RESEARCH ARTICLE

# Low-dose CT reconstruction using dataset-free learning

**Feng Wang** ⓘ *, Renfang Wang, Hong Qiu

College of Big Data and Software Engineering, Zhejiang Wanli University, Ningbo, Zhejiang, China

* wangf_721@zju.edu.cn

## Abstract

Low-Dose computer tomography (LDCT) is an ideal alternative to reduce radiation risk in clinical applications. Although supervised-deep-learning-based reconstruction methods have demonstrated superior performance compared to conventional model-driven reconstruction algorithms, they require collecting massive pairs of low-dose and norm-dose CT images for neural network training, which limits their practical application in LDCT imaging. In this paper, we propose an unsupervised and training data-free learning reconstruction method for LDCT imaging that avoids the requirement for training data. The proposed method is a post-processing technique that aims to enhance the initial low-quality reconstruction results, and it reconstructs the high-quality images by neural work training that minimizes the $\ell_1$-norm distance between the CT measurements and their corresponding simulated sinogram data, as well as the total variation (TV) value of the reconstructed image. Moreover, the proposed method does not require to set the weights for both the data fidelity term and the plenty term. Experimental results on the AAPM challenge data and LoDoPab-CT data demonstrate that the proposed method is able to effectively suppress the noise and preserve the tiny structures. Also, these results demonstrate the rapid convergence and low computational cost of the proposed method. The source code is available at https://github.com/linfengyu77/IRLDCT.

## Introduction

X-ray computed tomography (CT) is an essential imaging modality for clinical purposes, as it provides high-resolution images of the internal structure of the human body. However, X-ray radiation is known to be harmful to healthy tissues. In some major clinical tasks, a single CT scan can expose patients to radiation doses of up to 43 mSv [1], which may increase the risk of cancer. Consequently, reducing radiation dose while obtaining high-resolution images has become a significant area of research in CT scanning.

Currently, there are two primary strategies for reducing CT radiation dose: (1) decreasing the number of projection views and (2) lowering the X-ray tube current. This approach is commonly referred to as LDCT. LDCT algorithms can be broadly categorized into four groups: sinogram domain filtering, iterative reconstruction, and deep learning-based reconstruction.

Sinogram domain filtering methods exploit the distinct distributions of desired signals and noise in the sinogram domain to reconstruct CT images. This technique involves filtering out

components corresponding to artifacts or noise in the sinogram domain and then inverting the filtered sinogram data into the image domain using analytic algorithms. Numerous analytic filtering methods have been proposed based on the distribution of noise. For instance, filtered back projection (FBP) is a classical reconstruction method for CT images that performs high-pass filtering in the sinogram domain before back-projection. Sinogram domain filtering can produce high-quality CT images when the noise distribution is accurately characterized. However, determining this distribution can be challenging, particularly since artifacts or noise often correlate with image structures.

Compared with sinogram domain filtering methods, iterative reconstruction approaches are more flexible and stable. Iterative reconstruction approaches can be further divided into hybrid iterative reconstruction methods and model-based iterative reconstruction methods. Hybrid iterative reconstruction method produces an image by adjusting the statistical characters of the sinogram domain and the image domain. Model-based iterative reconstruction method utilizes the process of alternative performing the forward-projection (i.e., sinogram data generation) and back-projection (i.e., CT image reconstruction) to achieve iterative filtering in the sinogram domain and the image domain. Furthermore, the cost function of model-based iterative reconstruction method usually consists of a fidelity term with the noise model in the sinogram domain and a regularization term with the prior model in the image domain. The regularization term plays a vital role in reconstruction, and many regularizations have been proposed, such as total variation (TV) [2, 3], low-rank [4], non-local means (NLM) [5, 6], and dictionary learning [7]. The model-based iterative reconstruction method usually has better performance than hybrid iterative reconstruction method, but it is also computationally expensive. Additionally, model-based iterative reconstruction method requires manually designing the proper regularization and choosing the weight to obtain satisfactory reconstruction results.

In recent years, deep learning techniques have been widely employed in LDCT reconstruction, and they have demonstrated better performance than conventional LDCT reconstruction methods. Deep learning-based LDCT reconstruction methods can be categorized into four groups: sinogram domain processing (SDP), image domain processing (IDP), dual-domain processing (DDP), sinogram-image direct mapping (SIDP), and model-based deep learning (MBDL).

The SDP reconstruction algorithm aims to use a pre-trained neural network to inpaint the LDCT measurements into sinogram data that is very close to normal-dose CT (NDCT) measurements. For instance, [8] proposed a sinogram domain denoising approach using a convolutional neural network (CNN) with a filter loss function. Compared with image domain denoising methods, these approaches can easily estimate the noise level in the projection. Reference [9] proposed a sinogram data interpolation method by leveraging a conditional adversarial network (GAN). Although sinogram domain processing can correct errors in the sinogram domain, errors produced by the shortcomings of conventional methods can still negatively affect the final reconstructions.

In contrast to the SDP algorithm, IDP produces high-quality CT images by using a neural network to denoise the initial reconstructed images with artifacts. Most deep learning methods employ IDP to improve the quality of reconstructed images obtained using existing methods such as FBP [10, 11]. Reference [12] introduced a collaborative technique to train multiple Noise2Noise [13] generators simultaneously and learn the image representation from LDCT images. Reference [14] proposed Noise2Self that does not require any additional clean or noisy data. IDP is more straightforward compared to the SDP algorithm. Reference [15] proposed to a framework for sparse-view tomographic image reconstruction combining an early-stopped rapid iterative solver with a subsequent pre-trained neural network to complete the missing

iterations of rapid iterative solver. One main disadvantage of IDP is that it is difficult to recover information lost from the initial reconstructed images, which serve as inputs to the neural network.

DDP is a method that combines SDP and IDP. It leverages the advantages of both SDP and IDP to achieve higher-quality images compared to single-domain processing reconstruction methods. Reference [16] combined a deep convolutional neural network (CNN) with directional wavelet transform to extract the directional component of artifacts in low-dose CT images and exploit intra- and inter-band correlations. Reference [17] proposed a deep learning-based function optimization method for LDCT imaging, which incorporated the Radon inverse operator and disentangled each slice. To address of the limitation of acquiring independent noisy reference image of Noise2Noise [13, 18] proposed a method to generate both training inputs and training labels from the existing CT scans for count-domain and image-domain, which does not require any additional high-dose CT images or repeated scans. Although DDP can achieve good inversion results, it requires a larger training dataset due to its two training procedures: sinogram domain and image domain.

SIDP is an end-to-end reconstruction algorithm that directly transforms sinogram data into CT images. This method has the lowest complexity as it only requires training a neural network without extra processing such as sinogram data correction and inversion. For example, [19] presented a unified framework for image reconstruction called Automated Transform by Manifold Approximation (AUTOMAP), which directly converts sinogram data into CT images. Reference [20] proposed a direct reconstruction framework exclusively using deep learning architectures, which consists of denoising, reconstruction, and super resolution (SR). SIDP is a highly efficient reconstruction method but demands massive memory as the entire sinogram data needs to be fed into the neural network.

MBDL, also known as optimization unrolling scheme or plug-and-play, is an effective approach that replaces the parameters or regularization of conventional iterative schemes with learnable/pre-trained neural networks. Reference [21] unrolled the proximal gradient descent algorithm for iterative image reconstruction to finite iterations and replaced terms related to the penalty function with trainable CNN to reduce memory requirements and training time. Reference [22] incorporated the benefits from analytical reconstruction methods, iterative reconstruction methods, and DNNs. They unrolled proximal forward-backward splitting into iterative reconstruction updates of CT data fidelity and DNN regularization with residual learning. Reference [23] developed a unified reconstruction framework combining supervised and unsupervised learning, and physics and statistical models to enhance the accuracy and resolution of LDCT reconstruction images. By leveraging the advantages of deep learning and conventional methods, MBDL offers better interpretability than data-driven deep learning.

