Beyond Local Processing: Adapting CNNs for CT Reconstruction

Bassel Hamoud[1](✉), Yuval Bahat[2], and Tomer Michaeli[1]

[1] Technion–Israel Institute of Technology, Haifa, Israel
`bassel164@campus.technion.ac.il`
[2] Princeton University, Princeton, NJ, USA

Abstract. Convolutional neural networks (CNNs) are well suited for image restoration tasks, like super resolution, deblurring, and denoising, in which the information required for restoring each pixel is mostly concentrated in a small neighborhood around it in the degraded image. However, they are less natural for highly non-local reconstruction problems, such as computed tomography (CT). To date, this incompatibility has been partially circumvented by using CNNs with very large receptive fields. Here, we propose an alternative approach, which relies on the rearrangement of the CT projection measurements along the CNN's 3^{rd} (channels') dimension. This leads to a more local inverse problem, which is suitable for CNNs. We demonstrate our approach on sparse-view and limited-view CT, and show that it significantly improves reconstruction accuracy for any given network model. This allows achieving the same level of accuracy with significantly smaller models, and thus induces shorter training and inference times.

Keywords: CT reconstruction · Machine learning · ConvNets

1 Introduction

Deep learning has led to major leaps in our ability to solve complex inverse problems. In the context of image restoration, the architectures of choice are typically convolutional neural networks (CNNs). CNNs have pushed the state-of-the-art in tasks like super-resolution [11,13,33], denoising [12,16,31], deblurring [21,25], dehazing [30] and deraining [19]. These are tasks that are well suited for CNNs because of their local nature. Namely, good restoration in those problems can be achieved by predicting the value of each pixel from a local neighborhood around it in the input image, which is precisely how CNNs operate. However, there exist many important inverse problems that do not possess this locality

---

Y. Bahat—Part of the work was done while the author was affiliated with the Technion.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-25066-8_29.

L. Karlinsky et al. (Eds.): ECCV 2022 Workshops, LNCS 13803, pp. 513–526, 2023.
https://doi.org/10.1007/978-3-031-25066-8_29

property, and are thus a-priori less natural for CNNs. One particularly important example is computed tomography (CT), where cross-sections of the human body are reconstructed from partial measurements of their Radon transform (or sinogram).

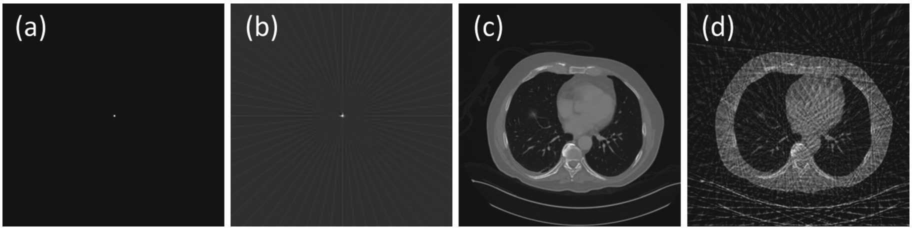

Fig. 1. Non-local nature of CT reconstruction. Ground-truth images (a & c) and their corresponding reconstructions (b & d) based on 30 projections, using FBP. Reconstruction artifacts exhibit a global structure, resembling streaks or rays, making them hard to tackle by CNNs, which operate locally.

CT reconstruction has seen significant progress over the years. Theoretically, infinitely many projection angles and noise-free measurements allow perfect reconstruction, *e.g.* via the *filtered back-projection* (FBP) method [6]. However, under practical capturing conditions, the reconstruction problem is ill-posed and the naive FBP method leads to poor reconstructions (Fig. 1). The effect becomes more severe when using low radiation doses and/or a small number of projections, which are desired for reducing scan times and exposure to ionizing X-ray radiation, but lead to noisy and under-sampled sinograms. Several works suggested to use CNNs for learning to reconstruct CT images from degraded sinograms. These methods first convert the low-quality sinogram into an image, and then feed it into the CNN. The conversion stage is done either with the FBP method [3,9] or using learned operations [8]. However, in both cases, the CNN is typically left with solving a very nonlocal reconstruction problem. This can be appreciated from Figs. 1(b), (d), which depict FBP reconstructions of a delta function and a chest scan, from 30 projection angles. As can be seen, the FBP-reconstructed images contain global streaking artifacts. Therefore, CNNs applied on such inputs, must have very large receptive fields [9] to ensure that they can access the information relevant for recovering each pixel. This comes at the cost of many learned parameters, and networks that are harder to train.

In this paper, we present an approach for adapting CNNs to the highly nonlocal CT reconstruction task, by modifying their input such that the information relevant for reconstructing each pixel is available in a small spatial neighborhood around it, across different channels. To this end, we begin by converting the captured sinogram into a series of per-projection images, which we then stack along the channels dimension, and feed into the CNN. The process is depicted in Fig. 2 and described in detail in Sect. 3.

