

Full Length Article

RBP-DIP: Residual back projection with deep image prior for ill-posed CT reconstruction[☆]Ziyu Shu^{*}, Alireza Entezari

CISE department, University of Florida, 32603, USA



ARTICLE INFO

Keywords:

Inverse problems
Deep learning
Computed tomography
Deep image prior
Neural networks

ABSTRACT

The success of deep image prior (DIP) in a number of image processing tasks has motivated their application in image reconstruction problems in computed tomography (CT). In this paper, we introduce a residual back projection technique (RBP) that improves the performance of deep image prior framework in iterative CT reconstruction, especially when the reconstruction problem is highly ill-posed. The RBP-DIP framework uses an untrained U-net in conjunction with a novel residual back projection connection to minimize the objective function while improving reconstruction accuracy. In each iteration, the weights of the untrained U-net are optimized, and the output of the U-net in the current iteration is used to update the input of the U-net in the next iteration through the proposed RBP connection. The introduction of the RBP connection strengthens the regularization effects of the DIP framework in the context of iterative CT reconstruction leading to improvements in accuracy. Our experiments demonstrate that the RBP-DIP framework offers improvements over other state-of-the-art conventional IR methods, as well as pre-trained and untrained models with similar network structures under multiple conditions. These improvements are particularly significant in the few-view and limited-angle CT reconstructions, where the corresponding inverse problems are highly ill-posed and the training data is limited. Furthermore, RBP-DIP has the potential for further improvement. Most existing IR algorithms, pre-trained models, and enhancements applicable to the original DIP algorithm can also be integrated into the RBP-DIP framework.

1. Introduction

CT reconstruction algorithms are essential in a number of biomedical imaging applications. The traditional filtered back projection (FBP) algorithm is widely used in commercial scanners as it is computationally efficient. However, its image quality is heavily affected by deviations from ideal imaging conditions (e.g., low-dose, few-view, and limited-angle). To improve image quality in these imaging applications, iterative reconstruction (IR) methods leverage mathematical models for acquisition as well as image priors into an optimization process that forms the basis for image reconstruction (Fessler, 2014). These methods have been widely studied in CT imaging and are considered the gold standard, their performance in practical imaging applications, robustness, and image quality issues are well understood (Willemink & Noël, 2019).

Neural network related methods have provided a new set of approaches in CT reconstruction. These methods can be used to improve image quality before (pre-processing, such as sinogram synthesis (Lee, Lee, Kim, Cho, & Cho, 2018)), during (end-to-end reconstruction (Zhu, Liu, Cauley, Rosen, & Rosen, 2018), plug in Shen, Ma, and

Jia (2019) and Shah and Hegde (2018)), or after (post processing, such as artifacts removal (Zhang et al., 2016)) reconstruction. The successful application of neural network methods in CT reconstruction is contingent upon access to high-quality datasets and their ability to learn priors that can generalize to instances outside the training dataset. While deep learning methods have had a significant impact in CT imaging research (Wang, Ye, Mueller, & Fessler, 2018), their application in reconstruction poses significant challenges as their lack of stability can lead to the introduction of false structures (hallucinations) that may lead to detrimental effects in medical imaging (Antun, Ren, Poon, Adcock, & Hansen, 2020; Bhadra, Kelkar, Brooks, & Anastasio, 2021).

This paper proposes the RBP-DIP framework that inherits the advantages of neural networks' hierarchical structure and conventional IR methods. RBP-DIP employs an untrained U-net as a deep image prior (DIP), combined with the newly proposed residual back-projection (RBP) connection, allowing for optimizing both the network weights

[☆] Manuscript submitted February 1st, 2024. This work was supported in part by the NSF [grant number CCF-2210866].

^{*} Corresponding author.

E-mail addresses: zs919@nyu.edu (Z. Shu), entezari@ufl.edu (A. Entezari).

and input in different ways (backpropagation through the neural network and conventional IR update through the RBP connection respectively) during each iteration to match the measurements. This provides the RBP-DIP with the following major advantages:

- No reliance on training data: the RBP-DIP algorithm employs an untrained neural network, and thus deviates from conventional neural network methodologies. RBP-DIP aligns more closely with IR techniques as it directly minimizes the loss function on inference data. Consequently, the RBP-DIP algorithm needs no training data, therefore circumventing interferences commonly associated with training data and the training process.
- Combination of DIP and IR: The original DIP algorithm can generate high-quality reconstruction under highly ill-posed conditions. However, it suffers from blurring and neural network-specific artifacts. While IR algorithms suffer from severe artifacts under ill-posed conditions, their stability and interpretability are what DIP lacks. The RBP-DIP algorithm, through the RBP connection, seamlessly combines the strengths of both, achieving superior performance across various conditions.
- Better reconstruction performance: The experiments show that the RBP-DIP provides significant improvements over conventional IR methods and other neural network models (both untrained and pre-trained) with a similar network structure under multiple conditions. The observed improvements are most significant under few-view and limited-angle conditions.
- Great potential for further improvement: The proposed framework has great potential for further improvement. It provides a new way of combining the advantages of deep image prior and conventional reconstruction techniques. In this paper, for simplicity, only the basic U-net, residual back projection, and objective function are used for deep image prior, RBP connection, and overall optimization. It is evident that more delicate network structures, IR update techniques, and objective functions can be implemented. Furthermore, the proposed method does not conflict with most existing DIP improvement algorithms. In that case, further improvement is feasible through collaboration with these methods.

The remainder of this paper is organized as follows: the details of the CT reconstruction problem and the related works are introduced in Section 2. The proposed framework is introduced in Section 3, and compared with conventional methods, pre-trained and untrained methods with a similar network structure on real CT images under various imaging conditions in Section 4. The results from our experiments are discussed in Section 5 followed by the conclusion Section 6.

2. Related works

2.1. Iterative CT reconstruction

IR approaches frame image reconstruction as an optimization problem, minimizing the inconsistency between the forward projection of the iterate and the measurements. X-ray physics, non-ideal effects, and image priors can be incorporated into the reconstruction algorithms by employing various constraints, regularizations, and forward models. The objective function for the CT optimization problem is often expressed by fidelity and regularization terms:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{g} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda R(\mathbf{x}), \quad (1)$$

where \mathbf{x} , \mathbf{g} , \mathbf{A} , and R indicate the image coefficients, sinogram measurements, forward model, and regularization respectively. With accurate forward models (Shu & Entezari, 2020, 2022a; Zhang & Entezari, 2019b) and effective priors (Bouman & Sauer, 1993; Gong & Zeng, 2020; Jin, Li, Chen, Zhang, & Xing, 2010; Kisner et al., 2012; Sidky, Kao, & Pan, 2006; Wang, Nakamoto, Zhang, & Liu, 2017; Zhang,

Bouman, Thibault, & Sauer, 2013), high-quality reconstruction imaging is possible under non-ideal conditions.

However, when the CT reconstruction is highly ill-posed (e.g., limited-angle and few-view CT reconstruction), the generation of high-quality reconstruction presents difficulties with sparsity-enhancing regularization like total variation, and this is the case even when the measurements are noise-free. In that case, researchers proposed using application-specific priors that are learned from a sample of images to further improve the reconstruction accuracy. Multiple neural network related approaches are developed following this line of thought, and will be discussed in the next section.

2.2. CT reconstruction using pre-trained neural networks

Zhu et al. (2018) proposed learning a mapping from the acquisition data to the reconstructed images directly. However, the use of fully connected layers makes the computational costs necessary for reconstruction at practical resolutions unacceptable. To circumvent the use of fully connected layers, researchers proposed incorporating additional information into the input to simplify the task so that fully convolutional neural networks are capable of being employed in image reconstruction problems. This is demonstrated by Chen et al. (2017), Ye, Buzzard, Ruby, and Bouman (2018), and Zhang et al. (2020), who suggested utilizing back projections, filtered back projections, and single-view back projections as inputs for high-quality CT reconstruction, respectively.

To further improve the accuracy of reconstructions, researchers proposed designing different neural networks for different kinds of CT reconstruction problems. For ill-posed CT reconstruction, Pan et al. (2024) designed a unique objective function and a convolution-assisted transformer that can simultaneously capture both local and long-range pixel interactions. Wu, Wang, Liu, Wang, and Zhang (2023) and Zhang, Mao, Wang, Guo, and Wu (2024) proposed improved wavelet denoising techniques to complement the score-based generative model. Li et al. and Wang et al. proposed introducing the diffusion models (Li et al., 2024; Wang, Li, & Wu, 2024). All these methods have achieved significant results.

Recently, more sophisticated models have been developed, which involve the collaboration of multiple neural networks working on various subtasks to achieve enhanced performance. Yin et al. (2019) introduced an approach that employs two neural networks for denoising in both the sinogram and the image domain. Hu et al. (2021) and Zhang et al. (2021) extended this concept by incorporating three neural networks. In addition to the two denoising neural networks described in Yin et al. (2019), an auxiliary image domain discriminator for real and reconstructed images was integrated to further refine the model. Similar architectures have also been explored in Ma et al. (2020), Wu et al. (2021) and Xie et al. (2020).

2.3. Challenges in pre-trained neural networks

Instability has been considered the greatest drawback for neural network image reconstruction. Neural network reconstruction algorithms are designed to learn the prior distribution of the imaged objects from training datasets implicitly or explicitly, so that high-quality reconstructions can be made even with limited and noisy measurements. However, the aforementioned circumstances are predicated upon the assumption that the inference data and training data follow the same distribution. This is particularly problematic for medical image reconstruction as the learned network may omit patient-specific features. Instead, it tends to substitute these unique characteristics with generic features that it has learned from training datasets. Several researchers have demonstrated the intrinsic instability of deep learning reconstruction methods, and features can be present in reconstructions that are hallucinated by the network, as investigated by Antun et al. (2020) and Gottschling, Antun, Hansen, and Adcock (2023).

Another problem is that current neural network related algorithms are difficult to implement on a large scale. On the one hand, training practical neural networks is challenging. For example, the collaboration of multiple sub-networks utilized in the aforementioned methods necessitates the use of complicated objective functions, which consequently exacerbate the intricacy of the training process. Additionally, to achieve optimal performance, a network must be retrained to learn the best prior for each distinct setting (e.g. reconstruction resolution, detector spacings, noise level). On the other hand, analyzing and improving an existing neural network is also resource-intensive. The absence of a systematic and efficient approach to enhance the performance of a neural network when it fails to meet expectations also presents a challenge. The cascade of multiple sub-networks in the aforementioned methods (Hu et al., 2021; Ma et al., 2020; Xie et al., 2020; Yin et al., 2019; Zhang et al., 2021) further complicates these challenges.