Recently, training dataset-free method have drawn much attention in LDCT imaging, which does not need to pre-train a neural network and works on a single image by utilizing the consistency between the CT measurements and sinogram data modeled on the reconstructed image. For instance, the deep image prior (DIP) [24], originally proposed for natural image denoising by using early stopping to fit the noisy image, has been widely exploited in medical imaging [25–27]. Also, DIP treats noise as i.i.d random noise rather than artifacts correlated to the entries of CT images. Reference [28] proposed an dataset-free reconstruction method based on Bayesian inference, which takes the $\mathcal{J}$-invariant transform of the FBP reconstructed image as the initial value. This method can reconstruct high-quality images from measurements; however, its reconstruction time is significantly higher than that of its competitors.

In this paper, we propose an iterative LDCT reconstruction method that ultizes neural network to improve the CT images reconstructed by FBP method without training data. During the iterative LDCT reconstruction, we minimize the loss, which consists of two components:

the $\ell_1$-norm distance between the CT measurements and the sinogram data modeled on the post-processed image, and the TV value of the post-processed image. We achieve this by training a neural network. The proposed method does not require collecting any training data and balancing the contribution of data fidelity and TV regularization in the loss. Once the network training is complete, the high-quality reconstructed results will be output immediately.

The rest of the paper is organized as follows: Methodology section describes how to build and solve the optimization problem. Experimental Results section presents the experimental setup and results using the 2016 Low-dose CT Grand Challenge data and the LoDoPaB-CT data [29]. Discussion and Conclusion section is the discussion and conclusion.

## Methodology

In this section, we introduce a proposed method for reconstructing LDCT from noisy measurements. This method utilizes a DNN to enhance the CT image reconstructed by the FBP method, without the need for training data.

### Problem setup

The forward formulation of LDCT can be formulated as

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

where $y$ represents the CT measurements, $A$ is the projection matrix of CT imaging, $\epsilon$ denotes the background contributions of scatter and electrical noise, and $x$ represents the ground-truth CT image. Typically, we can solve the inverse problem of Eq 2 by using the FBP method $\mathcal{F}$,

$$x_f = \mathcal{F}(y), \tag{2}$$

where $x_f$ is the reconstructed image by FBP. However, due to the low source intensity of X-ray and/or the random noise, the quality of $x_{fbp}$ is unsatisfactory, often suffering from noticeable streaky artifacts, random patterns, and low resolution.

Considering a DNN $\mathcal{NN}$ with parameters $\theta$ that can enhance the image's quality by $\mathcal{NN}(x_f; \theta)$, which means that we can re-formulate Eq 1 as

$$y = A\mathcal{NN}(x_f; \theta) + \epsilon, \tag{3}$$

According to Bayes's rule, we can obtain the posterior density of $\mathcal{NN}(x_f; \theta)$ by

$$p(\mathcal{NN}(x_f; \theta) \mid y) \propto p(y \mid \mathcal{NN}(x_f; \theta))p(y). \tag{4}$$

Supposing $p(y \mid \mathcal{NN}(x_f; \theta))$ as Gaussian distribution,

$$p(y \mid \mathcal{NN}(x_f; \theta)) = \mathcal{N}(A\mathcal{NN}(x_f; \theta), \Sigma_\epsilon). \tag{5}$$

where $\Sigma_\epsilon$ represents the covariance of the noise. Furthermore, taking the logarithm on both sides of Eq 5 then we obtain

$$\ln p(y \mid \mathcal{NN}(x_f; \theta)) \propto \frac{1}{2}(y - A\mathcal{NN}(x_f; \theta))^{\mathrm{T}}\Sigma_\epsilon^{-1}(y - A\mathcal{NN}(x_f; \theta)). \tag{6}$$

Taking the logarithm on both sides of Eq 4 and substituting $\ln p(\mathbf{y} \mid \mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}))$ with Eq 6, we obtain

$$
\begin{aligned}
\ln p(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}) \mid \mathbf{y}) \quad &\propto \ln p(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}) \mid \mathbf{y}) p(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})) \\
&\propto \frac{1}{2} (\mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}))^{\mathrm{T}} \boldsymbol{\Sigma}_\epsilon^{-1} \left( \mathbf{y} - \mathbf{A}\mathcal{NN}\left(\mathbf{x}_f; \boldsymbol{\theta}\right) \right) + \ln p(\mathbf{y}).
\end{aligned}
\tag{7}
$$

Therefore, we can obtain the maximum a posterior (MAP) objective,

$$
\begin{aligned}
\max\{\ln p(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}) \mid \mathbf{y})\} \quad &= \min\{-\ln p(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}) \mid \mathbf{y})\} \\
&\propto \min \left\{ \frac{1}{2} (\mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}))^{\mathrm{T}} \boldsymbol{\Sigma}_\epsilon^{-1} (\mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})) \right\}.
\end{aligned}
\tag{8}
$$

Assuming the noise is Gaussian independent and identically distributed (iid), i.e., $\boldsymbol{\Sigma}_\epsilon = \sigma_\epsilon^2 \mathbf{I}$. Furthermore, considering to regularize $\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})$ with $\mathcal{R}(\cdot)$, Eq 8 can be further rewritten as

$$
\begin{aligned}
\mathbf{x}^* = \arg \min_{\boldsymbol{\theta}} \left\{ \frac{1}{2\sigma_\epsilon^2} (\mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}))(\mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})) \right\} \\
\text{subject to} \quad \eta \mathcal{R}(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})),
\end{aligned}
\tag{9}
$$

where $\eta$ denotes the weight. To linearize this problem, we reformulate Eq 9 as

$$
\mathbf{x}^* = \arg \min_{\boldsymbol{\theta}} \parallel \mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}) \parallel_2^2 + \eta \mathcal{R}(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})).
\tag{10}
$$

In fact, the artifacts in LDCT images are highly correlated to the entries of CT images rather than random noise, and the results inverted through $\ell_2$-norm loss tend to be over-smoothed, which is not beneficial for preserving the tiny structures and/or sharp edges. Hence, we propose to optimize $\boldsymbol{\theta}$ by minimizing the $\ell_1$-norm misfit,

$$
\mathbf{x}^* = \arg \min_{\boldsymbol{\theta}} \parallel \mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}) \parallel_1 + \eta \mathcal{R}(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})).
\tag{11}
$$

Furthermore, we add the TV term of the reconstructed CT image $\hat{\mathbf{x}} = \mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})$ into Eq 11 as a smooth penalty to overcome the potential over-fitting induced by the noise in the CT measurements. Eq 11 thus becomes

$$
\begin{aligned}
\mathbf{x}^* = \arg \min_{\boldsymbol{\theta}} \{ \parallel \mathbf{y} - \mathbf{A}\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}) \parallel_1 + \nabla(\mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta})) \}, \\
\nabla(\hat{\mathbf{x}}) = \frac{1}{NM} \Big( \sum_{i=1}^{N} \sum_{j=1}^{M-1} \parallel \hat{\mathbf{x}}_{i,j+1} - \hat{\mathbf{x}}_{i,j} \parallel_1 + \sum_{i=1}^{N-1} \sum_{j=1}^{M} \parallel \hat{\mathbf{x}}_{i+1,j} - \hat{\mathbf{x}}_{i,j} \parallel_1 \Big),
\end{aligned}
\tag{12}
$$

where $\hat{\mathbf{x}} \in \mathrm{R}^{N \times M}$. Eq 12 can be solved by NN training, and we can derive $\mathbf{x}^*$ with the forward propagation of $\mathcal{NN}$ once $\boldsymbol{\theta}$ be optimized by

$$
\mathbf{x}^* = \mathcal{NN}(\mathbf{x}_f; \boldsymbol{\theta}^*).
\tag{13}
$$

## Solving the MAP

The proposed method can be considered as a kind of NN training-based reconstruction method, which optimizes the NN's parameters by minimizes the loss from both sinogram

domain and image domain. The proposed LDCT reconstruction method can be divided into two steps: (1) Reconstructing the initial CT image: The initial CT image is reconstructed using the the FBP method. Although this initial CT image may contain many artifacts due to the low intensity of X-ray, FBP provides fundamental information about the internal structure of the human body, which is helpful for enhancing the reliability of the inversion result by NN. Moreover, FBP performs much faster than iterative reconstruction approaches such as compressive sensing; (2) Post-processing the initial reconstruction result: Once the initial CT image is achieved, it will be fed into a pre-defined NN and will be improved through the NN training. To achieve $\boldsymbol{\theta}$, we establish the loss function for NN training based on Eq 12,