We bring our approach to bear on several sparse-view and limited-view CT reconstruction tasks, including under low dose and patient motion settings. These tasks are all highly relevant for the ongoing effort to reduce radiation exposure in patients undergoing CT scans, without compromising image quality and impairing medical diagnosis.

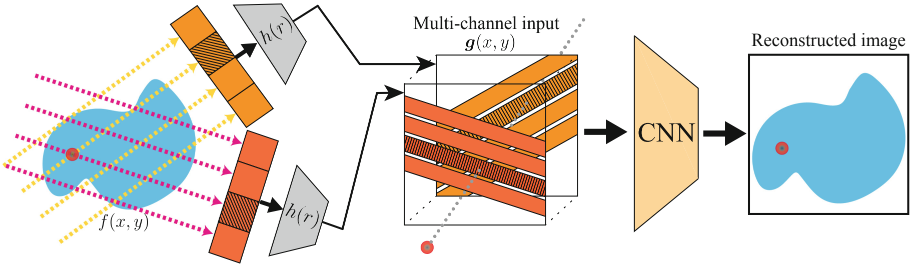

Fig. 2. Sketch of the proposed method. After data is acquired from N projection angles (left), we propose to arrange the filtered projections along the channel axis of a 3D tensor $\boldsymbol{g}(x, y)$ (middle), before feeding it to a reconstruction CNN to yield the reconstructed image (right).

2 Related Work

As mentioned above, the simple FBP method is not well suited for sparse-view and low-dose CT settings. Several more advanced techniques have been developed over the years. These can be broadly put in several categories, as follows.

Iterative Algorithms. These reconstruct the CT image by iteratively minimizing an objective function [2,26,32]. The algebraic reconstruction technique (ART) [7] and simultaneous ART (SART) [1] minimize a fidelity term that pushes the reconstructed image to be consistent with the sinogram. Other methods incorporate image priors over the reconstructed image by adding a regularization term, *e.g.* penalizing the ℓ_1 [20] or ℓ_0 [29] norm of the image gradients.

Radon Space. These methods operate on the sinogram before transforming it to an image. Some low-dose reconstruction methods proposed to denoise the sinogram using a MAP estimator based on low-dose CT noise statistics [24], or by using a total variation based denoising model tailored for Poisson noise [34]. For sparse-view CT, Lee et al. [14] trained a U-Net-like residual CNN on downsampled and interpolated sinogram patches, while Li et al. [15] learned a dictionary to fill in missing sinogram projections. However, common to all Radon space methods is the need to eventually transform the reconstructed sinogram into image space (e.g. using FBP), which typically results in introducing new artifacts.

Image Space. These techniques attempt to enhance a naïvely reconstructed image (obtained e.g. using FBP). Y. Chen et al. [4] employed dictionary learning to reduce the noise and streaking artifacts in low-dose CT images. Kang et al.

[10] applied a CNN to the wavelet coefficients of the reconstructed image, while an auto-encoder design was used in [3]. Yang et al. [28] trained a WGAN to improve perceptual quality. In terms of reconstruction performance, Jin et al. [9] achieved state of the art results by training a U-Net network to suppress the global streaking artifacts that arise in sparse-view reconstruction using FBP.

End-to-End. These methods process the sinogram, transform it into an image, and then operate further on the reconstructed image. He et al. [8] used fully connected layers to filter the sinogram, followed by a learned backprojection layer mapping into the image domain, and a residual CNN performing post-processing on the resulting image. Wang et al. [22] proposed to use two residual CNNs, one on the sinogram and one on the image obtained by transforming the processed sinogram into an image using FBP.

Our approach falls in the intersection of the Radon and image domains, as it proposes a novel intermediate representation, which makes the information relevant for reconstructing each pixel in the image locally accessible for subsequent processing. As such, it can be paired with almost any method operating in the image domain, including end-to-end methods.

3 Adapting Radon Space Representation for CNNs

CT scans are closely related to the Radon transform. Specifically, when using the simple setting of *parallel-beam* CT, the measurements correspond precisely to the Radon transform of the imaged slice, at the measured angles. When using the more popular *fan-beam* geometry, a narrow X-ray source illuminates a wide section of the body, and as a result the individual projections do not form horizontal lines in Radon space. However, those measurements can be converted into a regular sinogram by using a simple closed-form computation [5,23]. Therefore, in the interest of simplicity, we will describe our method for the parallel-beam CT setting, keeping in mind that it is applicable also to fan-beam CT in a straight-forward manner. We begin by describing the FBP method, and then explain how sinograms can be rearranged to better fit the local nature of CNNs.