2.4. CT reconstruction using untrained neural networks

To avoid reliance on training data, some researchers proposed methods that require no training data. Ulyanov, Vedaldi, and Lempitsky (2018) demonstrated that the hierarchical structure of convolutional networks inherently possesses the capability to capture abundant low-level image statistical priors, thereby enabling the generation of high-quality images. This property is often referred to as the deep image prior (DIP). Several researchers (Bauer, Leuschner, & Schmidt, 2020; Shu & Entezari, 2022b; Veen et al., 2018) leveraged the DIP property in their studies on inverse problems. These researchers advocated for minimizing the objective function by optimizing the weights of untrained networks during the reconstruction process. In doing so, the reconstruction is complete once the objective function is minimized. This method was further improved by Shu and Entezari (2022b), where the latent vector \mathbf{z} is updated together with the weights to enhance the convergence speed and reconstruction accuracy. Additionally, they proposed detailed reconstruction instructions and extra regularizations for more accurate reconstructions.

The primary challenge posed by these methods lies in their stability, given that the entire network is initialized randomly and can be trapped in a local minimum. Various strategies have been put forth to overcome this challenge. Veen et al. (2018) proposed running the same algorithm multiple times and choosing the best result; Bauer et al. (2020) recommended the incorporation of supplementary pre-trained neural networks for pre- and post-processing. However, the former approach relies on inefficient trial and error, while the latter reintroduces pre-trained models, contravening its initial objective.

3. Methods

3.1. CT reconstruction with deep image prior

As discussed before, DIP in CT reconstruction optimizes an objective function similar to the conventional IR method (Veen et al., 2018), which can be formulated as:

$$\begin{aligned} \mathbf{w}^* &= \arg \min_{\mathbf{w}} \|\mathbf{g} - \mathbf{A}G(\mathbf{w}; \mathbf{z})\|_2^2, \\ \mathbf{x}^* &= G(\mathbf{w}^*; \mathbf{z}), \end{aligned} \quad (2)$$

where $G(\mathbf{w}; \mathbf{z})$ represents an untrained convolutional neural network, it takes \mathbf{z} as the input and is parameterized by the weights \mathbf{w} . Here \mathbf{z} is a Gaussian vector that is randomly initialized and then fixed.

The key difference between the conventional IR method (Eq. (1)) and the DIP method (Eq. (2)) is that the DIP method uses the output of $G(\mathbf{w}; \mathbf{z})$ to represent the reconstruction result \mathbf{x} . DIP related methods utilize the untrained neural network G to minimize objective function by optimizing the network weights. In that case, the reconstruction result is constrained to lie in the space spanned by that untrained

Algorithm 1 RBP-DIP

Input: measure matrix \mathbf{A} , measurement \mathbf{g} , output \mathbf{x} (zero at the beginning), input \mathbf{z} (zero at the beginning), and step size β . The untrained model G indicates the U-net in Fig.1 without the RBP connection.

- 1: **Repeat:**
 - 2: $\mathbf{r} = \mathbf{A}^T(\mathbf{g} - \mathbf{A}\mathbf{x})$ /* Residual Back projection */
 - 3: $\mathbf{z} = \mathbf{z} + \beta \frac{\mathbf{r}}{\|\mathbf{r}\|_2}$ /* Network input modification */
 - 4: $\mathbf{z} = \frac{\mathbf{z}}{\|\mathbf{z}\|_2}$ /* Network input normalization */
 - 5: $\mathbf{x} = G(\mathbf{w}; \mathbf{z})$ /* U-net forward propagation */
 - 6: Update G , using $\|\mathbf{g} - \mathbf{A}\mathbf{x}\|_2^2$ as the loss function. /* U-net backpropagation */
 - 7: **Optional:** apply IR algorithm on \mathbf{x} .
-

neural network, so that the DIP property mostly ensures the generation of high-quality results, especially when the corresponding inverse problem is highly ill-posed.

It is worth mentioning that both DIP and IR methods minimize a similar loss function (regularizations can also be applied to Eq. (2) (Liu, Sun, Xu, & Kamilov, 2019)) on inference data directly. This ensures that the results of DIP are not inferior to those of IR under similar loss conditions. However, DIP related methods are susceptible to blurring and neural network specific artifacts. These shortcomings are demonstrated in Figs. 7d, and 8d, where the reconstruction results of the DIP method suffer from abnormal artifacts when the corresponding inverse problem is highly ill-posed, and lack image details even when the problem is well-posed (full-view CT reconstruction).

3.2. Residual back projection with deep image prior

Since its inception, researchers have been trying to further improve the DIP algorithm by adjusting its network input. Ulyanov et al. (2018) introduced small Gaussian noise into the network input in each iteration to improve the algorithm’s stability and noise resistance. Shu and Entezari’s research (Shu & Entezari, 2022b) showed that optimizing both the network input and network weights can further improve the convergence speed and reconstruction accuracy of the DIP method. Cui et al. (2022), in conditional DIP (CDIP), proposed using a reference image as the network input to guide the DIP reconstruction process. These studies imply that the network input can significantly impact the DIP algorithm, and using a fixed random vector may not be the best option. However, existing algorithms each have their shortcomings. The improvement provided by the method proposed in Shu and Entezari (2022b) is somewhat limited, mainly because the algorithm homogenizes the network input and network weights — they are both randomly initialized and updated via the neural network’s backpropagation algorithm. Given that the number of parameters for the network input is much smaller than that of the network weights, the impact of updating the network input is limited. CDIP algorithm’s effectiveness highly depends on the quality of the reference image. However, high-quality reference images similar to the ground truth image are often difficult to obtain (since the algorithm’s goal is to achieve high-quality reconstruction similar to the ground truth image). Furthermore, if the reference image itself contains artifacts, these artifacts will mislead the DIP algorithm.

Nevertheless, these algorithms have demonstrated that modifying the input of the DIP network can significantly affect the reconstruction process and its result. To further leverage this property while overcoming the challenges faced by the aforementioned algorithms, we propose the RBP-DIP framework. The basic idea of the RBP-DIP is to update the network input based on the network output in the previous iteration. RBP-DIP is similar to the method proposed in Shu and Entezari (2022b), as both involve updating the network input and network weights.

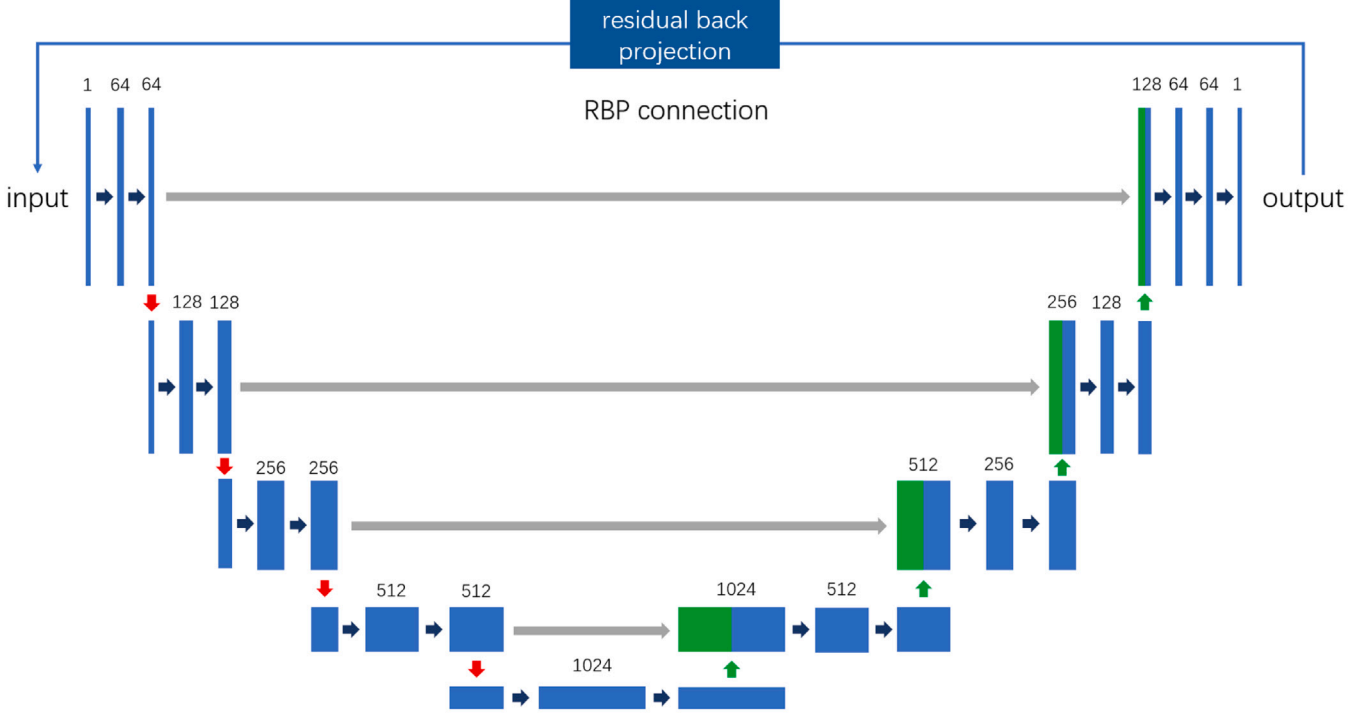


Fig. 1. The architecture of the RBP-DIP framework: In each iteration, a conventional IR algorithm can be used to calculate the necessary updates for the current reconstruction based on the neural network's output. This information is then passed to the neural network's input in the next iteration through the RBP connection, indirectly affecting the network's subsequent output. It is worth mentioning that the DIP algorithm, which is a highly nonlinear algorithm based on convolutional neural networks, uses different prior compared to conventional IR algorithms. This difference allows the information transmitted by the RBP connection to be more effectively utilized by the DIP algorithm, significantly enhancing its performance.

The difference is that RBP-DIP uses different update methods for the network input and network weights: network weights are updated through the highly nonlinear neural network by the backpropagation algorithm, while the network input is updated using a mostly linear IR algorithm. This addresses the homogeneity issue of network weight and network input in DIP methods. Furthermore, since different update algorithms are used, the information contained in the network input is unobtainable by the original DIP algorithm. Hence, the updated network input can also be regarded as a reference image to guide the reconstruction of the DIP algorithm. This also aligns with CDIP, but differs in that the reference image is automatically generated by RBP-DIP during reconstruction, eliminating the need for a high-quality reference image and avoiding potential interference from a suboptimal reference image.

The architecture of our framework, residual back projection with the deep image prior (RBP-DIP), is shown in Fig. 1. It contains a U-net and an RBP connection. The U-net, which has demonstrated its efficacy in various image processing problems, is used to implement the standard DIP algorithm in our approach. In each iteration, the input of the U-net is iteratively updated using the back projected residual between the sinogram of the current network output and the provided measurement via the RBP connection. This update incorporates the information from the conventional IR algorithm and can be used to enhance the output of the U-net. The RBP connection constitutes the primary distinction between our RBP-DIP model and the DIP model presented in Shu and Entezari (2022b) and Veen et al. (2018). Notably, the RBP connection does not participate in the backpropagation process. Also, batch normalization is still implemented in the RBP-DIP network. Although RBP-DIP directly optimizes the inference data, making its batch size equal to 1, batch normalization can still effectively address the issues of gradient explosion and gradient vanishing, making the neural network easier to converge.