$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{x}_f) = \frac{1}{NM} \parallel \boldsymbol{y} - A\mathcal{NN}(\boldsymbol{x}_f; \boldsymbol{\theta}) \parallel_1 + \nabla(\mathcal{NN}(\boldsymbol{x}_f; \boldsymbol{\theta})), \qquad (14)$$

and we use gradient descent-based optimization algorithms such as stochastic gradient descent to optimize $\boldsymbol{\theta}$ to minimize the loss function. It is worth noting that the proposed method does not require setting weights for both the data fidelity term and the regularization term, which significantly reduces the difficulty of manually setting the weights. In summary, the visual flowchart of the iterative reconstruction for LDCT is shown in Fig 1. Algorithm 1 further explains the construction algorithm in detail.

**Algorithm 1** Iterative reconstruction for LDCT imaging

```
Require: number of iterations: n; CT measurements: y; FPB operator: F;
projection matrix: A; learning rate: λ
Ensure: optimal θ
  initial i = 1, initial CT image x₀ = F(y)
  While i < = n
    reconstruct the CT image NN(x₀; θ)
    compute the loss L(y) = 1/NM ‖ y - ANN(x₀; θ) ‖₁ + ∇(NN(x₀; θ))
    update θ* ← θ - λ ∂L/∂θ
    i ← i + 1
  return x* = NN(x₀; θ*)
```

## NN architecture

To enhance the quality of the initial CT image, we have designed a DNN with a straightforward structure. As depicted in Fig 1, the network primarily consists of 2-D convolution, batch normalization, and LeakyReLU layers. The first layer is a convolution layer, followed by a LeakyReLU layer and a block composed of convolution, batch normalization (BN), and LeakyReLU. The convolution layers are employed for feature extraction, the BN layers for enhancing the stability of network training, and the LeakyReLU layers to ensure non-linearity throughout the network. LeakyReLU is defined as,

$$\text{LeakyReLU}(x) = \max(0, x) + \phi * \min(0, x). \qquad (15)$$

In the subsequent experimental test, we set the value of $\phi$ to 0.01.

## Experimental results

In this section, we evaluate the performance of the proposed method by comparing it with four representative methods: FBP, TV (post-processing and unsupervised method), DIP (unsupervised and data-free method), and RED-CNN (post-processing and supervised model).
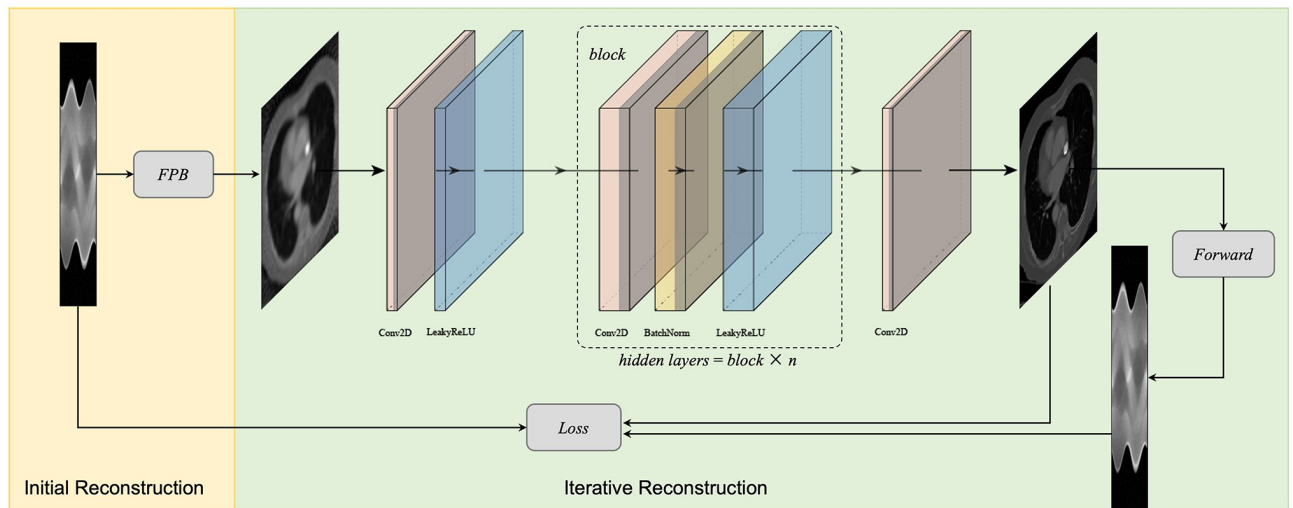
**Fig 1. Schematic diagram of the proposed method.**

## Parameter setting

The FBP, TV and DIP reconstruction are implemented by using Deep Inversion Validation Library (https://github.com/jleuschn/dival) and Operator Discretization Library (https://github.com/odlgroup/odl). For TV and DIP reconstruction, we use the parameters recommend by [30] (https://github.com/oterobaguer/dip-ct-benchmark). The NN training related tasks are all implemented on the PyTorch platform [31].

For TV reconstruction, the weight for $\ell_1$-norm term is set to $2.15 \times 10^{-7}$, and the number of iterations is set to 200, and we utilize the Douglas-Rachford Primal-Dual method as the solver. In addition, the initial reconstruction results for the TV method are obtained by the FBP method, and the parameters for the FBP method are the same as those for the proposed method, which means that both the TV method and the proposed method have the same input for the neural network. For DIP reconstruction, we use a learning rate of 0.0005, 6 scales, 1000 iterations for AAPM challenge data and 2000 iterations for LoDoPaB-CT data, and 128 channels for the U-Net at every scale. We adopt mean square error (MSE) as the loss function for both TV and DIP reconstruction. In the proposed method, we set the iterations to 2000, and save the result with the highest peak signal-to-noise ratio (PSNR). For all reconstruction methods, the filter and frequency scaling of FBP reconstruction are set to Hann and 0.8, respectively.

For RED-CNN training, we use the AAPM Challenge Data as the training dataset. We train the RED-CNN using full-dose CT scans from nine patients, reserving one patient (L067) for evaluation. In the training data generation process, we use a patch size of 64. The batch size for RED-CNN training is set to 32, the number of training epochs is 100, the loss function is MSE loss, and we use Adam optimizer with a learning rate of $10^{-5}$. We train three models for different low-dose levels by using pairs of FBP reconstructions of low-dose simulations and corresponding full-dose CT images.

There are 30 convolution layers in our NN, the size of filter kernels of the first convolution layer is $64 \times 1 \times 3 \times 3$, where the format is number of filters×number of channels × width × height. From the second to the penultimate convolution layer, we set the size of all filter kernels to $64 \times 64 \times 3 \times 3$. For the last convolution layer, the size of filter kernels is set to $1 \times 64 \times$

$3 \times 3$. We minimize the loss defined by Eq 8 by using the AdamW method with learning rate of $10^{-3}$.

## Data specification

To evaluate the effectiveness of the proposed method, we test its performance on on two data-sets: AAPM challenge data and LoDoPaB-CT data [29]. The AAPM challenge data consists of reconstructed simulated data from human abdomen CT scans provided by Mayo Clinic for the AAPM Low Dose CT Grand Challenge (https://www.aapm.org/GrandChallenge/LowDoseCT/). We use 1-mm slice thickness reconstructions with dimensions of 512 px × 512 px for RED-CNN training and performance comparison. The CT images form LoDoPaB-CT data are sampled from AAPM challenge data and have been cropped to dimensions of 362 px × 362 px. Additionally, these images have been subjected to dequantization noise uniformly distributed in [0, 1] for each pixel.

