Personalized Patient-adaptive Sparse-View CT Deep Reconstruction

Abstract. Image reconstruction is an important and challenging problem in medical imaging. Sparsely sampled measurements pose additional challenges due to the ill-posedness of the forward operator. Existing methods solve this problem by leveraging the prior knowledge of the images such as sparsity, smoothness or learned images distribution. However, these methods are subjected to bias due to regularization which hampers the reconstruction quality. To address this issue, we propose a novel corrector algorithm that can improve the reconstruction quality of an existing black-box reconstruction model without any training data. Our method is also flexible to any image modalities and adaptive to any existing supervised linear inverse problem solving algorithms. Specifically, we first train a neural representation algorithm that focuses on learning and correcting the residual signal, and then combine it with a back-projection algorithm to output the combined correction image. Empirically, we observe better performance of our method to other existing correction techniques in sparse-view CT reconstruction.

1 Introduction

Sparse-View CT reconstruction is a challenging problems. The forward Radon transform operator is well-known to be ill-posed, which is sensitive to noise with direct inversion. The null space is also very large that poses additional challenges to the exact recovery. Before the era of deep learning, methods focus on adding prior information to the loss function [23]. Prior information includes smoothness of image (TV-norm), sparsity in a transformed domain (e.g. wavelet transform) or some advanced techniques such as group sparsity.

With the advance of deep learning techniques, many models are able to generate realistic high-resolution images with the help of large training data-set and incorporating physical constraints in the deep learning model [22]. However deep-learning-based are susceptible to instability and network bias [7], which may cause performance to deteriorate with data distributional shift. They also require a vast amount of dataset for training and tuning, which may not be available in clinical practice.

Since there are already a lot of excellent models that train on large dataset, and give excellent results within the distribution, now a new problem arises: given a reasonable reconstruction model, can we further improve it in an unknown testing environment with very limited validation dataset?

This work studies how to remove the artifacts and recover fine details from the outputs of an existing deep-learning-based sparse-view CT reconstruction algorithm with limited training data.

In this work, we propose to address the challenges above by applying corrector algorithms to remove the bias and recover fine details from the an existing deep learning model output. We consider the challenging case that the noise level is unknown in the testing environment, and only one ground truth image is available in the validation set for us to tune our algorithms. We propose a novel Residual-NeRP-Ensemble method that tackles the challenges very well. Our contributions can be summarized as below.

- We propose a novel corrector algorithm incorporating implicit regularization, prior knowledge and data consistency through neural representation learning to improve existing sparse-view ct reconstruction algorithms. Our method requires no training data from external subjects and can be easily generalized across different imaging modalities.
- Our algorithm is flexible and could be adaptive to improve any existing supervised linear inverse problem solving algorithms, and may be extended to unsupervised algorithms.

- We perform extensive experiments on the performance of different corrector algorithms on sparse-view ct reconstruction tasks and demonstrate excellent performance of our proposed method. To our knowledge, we are the first to study the performance of a variety of correction algorithms on an existing inverse problem solver.

2 Related Works

There are a lot of work trying to solve the sparse-view ct reconstruction problem. Bora et al [23] proposed an unsupervised generative framework (CSGM) that enables optimizing in the latent space of a pretrained deep generative space. They argued that the trained latent space provides a very strong prior comparing to other methods such as Lasso-DCT. However, their results show poor image quality when number of projection increases and cannot be applied for CT reconstruction. Dhar et al [19]. proposed to improve the CSGM model by adding a sparse deviation to it and train simultaneously. However, their results were still far from satisfactory. Hussein et al [15] proposed an image-adaptive approach that use the CSGM results as the pretrained weight and unfreezes the weights for a step-2 optimization. They also propose a simple but effective way to improve the CSGM result: projecting the CSGM output to $R(A^T)$. Their results show significant improvement over the CSGM outputs, but still not satisfying the resolution requirement for medical imaging. In recent years, score-based generative models show excellent performance of generative image qualities. Jalal et al. proposed a score-based compressed sensing MRI framework that shows excellent performance and robust to distributional shift [7]. Song et al. [3] further improved the score-based method by simultaneously do the unconditional sampling and conditional correction to make sure that the samples are close to the posterior distribution $p(x|y)$. However, their methods require a lot of training data and can be very sensitive to hyper-parameter tuning.