An intuitive explanation of RBP-DIP is that it uses a conventional IR method to update network input z , and uses the network output x as the

input of the conventional IR method. Such an architecture is designed to take advantage of both the conventional IR method and the DIP method. The DIP property can help the untrained U-net generate high-quality images that are free from most artifacts existing in conventional IR methods, especially when the corresponding inverse problem is highly ill-posed. However, the DIP method is always considered unstable and may generate neural network-specific artifacts, as the whole neural network is randomly initialized and cannot be clearly explained. On the contrary, although the conventional IR methods are unable to generate accurate reconstruction results when the inverse problem is highly ill-posed, its convergence rate, stability, and robustness have already been well-analyzed. Thus, it can be used to guide the DIP method, correct the aforementioned network-specific artifacts, fine-tune the reconstruction result, and increase the robustness of the whole framework. As a result, the integration of the RBP connection and DIP method can amalgamate the advantages of both methods and mitigate each other's shortcomings.

The details of RBP-DIP are shown in Algorithm 1, where lines 2–4 correspond to the RBP connection. r in line 2 indicates the back-projected residual, which is the derivative of the data inconsistency $\|g - Ax\|^2$ with respect to x and is also a major component of most IR algorithms' updating equation. In the RBP-DIP algorithm, instead of updating x directly, r is used by the RBP connection to update the network's input z . The reason is that directly updating x will result in streak artifacts, as shown in Fig. 3. Updating z allows the deep image prior to help eliminate these artifacts. Additionally, it is postulated that the iteratively updated vector z can more effectively guide the model since it contains the information provided by the IR algorithm embedded in the RBP connection. Lines 5 and 6 correspond to the optimization of the untrained U-net using the backpropagation algorithm.

This paper focuses on demonstrating the advantages of the proposed RBP connection and the corresponding RBP-DIP framework over the original DIP algorithm. It also shows the great potential of replacing

the original DIP algorithm with the RBP-DIP algorithm in most DIP-related algorithms. Therefore, we propose using the basic setup similar to the original DIP algorithm. Thus, the objective function of RBP-DIP is just the data inconsistency $\|g - Ax\|_2^2$, and the IR algorithm integrated into the RBP connection is a simple back projection $r = A^T(g - Ax)$. We believe that if the RBP-DIP framework can succeed in the simplest setup, then it can achieve even better improvements with more sophisticated objective functions, IR algorithms (in the RBP connection), and network structures.

It is worth mentioning that substantial fluctuations in the input to the U-net could undermine its stability and overall performance. Consequently, we normalize the input z in line 4 to let the network input always be a unit vector. Moreover, the magnitude of RBP updates necessitates regulation, as the RBP updates would otherwise dominate the reconstruction process and lead to artifacts commonly seen in conventional IR methods. In that case, the residual r is also normalized in lines 3. Together with the function β , the RBP update in line 3 is a summation of the previous unit vector z and another vector of length β , which can be expressed as:

$$\beta = \frac{10^{-3}}{1 + e^{-\left(\frac{n}{n_s} - n_c\right)}}, \quad (3)$$

where n indicates the current number of iterations. β is a sigmoid function centered at n_c and stretched by the factor n_s . In the initial phase of the reconstruction procedure ($\frac{n}{n_s} \ll n_c$), the DIP property plays a crucial role in generating high-quality preliminary reconstruction outcomes. During this period, the RBP connection is suppressed by the factor β to avoid introducing artifacts commonly associated with IR methods. As the reconstruction progresses and $\frac{n}{n_s} \gg n_c$, the RBP connection's strength is augmented by the increasing β . In this later stage, the RBP connection utilized its embedded IR method to further refine the reconstruction without engendering additional artifacts, given that a high-quality preliminary reconstruction has already been generated. To achieve optimal performance, it is imperative for n_c to be sufficiently large to yield a smooth function, preventing significant disruptions to the network stability caused by input updates. The selection of n_c mainly depends on the nature of the inverse problem itself. In instances where the inverse problem is highly ill-posed, such as few-view and limited-angle CT reconstruction, conventional iterative reconstruction (IR) methods implemented in the proposed RBP connection may result in pronounced artifacts. In such cases, it is advisable to use a larger n_c . This adjustment gives the algorithm more time to leverage the properties of DIP to obtain a better initial guess. Conversely, if the inverse problem is closer to being well-posed or over-determined, a smaller n_c enables the RBP to take in the optimization process more promptly. In this paper, these parameters are approximately determined through our experiments and ensure that all algorithms perform well. While adjusting hyperparameters for different application scenarios can further enhance the performance, the gains from such fine-tuning are insignificant compared to the improvements achieved by RBP-DIP based on DIP. Moreover, this paper primarily aims to propose an RBP-DIP framework based on the original DIP algorithm, rather than optimizing the network for a specific image reconstruction task. As a result, we have set the parameters to $n_s = 250$, $n_c = 10$, and the total number of iterations is 5000. By employing these values, the RBP connection is suppressed during the first half of the reconstruction process and gradually augmented during the second half.

RBP-DIP, being related to DIP, currently lacks comprehensive theoretical proof of convergence. However, numerous experiments (Bagger et al., 2020; Gong, Catana, Qi, & Li, 2018; Mataev, Milanfar, & Elad, 2019; Shu & Entezari, 2022b; Ulyanov et al., 2018; Veen et al., 2018) by different researchers have also shown that the DIP network converges in most cases. The main difference between the RBP-DIP and the original DIP algorithm lies in the iterative updates to the network input z . Several studies (Cui et al., 2022; Mataev et al., 2019; Shu & Pan, 2024; Ulyanov et al., 2018) have shown that this not only

retains the stability and convergence performance of the algorithm but also improves the algorithm's noise resistance. In fact, in many DIP-related papers, including the original DIP paper, modifications are made by authors to the vector z in each iteration, such as adding a small Gaussian noise vector, to enhance the robustness of the algorithm. The RBP-DIP algorithm updates z with more meaningful information to achieve better performance, as demonstrated by our experiments. Furthermore, to mitigate potential adverse effects, we propose a conventional IR algorithm at line 7 as an optional post-processing step. This aims to ensure convergence. However, in all subsequent experiments presented in this paper, we omit the post-processing step and still observed excellent results.

4. Experiments and results

In this section, we first conduct experiments showing the effectiveness of the proposed RBP connection. We also compare the differences between the proposed framework and the well-established IR algorithm ASD-POCS (Sidky & Pan, 2008) under multiple conditions. In that case, the effect of employing an untrained neural network for iterative reconstruction can be assessed. In addition, the proposed framework undergoes comparative analysis against the well-known DIP model (Bagger et al., 2020), as well as two widely recognized pre-trained models MED50 (Jin, McCann, Froustey, & Unser, 2017) and RED-CNN (Chen et al., 2017). The structural similarities among these models show the merits of the proposed method while eliminating model complexity bias. The aforementioned methods are first tested under parallel-beam geometry with the reconstruction resolution 512×512 . The Gram filtering method proposed by Shu and Entezari (2022a) is used to calculate the forward and back projection exactly and efficiently. Next, 512×512 reconstructions under fan-beam geometry are conducted, where the ASTRA toolbox's (Aarle et al., 2015) line fanflat projector is used for forward and back projection. Images from LIDC-IDRI (Armato III et al., 2011) (The Lung Image Database Consortium image collection) dataset are used in our experiments. To address potential issues related to inverse crime, sinogram data g is acquired from upsampled (1024×1024) ground truth images. Finally, we tested our method on the real CT data from Finnish Inverse Problems Society (Bubba et al., 2017; Hämäläinen et al., 2015). The signal-to-noise ratio (dB) and the structural similarity index (SSIM) are calculated to evaluate the accuracy of reconstructions. However, in this paper, we only plot the SNR curve since the trends of both metrics are highly consistent.

It is worth mentioning that we have also tried other forward and back projection methods, including lookup table (Ha & Mueller, 2017), box spline (Zhang & Entezari, 2019a), and separable footprint (Long, Fessler, & Balter, 2010). The experiments show that the reconstruction accuracy of the RBP-DIP algorithm indeed varies depending on the accuracy of the chosen projection operator, but it is always significantly higher than that of conventional IR and DIP algorithms using the same projection operator.

The aforementioned neural networks are implemented in Python with the PyTorch library. The RMSProp algorithm serves as the optimizer, with a learning rate set to 1×10^{-4} , which decreases by a factor of 0.9 every 250 iterations. The hyperparameters utilized in the RBP-DIP framework are assigned the values $n_s = 250$, and $n_c = 10$. We have tried using methods such as uniform, normal, Xavier uniform, and Xavier normal distributions for network parameter initialization but found no significant differences. Therefore, the experiments in our paper all used PyTorch's default uniform initialization method.

It is worth mentioning that each pre-trained model is retrained for different settings such as the number of views, and projection angular range. Moreover, during the training procedure, test images are employed for the direct assessment of the pre-trained models, while validation images are not utilized in this context. Although this training strategy is normally unreasonable, the purpose of our experiment is not to train a stable pre-trained model. Instead, it is to show that in our experiments, the best performance a neural network can achieve on test images is still far below expectations.

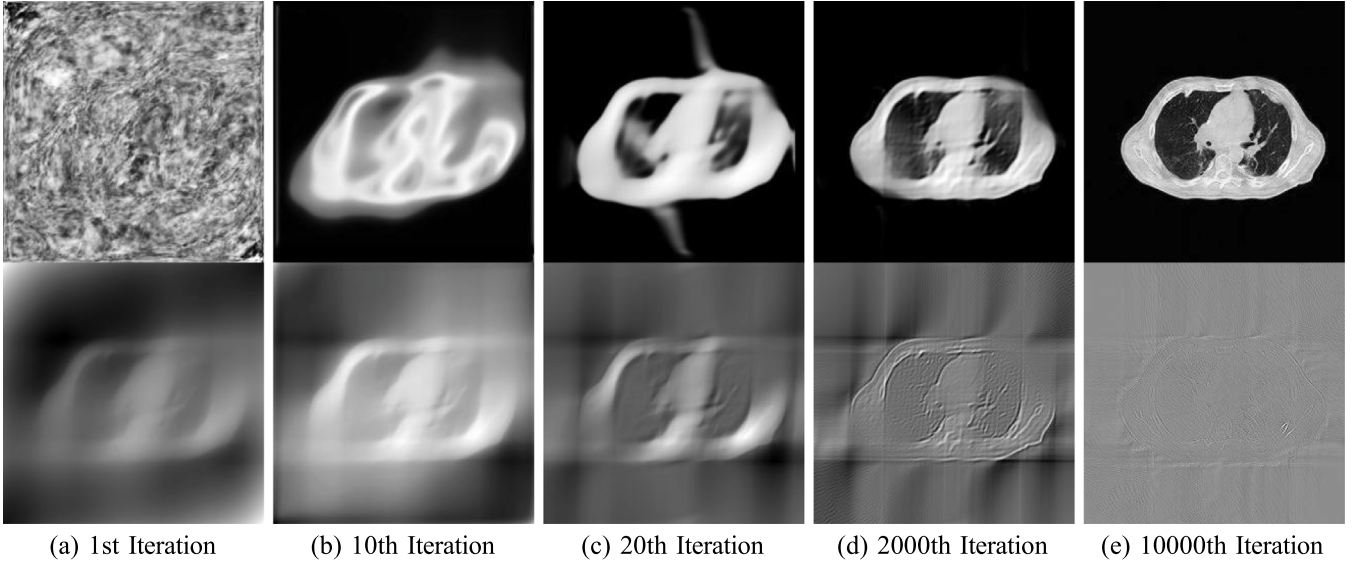


Fig. 2. The output (first row) and the corresponding input z (second row) at different iteration. The number of views is 90, uniformly distributed from 0° to 90° . Each figure in the second row has its histogram individually adjusted for better visualization.