3.1 The Filtered Backprojection Algorithm

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a real-valued image. Its Radon transform [17] is the function $s : [0, \pi) \times \mathbb{R} \to \mathbb{R}$ obtained from the linear transformation

$$s(\theta, r) = \int_{-\infty}^{\infty} f(z \sin\theta + r\cos\theta, -z\cos\theta + r\sin\theta)\, dz. \tag{1}$$

That is, each point (θ, r) in the sinogram corresponds to an integral of f over the straight line $\{(x, y) : x\cos\theta + y\sin\theta = r\}$, which lies at a distance r from the origin and forms an angle θ with the vertical axis. In the context of CT, f corresponds to the attenuation coefficients across a slice within the imaged object, and s is obtained by measuring the amount of radiation that passes through lines within f (see Fig. 2). Under mild assumptions, f can be recovered from s using the inverse Radon transform, which can be expressed as [17]

$$f(x,y) = \frac{1}{2\pi}\int_0^{\pi} \tilde{s}(\theta, x\cos\theta + y\sin\theta)\, d\theta, \tag{2}$$

where $\tilde{s}(\theta,\cdot) = s(\theta,\cdot) * h$. Here, $*$ denotes 1D convolution and $h(r)$ is the Ram-Lak filter [18] (ramp filter). In words, each projection is filtered with $h(r)$ and then backprojected onto the image space at its corresponding angle.

In practical settings, the sinogram $s(\theta, r)$ is measured only on a discrete set of angles θ. In such cases, f generally cannot be perfectly reconstructed. One approach to obtain an approximation of f is to discretize (2). Specifically, assuming N uniformly spaced angles $\{\theta_n\}_{n=1}^N$ over the interval $[0,\pi)$, one can approximate f as

$$f(x,y) \approx \frac{1}{2\,N}\sum_{n=1}^{N} \tilde{s}(\theta_n, x\cos\theta_n + y\sin\theta_n), \tag{3}$$

This is the filtered back-projection (FBP) algorithm of Feldkamp et al. [6].

The FBP method produces satisfactory reconstructions when the angles are sufficiently densely spaced. However, it leads to highly nonlocal artifacts when the projections are sparse, as illustrated in Fig. 1. To gain intuition into the nonlocal nature of the problem, consider the setting of Figs. 1(a), (b). Here, we measure N projections along angles $\{\theta_n\}_{n=1}^N$ of a Dirac delta function that is located at (x_0, y_0) in the image domain,

$$f(x,y) = \delta(x - x_0, y - y_0). \tag{4}$$

The resulting sinogram is given by (see Supp. A.1)

$$s(\theta_n, r) = \delta(r - (x_0\cdot\cos\theta_n + y_0\cdot\sin\theta_n)), \quad n = 1, 2, ..., N. \tag{5}$$

Namely, it is a sinusoid in Radon space. As we show in Supp. A.2, reconstructing the image using FBP yields an approximation of f in the form of a star-like pattern with N rays that are inclined at angles $\{\theta_n\}_{n=1}^N$ and intersect at (x_0, y_0),

$$f(x,y) \approx \frac{1}{2N}\sum_{n=1}^{N} g\big((x - x_0)\cos\theta_n + (y - y_0)\sin\theta_n\big). \tag{6}$$

Here, $g(r) = k_{\max}^2\left(\operatorname{sinc}(k_{\max} r) - \frac{1}{4}\operatorname{sinc}^2\left(\frac{k_{\max} r}{2}\right)\right)$. These artifacts can be suppressed using a CNN, however their non-local nature necessitates architectures with very large receptive fields.

3.2 Localizing the Inverse Transformation

To be able to effectively exploit the power of CNNs for improving CT reconstruction, here we propose to rearrange the sinogram data into a new multi-channel image-domain representation $\boldsymbol{g}(x,y)$, before feeding it into the CNN. The key idea is to make each (vector-valued) pixel in $\boldsymbol{g}(x,y)$ contain all the sinogram

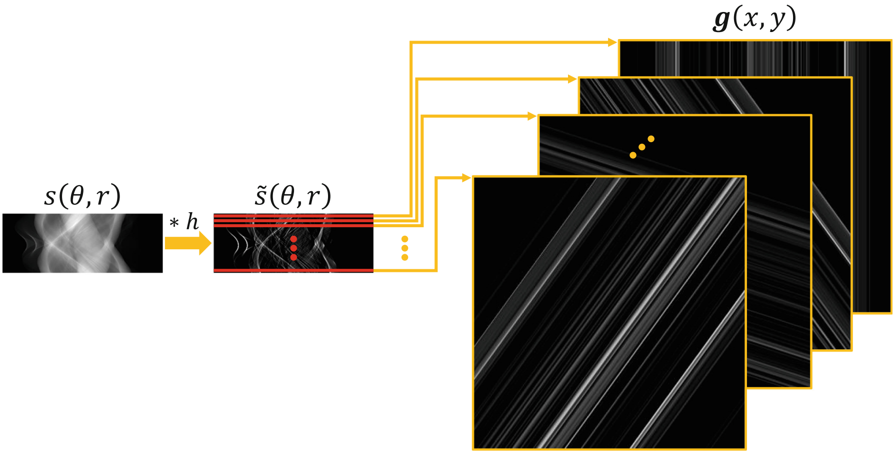

Fig. 3. Our proposed method (CBP). The captured sinogram's rows are first filtered with the Ram-Lak filter h to yield $\tilde{s}$. The multi-channel tensor $\boldsymbol{g}(x, y)$ is then constructed by smearing each row of the filtered sinogram $\tilde{s}$ along its corresponding angle to form a 2D image. The smeared images are finally arranged along the channel dimension to form $\boldsymbol{g}$.