For sinogram data simulation, we construct a 2-D fan-beam geometry with 1000 angles, 1000 pixels, source to axis distance 500 mm, and axis to detector distance 500 mm [32]. The LDCT image are simulated by adding Poisson noise with $I_i = [1\,e\,3, 1\,e\,4, 5\,e\,4]$ following the Poisson distribution according to the process of photon generation, attenuation, and detection, which can be expressed as,

$$\boldsymbol{y}_i \sim \mathrm{Poisson}\left\{ I_i e^{-[\boldsymbol{Ax}]_i} + \boldsymbol{\sigma}_i \right\}, i = 1, \ldots, m, \tag{16}$$

where $I_i$ denotes the source intensity of the $i$-th X-ray, $\boldsymbol{y}_i$ represents the CT measurements produced by the $i$-th X-ray, $\boldsymbol{A}$ is the projection matrix of CT imaging, $\boldsymbol{\sigma}_i$ denotes the background contributions of scatter and electrical noise, and $\boldsymbol{x}$ represents the full-dose CT image. Additionally, the full-dose CT images $x$ are normalized before sinogram simulation by

$$x = \frac{x - \min(x)}{\max(x) - \min(x)}. \tag{17}$$

## Quantitative indices

We adopt two quantitative indices, PSNR and structural similarity index (SSIM), to quantify the quality of the reconstructed CT images. The PSNR expresses the ratio between the maximum possible power of a signal and the power of corrupting noise, which is measured by the mean squared error (MSE),

$$\mathrm{PSNR}(\tilde{x}, x) = 10 \log_{10}\left( \frac{\max(x)^2}{\mathrm{MSE}(\tilde{x}, x)} \right),$$

$$\mathrm{MSE}(\tilde{x}, x) = \frac{1}{n} \sum_{i=1}^{n} \| \tilde{x}_i - x_i \|^2, \tag{18}$$

where $x$ and $\tilde{x}$ denotes the ground truth image and the reconstruction, respectively, and $n$ is the number of pixels in the reconstructed image. A higher PSNR value indicates better reconstruction quality.

The SSIM, which lies in the range [0, 1], is used to measure the similarity between the ground-truth image and the reconstruction image,

$$\text{SSIM}(\tilde{x}, x) = \frac{1}{M} \sum_{j=1}^{M} \frac{(2\tilde{\mu}_j \mu_j + C_1)(2\Sigma_j + C_2)}{(\tilde{\mu}_j^2 + \mu_j^2 + C_1)(\tilde{\sigma}_j^2 + \sigma_j^2 + C_2)}, \tag{19}$$

where $\tilde{\mu}_j$ and $\mu_j$ are the average pixel intensities, $\tilde{\sigma}_j^2$ and $\sigma_j^2$ represent the variances, and $\Sigma_j$ is the the covariance of $\tilde{x}, x$ at the $j$-th local window. The constants $2C_1 = (K_1 L)^2$ and $C_2 = (K_2 L)^2$ tend to be zero to avoid instability. Following [33, 34], we choose $K_1 = 0.01$, $K_2 = 0.03$, $L = \max(x) - \min(x)$, and the window size is $7 \times 7$. A higher SSIM value indicates better reconstruction quality.

### Reconstruction results

**AAPM challenge data.** We randomly select three full-dose CT images from AAPM challenge data to evaluate effectiveness of the propose method with the X-ray source intensity $I_i = [1e3, 1e4, 5e4]$. From Figs 2–4, we can observe that the quality of the FBP reconstruction images degraded significant as the X-ray source intensity decreased, resulting in amplified noise and artifacts distributed throughout the entire image. As a post-processing method, TV achieves higher quality images by post-processing the reconstructed images through FBP. Another post-processing and supervised method RED-CNN, can effectively remove noise and artifacts, but it tends to smooth out some tiny structures. Although DIP is unsupervised and takes random noise as input, it can effectively remove noise while producing images with higher resolution than RED-CNN. Comparing the reconstructed results by different methods, we can see that the proposed method achieves the best performance in terms of noise and artifacts attenuation and preservation of tiny structures.
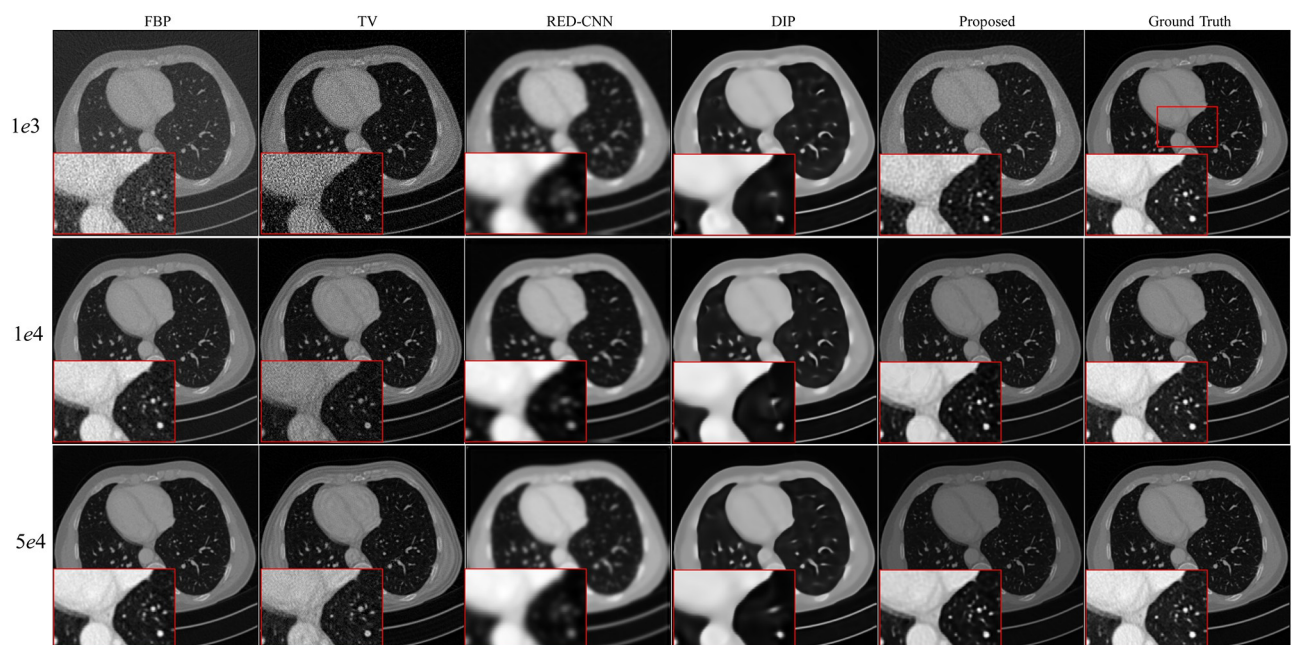


**Fig 2. Reconstruction results of case AAPM-1 at different dose levels by different methods.** Zoomed parts over the region of interest (ROI) marked by the red box in the ground-truth image.