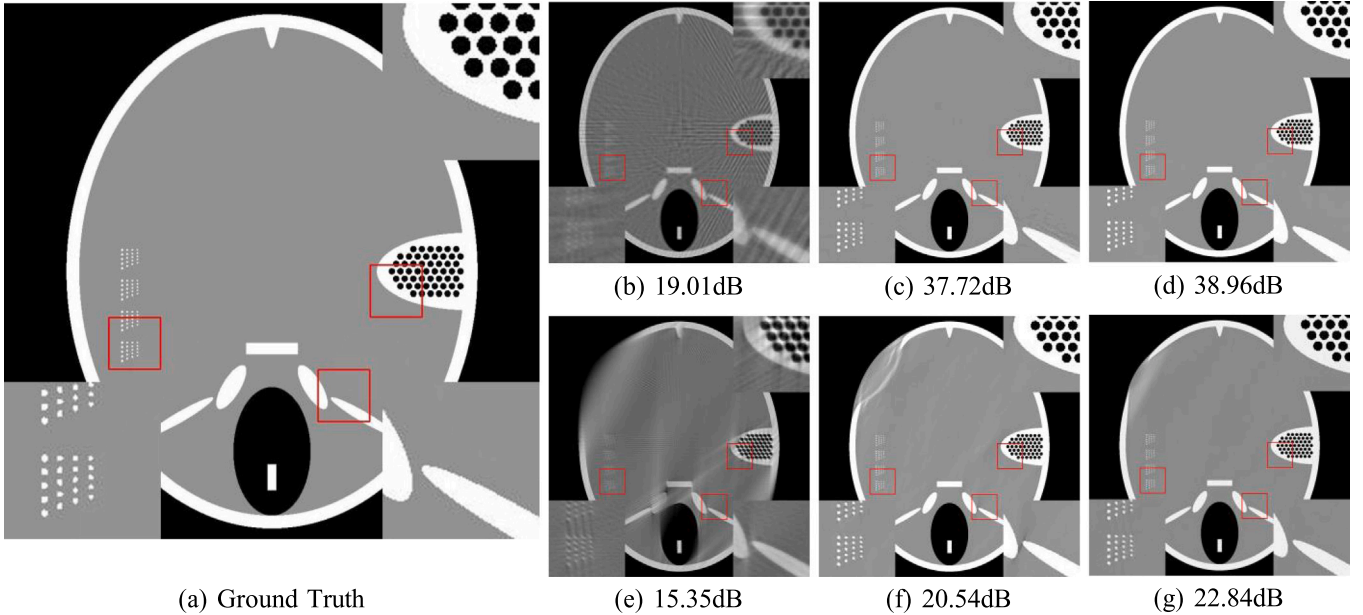


Fig. 3. The few-view (first row, 60 views which are evenly distributed from 0° to 360°) and limited-angle (second row, 120 views which are evenly distributed from 0° to 120°) reconstruction results of the Forbild phantom for ASD-POCS (b and e), DIP (c and f), and RBP-DIP (d and g). The proposed RBP-DIP method can generate high-quality reconstruction results with minimal artifacts.

4.1. Effect of residual back projection in reconstruction

To demonstrate the efficacy of the RBP connection and show the details of the RBP-DIP’s reconstruction process, the procedure of a limited-angle CT reconstruction is shown in Fig. 2. Its first row illustrates the results of reconstruction across different iterations, whereas the second row presents the respective inputs fed into the U-net, which are updated through the RBP connection. Here, the number of views is set to 90, uniformly distributed from 0° to 90° . This scenario presents a challenging limited-angle CT reconstruction problem.

In the first iteration, as depicted in Fig. 2a, the input image undergoes an update via the RBP connection prior to being fed to the U-net. Thus, this input is the normalized first iteration output of the implemented IR algorithm (normalized back projection image in our

case). The output appears completely randomized since the entire neural network is randomly initialized.

In the 10th and 20th iterations (Fig. 2b and c), the DIP property effectively expedites the recovery of the object over its support. Of note are lack of artifacts commonly caused by having missing views in the data. The input images highlight the region which can be relatively accurately reconstructed by conventional IR methods. This can be used to guide the model in the later iteration. In our experiment, the model capitalizes on the input images more when reconstructing the upper-left and lower-right segments of the image, while relying predominantly on the DIP property for the reconstruction of the upper-right and lower-left parts.

In the 2000th iteration, as depicted in Fig. 2d, the reconstruction result becomes relatively artifact-free. At this stage, the network input primarily emphasizes the edges to help the method improve the

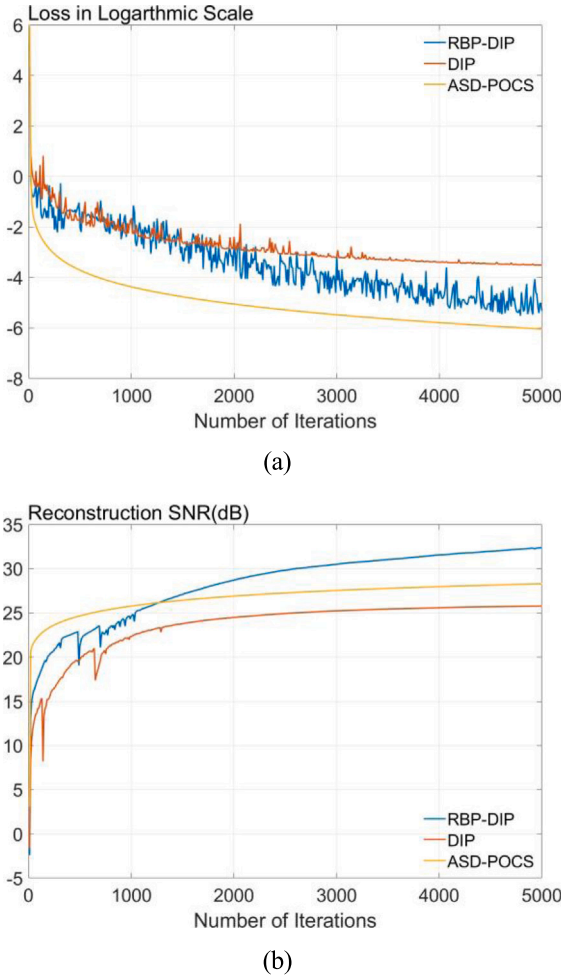


Fig. 4. The reconstruction loss (l_2 norm in logarithmic scale) and SNR (dB) of different methods under few-view condition (number of views is 90, uniformly distributed from 0° to 180°). On the one hand, RBP-DIP converges faster than DIP, demonstrating the effectiveness of our proposed RBP connection in improving the original DIP method. On the other hand, RBP-DIP achieves higher SNR with a loss larger than the ASD-POCS method. This demonstrates the superiority of the RBP-DIP algorithm, as such improvement is not merely caused by the further minimization of the loss function, but rather from the utilization of a better prior (the successful combination of conventional IR and deep image prior).

supporting area. Moreover, the RBP connection can rectify artifacts specific to convolutional neural networks. Evidence of this can be observed in the second row of Fig. 2c, d, and e, which display distinct horizontal and vertical patterns. These patterns are mainly caused by the convolution operation in the U-net. In other words, the DIP and RBP parts of the proposed framework are able to mutually rectify each other’s errors. Consequently, a high-quality reconstruction result is attainable, as shown in Fig. 2e.

4.2. Effect of residual back projection with deep image prior

A direct comparison among the proposed RBP-DIP, the widely used IR method ASD-POCS, and the current untrained method (Bagner et al., 2020) is made to show the advantages of the RBP-DIP method. Given that the number of views is set to 90, uniformly distributed from 0° to 180° , the scenario presents a challenging few-view CT reconstruction problem. The reconstruction performance in terms of reconstruction loss and SNR is shown in Fig. 4. From Fig. 4a, it is evident that compared to DIP, RBP-DIP can better minimize the objective function. Although the ASD-POCS method can further minimize the objective

function, Fig. 4b shows that RBP-DIP achieves the highest SNR. This indicates that when the problem is highly ill-posed, where multiple solutions can minimize the objective function, the result chosen by RBP-DIP is better than others. Another thing worth mentioning is that the loss of RBP-DIP shown in Fig. 4a is fluctuating, which is in line with our analysis in Section 3.2. This implies that using the conventional IR method to fine-tune the output RBP-DIP can further improve its reconstruction results.

For a more direct comparison, the three aforementioned methods are tested on the Forbild phantom under few-view (60 views which are evenly distributed from 0° to 360°), and limited-angle (120 views which are evenly distributed from 0° to 120°). From Fig. 3, it is clear that both the DIP and RBP-DIP exhibit superior performance compared to ASD-POCS, particularly in the reconstruction of detailed structures. The magnified views of these structures are available in the bottom left and top right corners of these images. In contrast to DIP, RBP-DIP is better at mitigating some neural network-specific artifacts, with the corresponding magnified view presented in the bottom right corner of these images.

To better evaluate the effectiveness of the RBP-DIP framework in real-world scenarios, multiple experiments using real CT images are conducted in the following sections.

4.3. Few-view CT reconstruction

In this section, the reconstruction performance of our proposed framework under few-view conditions will be tested. For the parallel and fan beam geometry, the number of views increases from 30 to 180, uniformly distributing from 0° to 180° and 0° to 360° respectively. Such settings provide a complete benchmark of reconstruction performance, ranging from extremely sparse to relatively complete, full-view CT reconstruction. The experiment results are shown in Fig. 5. Additionally, the ground truth, few-view (30 views), and full-view (180 views) CT reconstruction results of different methods are shown in the first and third rows of Fig. 7 (parallel-beam, LIDC-IDRI dataset), and Fig. 8 (fan-beam, LIDC-IDRI dataset).

4.4. Limited-angle CT reconstruction

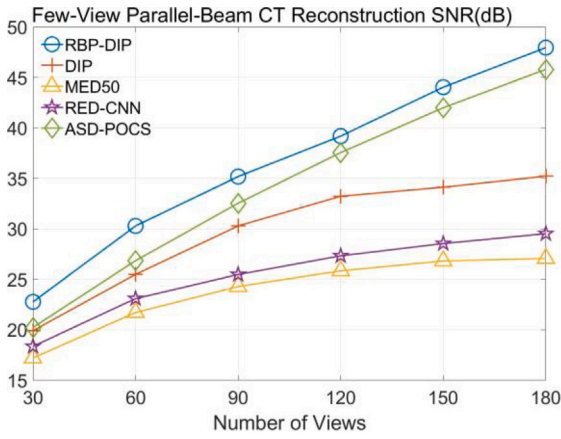
To test the proposed framework’s performance on limited-angle reconstruction, we redo the experiment in the above section with the angular range changing from 0° – 90° to 0° – 165° for parallel-beam geometry and 0° – 90° to 0° – 180° for fan-beam geometry, one projection per degree. The experiment results are shown in Fig. 6. Also, the ground truth and the limited-angle CT reconstruction results of different methods are shown in the row of Fig. 7 (parallel-beam, LIDC-IDRI dataset), and Fig. 8 (fan-beam, LIDC-IDRI dataset).