entries affected by the corresponding pixel in $f(x, y)$. Specifically, note from (1) that a pixel $f(x_0, y_0)$ affects the sinogram entries at all locations (θ, r) satisfying $x_0 \cos\theta + y_0 \sin\theta = r$. Since we have N discrete angles, this affected set contains precisely N sinogram entries (one r for each angle θ_n). We collect those N entries from the filtered sinogram $\tilde{s}$ and arrange them along the channel dimension of $\boldsymbol{g}(x, y)$. Namely, for each pixel location (x, y) and channel $n \in \{1, ..., N\}$, we set

$$[\boldsymbol{g}(x, y)]_n = \tilde{s}(\theta_n, r_n), \tag{7}$$

where $r_n = x \cos\theta_n + y \sin\theta_n$. Accordingly, we term our method *Channel Back Projection* (CBP).

It is instructive to note that the nth channel of $\boldsymbol{g}(x, y)$ contains a pattern that changes only along the direction θ_n. Particularly, for a fixed θ_n, if we walk along any line of the form $(x, y) = (r \cos\theta_n, r \sin\theta_n)$ by varying $r \in \mathbb{R}$, then the pattern we encounter is precisely $\tilde{s}(r, \theta_n)$. Thus, each channel of $\boldsymbol{g}(x, y)$ can be formed by taking all entries of $\tilde{s}$ corresponding to a single angle, and "smearing" them along that angle to form a 2D image. This is illustrated in Fig. 3.

An additional important observation is that averaging the entries of $\boldsymbol{g}(x, y)$ along the channel direction yields the naive FBP reconstruction (3) (up to a factor of 1/2). This suggests that our representation $\boldsymbol{g}(x, y)$ generally contains more information than the FBP representation. As we now illustrate, this turns out to allow a significant reduction in the CNN's receptive field (and thus number of parameters) while retaining good reconstruction accuracy.

4 Experiments

To demonstrate the advantages of the proposed approach, we evaluate it on the task of sparse-view CT reconstruction in noiseless, noisy, and motion blur settings (we study both sensor blur and patient movement blur). These settings are often encountered when working with low radiation doses. We also evaluate our method on the task of limited-view CT reconstruction, where the projection angles are spread over $[0, \theta_{\max})$ with a $\theta_{\max}$ that is strictly smaller than 180°. Such settings arise, *e.g.* when imaging an object with periodic motion such as coronary vasculature.

We use the DeepLesion [27] dataset, which contains clean full-dose ground-truth (GT) CT images spanning the whole body. We used 5152 CT images for training and 960 for testing. The train and test sets contain CT scans of different patients (74 and 13 patients, respectively). Our approach can be paired with any method operating in the image domain. Here, we demonstrate it with the deep U-Net architecture proposed in [9], which achieves state of the art results in sparse-view CT reconstruction. The architecture is shown in Fig. 4. We next elaborate on each of the reconstruction tasks we consider.

4.1 CT Reconstruction Tasks

We simulate four sparse-view CT reconstruction tasks, in which we use 30 projection directions uniformly spread in $[0°, 180°)$, and two limited-view CT reconstruction tasks where the projections are limited to $[0°, 120°)$ and $[0°, 90°)$.

1. **Sparse-view CT:**
   (a) **Noise-free:** The 30-view sinogram $s(\theta, r)$ serves as the measurements.
   (b) **Noisy:** A noisy version of the 30-view $s(\theta, r)$ serves as the measurements. This is done by converting $s(\theta, r)$ into intensity measurements as $I(\theta, r) = I_0 \exp\{-s(\theta, r)\}$, contaminating it by Poisson noise as $I_{\mathrm{n}}(\theta, r) \sim \mathrm{Poisson}(I(\theta, r))$, and converting it back into a noisy sinogram $s_{\mathrm{n}}(\theta, r) = -\log(I_{\mathrm{n}}(\theta, r)/I_0)$. Here I_0 is the X-ray source's intensity, which we set to 10^5.
   (c) **Sensor-motion blur:** Here we simulate motion blur resulting from the rotation of the X-ray tube and the sensor array during the scan. Each of the 30 projections in $s(\theta, r)$ is calculated as the average of 12 projections spanning 6°, without overlap.
   (d) **Patient-motion blur:** We consider a simplified case of rigid patient motion. Assuming short exposure times, we focus on motion in straight lines. To simulate this, we draw a random direction and a random length l for each GT image. We then sequentially compute the 30 projections of the corresponding degraded sinogram, while uniformly translating the image between $-\ell/2$ and $\ell/2$, along the chosen direction. Thus, each projection within the sinogram corresponds to a different shift of the image (projection $j \in \{0, 1, ..., 29\}$ corresponds to a shift of $-\ell/2 + j\ell/30$).
2. **Limited-view CT:** The measurements here consist of 120 or 90 projections spread uniformly over $[0°, 120°)$ or $[0°, 90°)$, respectively.