Another approach to address the challenge is the direct supervised approach (i.e. pairing each ground truth image with the signal). Many existing works show superior performance of reconstruction quality. Wei et al. [4] proposed a two-step ct reconstruction algorithm based on adversarial training that achieves very high image quality for sparse-view reconstruction with only 23 projection angles. Also, other works focus on incorporating the geometry constraints of the measurement process into the model. Shen et al. proposed a geometry informed framework for 3D reconstruction. [21]. Chen et al. [9] proposed a deep decomposition learning framework for separately learning null space, row space and noise portion of the signal. Zhang et al. [1] proposed a novel idea that use a existing reconstruction model as the prior image. Then apply the PICCS (Prior image constrained compressed sensing) framework to improve the results from that existing model. However, all of these works require a large dataset for training, and may not be able to adapt to a new data modality or distributional shift.

With the advance of neural representation learning in recent years, some emerging work tries to solve the sparse-view ct reconstruction problem with neural representation learning. Shen et al. proposed to use an MLP with SIREN [13] layers for sparse-view CT reconstruction only with one prior image and the ground truth signal. However, this work requires a prior scan of the same patient, which may not be available in real clinical settings. Sun et al [6] tries to inpaint the sinogram and gives reasonable results. Lindell et al. [10] tries to further improve the SIREN network by an automatic integration network that inpaints the sinogram. However, these sinogram completion quality does not satisfy the requirement of medical imaging.

3 Methods

3.1 Residual-NeRP

The classical Pior-Image Constraint Compressed Sensing (PICCS) formula is given by

$$\hat{x} = \arg\min_{x} \lambda||Ax - y||_2^2 + \alpha||\Psi_1(x - x_p)||_1 + (1-\alpha)||\Psi_2(x)||_1$$

We can let x_p be the image reconstructed by FBP-UNet.

We have $y = Ax_{true} + \epsilon$, where ϵ is noise. x_{true} is the ground truth image. The above formula can be reframed as

$$\hat{x} = \arg\min_{x} \lambda||A(x - x_p) + (Ax_p - y)||_2^2 + \alpha||\Psi_1(x - x_p)||_1 + (1-\alpha)||\Psi_2((x - x_p) + x_p)||_1$$

Let $x - x_p$ be the residual image that we are interested in. We can use neural representation to represent the residual image using neural representation learning. Let M_θ be the output of the neural representation, we can train it to represent $x - x_p$. We do this because the we assume that prior model has already learned the structure of the image, and we only need to learn the remaining structure that the prior model fails to learn. Hence the objective in our case is given by

$$\hat{x} = \arg\min_{\theta} \lambda||AM_\theta + (Ax_p - y)||_2^2 + \alpha||\Psi_1(M_\theta)||_1 + (1-\alpha)||\Psi_2((M_\theta) + x_p)||_1$$

In our case, we let Ψ_1 be identity operator and Ψ_2 be the TV operator in constrast to both being TV operator as in (Zhang, 2021)'s work.

Lemma 1

The neural representation output can be designed to be band-limited, which is determined by the network architecture. $\mathcal{F}M_\theta(\omega)$ is compacted supported.

By Lemma 1, we impose implicit constraint on the residual image by neural representation learning. We can also add explicit constraints directly since the gradient of the NeRP is easy to compute.

We observe that classical methods need lots of hyper-parameters tuning carefully to achieve the best performance when noise level changes, but the optimal hyper-parameters of Residual-NeRP are insensitive to noise levels. This may be explained by that Residual-NeRP has already imposed implicit regularization in the network. This property makes Residual-NeRP suitable for unknown noise levels. Nevertheless, when the SNR of residual signal becomes closer to 1, corrector algorithms may not be suitable in those senarios since the majority of the improvement comes from the data-fidelity term.

