

Research Article

NeuRec: Incorporating Interpatient prior to Sparse-View Image Reconstruction for Neurorehabilitation

Cong Liu ^{1,2,3} Qingbin Wang ⁴ and Jing Zhang ¹

¹Faculty of Business Information, Shanghai Business School, Shanghai 200235, China

²The Affiliated Changzhou No. 2 People's Hospital of Nanjing Medical University, Changzhou 213003, China

³Center of Medical Physics, Nanjing Medical University, Changzhou 213003, China

⁴Ruijin Hospital, Shanghai Jiao Tong University, Shanghai 200031, China

Correspondence should be addressed to Jing Zhang; zhangj@sbs.edu.cn

Received 21 February 2022; Accepted 31 March 2022; Published 9 May 2022

Academic Editor: Boya Nugraha

Copyright © 2022 Cong Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical imaging technologies such as computed tomography (CT) and magnetic resonance imaging (MRI) imaging are indispensable for contemporary neurorehabilitation diagnostics, intervention, and monitoring. It would be desirable to reconstruct images from sparse measurements to reduce the ionizing radiation and motion artifacts. Although recent coordinate-based representation methods have shown promise advances for sparse-view reconstruction, they overfit a single MLP on a single patient. In this work, we generalize it across many patients by incorporating an interpatient prior into the ill-posed inverse/reconstruction problem, which is the missing ingredient in the previous works. The experiment demonstrates that our method significantly improves image quality over the state-of-the-art both qualitatively and quantitatively. Thus, our method provides a powerful and principled means to deal with the measurement-scarce problem.

1. Introduction

The majority of patients that enter neurorehabilitation therapies will have undergone some kind of medical imaging [1]. Computed tomography (CT) and magnetic resonance imaging (MRI) are the two most common modalities for neuroimaging. CT scans of the brain are much faster than MRI making it the ideal choice in cases of trauma and other acute neurological emergencies, where a rapid decision regarding appropriate medical management is imperative. The soft-tissue details generated by a CT are usually acceptable for clinicians to assess the acute damage, while MRI has a much greater range of available soft-tissue contrast. Although their broad application in daily rehabilitation practice, CT, and MRI have had their own weakness since their early days, CT imaging involves the use of x-rays, which are a form of ionizing radiation. Exposure to ionizing radiation is known to increase the risk of cancer. MRI requires a longer scan time than CT to collect sufficient data to form an image thereby being sensitive to excessive patient motion during

the actively scanning. The effects of motion include blurring and ghosting in the resulting image [2]. One possible method of addressing these issues is to reconstruct CT or MRI from partial measurements such as sparse-view projections/sinograms and low-resolution k -spaces. The sparse-view restriction reduces the requirements of the number of measurements thereby lowering the ionizing radiation for patients and suppressing motion artifacts in clinical MRI.

However, reconstructing images from limited data, as well as the ill-posed inverse problem, is very challenging to be solved. For example, for sparse-view sinograms, the corresponding projection matrix or Radon transform has more columns than rows. It can be shown that null vectors always exist for this underdetermined linear system, which means the solution for this system is not unique; more than one image may yield the same measurement. We can obtain a reasonable solution from the least-square minimum norm, but the solution is usually smoothed version of the original because the null vectors that are not measurable are disregarded. To obtain a better reconstruction, the conventional

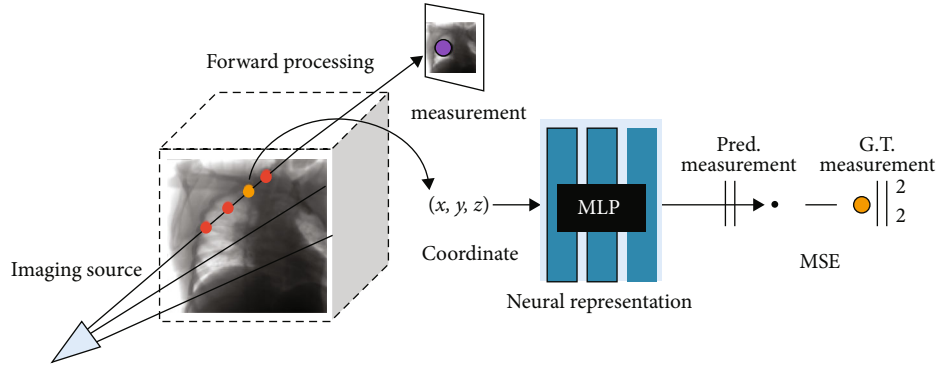


FIGURE 1: A schematic diagram of the neuRec.