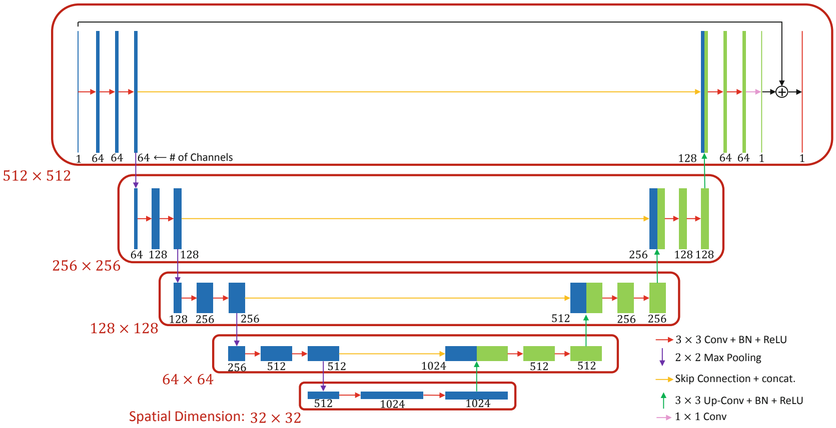

Fig. 4. U-Net architecture. The original architecture of [9], comprising five scales consisting of 34.5×10^6 parameters, and inducing a receptive field of 203×203.

4.2 Reconstruction Methods

We compare our approach to other *CNN based methods*. These achieve far better reconstruction accuracy than classical iterative techniques (see comparison in Supp. C. Particularly, in all tasks, we compare the performance of our method (CBP) to that of feeding the CNN with the image reconstructed using FBP. Moreover, to evaluate the significance of allowing *deep* network processing over the localized input, we additionally compare to a degenerate variant of our method, that learns a weighted linear combination of channels $[\boldsymbol{g}(x, y)]_n$, $n = 1, \ldots, N$, before feeding the resulting *single-channel* image to the CNN. This variant is equivalent to learning a *sinusoidal back projection* (SBP) layer, as proposed in [8], which allows a degree of freedom over using the FBP as input, but does not allow subsequent network layers to access the localized information.

We experiment with different model sizes corresponding to different receptive field sizes, to show how using our method for localizing the input's arrangement alleviates the need for large receptive fields. To this end, we use the original U-Net architecture of [9] (Fig. 4), as well as variants having reduced number of scales, and thus smaller receptive fields and fewer learned parameters. The full architectures and parameters of the variants are provided in Supp. B.

In all experiments, we train the network until convergence by minimizing the mean square error (MSE) between the network's outputs and the corresponding GT CT images, as proposed in [9]. We employ the Adam optimizer with the default settings and batch size of 32. We initialize the learning rate to 10^{-4} and automatically decrease it by half whenever the loss reaches a plateau.

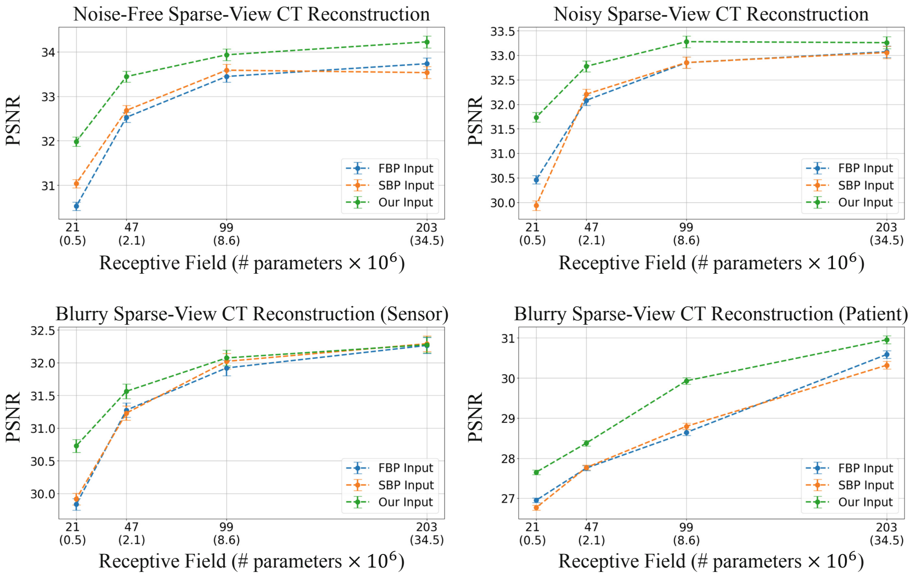

Fig. 5. Reconstruction accuracy vs. receptive field size in four sparse-view CT reconstruction tasks. Our CBP input (green) consistently improves performance over the FBP and SBP inputs, and the effect is stronger for smaller models (with smaller receptive fields). These models struggle to handle global artifacts. Error bars correspond to standard error of the mean (SEM) and horizontal axes correspond to receptive field sizes and number of parameters (in parentheses). Please refer to Supp. C for SSIM comparisons, which reveal the same behavior.