3.2 Residual Ensemble

Ensemble methods achieves a great success in machine learning applications, but there is few study about ensembling in compressive sensing. Instead of ensembling the full reconstruction images, we ensemble the residual reconstruction images. Since different compressive sensing methods use different regularizations, and it is possible to gather information from all of these methods. We propose that we can improve the model performance by using a linear combination of the residual reconstruction images from different methods. Suppose we have n models. Let s_i be the residual image reconstruction from the i_{th} model. The final reconstruction image we propose is given by

$$x_{final} = x_p + \sum_{i=1}^{n} \alpha_i s_i$$

where α_i is the weight of each residual image reconstruction which is between 0 and 1.

Lemma 2 Optimal ensemble weight

Suppose that the residual image reconstructions are random variables that are independent of each other, and their expectations are orthogonal to each other. Let l denote the ground truth residual image; let $\sigma_1^2, ..., \sigma_n^2$ denote the variance; let $e_1, ..., e_n$ denote the expectations of the reconstructed residual images. Then the optimal weight $\hat{\alpha_i}$ is given by

$$\hat{\alpha_i} = \frac{l^T e_i}{e_i^T e_i + \sigma_i^2}$$

Suppose that we are only ensembling two residual images, then the optimal weight of $\hat{\alpha_i}$ is given by

$$\hat{\alpha_1} = \frac{4l^T e_1(e_2^T e_2 + \sigma_2^2) - 2E[s_1^T s_2]l^T e_2}{4(e_1^T e_1 + \sigma_1^2)(e_2^T e_2 + \sigma_2^2) - E[s_1^T s_2]^2}$$

We observe that when the bias decreases, the optimal weight increases; when the variance increases, the optimal weight decreases. When two residual image reconstructions are positively correlated with each other, if one is closer to the ground truth residual image, the weight of that residual image reconstruction is higher than the other one. Practically, we tune the weights based on the validation set, and may adjust the weight based on the change in data distribution. If we have the knowledge that the performance of an estimator will deteriorate in the testing environment, we may decrease the weight of that estimator and increase the weight of other estimators. For example, we know ridge regression is sensitive to hyper-parameter changes. When we are in a testing environment with unknown noise level, we expect ridge regression to perform worse than in the validation set. Hence, we can decrease the weight of ridge regression estimator and increase the weights of other estimators.

Lemma 3 Variance reduction of residual signal for radon transform

$$Var(A^T y) = \frac{1}{n} Var(y)$$

where n is the size of the image

By this idea, we propose to directly apply ridge regression on the residual signal. The results demonstrates good performance. We then ensemble the residual-ridge output to other residual reconstructions.

3.3 Residual-NeRP Ensemble

The final framework is demonstrated in the graph below. We first use a existing model to produce a prior image. Then we compute the residual signal. With the residual signal, we perform the residual image reconstruction by Residual-NeRP, Residual-Ridge and PICSS. Finally, we ensemble the residual reconstructions together using the Residual Ensemble technique above.