methods [3–6] add the prior information of the unknown images such as sparse prior (e.g., total variation) and redundant prior (nonlocal mean) to the null space. These methods have been shown to achieve accurate recovery of the real images under some circumstances. Machine learning methods [7–9] especially the convolutional neural networks (CNNs) have recently emerged as the state of the art for solving the inverse problem though clinical-sized applications have yet to be demonstrated.

The CNN-based reconstruction methods can be categorized into three main directions. The first one is the learned postprocessing methods [10] where a learned operator is applied to the images that are reconstructed by conventional methods such as filtered back-projection (FBP) [11]. This operator performs the enhancement, denoising, and sharpness to improve the image quality. This type of method is relatively easy to be implemented because the physics principles are not involved. The second method is to learn the prior and use this prior in the classical iterative reconstruction formulation [12]. This kind of priors is learned from data, rather hand-designed, based on a large training patient dataset. The learned priors introduce the data priors to constraint the null space showing promise results though clinical-sized applications have yet to be demonstrated. The final method directly learns the reconstruction from the partial measurements in a data-driven style [13]. The partial measurements are directly mapped to the images by supervised training on a large dataset consisting of the pairs of measurements and images. Although these methods show outstanding performance for medical image reconstruction, they also have some shortages. For example, the requirement of a large-scale training dataset may prevent the CNNs from generalizing to the unseen view angles. Another issue is their memory-hungry feature inheriting from the dense voxel-based representation. While dense voxel-based representations are fast to query, they are memory inefficient, and 3D CNNs, potentially operating on these volumes, are computationally heavy.

To overcome the aforementioned problem, we propose neuRec (neural reconstruction) framework which is based on the recent implicit neural representation [14]. Instead of storing human tissues using voxels, it defines a coordinate-based multilayer perceptron (MLP) to approximate the quantities of interest, e.g., electronic density in CT. The

MLP maps continuous coordinates to voxel values, which is in contrast to conventional CNNs methods that propagate and backpropagate the 3D voxels tensor through the entire network. The coordinate-based MLP representation offers potential compact representation in terms of memory efficiency. Moreover, theoretically, it presents voxels at infinite resolution because continuous coordinate is used as the input of the network. The continuous assumption is also a useful characteristic for the sparse-view reconstruction, which enforces spatial coherence to the null space. The spatially nearby voxels are constraint to be associated with the same value implicating human tissues change slowly over space. The spatial coherence prior can also be combined with other priors for medical image reconstruction as shown in recent works [15, 16]. However, only interpatient priors such as previous CT are employed in these works. Our neuRec differs from them in that the interpatient prior is exploited for the patient in hand.

2. Methods

2.1. NeuRec. As illustrated in Figure 1, we aim to learn a coordinate-based representation $f_\theta : X \rightarrow Y$ where the voxel coordinate $x \in X$ is mapped to voxel value $y \in Y$, and the mapping is parameterized with θ . Given the limited number of physic measurements z and their corresponding geometry specifics, we can approximate the measurements using linear perturbation theory with forwarding process $z = Fy$, where F is the linearized measurement operator or Jacobian. If the measurement is performed M times and the image has N voxels, then F has the shape of $M \times N$. For sparse-view measurements, the transform F is a short, fat matrix, i.e., $M \ll N$.