4.3 Results

We present quantitative comparisons in Figs. 5 and 6, demonstrating the advantage of our method across all tasks and model sizes. This advantage is typically bigger with smaller models, whose smaller receptive field does worse at handling

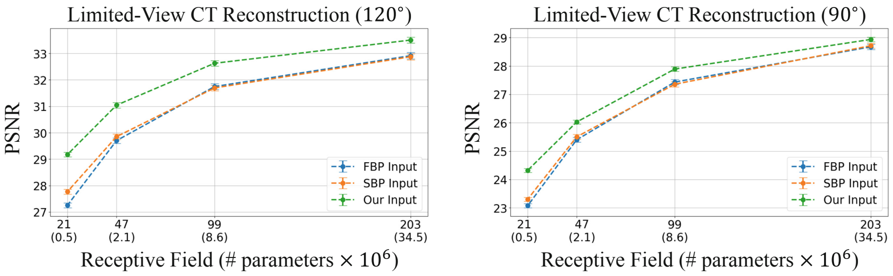

Fig. 6. Reconstruction accuracy vs. receptive field in limited-view CT. Our method is consistently advantageous over operating on FBP or SBP inputs, and the effect is stronger for smaller models, whose small receptive fields cannot handle global artifacts. Error bars correspond to standard error of the mean (SEM). Please refer to Supp. C for SSIM comparisons, which reveal the same behavior.

global artifacts. Note the essential role of allowing deep learning over the localized input (CBP), compared with the inefficacy of merely learning a weighted combination of the filtered projections as in [8] (SBP). The behavior seen in these plots exists also when quantifying performance using SSIM rather than PSNR (see Supp. C). Figures 7 and 8 present a qualitative comparison, which shows that our method (right) does better at removing global (smear-like) artifacts, while recovering finer details from the GT images (left) and inducing less artifacts compared to the models fed with FBP or SBP.

These experiments illustrate that our CBP approach allows reducing model size, while maintaining reconstruction performance. Note, however, that when we change the model size we change both the receptive field and the number of parameters. We next perform an additional ablation experiment to decompose

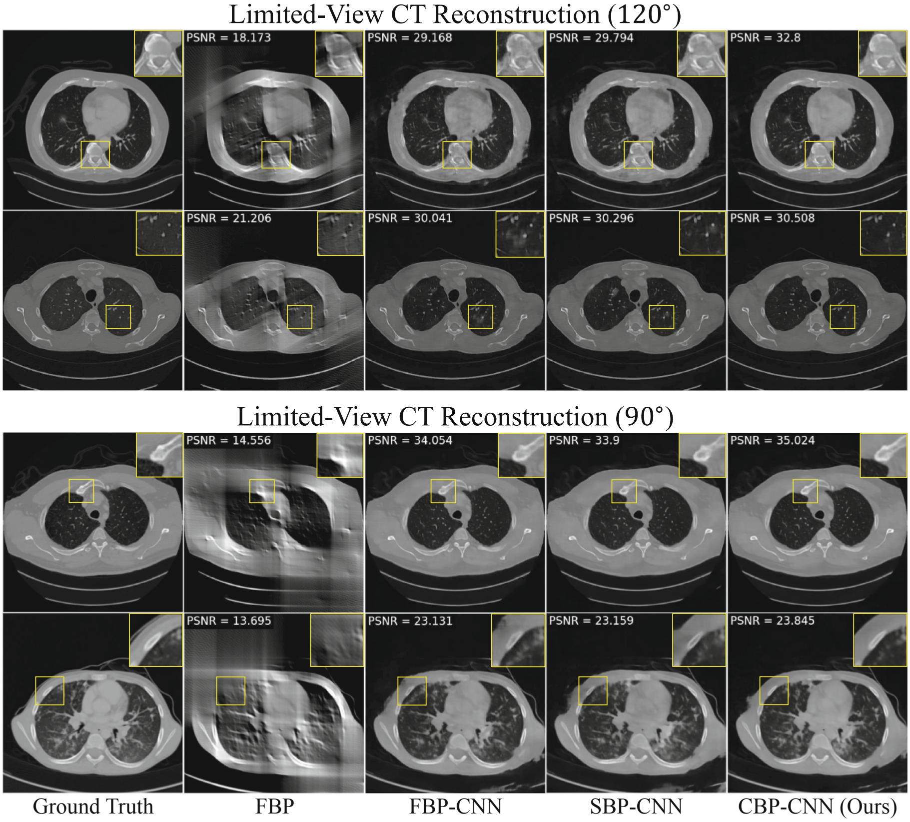

Fig. 7. Visual comparisons for limited-view CT reconstruction. Degraded images (2$^{\mathrm{nd}}$ col.) are reconstructed using a 21 × 21 receptive field U-Net, which is fed with FBP-reconstructed images (3$^{\mathrm{rd}}$ col.), with the outputs of the SBP layer of [8] (4$^{\mathrm{th}}$ col.), or with our proposed tensor $\boldsymbol{g}$ (last col.). Our method leads to more reliable reconstructions, with less artifacts.