4.5. Non-ideal factors

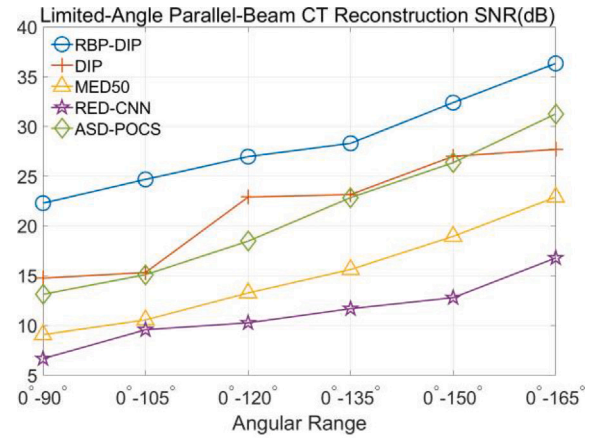
As previously mentioned, one of the most significant advantages of our proposed method lies in the elimination of the training process. This advantage precludes any potential hindrances arising from an insufficient training procedure or inconsistencies between inference input and training dataset. To examine this assertion, experiments under multiple non-ideal conditions have been conducted.

4.5.1. Rotation

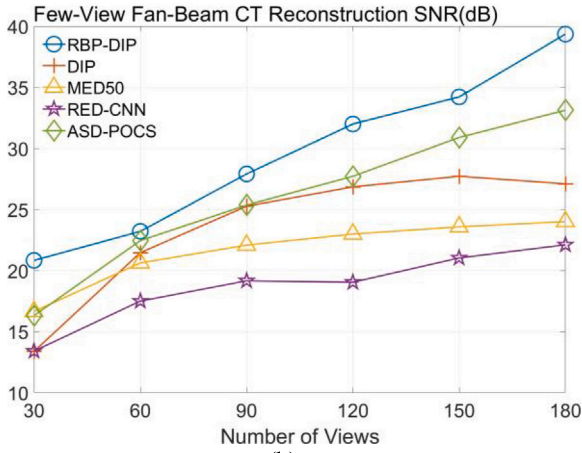
The few-view CT reconstruction experiment is replicated, utilizing an identical test image rotated by 30° degrees to simulate a rudimentary pose alteration of the patient. The reconstruction performances are shown in Fig. 9, and the reconstruction results corresponding to 30 views are shown in Fig. 10. A similar experiment is also conducted under the limited-angle condition, yielding similar results.



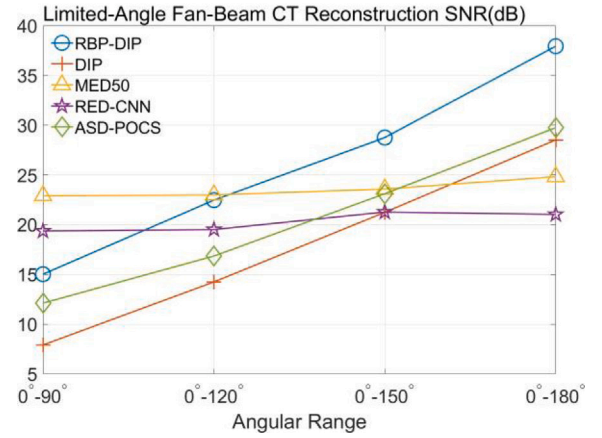
(a)



(a)



(b)



(b)

Fig. 5. The performance of few-view CT reconstruction for different methods under different numbers of views and geometries. The RBP-DIP method outperforms all other methods in all cases.

4.5.2. Quantum noise

To evaluate the efficacy of the proposed algorithm under conditions of reduced dosage, the previously described reconstruction experiments are replicated utilizing sinogram data contaminated with noise following a Poisson distribution. The mean count of X-ray photons detected by the i th detector can be expressed as:

$$E_i = I_0 e^{-\mu g_i}.$$

In this experiment, a multiplication factor denoted by $\mu = 0.183/\text{cm}$ was employed. This factor corresponds to an energy level of 80 keV, which approximates the mean energy of the beam utilized in a diagnostic CT scan conducted at a voltage of 120 kVp (Yu et al., 2012). Throughout the experiment, I_0 , the blank measurement, is assigned values ranging from 10^2 to 10^8 , which correspond to sinogram SNR (dB) values of 25.6, 35.6, ..., 85.6. It is worth mentioning that analogous experiments were conducted under both few-view and limited-angle scenarios, with the proposed method demonstrating a marked superiority over other methods across all conditions. This enhancement, however, is primarily attributed to the exceptional performance of the proposed method under sparse-measurement circumstances, rather than an increased resistance to noise. To facilitate an equitable comparison, the results displayed in Fig. 11 are solely derived from the full-view scenario, which encompasses an angular range of 0° - 180° with one view per degree. It is worth mentioning that this experiment solely focuses on the resilience of the RBP-DIP framework itself against noise, thus not

Fig. 6. The performance of limited-angle CT reconstruction for different methods under different angular ranges and geometries. The RBP-DIP method outperforms all other methods in most cases.

employing any existing methods such as total variation regularization and DeepRED (Mataev et al., 2019) for noise mitigation. However, these methods can be easily integrated into the RBP-DIP framework, as it shares a similar workflow with DIP.

4.6. Real CT reconstruction

To further verify the effectiveness of the RBP-DIP framework, real CT data from the Finnish Inverse Problem Society (Bubba et al., 2017; Hämäläinen et al., 2015) was used to validate our algorithm. The few-view and limited-angle reconstruction results are shown in Fig. 12. Due to the lack of available training sets, pre-trained models were not included in the experiments. Different from the previous experiment, data from Finnish Inverse Problem Society is highly noisy. However, it is still evident that the RBP-DIP outperforms all other algorithms in all experiments, even under high noise and highly ill-posed conditions.

5. Discussion

5.1. Comparisons on DIP, IR, and RBP-DIP

It is worth mentioning that both DIP and RBP-DIP are iterative reconstruction methods since all these methods reconstruct images by minimizing an objective function on inference data iteratively. This minimization process offers a dependable lower bound, ensuring that

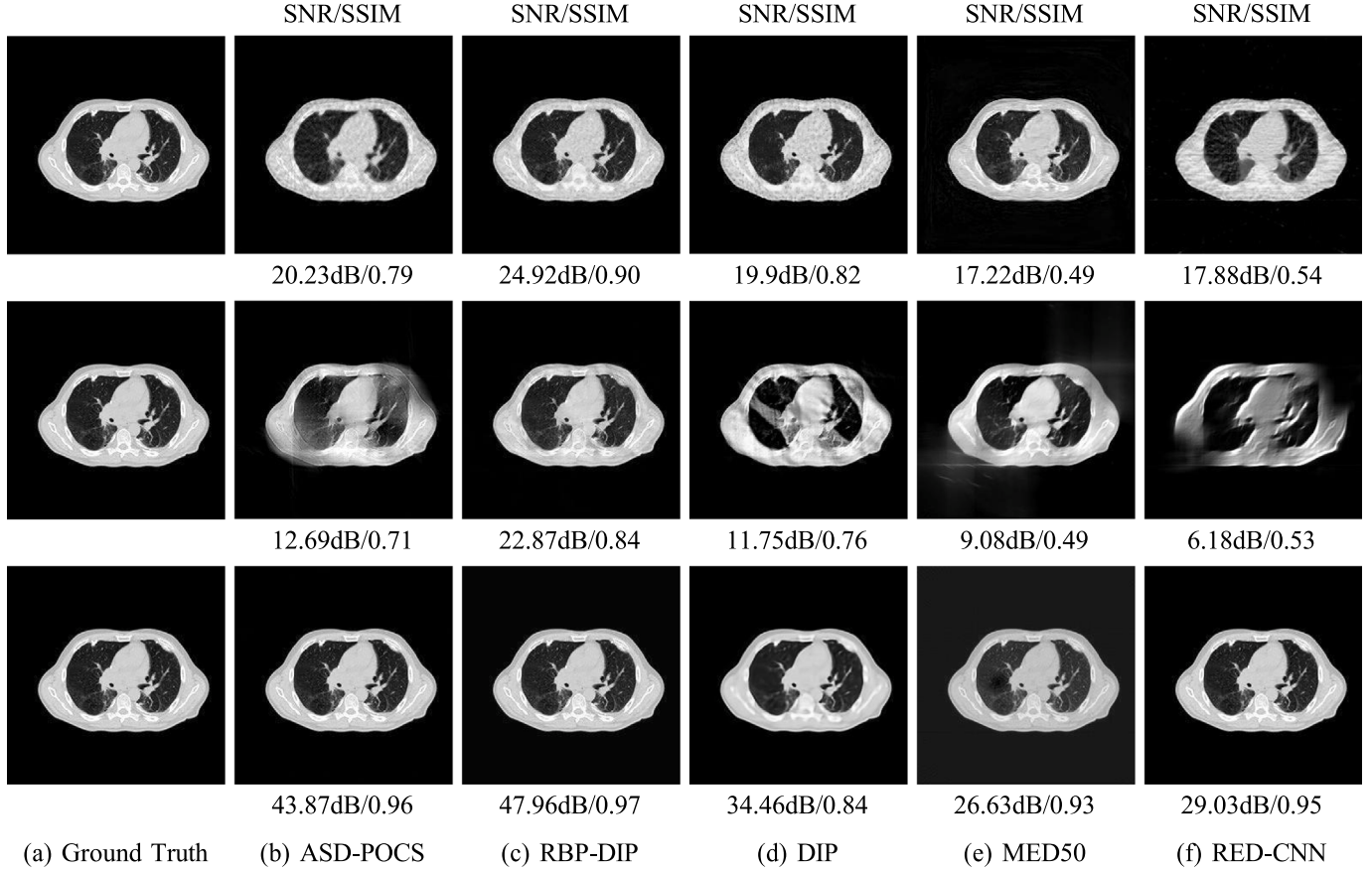


Fig. 7. The parallel-beam reconstruction results of LIDC-IDRI dataset for different methods under few-view (first row, 30 projections distributed uniformly from 0° to 180°), limited-angle (second row, 90 projections distributed uniformly from 0° to 90°) and full-view (third row, 180 projections distributed uniformly from 0° to 180°) conditions. The RBP-DIP method outperforms all other methods in all cases.

their performance does not fall below that of conventional IR methods if both obtain similar loss. The distinction between conventional IR and DIP related methods lies in the use of an untrained convolutional neural network.