The reconstruction or inverse problem is to the unknown image y when z and F are given. To solve y , we turn to an optimal mapping f_θ^* that can be obtained by the lowest average mean-squared distance to the true measurement z :

$$\operatorname{argmin}_\theta E_{(x,z) \sim \mu} [\|F(f_\theta(x)) - z\|_2^2], \quad (1)$$

where $f_\theta(x)$ predicts the voxel values that are then forwarded to predicting measurements $F(f_\theta(x))$ and are

compared with the ground truth measurements z . Here, the expectation is taken w.r.t the joint probability distribution μ of random variables (x, z) . Since we generally do not have access to the true distribution μ , we then replace the expectation in Equation (1) with its empirical estimator

$$\theta^* = \operatorname{argmin}_{\theta} \frac{1}{K} \sum_{k=1}^K [\|F(f_{\theta}(x_k)) - z_k\|_2^2], \quad (2)$$

where k denotes the individual sample for a patient and K is the number of all samples. After obtaining the optimal f_{θ} , we can easily compute the corresponding optimal voxel values with the mapping $f_{\theta}^*(x)$. As we mentioned before, the coordinate-based representation has a build-in spatial coherence prior. This can be explained by noticing Equation (2), where the coordinate x is assumed to follow a continuous distribution though the discrete samples are drawn from this distribution to train the f_{θ} .

To further constraint the sparse-view inverse problem, we introduce the interpatient priors in the next. To put it simply, we train the coordinate-based representation f_{θ} on a large patient dataset such that the resulting θ is set at an appropriate point where the pretrained model can generalize to the patient in hand. To adapt the coordinate-based representation f_{θ} to variable voxel values and positions, we adopt the Generative Latent Optimization (GLO) [17] in which each patient is assigned a corresponding low-dimensional latent vector that encodes the patient-dependent specifics. We then change the input of the MLP mapping $f_{\theta}(x_k)$ with a patient-dependent variable e_i to have

$$f_{\theta, e_i}(x_k^i), \quad (3)$$

where i denotes the patient index and k index the training samples for the patient i . To reconstruct the voxels for a patient i , the latent vector e_i is also optimized so that the prediction from the MLP $f_{\theta, e_i}(x_k^i)$ matches the measurement z_k^i of this patient. With the low-dimensional latent vector, we condition the coordinate-based representation f on a patient to enable the network to explain certain voxel variations derived from this patient. The network parameter θ is trained and shared among all patients and serves as an initial weight for a new patient. Moreover, the shared parameter θ acts as a strong prior that encodes the common human anatomy structures, which enables both better generalization and faster convergence during optimization.

Autoencoders or encoder-decoder networks are widely used for representation learning as their bottleneck features tend to form natural latent variable representations [18, 19]. The Generative Latent Optimization used here is different from the autoencoders in that only the decoder part of the latter is retained for training and inference. The encoder part of the autoencoders is dropped as it may waste computational resources. Without the encoder, the latent space can still emerge as we search for an optimal latent variable representation during the training [20].

We follow the classical ResNets [21] to implement our MLP network except by using feedforward layers in place

of convolutional layers, i.e., residual multilayer perceptron (ResidualMLP). The inputs are fed to a sequence of MLP blocks to produce output embeddings. Similar to ResNets, each block is paralleled with a skip-connection [21] to improve the convergence. We implemented ResidualMLP with six residual blocks and condition the network on the latent vector of dimension 16. The shortcut connections are used to connect the adjacent blocks. Each block consists of the affine layer defined by the weight matrix and the biases applied on its input, followed by a sine nonlinearity applied to each component of the resulting vector. We used 256 neurons for each hidden layer. To model the variation across the patients, a small MLP is constructed with 2 affine layers of 256 neurons so the dimension of the resulting feature vector is the same as the main network. We then add the resulting feature vector to each block's outputs to adapt the main network for the patient-specific. The output dimension of the last layer is 1.

2.2. Data. We used two datasets in this study. Dataset 1 contains a set of MRI data volumes produced by an MRI simulator [22]. These data are usually used by the neuroimaging study to evaluate the performance of various image analysis methods in a setting where the truth is known. Dataset 2 consists of customized versions of Shepp-Logan phantom [23]. The radians, angles, and positions of 3D ellipses are generated randomly. The Shepp-Logan phantom is often used in CT reconstruction to evaluate the performance of the reconstruction algorithms.

2.3. Implementation Details. We first train the network and the latent vector with all training samples from two datasets in the proposed neuRec way, then finetune the network from the pretrained checkpoint on the remaining test patients to evaluate the performance. The latent vector is initiated randomly and not optimized during the finetune. Two NVIDIA TITAN 3090 GPUs and PyTorch [24] deep learning framework are used to develop codes. For training the model, the measurements are centered and normalized by the mean and standard deviation across the training samples. We implement the details suggested in literature [25] to boost the performance, i.e., the Adam optimizer is used with a custom learning rate schedule that warms up linearly from 0 to $1e-4$ for the first 100 iterations, then decays proportional to the square root of the step count. The training takes about 17 hours due to the large patient samples, and the inference takes only several milliseconds due to the simplicity of the network.