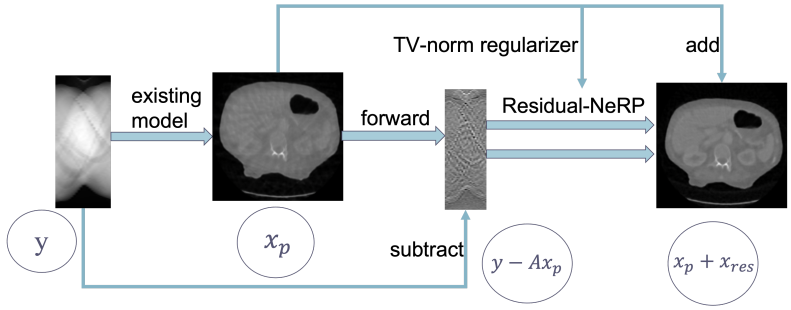

Fig 1: An overview of our method

4 Experiments

The CT scans of multiple anatomic sites, including head, chest, and abdomen from 50 patients are used in this study. The full-dose images are used as the ground truth for evaluation. We select images from 40 patients for training. Images from the remaining 10 patients are used for testing. For the training set, 3 images from each anatomic region are randomly sampled for tuning hyperparameters of the corrector algorithms. Images are normalized to the interval of [0,1]. CT measurements (sinograms) are simulated with parallel-beam geometry using 25 projection angles distributed evenly across 180 degrees using "torch-radon" package. An additional independent Gaussian noise was added to the sinograms. The standard deviation of the noise was set to 0.003 for the training set and 0.001 in the testing set. All models do not have any information about the noise level. A deep-learning-based sparse-view reconstruction model (FBP-Convolutional neural network) was trained on data from 40 patients. 10 slices from each anatomic site are randomly sampled for each patient in the testing set. The performance of four corrector algorithms (back-projection, ADMM-CNN, PICCS, Residual-NeRP ensemble) together with the trained deep-learning model was evaluated on the 300 images from the 10 remaining patients. The reconstruction image quality is quantified by PSNR (peak-signal-to-noise-ratio) and SSIM (structural-similarity-index). The results are summarized in Table 1. Each anatomic site has 100 images for evaluation.

Table 1: Results

Method	Abdominal		Head		Chest	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
FBP-Conv	32.78	0.938	31.91	0.919	29.78	0.885
+BP	32.28	0.918	31.76	0.919	30.30	0.888
+ADMM	32.84	0.942	33.40	0.945	29.67	0.892
+PICCS	34.70	0.959	35.44	0.964	31.12	0.913
Ours	**35.75**	**0.968**	**36.85**	**0.969**	**31.54**	**0.920**

We also perform the qualitative analysis. We found that our prior CNN model distorts the details of the images significantly and introduce a artificial noise on the image. This may due to the network bias of the CNN model. PICSS recovers the image well but due to the high TV-norm constraint, it blur out some details of the images. Residual-NeRP does not need that amount of regularization, so it recovers the better details of the images. The ensemble gives us an even better reconstruction results.

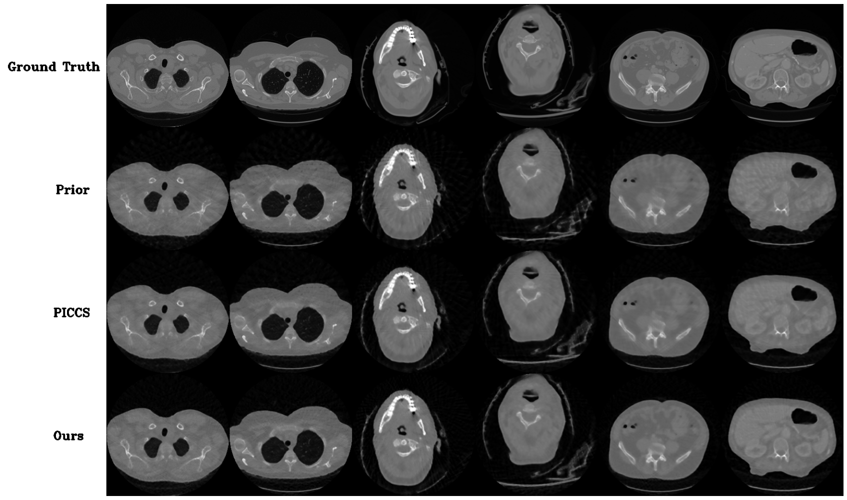

Fig 2: Visualization of our result

5 Conclusion

Residual-NeRP Ensemble, PICCS, and ADMM-CNN are able to improve the reconstruction quality of the given deep-learning model. Residual-NeRP ensemble gives the greatest improvement in terms of PSNR and SSIM on all anatomic sites of the 10 patients in the evaluation set with an average gain of 3.22db in PSNR. Residual-NeRP ensemble is able to recover the fine details and remove the artifacts from the given deep-learning reconstruction. This work indicates the potential of removing bias and recovering fine details from a deep-learning-based CT reconstruction algorithm. The corrector algorithms look promising in improving the reconstruction quality of deep-learning-based models. This work can be extended to 3D cone-beam CT reconstruction and accelerated MRI reconstruction in future investigations.

References

1 Zhang, Chengzhu, Yinsheng Li, and Guang-Hong Chen. "Accurate and robust sparse-view angle CT image reconstruction using deep learning and prior image constrained compressed sensing (DL-PICCS)." Medical Physics 48.10 (2021): 5765-5781.

2 Shen, Liyue, John Pauly, and Lei Xing. "NeRP: Implicit Neural Representation Learning with Prior Embedding for Sparsely Sampled Image Reconstruction." arXiv preprint arXiv:2108.10991 (2021).