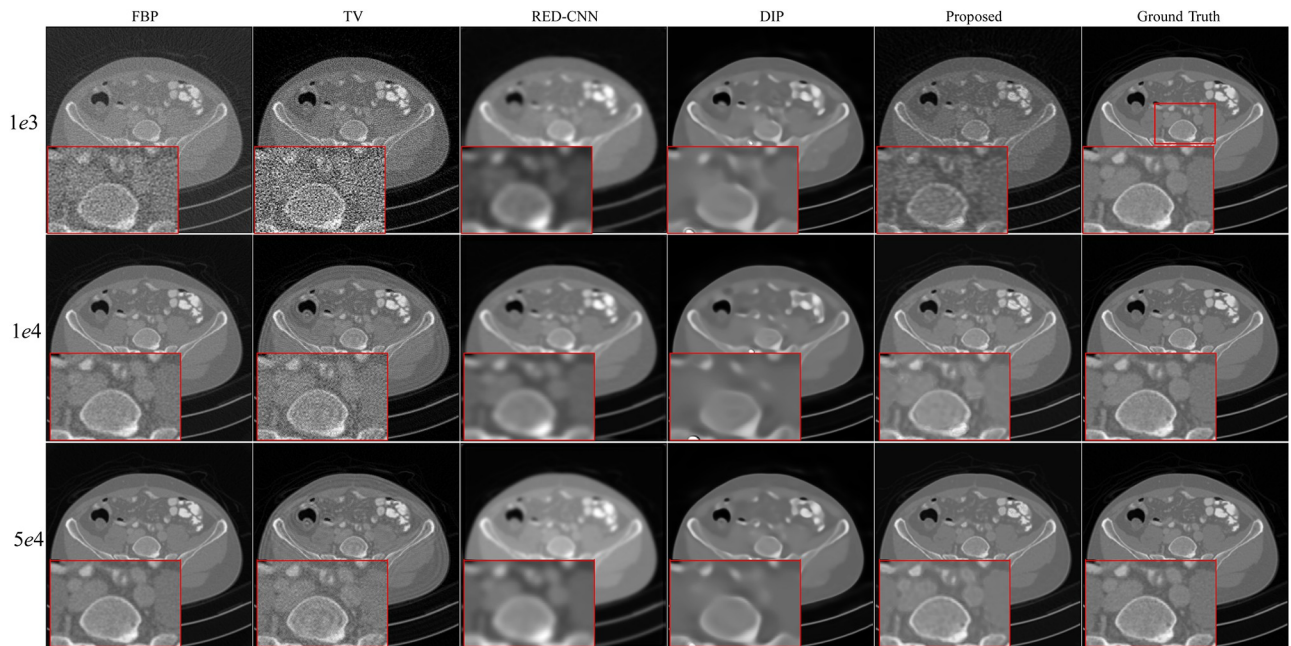https://doi.org/10.1371/journal.pone.0304738.g002

**Fig 3. Reconstruction results of case AAPM-2 at different dose levels by different methods.** Zoomed ROI images from the ground-truth image.

To better illustrate the effectiveness of the proposed method, we further demonstrate the zoomed-in results corresponding to the red box in each ground truth. As shown in Figs 2–4, the reconstructed results by FBP and TV are contaminated by noise and artifacts. Although RED-CNN and DIP can suppress the noise, many valuable details are smoothed out. In
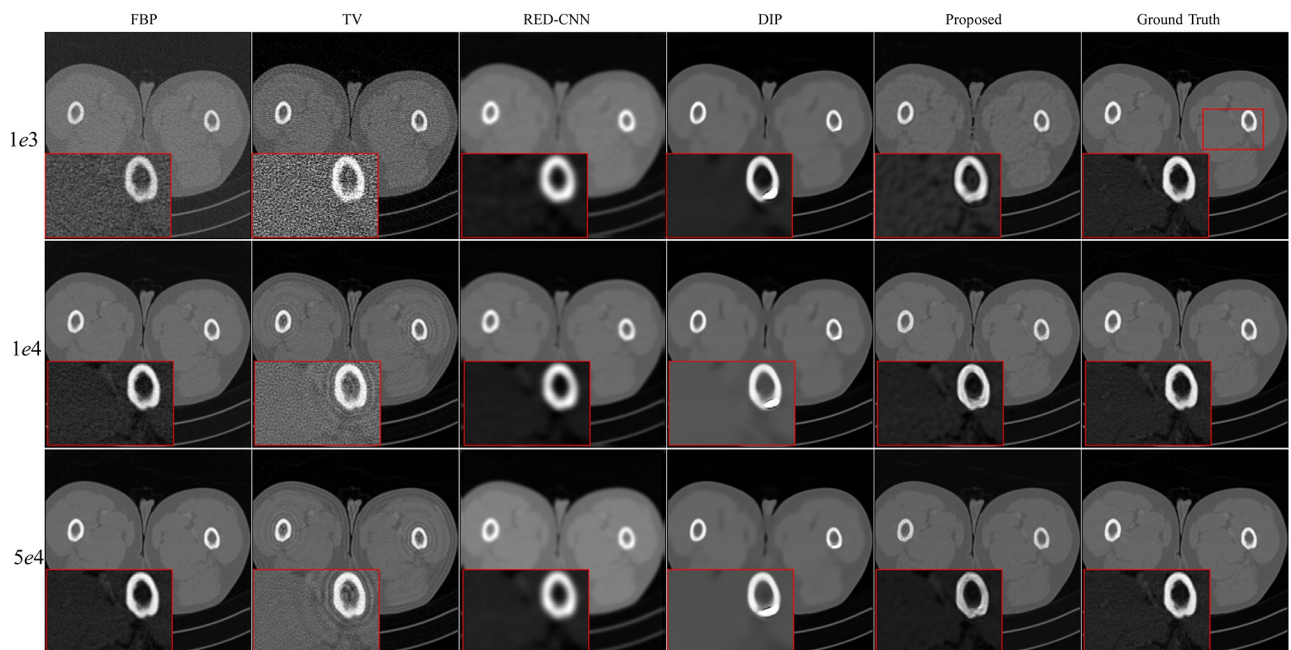


**Fig 4. Reconstruction results of case AAPM-3 at different dose levels by different methods.** Zoomed ROI images from the ground-truth image.

comparison, the proposed method achieves better reconstruction accuracy than the competitive methods. It is worth noting that although the ground-truth images are norm-dose CT images, slight noise and artifacts still remain in them. Furthermore, the reconstructed results by the proposed method outperform the ground-truth images in terms of resolution, particularly with $I_i = 1e4$ and $5e4$.

**LoDoPab-CT data.** For the LoDoPab-CT data, the reconstruction results are shown in Fig 5. From Fig 5, we can observe that the performance of each reconstruction method is similar to their performance for the above AAPM challenge data reconstruction. The reconstructed results by FBP and TV suffer from noise and artifacts, although TV can suppress a lot of noise. The textures and edges in the reconstructed results by RED-CNN are smoothed out, whereas DIP can remove noise and preserve tiny structures more effectively. The proposed method achieves the best performance with regard to noise suppression and preservation of tiny structures. Furthermore, the reconstruction errors (Fig 6) further demonstrate that FBP method sacrifices a lot of useful information. TV and RED-CNN can effectively improve the reconstructed results by FPB; however, TV can not preserver edges well, and RED-CNN tends to smooth edges and textures. DIP has slighter residual errors in terms of edges and textures. Compared with the competitive methods, the proposed method significantly decreases the reconstruction errors by FBP and achieves minimal reconstruction errors.

To quantitatively analyze the performance of our method, we calculate the the PSNR and SSIM values of the above reconstruction results, including the AAPM challenge data and the LoDoPab-CT data. As shown in Table 1, our method achieves the highest PSNR and SSIM among the five approaches, except for the reconstruction task of AAPM-2 with respect of the SSIM under $I_i = 1e3$ and of AAPM-3 with respect of the PSNR under $I_i = 1e3$. Specifically, the SSIM and PSNR of DIP are 0.06 and 0.24 dB higher than those of the proposed method.