2.4. Quality Evaluation Metrics. For reconstructions we compute the mean squared error (MSE), structured similarity (SSIM) index, and peak signal-to-noise ratio (PSNR) values with respect to the ground truth image (higher is better).

3. Results and Discussion

3.1. MRI Reconstruction. In order to verify the quality of the new method in MRI reconstruction, it is compared with the other four methods: filtered back projection (FBP) [26],

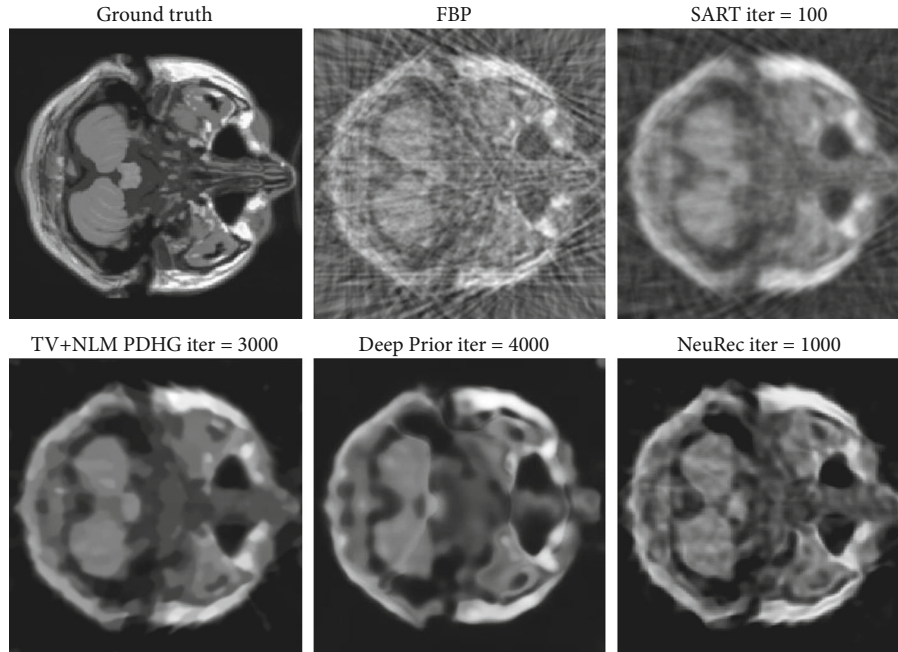


FIGURE 2: MRI reconstruction comparison.

TABLE 1: Comparison of the neuRec in MRI reconstruction with state-of-the-art methods.

Metrics	neuRec	FBP	SART	TV-NLM-PDHG	Deep prior
SSIM	0.873	0.201	0.623	0.718	0.843
PSNR	31.6	15.3	22.6	25.3	27.6

simultaneous algebraic reconstruction technique (SART) [27], total variation (TV), and nonlocal mean (NLM) priors solved by the primal-dual algorithm (TV-NLM-PDHG) [28], and Deep Prior [29]. The FBP reconstruction is used as the representation of classical analysis methods. We use the Hann filter for its filter stage. SART is superior to the standard algebraic reconstruction technique and is useful in cases when the projection data is limited. The TV corresponds to the local structure prior, while NLM aims to regularize the image statistics at a global level. PDHG iterative reconstruction is performed using 1000 iterations of the classical PDHG algorithm. Deep prior is a CNN-based prior for the inverse problem which captures the low-level image statistics in the structure of the network.

As illustrated in Figure 2, the classical FBP method performs worst because the analytical method assumes a complete measurement is available which prevents it from a plausible reconstruction. The SART method outperforms the FBP qualitatively, as it suppresses the artifacts effectively and produces a clearer image. Compared with the last two methods, the TV-NLM-PDHG algorithm produces a better reconstruction because the local- and global-prior are introduced to make up the null space. Finally, our neuRec method shows unprecedented qualitative results as high-frequency features such as the hippocampus are distinctly observed. The quantitative results in Table 1 also suggest a

significant increase of PSNR and SSIM values is obtained with our neuRec.