Conventional IR methods, even those with the help of regularizations such as total variation, are prone to artifacts when constrained by few-view and limited-angle conditions. However, as illustrated in the first and second rows of Figs. 7b, and 8b, these images still contain meaningful information which can be used to guide DIP related reconstruction methods, despite the presence of artifacts.

DIP related methods, which leverage the hierarchical structure of neural networks as a powerful prior, can better handle the aforementioned challenge. However, the original DIP method has its own limitations. It cannot generate detailed images or effectively enhance its accuracy as the number of measurements increases. For instance, in Fig. 5a, the ASD-POCS algorithm achieves an approximate SNR gain of 25 dB when the number of views increases from 30 to 180, while the DIP method only attains an approximate 15 dB gain. This problem is also shown in the last row of Figs. 7, and 8. Moreover, the DIP method may produce neural network specific artifacts, as shown in Figs. 7d, 8d, and 12c. These artifacts are particularly problematic as they are often considered more undesirable than streak artifacts. Radiologists, with their professional experience, can interpret and account for streak artifacts, whereas network specific artifacts may prove more challenging to identify and address.

The proposed RBP-DIP framework combines the two approaches utilizing the newly devised RBP connection so that inherits the benefits of both methods. In Fig. 4, the RBP-DIP method's attainment of a 5 dB SNR enhancement over the ASD-POCS method, despite exhibiting a larger loss. This implies that the improvement is not caused by the

further minimization of the loss function but by the ability to choose a more reasonable reconstruction result from all candidates with similar losses. This is precisely the direct evidence of the successful combination of conventional IR and deep image prior in the proposed RBP-DIP framework. Subsequently, the improvement surpassing the original DIP method indicates the efficacy of the RBP connection. Moreover, by employing the RBP connection, neural network specific artifacts can be rectified effectively. As a result, substantial advancements can be shown in Figs. 5, 6, 7, and 8.

5.2. Comparing with pre-trained models

We have also conducted comparative analyses with pre-trained models MED50 and RED-CNN under multiple conditions. These models share similar network structures and complexities with the RBP-DIP framework. We have opted not to examine models with more complicated networks, as the proposed method is untrained, necessitating a focus on the structural properties of the networks for equitable comparisons. Theoretically, it is expected that pre-trained models should obtain better reconstruction results than that of the proposed method, especially when the number of measurements is not enough, as they can acquire extra information from training datasets. However, as shown in Figs. 5 and 6, RBP-DIP outperforms the two pre-trained models in most cases, with the only exception being limited-angle fan-beam reconstruction where the angular range is 0° - 90° .

Pre-trained models' low performances may be caused by the limited number of views. In the original research, the network inputs are high-quality FBP images (with 400 views), which possess sufficient projections and relatively minimal artifacts. In that case, the

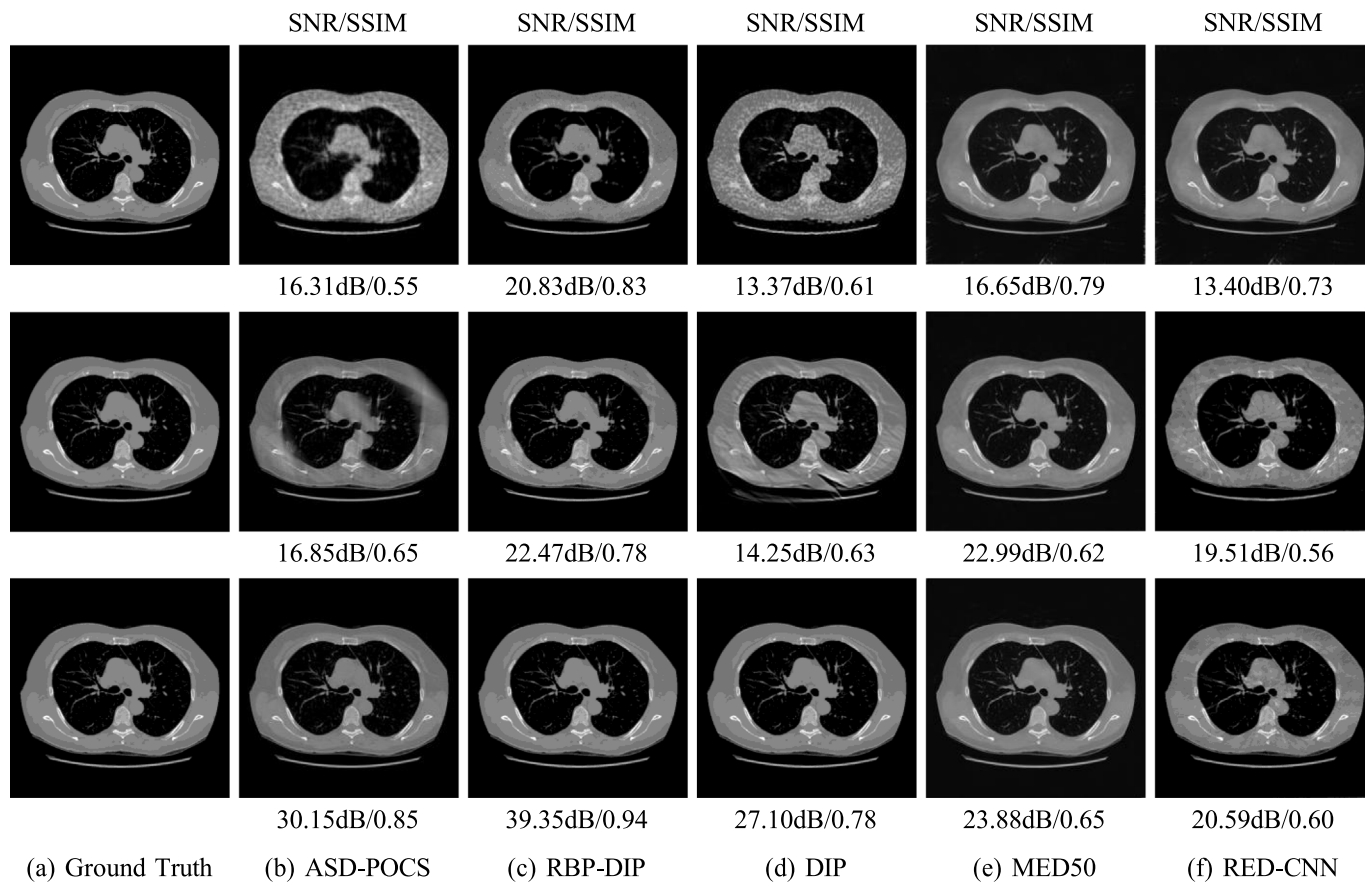


Fig. 8. The fan-beam reconstruction results of LIDC-IDRI dataset for different methods under few-view (first row, 30 projections distributed uniformly from 0° to 360°), limited-angle (second row, 120 projections distributed uniformly from 0° to 120°) and dense-view (third row, 180 projections distributed uniformly from 0° to 360°) conditions. The RBP-DIP method outperforms all other methods in most cases.

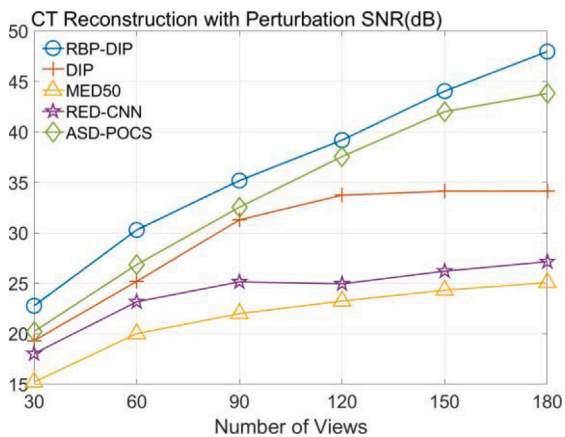


Fig. 9. The performance of few-view CT reconstruction with perturbation (rotated by 30°) for different methods.

inconsistencies between the input and the corresponding output are relatively small, so that pre-trained models can learn the correct prior with limited training data. In our experiments, the FBP images always suffer from severe artifacts, and the aforementioned differences are relatively large. As a result, pre-trained models cannot be trained properly without a large, high-quality training dataset. In fact, some training instances in the current training dataset may even downgrade the performance of pre-trained models due to discrepancies in postures and CT slices, as illustrated in Fig. 13.

Also, Figs. 5 and 6, particularly Fig. 6b, show that an increasing number of views does not sufficiently lead to improved performance of pre-trained models. The reason is that pre-trained models aim to learn the mapping between the input and the corresponding output from training datasets, rather than actually solve the corresponding inverse problem. Increasing the number of views cannot directly strengthen this mapping. Conversely, the proposed RBP-DIP directly minimizes the inconsistency between the ground truth and reconstructed images under the same measurements. Increasing the number of views reduces the dimensions of the solution space and thus benefits both network optimization and the iterative reconstruction (IR) algorithms integrated in the RBP connection. This outcome is further verified in the last row of Figs. 7, and 8, where RBP-DIP considerably outperforms other methods.

5.3. Non-ideal factors

In practical applications, the imaged object might exhibit inconsistencies with the training dataset due to various non-ideal factors. The 30° rotation operation in Section 4.5 serves as a simplified simulation of such discrepancies, which may arise from differing poses of patients. The impact on reconstruction accuracy is shown in Fig. 9. The comparison between Figs. 5b and 9 reveals that the pre-trained MED50 method is predominantly affected by a considerable decline in SNR. The reconstruction results are also presented in Fig. 10. In contrast to the first row of Fig. 7, the MED50 reconstruction result lacks a completely black background and exhibits horizontal artifacts in the non-empty region. This observation suggests that MED50 cannot effectively handle the perturbations in the inference data. A similar result is also obtained when adding quantum noise to the sinogram.

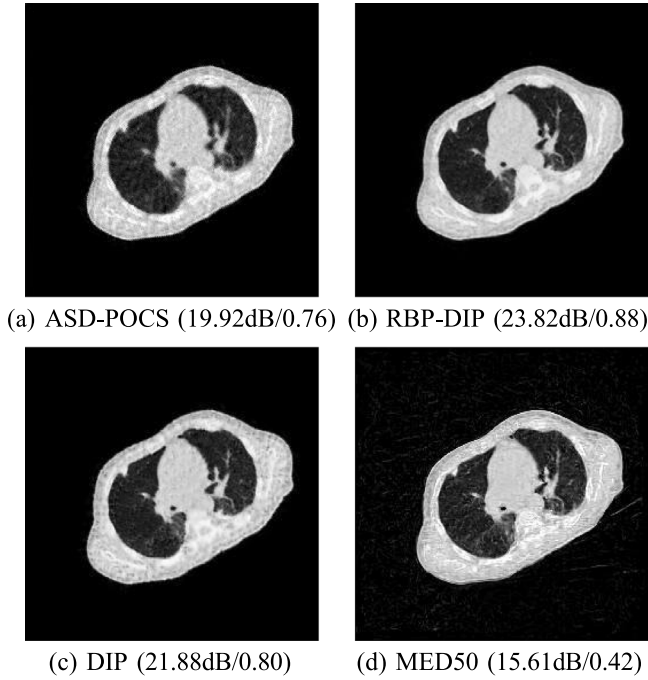


Fig. 10. The reconstruction results and corresponding SNR/SSIM of the few-view (30 views) CT with perturbation (rotate by 30°) using different methods. The pre-trained model suffers from artifacts caused by the rotation.