In addition, we take the evolution curves of PSNR and SSIM versus iteration of LoDoPab-CT data reconstruction as an example to illustrate the convergence of the proposed method.
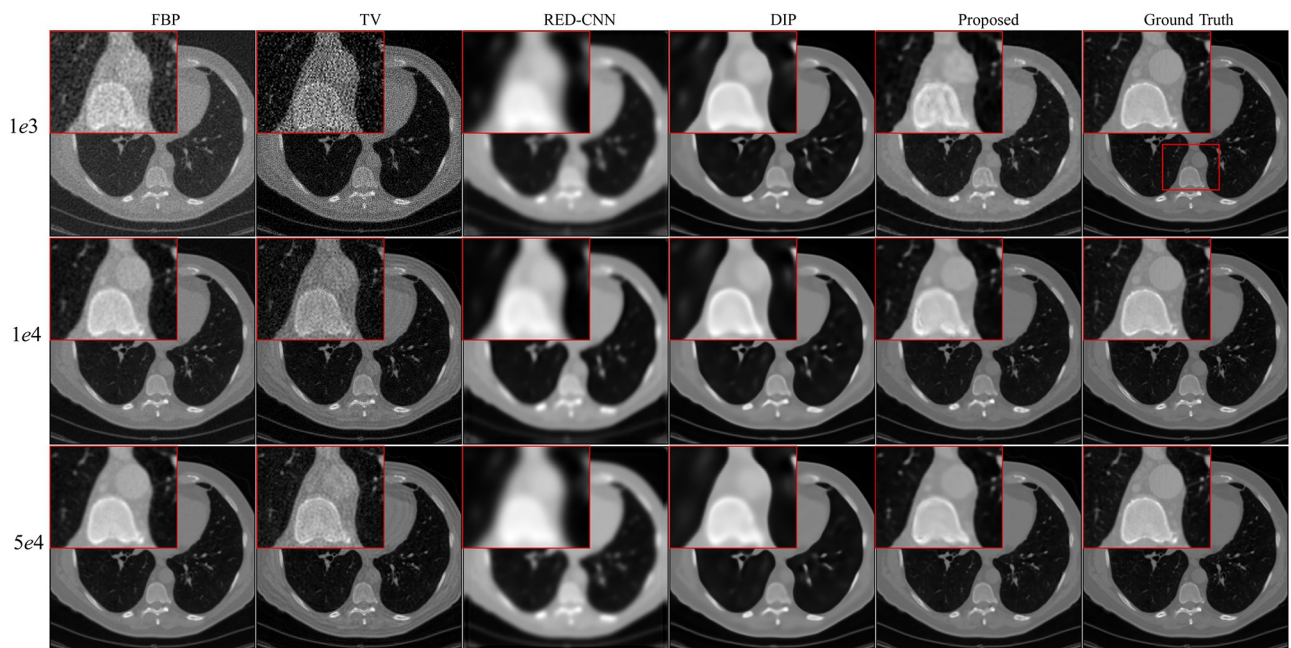


**Fig 5. Reconstruction results of LoDoPab-CT data at different dose levels by different methods.** Zoomed ROI images from the ground-truth image.
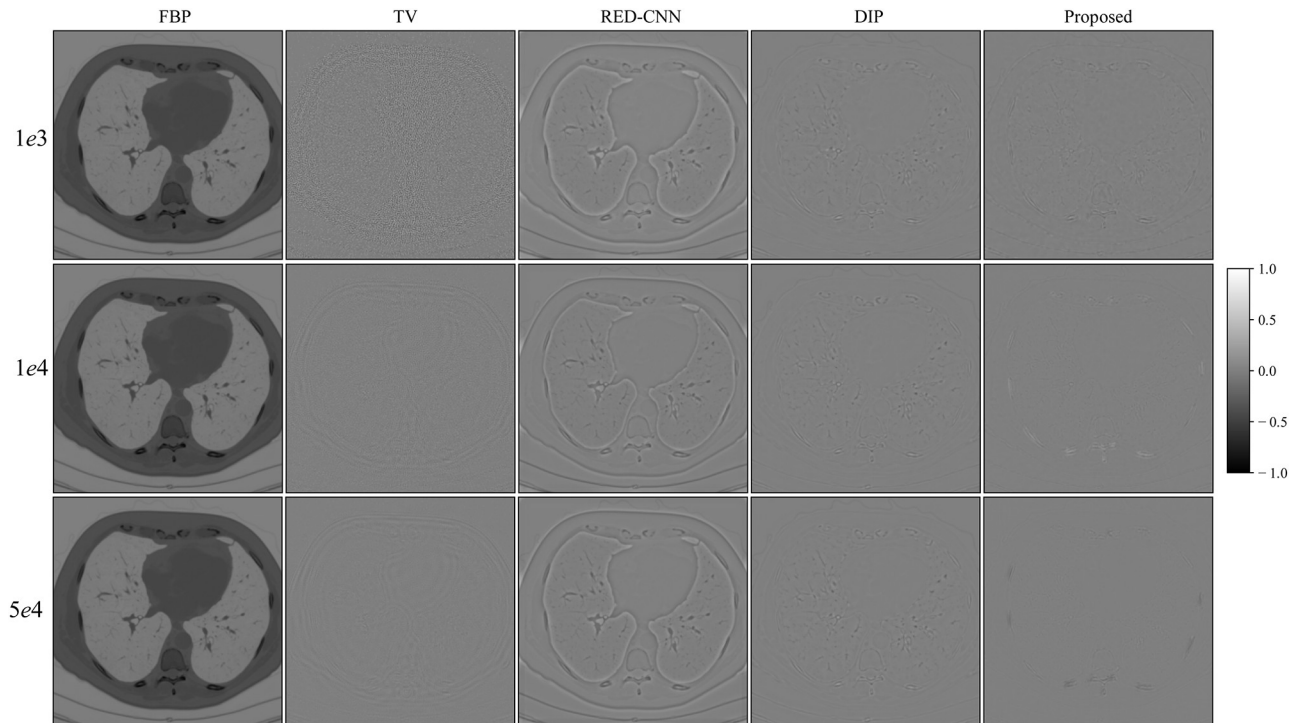
**Fig 6. Reconstruction errors of LoDoPab-CT data at different dose levels by different methods.**

As shown in Fig 7, the PSNR and SSIM increase while the loss decreases rapidly, which reveals that the proposed method can converge quickly. Specifically, the curves of PSNR and SSIM begin to converge after about 250 iterations, and the curves of loss start to converge after about 100 iterations. Although there are some fluctuations in these curves since the measurements contain noise, they converge quickly again, which indicates the good robustness of our method. Table 2 lists the computation time of different method on a single GPU (Nvidia Tesla K80), it can be seen that FBP, TV and the proposed method have great disadvantages in terms