3.2. CT Reconstruction. To demonstrate the impact of interpatient prior, we carry on the experiments with the prior and without the prior for CT sparse-view reconstruction. Only 10 sinograms of the 3D Shepp-Logan phantom are used to push the neuRec to its limitation. Figure 3 shows the MSE with respect to the test iteration for neuRec with and without the interpatient prior. From the curves, we find neuRec with interpatient prior enables faster and more accurate reconstruction than neuRec without interpatient prior. To verify this conclusion, we further plot the 100-iteration intermediate results in Figure 4. As we can see from the figure, neuRec with interpatient prior achieves a much better reconstruction than neuRec without interpatient prior, even though only 10 sinograms are used. The sparse-view reconstruction, i.e., the ill-posed inverse problem is very challenging to be solved. As illustrated in the figure, the results produced by NeuRec without interpatient prior tend to be smoothed because the least-square minimum in Equation (2) estimates the average of the sparse samples. On the contrary, with interpatient prior, the network parameter is trained and shared among all patients and serves as an initial weight for a new patient. The shared parameter acts as a strong prior that encodes the common human anatomy structures,

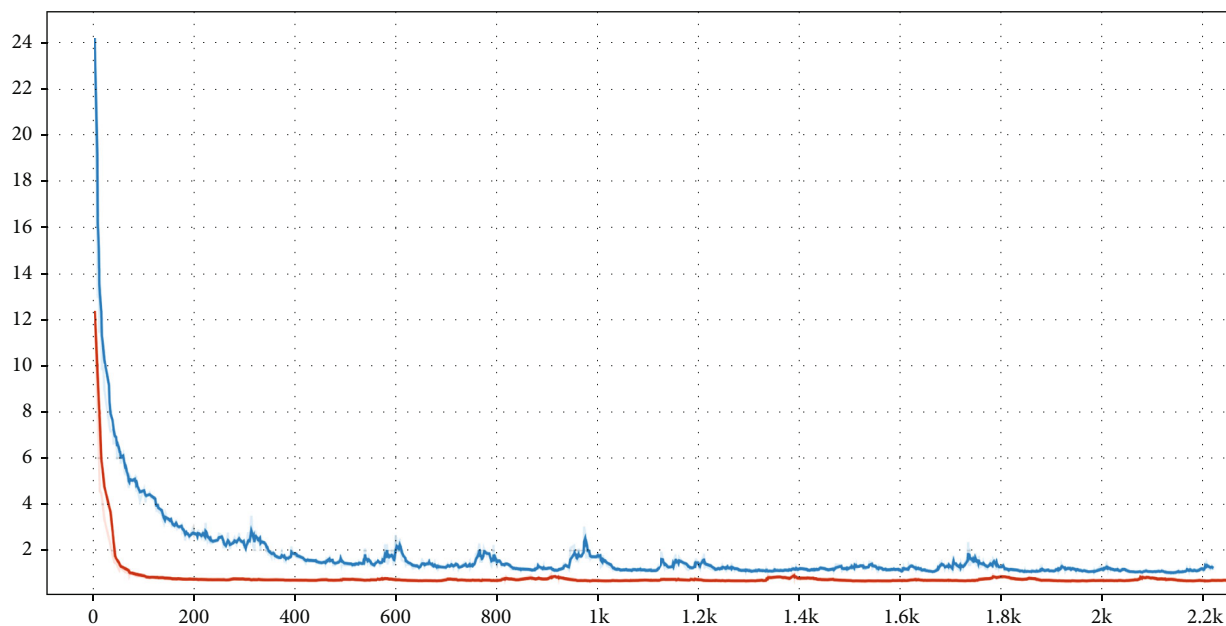


FIGURE 3: Ablation study. The MSE of neuRec with interpatient (red) and without interpatient prior (blue) is compared.