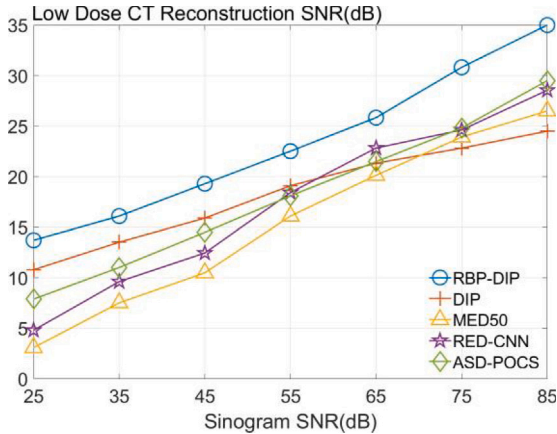


Fig. 11. Impact of quantum noise on reconstruction SNR(dB). The RBP-DIP is the least susceptible to noise.

The proposed RBP-DIP framework does not necessitate any training images and attains the highest reconstruction accuracy in both experiments. Moreover, additional constraints or regularizations can be directly incorporated into the objective function or indirectly integrated into the RBP connection, enabling enhanced handling of various factors without retraining.

5.4. Future works and prospects

This study shows that the RBP-DIP framework shows improvements over the original DIP method and two pre-trained methods with similar network structures. Also, the RBP-DIP framework has great potential for further improvement. The RBP connection in this paper is implemented by a basic IR algorithm with the objective function $\|g - Ax\|^2$ without any additional constraints or regularizations. However, it is evident that RBP-DIP is compatible with all other IR methods, constraints, and regularizations. Notably, these techniques can be incorporated into both the

objective function and the RBP connection. In contrast to pre-trained models, these adjustments can be adapted on demand (e.g., augmenting the total variation regularization's weight when dealing with noisy input data). Moreover, the proposed framework employs a basic U-net structure, but more sophisticated models such as U-net++ (Zhou, Rahman Siddiquee, Tajbakhsh, & Liang, 2018) and ResUnet (Zhang, Liu, & Wang, 2018) might be integrated for further improvement.

As an enhancement of the DIP method, the challenges faced by the RBP-DIP algorithm are similar to the DIP:

- **Computational time:** Unlike the inference process of pre-trained neural networks, where the network is initialized with learned weights, the RBP-DIP network starts from a random initialization, which implies that the proposed method is time-consuming. The reconstruction process of the RBP-DIP algorithm is the same as the training process of an ordinary neural network. The only difference is that the DIP and RBP-DIP algorithms need to be optimized on a single inference case, while ordinary neural networks need to be optimized on the entire training set. Considering the powerful parallel processing capabilities of current GPUs, the time to perform one forward and backward pass on a single case is almost the same as that for an entire batch of cases. In that case, the reconstruction speed of the RBP-DIP algorithm is $\frac{N}{M}$ times the training speed of a neural network with a training set size of N , a batch size of M , and the same network structure. This also matches our experimental results. In our future research, we will dedicate our efforts to exploring strategies such as employing a warm start and simplifying the network architecture to accelerate the RBP-DIP framework. Existing frameworks such as ADMM can also accelerate our method, which has been proven by researchers such as Mataev et al. (2019).
- **Noise sensitivity:** Although superior to the original DIP algorithm, RBP-DIP remains relatively sensitive to noise. Thanks to the similarity in the workflows of the RBP-DIP and DIP, most noise mitigation methods (Jo, Chun, & Choi, 2021; Liu et al., 2019; Mataev et al., 2019) applicable to DIP are also suitable for RBP-DIP. Therefore, in future research, we will dedicate efforts to integrating the RBP-DIP algorithm with existing noise reduction techniques to further enhance its performance.

6. Conclusion

In this paper, we present an untrained neural network related framework, RBP-DIP, for X-ray CT reconstruction. Remarkable reconstruction results are achieved using RBP-DIP without any training process or extra regularization, particularly under highly ill-posed conditions. The improvements are primarily attributable to the employment of the untrained neural network and the RBP connection, which utilize the advantages of DIP and conventional IR algorithms, respectively.

In comparison with conventional IR methods, RBP-DIP attains a higher reconstruction SNR, especially when the corresponding inverse problem is highly ill-posed. In comparison with the original DIP method, RBP-DIP exhibits superior accuracy and is free from neural network specific artifacts. These findings suggest that our proposed methodology successfully utilizes the advantages of both IR and DIP methods.

Comprehensive comparisons are also conducted with pre-trained models possessing similar network structures to RBP-DIP. The results indicate that obtaining a well-trained model necessitates a substantial, high-quality training dataset. The experiment subjected to perturbations reveals that a pre-trained network is biased toward training data. Such a problem is even more severe in the field of medical imaging, as deviation from normal is the key to diagnosis. In contrast, our untrained RBP-DIP leverages the hierarchical properties of convolutional neural networks without capitalizing on what the training data deems normal. These significant improvements further substantiate the superiority of RBP-DIP.

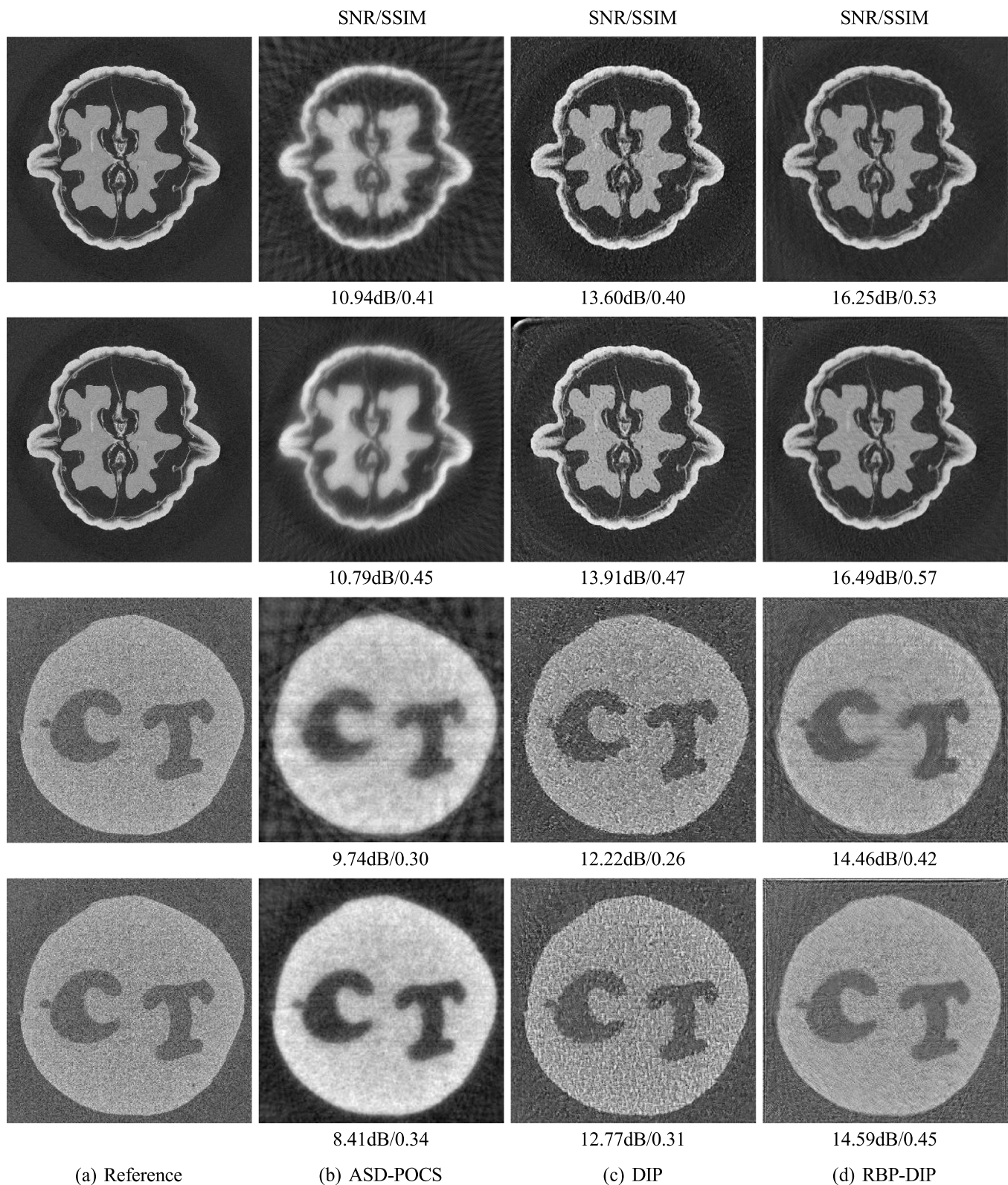


Fig. 12. The fan-beam reconstruction results of a walnut and a carved cheese. The data of the reference images is from the Finnish Inverse Problems Society. The first and third rows correspond to few-view CT reconstructions (30 projections distributed uniformly from 0° to 360°). The second and fourth rows correspond to limited-view CT reconstructions (60 projections distributed uniformly from 0° to 120°). Each column shows the reconstruction results of different methods. The first column shows the reference images which are generated by the FBP algorithm using full-view and high-resolution data.

The proposed framework can handle multiple geometries and non-ideal factors. Such versatility exists thanks to the U-net architecture's widespread use in 2D and 3D images and the thorough analysis of IR

algorithms in the RBP connection. In fact, the RBP-DIP framework is not limited to CT and has the potential to solve other reconstruction problems, as it only requires an untrained model for DIP image generation

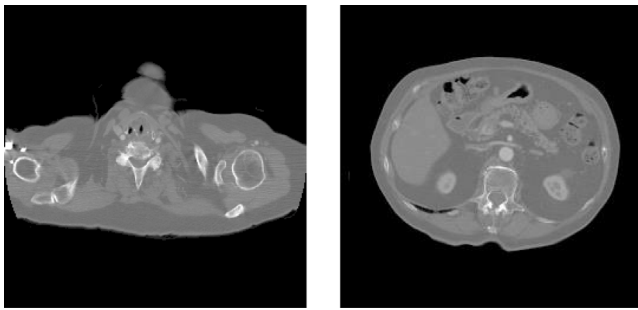


Fig. 13. Training instances in LIDC-IDRI dataset. Some of the training instances may even downgrade the performance of pre-trained models due to the inconsistency in postures and CT slices.

and a conventional IR method for residual back projection. Thanks to its concise architecture, further improvements such as employing more delicate IR methods and neural networks are also possible. Additional priors can also be incorporated as needed in both the RBP connection and the objective function.

In conclusion, the proposed framework further improves the original DIP algorithm for CT reconstruction, particularly under conditions of few views and limited angles.