**Table 1. Quantitative results (PNSR/SSIM) of different algorithms.**

| Data | $I_i$ | FBP | TV | RED-CNN | DIP | Proposed |
|------|-------|-----|-----|---------|-----|----------|
| AAPM-1 | 1e3 | 10.92/0.09 | 14.51/0.19 | 23.54/0.64 | 29.83/0.81 | **32.37/0.82** |
| | 1e4 | 10.92/0.09 | 21.19/0.37 | 25.39/0.74 | 30.22/0.82 | **35.99/0.92** |
| | 5e4 | 10.92/0.09 | 26.25/0.54 | 24.23/0.69 | 30.36/0.82 | **37.85/0.94** |
| AAPM-2 | 1e3 | 9.75/0.12 | 12.86/0.16 | 24.89/0.68 | 31.46/**0.85** | **31.56**/0.79 |
| | 1e4 | 9.75/0.12 | 19.32/0.31 | 27.49/0.79 | 30.33/0.83 | **35.30/0.92** |
| | 5e4 | 9.75/0.12 | 25.21/0.49 | 25.89/0.75 | 31.94/0.85 | **38.06/0.94** |
| AAPM-3 | 1e3 | 10.89/0.13 | 13.30/0.17 | 25.71/0.73 | **35.71/0.90** | 35.47/**0.90** |
| | 1e4 | 10.89/0.13 | 19.99/0.30 | 28.35/0.83 | 36.24/0.91 | **38.07/0.93** |
| | 5e4 | 10.89/0.13 | 26.03/0.47 | 26.64/0.80 | 35.20/0.90 | **37.72/0.94** |
| LoDoPab-CT | 1e3 | 10.65/0.09 | 15.72/0.22 | 24.33/0.65 | 32.76/0.88 | **33.25/0.89** |
| | 1e4 | 10.65/0.09 | 24.26/0.45 | 26.12/0.73 | 33.14/0.88 | **35.84/0.94** |
| | 5e4 | 10.65/0.09 | 29.24/0.65 | 24.95/0.70 | 32.82/0.87 | **36.39/0.95** |
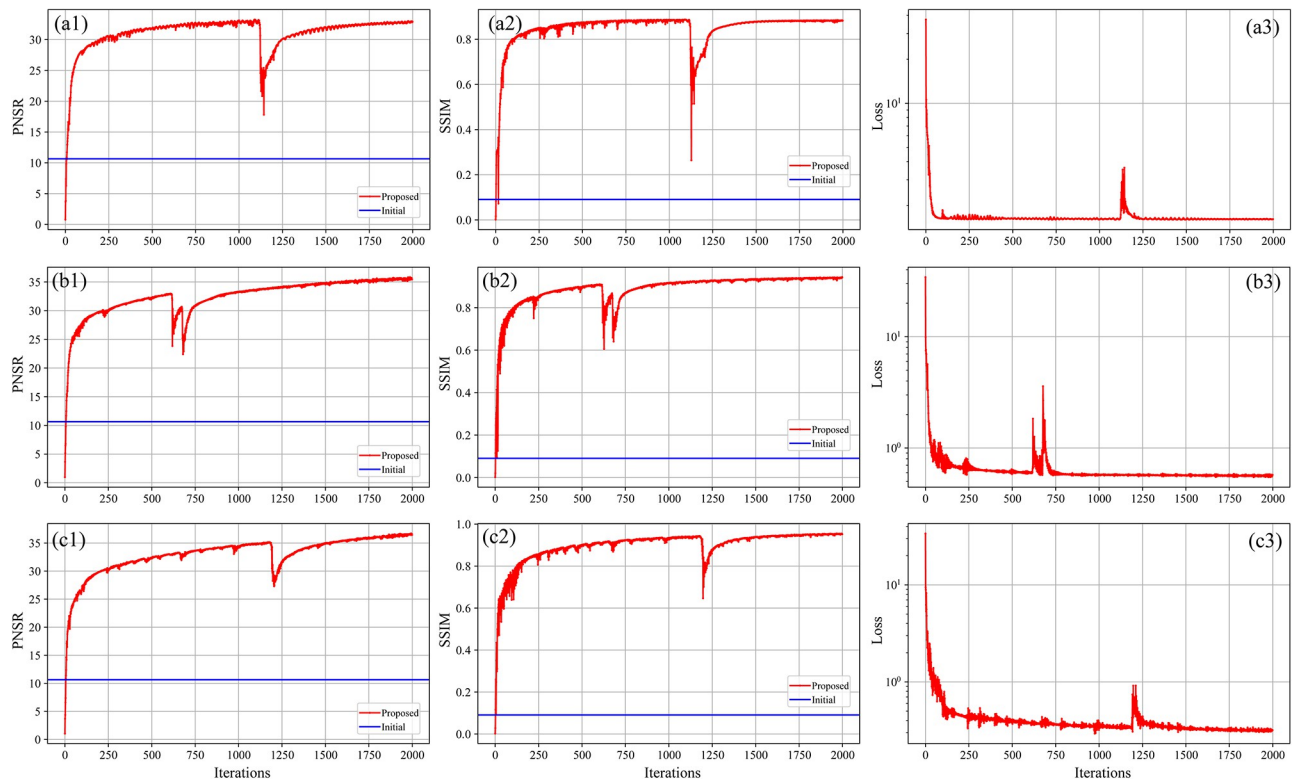
**Fig 7. Convergence analysis of the proposed method in LoDoPab-CT data reconstruction.** (a1), (a2) and (a3) represent the PSNR, SSIM, and loss curve, respectively, for $I_i = 1e3$. (b1), (b2) and (b3) depict the PSNR, SSIM, and loss curve, respectively, for $I_i = 1e4$. (c1), (c2) and (c3) is the PSNR, SSIM, and loss curve, respectively, for $I_i = 5e4$.

https://doi.org/10.1371/journal.pone.0304738.g007

**Table 2. Computation time of different algorithms for LoDoPab-CT data reconstruction.**

| Method | GPU time | |
|---|---|---|
| | Training | Reconstruction |
| FBP (s) | / | 0.1 |
| TV (s/iteration) | / | 0.07 |
| RED-CNN (s/epoch) | 310 | 5.1 |
| DIP (s/iteration) | / | 3.11 |
| Proposed (s/iteration) | / | 1.02 |

https://doi.org/10.1371/journal.pone.0304738.t002

of reconstruction time. Although RED-CNN only need one inference to reconstruct the high-quality image, the process of NN training is time consuming. Therefore, one can set a larger number of iterations to ensure that good reconstruction results can be obtained due to the rapid convergence and low computational cost of the proposed method.

## Discussion and conclusion

For the initial LDCT reconstruction, we utilize the results reconstructed by FBP as the initial model for the proposed method. FBP can extract fundamental information about the internal structure of the human body, despite potential contamination from artifacts caused by the low

intensity of X-ray. This is crucial for neural network-based LDCT imaging, as the black-box nature of these networks can significantly decrease the reliability of LDCT reconstruction results. It's also important to note that the quality of the initial reconstructed image can impact the performance of the proposed method. One could substitute the FBP input with a high-quality image to further improve resolution. Additionally, FBP often performs much faster than iterative reconstruction approaches such as compressive sensing, which aids in enhancing inversion efficiency. Although our proposed method can converge rapidly, fluctuations due to noise in measurements might negatively impact the reconstruction efficiency. In future work, we aim to investigate better regularization techniques to promote convergence stability.

In this work, we propose an unsupervised and training data-free method for LDCT imaging. The proposed method aims to improve the initial reconstruction results with low quality, which reconstructs the high-quality image by DNN training without any training samples. We implement the DNN training by minimizing the $\ell_1$-norm distance between the CT measurements and their corresponding simulated sinogram data on the reconstructed image and the TV value of the reconstructed image. Notably, the proposed method dose not need to set weights for both the data fidelity term and the regularization term, which significantly reduces the difficulty of manually setting the weights. Experimental results on the AAPM challenge data and LoDoPab-CT data demonstrate that the proposed method could achieve better performance than the representative non-learning methods and supervised method, with higher resolution and lower computational cost. The proposed method can be implemented flexible and has the potential to be applied to other medical image reconstruction problems, including sparse-view CT reconstruction and image reconstruction from sparse samples in MRI. These applications are particularly useful when collecting training samples is either expensive or difficult.

## Author Contributions

**Conceptualization:** Feng Wang.

**Funding acquisition:** Renfang Wang.

**Investigation:** Feng Wang, Renfang Wang, Hong Qiu.

**Methodology:** Feng Wang, Renfang Wang, Hong Qiu.

**Software:** Feng Wang.

**Validation:** Feng Wang.

**Visualization:** Feng Wang.

**Writing – original draft:** Feng Wang, Renfang Wang, Hong Qiu.

**Writing – review & editing:** Feng Wang, Renfang Wang, Hong Qiu.

## References

1. Smith-Bindman R, Lipson J, Marcus R, Kim KP, Mahesh M, Gould R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Archives of Internal Medicine. 2009; 169(22):2078–2086. https://doi.org/10.1001/archinternmed.2009.427 PMID: 20008690

2. Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Physics in Medicine and Biology. 2008; 53(17):4777–4807. https://doi.org/10.1088/0031-9155/53/17/021 PMID: 18701771

3. Kim K, Ye JC, Worstell W, Ouyang J, Rakvongthai Y, Fakhri GE, et al. Sparse-View Spectral CT Reconstruction Using Spectral Patch-Based Low-Rank Penalty. IEEE Transactions on Medical Imaging. 2015; 34(3):748–760. https://doi.org/10.1109/TMI.2014.2380993 PMID: 25532170

4. Cai JF, Jia X, Gao H, Jiang SB, Shen Z, Zhao H. Cine Cone Beam CT Reconstruction Using Low-Rank Matrix Factorization: Algorithm and a Proof-of-Principle Study. IEEE Transactions on Medical Imaging. 2014; 33(8):1581–1591. https://doi.org/10.1109/TMI.2014.2319055 PMID: 24771574