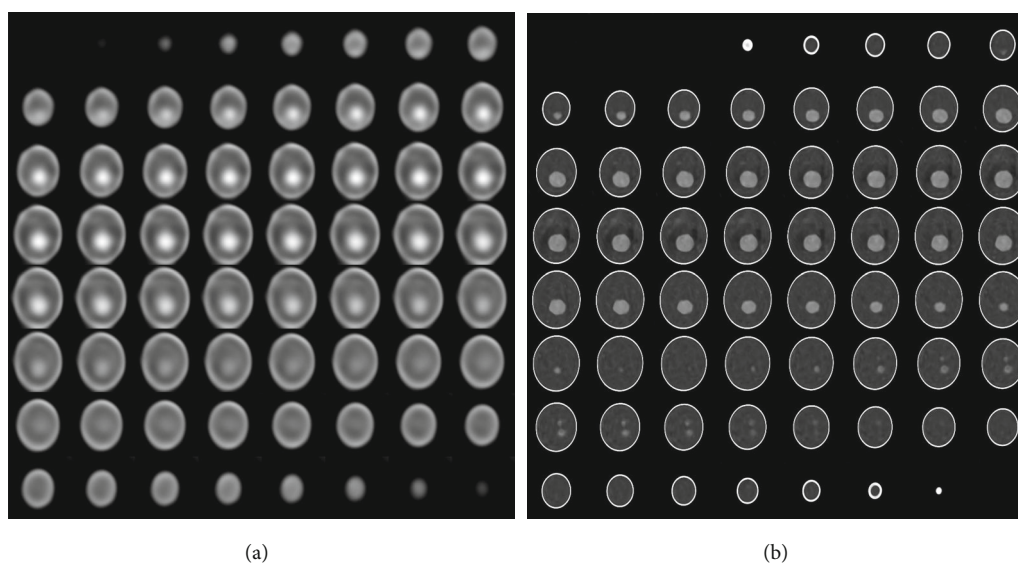


FIGURE 4: Qualitative ablation study. Reconstruction of neuRec without interpatient (a) and with interpatient (b).

which enables both better generalization and faster convergence during optimization.

4. Conclusions

This study proposes and evaluates a novel deep learning-based method neuRec for the sparse-view reconstruction of medical images, such as MRI and CT. It solves this ill-posed inverse problem through a coordinate-based neural network representation and the introduction of interpatient prior that is not exploited in other works. We demonstrate its performance on two datasets where it outperformed several state-of-the-art reconstruction methods. Although we presented results for CT and MRI reconstruction from

sinograms and k -spaces, neuRec can be used whenever the rendering or measurement process is differentiable. A possible line of research is to investigate interpatient priors for solving the more aggressive inverse problem.

Data Availability

Research data are stored in an institutional repository and will be shared upon request to the corresponding author.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Natural Science Foundation of Shanghai (20ZR1440300) and China Postdoctoral Science Foundation (2021M690481).

