconstruction RMSE (mean±std) (HU) of FBP, RL-Ramp and resFBP under Poisson noise and Gaussian noise training dataset. It can be observed that regardless of the training dataset used, resFBP achieves lower RMSE mean and RMSE standard deviation than PL-Ramp and has great superiority over FBP under the same condition. For those two filter learning methods, the models trained with Poisson noise have better performance than those trained with Gaussian noise when tested on Gaussian noise data.

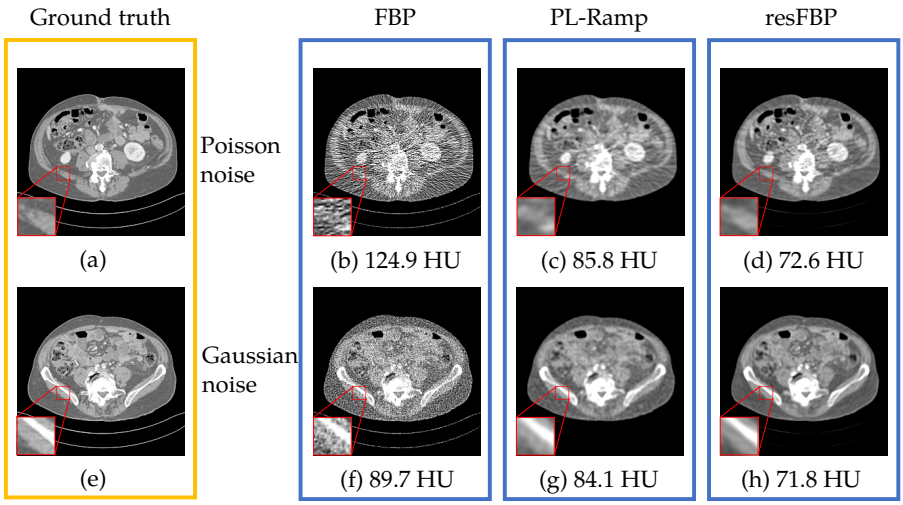


Fig. 3. Reconstruction results of two sample slices from the test patient, in the window of [-300,300] HU. The RMSE value of each reconstruction result is displayed. (a)(e) are ground truths. (b)(c)(d) are tested with Poisson noise data, and (f)(g)(h) are tested with Gaussian noise data. (a), (b), (e) and (f) use FBP with Ram-Lak filter. (c) and (g) use PL-Ramp. (d) and (h) use resFBP.

The reconstruction results of two sample slices from the test patient are visualized in the window of [-300,300] HU in Fig. 3. The images are reconstructed from Poisson noise data (the first row) and Gaussian noise data (the second row). Ground truths from original CT slice images are shown in Fig. 3(a) and (e) for reference. Fig. 3(b) and (f) show the results of FBP, which produced severe artifacts due to the high-pass filter designed in FBP. Fig. 3(c) and (g) show the results of PL-ramp, while Fig. 3(d) and (h) show the results of resFBP. The images reconstructed by PL-Ramp are much better than that by FBP due to its capability of learning from data, however, resFBP reconstructs images with sharper boundaries and lower reconstruction error. Therefore, resFBP is less sensitive to noise.

4 Discussion

With the help of the residual neural network, resFBP can achieve lower reconstruction error with sharper boundaries compared to PL-Ramp on Poisson noise data and Gaussian noise data. Due to regularization terms, resFBP is more robust than PL-Ramp and less likely to suffer from over-fitting. In consideration of the complexity of designing, this method is better than the analytical filter kernel design method. Meanwhile, it can still provide a good filter by learning in a noisy situation, where it is difficult to design analytic filters. In addition, the reconstruction by resFBP can serve as a better prior for deep learning-based post-processing methods. Robustness against noise on real datasets and the contribution of different modules remain to be further studied.

References

1. Ramachandran, G, N, et al. Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms. Proc Natl Acad Sci USA. 1971;68(9):2236–2240.
2. Shepp LA, Logan BF. The Fourier reconstruction of a head section. IEEE Trans Nucl Sci. 1974;NS21(3):21–43.
3. Wei Y, Bonse U, Wang G, et al. CT reconstruction filter design in the real space. Proc SPIE. 2004;5535:628–635.
4. Shi H, Luo S. A novel scheme to design the filter for CT reconstruction using FBP algorithm. Biomed Eng Online. 2013;12(1):50–50.
5. Wang B, Liu H. FBP-Net for direct reconstruction of dynamic PET images. Phys Med Biol. 2020;.
6. Syben C, Stimpel B, Breininger K, et al. Precision learning: reconstruction filter kernel discretization. arXiv preprint arXiv:171006287. 2017;.
7. He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks. European Conference on Computer Vision. 2016; p. 630–645.
8. Syben C, Michen M, Stimpel B, et al. Technical note: PYRONN: python reconstruction operators in neural networks. Med Phys. 2019;46(11):5110–5115.
9. Mccollough CH, Bartley AC, Carter RE, et al. Low-dose CT for the detection and classification of metastatic liver lesions: results of the 2016 low dose CT grand challenge. Med Phys. 2017;44(10).