5. Ma J, Zhang H, Gao Y, Huang J, Liang Z, Feng Q, et al. Iterative image reconstruction for cerebral perfusion CT using a pre-contrast scan induced edge-preserving prior. Physics in Medicine and Biology. 2012; 57(22):7519–7542. https://doi.org/10.1088/0031-9155/57/22/7519 PMID: 23104003

6. Zhang Y, Xi Y, Yang Q, Cong W, Zhou J, Wang G. Spectral CT Reconstruction With Image Sparsity and Spectral Mean. IEEE Transactions on Computational Imaging. 2016; 2(4):510–523. https://doi.org/10.1109/TCI.2016.2609414 PMID: 29034267

7. Qiong Xu, Hengyong Yu, Xuanqin Mou, Lei Zhang, Jiang Hsieh, Ge Wang. Low-Dose X-ray CT Reconstruction via Dictionary Learning. IEEE Transactions on Medical Imaging. 2012; 31(9):1682–1697. https://doi.org/10.1109/TMI.2012.2195669

8. Yuan N, Zhou J, Gong K, Qi J. Low-dose CT count-domain denoising via convolutional neural network with filter loss. In: Bosmans H, Chen GH, Gilat Schmidt T, editors. Medical Imaging 2019: Physics of Medical Imaging. San Diego, United States: SPIE; 2019. p. 26. Available from: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10948/2513479/Low-dose-CT-count-domain-denoising-via-convolutional-neural-network/10.1117/12.2513479.full.

9. Liu J, Li J. Sparse-sampling CT Sinogram Completion using Generative Adversarial Networks. In: 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Chengdu, China: IEEE; 2020. p. 640–644. Available from: https://ieeexplore.ieee.org/document/9263571/.

10. Jin KH, McCann MT, Froustey E, Unser M. Deep Convolutional Neural Network for Inverse Problems in Imaging. IEEE Transactions on Image Processing. 2017; 26(9):4509–4522. https://doi.org/10.1109/TIP.2017.2713099

11. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, et al. Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network. IEEE Transactions on Medical Imaging. 2017; 36(12):2524–2535. https://doi.org/10.1109/TMI.2017.2715284 PMID: 28622671

12. Hasan AM, Mohebbian MR, Wahid KA, Babyn P. Hybrid-Collaborative Noise2Noise Denoiser for Low-Dose CT Images. IEEE Transactions on Radiation and Plasma Medical Sciences. 2021; 5(2):235–244. https://doi.org/10.1109/TRPMS.2020.3002178

13. Lehtinen J, Munkberg J, Hasselgren J, Laine S, Karras T, Aittala M, et al. Noise2Noise: Learning Image Restoration without Clean Data; 2018. Available from: http://arxiv.org/abs/1803.04189.

14. Hendriksen AA, Pelt DM, Batenburg KJ. Noise2Inverse: Self-supervised deep convolutional denoising for tomography. IEEE Transactions on Computational Imaging. 2020; 6:1320–1335. https://doi.org/10.1109/TCI.2020.3019647

15. Evangelista D, Morotti E, Loli Piccolomini E. RISING: A new framework for model-based few-view CT image reconstruction with deep learning. Computerized Medical Imaging and Graphics. 2023; 103:102156. https://doi.org/10.1016/j.compmedimag.2022.102156 PMID: 36528018

16. Kang E, Min J, Ye JC. A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction. Medical Physics. 2017; 44(10):e360–e375. https://doi.org/10.1002/mp.12344 PMID: 29027238

17. Zheng A, Gao H, Zhang L, Xing Y. A dual-domain deep learning-based reconstruction method for fully 3D sparse data helical CT. Physics in Medicine and Biology. 2020; 65(24):245030. https://doi.org/10.1088/1361-6560/ab8fc1 PMID: 32365345

18. Yuan N, Zhou J, Qi J. Half2Half: deep neural network based CT image denoising without independent reference data. Physics in Medicine & Biology. 2020; 65(21):215020. https://doi.org/10.1088/1361-6560/aba939 PMID: 32707565

19. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature. 2018; 555(7697):487–492. https://doi.org/10.1038/nature25988 PMID: 29565357

20. Kandarpa VSS, Bousse A, Benoit D, Visvikis D. DUG-RECON: A Framework for Direct Image Reconstruction Using Convolutional Generative Networks. IEEE Transactions on Radiation and Plasma Medical Sciences. 2021; 5(1):44–53. https://doi.org/10.1109/TRPMS.2020.3033172

21. Wu D, Kim K, Li Q. Computationally efficient deep neural network for computed tomography image reconstruction. Medical Physics. 2019; 46(11):4763–4776. https://doi.org/10.1002/mp.13627 PMID: 31132144

22. Cheng W, Wang Y, Li H, Duan Y. Learned Full-Sampling Reconstruction From Incomplete Data. IEEE Transactions on Computational Imaging. 2020; 6:945–957. https://doi.org/10.1109/TCI.2020.2996751

23. Ye S, Li Z, McCann MT, Long Y, Ravishankar S. Unified Supervised-Unsupervised (SUPER) Learning for X-Ray CT Image Reconstruction. IEEE Transactions on Medical Imaging. 2021; 40(11):2986–3001. https://doi.org/10.1109/TMI.2021.3095310 PMID: 34232871

24. Ulyanov D, Vedaldi A, Lempitsky V. Deep Image Prior. International Journal of Computer Vision. 2020; 128(7):1867–1888. https://doi.org/10.1007/s11263-020-01303-4

25. Gong K, Catana C, Qi J, Li Q. PET Image Reconstruction Using Deep Image Prior. IEEE Transactions on Medical Imaging. 2019; 38(7):1655–1665. https://doi.org/10.1109/TMI.2018.2888491 PMID: 30575530

26. Yokota T, Kawai K, Sakata M, Kimura Y, Hontani H. Dynamic PET Image Reconstruction Using Non-negative Matrix Factorization Incorporated With Deep Image Prior. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV); 2019. p. 3126–3135.

27. Yoo J, Jin KH, Gupta H, Yerly J, Stuber M, Unser M. Time-Dependent Deep Image Prior for Dynamic MRI. IEEE Transactions on Medical Imaging. 2021; 40(12):3337–3348. https://doi.org/10.1109/TMI.2021.3084288 PMID: 34043506

28. Ding Q, Ji H, Quan Y, Zhang X. A dataset-free deep learning method for low-dose CT image reconstruction. Inverse Problems. 2022; 38(10):104003. https://doi.org/10.1088/1361-6420/ac8ac6

29. Leuschner J, Schmidt M, Baguer DO, Maass P. LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction. Scientific Data. 2021; 8(1):109. https://doi.org/10.1038/s41597-021-00893-z PMID: 33863917

30. Baguer DO, Leuschner J, Schmidt M. Computed tomography reconstruction using deep image prior and learned reconstruction methods. Inverse Problems. 2020; 36(9):094004. https://doi.org/10.1088/1361-6420/aba415

31. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Advances in Neural Information Processing Systems 32. Curran Associates, Inc.; 2019. p. 8024–8035. Available from: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

32. He Z, Zhang Y, Guan Y, Guan B, Niu S, Zhang Y, et al. Iterative Reconstruction for Low-Dose CT Using Deep Gradient Priors of Generative Model. IEEE Transactions on Radiation and Plasma Medical Sciences. 2022; 6(7):741–754. https://doi.org/10.1109/TRPMS.2022.3148373

33. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing. 2004; 13(4):600–612. https://doi.org/10.1109/TIP.2003.819861 PMID: 15376593

34. Leuschner J, Schmidt M, Ganguly PS, Andriiashen V, Coban SB, Denker A, et al. Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications. Journal of Imaging. 2021; 7(3):44. https://doi.org/10.3390/jimaging7030044 PMID: 34460700