References

- [1] E. A. Wilde, J. V. Hunter, and E. D. Bigler, "A primer of neuroimaging analysis in neurorehabilitation outcome research," *NeuroRehabilitation*, vol. 31, no. 3, pp. 227–242, 2012.
- [2] M. Zaitsev, J. Maclaren, and M. Herbst, "Motion artifacts in MRI: a complex problem with many partial solutions," *Journal of Magnetic Resonance Imaging*, vol. 42, no. 4, pp. 887–901, 2015.
- [3] H. Kudo, T. Suzuki, and E. A. Rashed, "Image reconstruction for sparse-view CT and interior CT-introduction to compressed sensing and differentiated backprojection," *Quantitative Imaging in Medicine and Surgery*, vol. 3, no. 3, pp. 147–161, 2013.
- [4] G. Zang, M. Aly, R. Idoughi, P. Wonka, and W. Heidrich, "Super-resolution and sparse view CT reconstruction," *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 137–153, Munich, Germany, September 2018.
- [5] R. Zhang, J.-B. Thibault, C. A. Bouman, K. D. Sauer, and J. Hsieh, "Model-based iterative reconstruction for dual-energy X-ray CT using a joint quadratic likelihood model," *IEEE Transactions on Medical Imaging*, vol. 33, no. 1, pp. 117–134, 2014.
- [6] J. Du, G. Zang, B. Mohan et al., "Study of spray structure from non-flash to flash boiling conditions with space-time tomography," *Proceedings of the Combustion Institute*, vol. 38, no. 2, pp. 3223–3231, 2021.
- [7] R. Anirudh, H. Kim, J. J. Thiagarajan, K. A. Mohan, K. Champley, and T. Bremer, "Lose the views: limited angle CT reconstruction via implicit sinogram completion," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6343–6352, Salt Lake City, USA, June 2018.
- [8] Z. Liu, T. Bicer, R. Kettimuthu, D. Gursoy, F. De Carlo, and I. Foster, "TomoGAN: low-dose synchrotron x-ray tomography with generative adversarial networks: discussion," *JOSA A*, vol. 37, no. 3, pp. 422–434, 2020.
- [9] D. Hu, J. Liu, T. Lv et al., "Hybrid-domain neural network processing for sparse-view CT reconstruction," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 1, pp. 88–98, 2021.
- [10] E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Medical Physics*, vol. 44, no. 10, pp. e360–e375, 2017.
- [11] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *Josa a*, vol. 1, no. 6, pp. 612–619, 1984.
- [12] T. Meinhardt, M. Moller, C. Hazirbas, and D. Cremers, "Learning proximal operators: using denoising networks for regularizing inverse imaging problems," *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1781–1790, Venice, Italy, October 2017.
- [13] H. Wei, F. Schiffrers, T. Würfl et al., "2-step sparse-view ct reconstruction with a domain-specific perceptual network," 2020, <http://arxiv.org/abs/2012.04743>.
- [14] S. A. Eslami, D. Jimenez Rezende, F. Besse et al., "Neural scene representation and rendering," *Science*, vol. 360, no. 6394, pp. 1204–1210, 2018.
- [15] G. Zang, R. Idoughi, R. Li, P. Wonka, and W. Heidrich, "IntraTomo: self-supervised learning-based tomography via sinogram synthesis and prediction," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 1960–1970, Montreal, Canada, October 2021.
- [16] L. Shen and L. Xing, "NeRP: implicit neural representation learning with prior embedding for sparsely sampled image reconstruction," 2022, <http://arxiv.org/abs/2108.10991>.
- [17] P. Bojanowski, A. Joulin, D. Lopez-Paz, and A. Szlam, "Optimizing the latent space of generative networks," 2017, <http://arxiv.org/abs/1707.05776>.
- [18] T. Bagautdinov, C. Wu, J. Saragih, P. Fua, and Y. Sheikh, "Modeling facial geometry using compositional vaes," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 3877–3886, Salt Lake City, USA, June 2018.
- [19] O. Litany, A. Bronstein, M. Bronstein, and A. Makadia, "Deformable shape completion with graph convolutional autoencoders," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1886–1895, Salt Lake City, USA, June 2018.
- [20] R. Martin-Brualla, N. Radwan, M. S. Sajjadi, J. T. Barron, A. Dosovitskiy, and D. Duckworth, "Nerf in the wild: neural radiance fields for unconstrained photo collections," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 7210–7219, Nashville, USA, June 2021.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778, Las Vegas, USA, June 2016.
- [22] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "Brainweb: Online Interface to a 3D MRI Simulated Brain Database *NeuroImage*," vol. 5, Article ID 425, 1997.
- [23] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, pp. 21–43, 1974.
- [24] A. Paszke, S. Gross, F. Massa et al., "PyTorch: an imperative style, high-performance deep learning library," *Advances in Neural Information Processing Systems*, 2019, pp. 8024–8035, Vancouver, Canada, Dec. 2019.
- [25] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," *International conference on machine learning*, 2020, Vienna, Austria, July 2020, 2020.
- [26] X. Pan, E. Y. Sidky, and M. Vannier, "Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?," *Inverse Problems*, vol. 25, no. 12, article 123009, 2009.
- [27] D. Ji, G. Qu, and B. Liu, "Simultaneous algebraic reconstruction technique based on guided image filtering," *Optics Express*, vol. 24, no. 14, pp. 15897–15911, 2016.
- [28] Y. Wen, R. H. Chan, and T. Zeng, "Primal-dual algorithms for total variation based image restoration under Poisson noise," *SCIENCE CHINA Mathematics*, vol. 59, no. 1, pp. 141–160, 2016.
- [29] M. Gadelha, R. Wang, and S. Maji, "Shape reconstruction using differentiable projections and deep priors," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 22–30, Seoul, Korea, Oct. 2019.