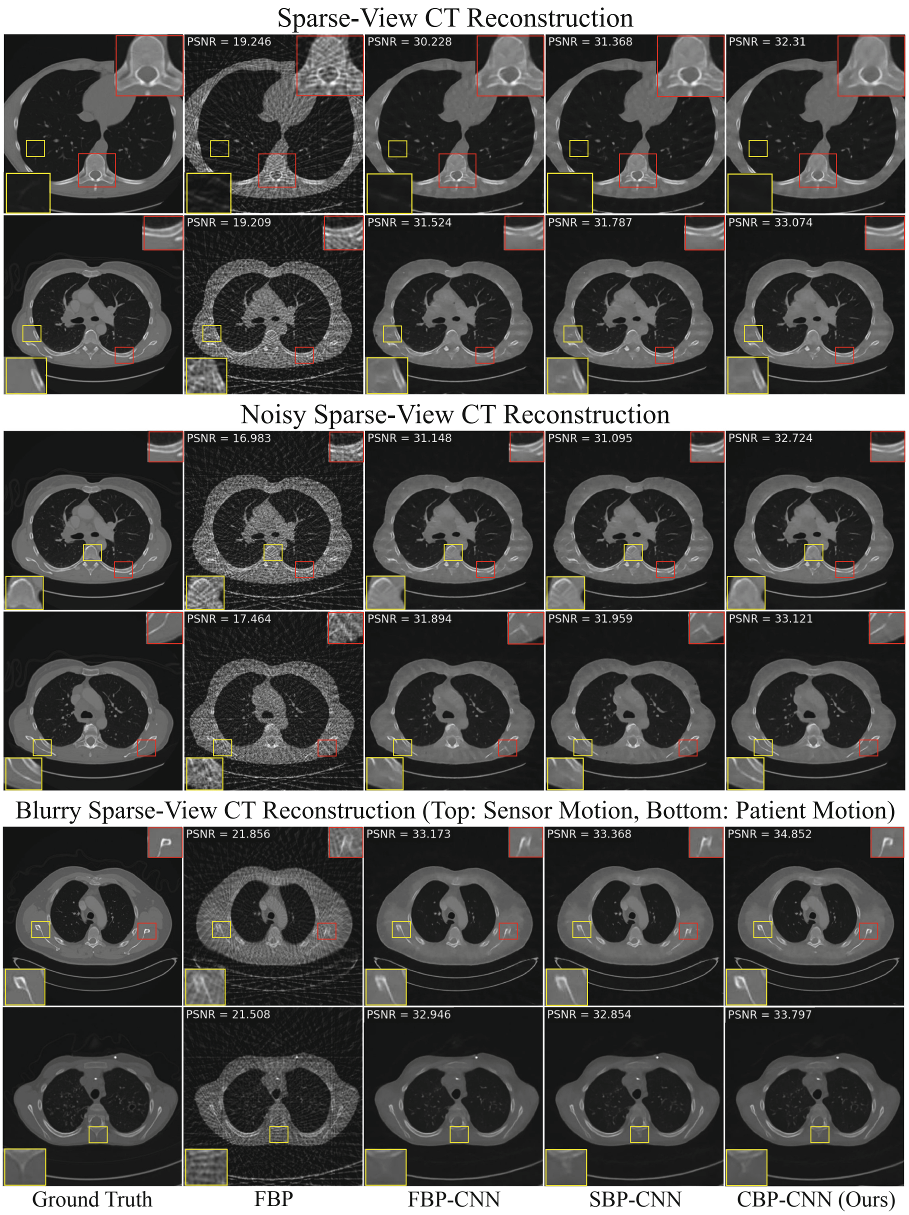

Fig. 8. Visual comparisons for sparse-view CT reconstruction. Degraded images (2nd col.) are reconstructed using a 47 × 47 receptive field U-Net, which is fed with the FBP-reconstructed images (3rd col.), with the outputs of the SBP layer of [8] (4th col.), or with our proposed tensor $\boldsymbol{g}$ (last col.). Our method leads to more reliable reconstructions, with less artifacts.

the effect of model size into the specific effects of number of learned parameters and receptive field size. To this end, we train six network variants on noise-free sparse-view CT reconstruction, keeping a fixed (21×21) receptive field size, while varying the number of channels to induce different numbers of learned parameters. In particular, we use our 0.5 million parameter (21 × 21 receptive field) variant of U-Net (denoted U-Net-S2, see Supp. B) and five sub-variants thereof, modified to have 0.5, 1.5, 2, 2.5, or 3 times the number of channels (resulting in 0.1, 1.1, 1.9, 3, or 4.3 million parameters, respectively).

Figure 9 shows that the advantage of our approach is unchanged when merely varying the number of learned parameters, without changing the receptive field size. This supports our hypothesis that the improved performance of larger U-net models in Figs. 5 and 6 should be attributed to their larger receptive fields, allowing them to partially circumvent the incompatibility of CNNs to the nonlocal task at hand. Interestingly, Fig. 9 also indicates that the 0.5 million parameter model that uses our CBP performs better than 4.3 million parameter models that use the FBP and SBP representations. Namely, CBP achieves better performance with 8.6× less parameters, illustrating its effectiveness in reducing model sizes.