3 Song, Yang, et al. "Solving Inverse Problems in Medical Imaging with Score-Based Generative Models." arXiv preprint arXiv:2111.08005 (2021).

4 Haoyu Wei, Florian Schiffers, Tobias Würfl, Daming Shen, Daniel Kim, Aggelos K Katsaggelos, and Oliver Cossairt. 2-step sparse-view ct reconstruction with a domain-specific perceptual network. arXiv preprint arXiv:2012.04743, 2020.

5 Lindell, David B., et al. ”BACON: Band-limited Coordinate Networks for Multiscale Scene Representation.” arXiv preprint arXiv:2112.04645 (2021).

6 Sun, Yu, et al. ”Coil: Coordinate-based internal learning for imaging inverse problems.” arXiv preprint arXiv:2102.05181 (2021).

7 Jalal, Ajil, et al. ”Robust compressed sensing mri with deep generative priors.” Advances in Neural Information Processing Systems 34 (2021).

8 Antun, Vegard, et al. ”On instabilities of deep learning in image reconstruction and the potential costs of AI.” Proceedings of the National Academy of Sciences 117.48 (2020): 30088-30095.

9 Chen, Dongdong, and Mike E. Davies. ”Deep decomposition learning for inverse imaging problems.” European Conference on Computer Vision. Springer, Cham, 2020.

10 Lindell, David B., Julien NP Martel, and Gordon Wetzstein. ”Autoint: Automatic integration for fast neural volume rendering.” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.

11 Ganaie, M. A., and Minghui Hu. ”Ensemble deep learning: A review.” arXiv preprint arXiv:2104.02395 (2021).

12 Xiang, Jinxi, Yonggui Dong, and Yunjie Yang. ”FISTA-net: Learning a fast iterative shrinkage thresholding network for inverse problems in imaging.” IEEE Transactions on Medical Imaging 40.5 (2021): 1329-1339.

13 Sitzmann, Vincent, et al. ”Implicit neural representations with periodic activation functions.” Advances in Neural Information Processing Systems 33 (2020): 7462-7473.

14 Tancik, Matthew, et al. ”Fourier features let networks learn high frequency functions in low dimensional domains.” Advances in Neural Information Processing Systems 33 (2020): 7537-7547.

15 Hussein, Shady Abu, Tom Tirer, and Raja Giryes. ”Image-adaptive GAN based reconstruction.” Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. No. 04. 2020.

16 Lee, Jaeho, et al. ”Meta-Learning Sparse Implicit Neural Representations.” Advances in Neural Information Processing Systems 34 (2021).

17 Ghani, Muhammad Usman, and W. Clem Karl. ”Deep learning-based sinogram completion for low-dose CT.” 2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). IEEE, 2018.

18 Gilton, Davis, Greg Ongie, and Rebecca Willett. ”Neumann networks for linear inverse problems in imaging.” IEEE Transactions on Computational Imaging 6 (2019): 328-343.

19 Dhar, Manik, Aditya Grover, and Stefano Ermon. ”Modeling sparse deviations for compressed sensing using generative models.” International Conference on Machine Learning. PMLR, 2018.

20 Van Veen, Dave, et al. ”Compressed sensing with deep image prior and learned regularization.” arXiv preprint arXiv:1806.06438 (2018).

21 Shen, Liyue, Wei Zhao, and Lei Xing. ”Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning.” Nature biomedical engineering 3.11 (2019): 880-888.

22 Shen, Liyue, et al. ”A Geometry-Informed Deep Learning Framework for Ultra-Sparse 3D Tomographic Image Reconstruction.” arXiv preprint arXiv:2105.11692 (2021).

23 Bora, Ashish, et al. ”Compressed sensing using generative models.” International Conference on Machine Learning. PMLR, 2017.
24 Beck, Amir, and Marc Teboulle. ”A fast iterative shrinkage-thresholding algorithm for linear inverse problems.” SIAM journal on imaging sciences 2.1 (2009): 183-202.
25 Candès, Emmanuel J., and Michael B. Wakin. ”An introduction to compressive sampling.” IEEE signal processing magazine 25.2 (2008): 21-30.