CRediT authorship contribution statement

Ziyu Shu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alireza Entezari:** Supervision, Funding acquisition, Writing – review & editing,.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alireza Entezari reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Antun, V., Renna, F., Poon, C., Adcock, B., & Hansen, A. C. (2020). On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proceedings of the National Academy of Sciences*, 117(48), 30088–30095.

Armato III, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., et al. (2011). The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical Physics*, 38(2), 915–931.

Baquer, D. O., Leuschner, J., & Schmidt, M. (2020). Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems*, 36(9), Article 094004.

Bhadra, S., Kelkar, V. A., Brooks, F. J., & Anastasio, M. A. (2021). On hallucinations in tomographic image reconstruction. *IEEE Transactions on Medical Imaging*, 40(11), 3249–3260.

Bouman, C., & Sauer, K. (1993). A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Transactions on Image Processing*, 2(3), 296–310.

Bubba, T. A., Juvonen, M., Lehtonen, J., März, M., Meaney, A., Purisha, Z., et al. (2017). Tomographic X-ray data of carved cheese. arXiv preprint arXiv:1705.05732.

Chen, H., Zhang, Y., Zhang, W., Liao, P., Li, K., Zhou, J., et al. (2017). Low-dose CT denoising with convolutional neural network. In *2017 IEEE 14th international symposium on biomedical imaging* (pp. 143–146). IEEE.

Cui, J., Gong, K., Guo, N., Kim, K., Liu, H., & Li, Q. (2022). Unsupervised PET logan parametric image estimation using conditional deep image prior. *Medical Image Analysis*, 80, Article 102519.

Fessler, J. A. (2014). Vol. 2, *Comprehensive biomedical physics* (pp. 263–295). Elsevier.

Gong, K., Catana, C., Qi, J., & Li, Q. (2018). PET image reconstruction using deep image prior. *IEEE Transactions on Medical Imaging*, 38(7), 1655–1665.

Gong, C., & Zeng, L. (2020). Self-guided limited-angle computed tomography reconstruction based on anisotropic relative total variation. *IEEE Access*, 8, 70465–70476.

Gottschling, N. M., Antun, V., Hansen, A. C., & Adcock, B. (2023). The troublesome kernel – On hallucinations, no free lunches and the accuracy-stability trade-off in inverse problems.

Ha, S., & Mueller, K. (2017). A look-up table-based ray integration framework for 2-D/3-D forward and back projection in X-ray CT. *IEEE Transactions on Medical Imaging*, 37(2), 361–371.

Hämäläinen, K., Harhanen, L., Kallonen, A., Kujanpää, A., Niemi, E., & Siltanen, S. (2015). Tomographic X-ray data of a walnut. arXiv preprint arXiv:1502.04064.

Hu, D., Zhang, Y., Liu, J., Du, C., Zhang, J., Luo, S., et al. (2021). SPECIAL: single-shot projection error correction integrated adversarial learning for limited-angle CT. *IEEE Transactions on Computational Imaging*, 7, 734–746.

Jin, X., Li, L., Chen, Z., Zhang, L., & Xing, Y. (2010). Anisotropic total variation for limited-angle CT reconstruction. In *IEEE nuclear science symposium & medical imaging conference* (pp. 2232–2238). IEEE.

Jin, K. H., McCann, M. T., Froustey, E., & Unser, M. (2017). Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9), 4509–4522.

Jo, Y., Chun, S. Y., & Choi, J. (2021). Rethinking deep image prior for denoising. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 5087–5096).

Kisner, S. J., Haneda, E., Bouman, C. A., Skatter, S., Kourinny, M., & Bedford, S. (2012). Model-based CT reconstruction from sparse views. In *Second international conference on image formation in x-ray computed tomography* (pp. 444–447).

Lee, H., Lee, J., Kim, H., Cho, B., & Cho, S. (2018). Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3(2), 109–119.

Li, Z., Chang, D., Zhang, Z., Luo, F., Liu, Q., Zhang, J., et al. (2024). Dual-domain collaborative diffusion sampling for multi-source stationary computed tomography reconstruction. *IEEE Transactions on Medical Imaging*.

Liu, J., Sun, Y., Xu, X., & Kamilov, U. S. (2019). Image restoration using total variation regularized deep image prior. In *ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing* (pp. 7715–7719). Ieee.

Long, Y., Fessler, J. A., & Balter, J. M. (2010). 3D forward and back-projection for X-ray CT using separable footprints. *IEEE Transactions on Medical Imaging*, 29(11), 1839–1850.

Ma, Y., Wei, B., Feng, P., He, P., Guo, X., & Wang, G. (2020). Low-dose CT image denoising using a generative adversarial network with a hybrid loss function for noise learning. *IEEE Access*, 8, 67519–67529.

Mataev, G., Milanfar, P., & Elad, M. (2019). DeepRED: Deep image prior powered by RED. In *Proceedings of the IEEE/CVF international conference on computer vision workshops*.

Pan, J., Yu, H., Gao, Z., Wang, S., Zhang, H., & Wu, W. (2024). Iterative residual optimization network for limited-angle tomographic reconstruction. *IEEE Transactions on Image Processing*.

Shah, V., & Hegde, C. (2018). Solving linear inverse problems using gan priors: An algorithm with provable guarantees. In *2018 IEEE international conference on acoustics, speech and signal processing* (pp. 4609–4613). IEEE.

Shen, C., Ma, G., & Jia, X. (2019). Low-dose CT reconstruction assisted by a global CT image manifold prior. Vol. 11072, In *15th international meeting on fully three-dimensional image reconstruction in radiology and nuclear medicine*. International Society for Optics and Photonics, Article 1107205.

Shu, Z., & Entezari, A. (2020). Gram filtering and sinogram interpolation for pixel-basis in parallel-beam X-ray CT reconstruction. In *2020 IEEE 17th international symposium on biomedical imaging* (pp. 624–628). IEEE.

Shu, Z., & Entezari, A. (2022a). Exact gram filtering and efficient backprojection for iterative CT reconstruction. *Medical Physics*, 49(5), 3080–3092.

Shu, Z., & Entezari, A. (2022b). Sparse-view and limited-angle CT reconstruction with untrained networks and deep image prior. *Computer Methods and Programs in Biomedicine*, 226, Article 107167, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S016926072200548X>.

Shu, Z., & Pan, Z. (2024). SDIP: Self-reinforcement deep image prior framework for image processing. arXiv preprint arXiv:2404.12142.

Sidky, E. Y., Kao, C.-M., & Pan, X. (2006). Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *Journal of X-Ray Science and Technology*, 14(2), 119–139.

Sidky, E. Y., & Pan, X. (2008). Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in Medicine and Biology*, 53(17), 4777.

Ulyanov, D., Vedaldi, A., & Lempitsky, V. (2018). Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 9446–9454).

- Van Aarle, W., Palenstijn, W. J., De Beenhouwer, J., Altantzis, T., Bals, S., Batenburg, K. J., et al. (2015). The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography. *Ultramicroscopy*, 157, 35–47.
- Veen, D. V., Jalal, A., Soltanolkotabi, M., Price, E., Vishwanath, S., & Dimakis, A. G. (2018). Compressed sensing with deep image prior and learned regularization.
- Wang, Y., Li, Z., & Wu, W. (2024). Time-reversion fast-sampling score-based model for limited-angle ct reconstruction. *IEEE Transactions on Medical Imaging*.
- Wang, T., Nakamoto, K., Zhang, H., & Liu, H. (2017). Reweighted anisotropic total variation minimization for limited-angle CT reconstruction. *IEEE Transactions on Nuclear Science*, 64(10), 2742–2760.
- Wang, G., Ye, J. C., Mueller, K., & Fessler, J. A. (2018). Image reconstruction is a new frontier of machine learning. *IEEE Transactions on Medical Imaging*, 37(6), 1289–1296.
- Willemink, M. J., & Noël, P. B. (2019). The evolution of image reconstruction for CT? from filtered back projection to artificial intelligence. *European Radiology*, 29, 2185–2195.
- Wu, W., Hu, D., Niu, C., Yu, H., Vardhanabhuti, V., & Wang, G. (2021). DRONE: Dual-domain residual-based optimization network for sparse-view CT reconstruction. *IEEE Transactions on Medical Imaging*, 40(11), 3002–3014.
- Wu, W., Wang, Y., Liu, Q., Wang, G., & Zhang, J. (2023). Wavelet-improved score-based generative model for medical imaging. *IEEE Transactions on Medical Imaging*.
- Xie, H., Shan, H., Cong, W., Liu, C., Zhang, X., Liu, S., et al. (2020). Deep efficient end-to-end reconstruction (DEER) network for few-view breast CT image reconstruction. *IEEE Access*, 8, 196633–196646.
- Ye, D. H., Buzzard, G. T., Ruby, M., & Bouman, C. A. (2018). Deep back projection for sparse-view CT reconstruction. In *2018 IEEE global conference on signal and information processing* (pp. 1–5). IEEE.
- Yin, X., Zhao, Q., Liu, J., Yang, W., Yang, J., Quan, G., et al. (2019). Domain progressive 3D residual convolution network to improve low-dose CT imaging. *IEEE Transactions on Medical Imaging*, 38(12), 2903–2913.
- Yu, Z., Noo, F., Dennerlein, F., Wunderlich, A., Lauritsch, G., & Hornegger, J. (2012). Simulation tools for two-dimensional experiments in x-ray computed tomography using the FORBILD head phantom. *Physics in Medicine and Biology*, 57(13), N237.
- Zhang, R., Bouman, C. A., Thibault, J.-B., & Sauer, K. D. (2013). Gaussian mixture Markov random field for image denoising and reconstruction. In *2013 IEEE global conference on signal and information processing* (pp. 1089–1092). IEEE.
- Zhang, K., & Entezari, A. (2019a). Box spline projection in non-parallel geometry. In *2019 IEEE 16th international symposium on biomedical imaging* (pp. 1844–1847). IEEE.
- Zhang, K., & Entezari, A. (2019b). A convolutional forward and back-projection model for fan-beam geometry. arXiv preprint arXiv:1907.10526.
- Zhang, Q., Gao, J., Ge, Y., Zhang, N., Yang, Y., Liu, X., et al. (2020). PET image reconstruction using a cascading back-projection neural network. *IEEE Journal of Selected Topics in Signal Processing*, 14(6), 1100–1111.
- Zhang, Y., Hu, D., Zhao, Q., Quan, G., Liu, J., Liu, Q., et al. (2021). CLEAR: comprehensive learning enabled adversarial reconstruction for subtle structure enhanced low-dose CT imaging. *IEEE Transactions on Medical Imaging*, 40(11), 3089–3101.
- Zhang, H., Li, L., Qiao, K., Wang, L., Yan, B., Li, L., et al. (2016). Image prediction for limited-angle tomography via deep learning with convolutional neural network.
- Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), 749–753.
- Zhang, J., Mao, H., Wang, X., Guo, Y., & Wu, W. (2024). Wavelet-inspired multi-channel score-based model for limited-angle CT reconstruction. *IEEE Transactions on Medical Imaging*.
- Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support* (pp. 3–11). Springer.
- Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R., & Rosen, M. S. (2018). Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697), 487–492.