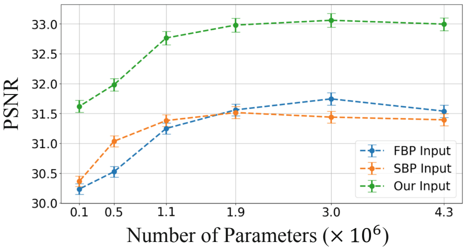

Fig. 9. Reconstruction performance vs. model size, using a fixed receptive field. With a fixed receptive field size, the advantage of our approach (green) remains unchanged as the number of learned model parameters increases. This suggests that the improved performance of larger U-net models (right points in Fig. 5) is primarily due to their larger receptive fields, which compensate for the non-local nature of the CT reconstruction problem. Error bars correspond to standard error of the mean (SEM).

5 Conclusion

Deep CNNs have demonstrated unprecedented performance in many image restoration tasks. However, their operation is inherently local, making them less suitable for handling CT reconstruction tasks, which suffer from global artifacts. While existing works circumvent this incompatibility using deeper networks with very large receptive fields, we inherently address the problem by proposing a new locality-preserving representation for the CNN's input data. We evaluate our approach across different CT reconstruction tasks, showing how it can improve a network's performance, or alternatively allow reducing the model size, while maintaining reconstruction quality.

Acknowledgement. This research was partially supported by the Ollendorff Miverva Center at the Viterbi Faculty of Electrical and Computer Engineering, Technion. Yuval Bahat is funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 945422.

References

1. Andersen, A., Kak, A.: Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm. Ultrasonic Imaging **6**(1), 81–94 (1984)
2. Cai, J.F., Jia, X., Gao, H., Jiang, S.B., Shen, Z., Zhao, H.: Cine cone beam ct reconstruction using low-rank matrix factorization: algorithm and a proof-of-princple study (2012)
3. Chen, H., et al.: Low-dose CT with a residual encoder-decoder convolutional neural network. IEEE Trans. Med. Imaging **36**(12), 2524–2535 (2017)
4. Chen, Y.: Artifact suppressed dictionary learning for low-dose CT image processing. IEEE Trans. Med. Imaging **33**(12), 2271–2292 (2014)
5. Dreike, P., Boyd, D.P.: Convolution reconstruction of fan beam projections. Comput. Graph. Image Process. **5**(4), 459–469 (1976)
6. Feldkamp, L.A., Davis, L.C., Kress, J.W.: Practical cone-beam algorithm. J. Opt. Soc. Am. A-optics Image Sci. Vision **1**, 612–619 (1984)
7. Gordon, R., Bender, R., Herman, G.T.: Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography. J. Theor. Biol. **29**(3), 471–481 (1970)
8. He, J., Wang, Y., Ma, J.: Radon inversion via deep learning. IEEE Trans. Med. Imaging **39**(6), 2076–2087 (2020)
9. Jin, K.H., McCann, M.T., Froustey, E., Unser, M.: Deep convolutional neural network for inverse problems in imaging. IEEE Trans. Image Process. **26**(9), 4509–4522 (2017)
10. Kang, E., Min, J., Ye, J.C.: A deep convolutional neural network using directional wavelets for low-dose x-ray CT reconstruction. Med. Phys. **44**(10), e360–e375 (2017)
11. Lai, W.S., Huang, J.B., Ahuja, N., Yang, M.H.: Fast and accurate image super-resolution with deep laplacian pyramid networks. IEEE Trans. Pattern Anal. Mach. Intell. **41**(11), 2599–2613 (2018)
12. Laine, S., Lehtinen, J., Aila, T.: Self-supervised deep image denoising. CoRR abs/1901.10277 (2019)
13. Ledig, C., et al.: Photo-realistic single image super-resolution using a generative adversarial network. CoRR abs/1609.04802 (2016)
14. Lee, H., Lee, J., Kim, H., Cho, B., Cho, S.: Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. IEEE Trans. Radiat. Plasma Med. Sci. **3**(2), 109–119 (2019)
15. Li, S., Cao, Q., Chen, Y., Hu, Y., Luo, L., Toumoulin, C.: Dictionary learning based sinogram inpainting for ct sparse reconstruction. Optik - Int. J. Light Electron Optics **125**, 2862–2867 (2014)
16. Mao, X., Shen, C., Yang, Y.: Image denoising using very deep fully convolutional encoder-decoder networks with symmetric skip connections. CoRR abs/1603.09056 (2016)
17. Radon, J.: On the determination of functions from their integral values along certain manifolds. IEEE Trans. Med. Imaging **5**(4), 170–176 (1986)

18. Ramachandran, G.N., Lakshminarayanan, A.V.: Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of fourier transforms. Proc. Natl. Acad. Sci. USA **68**(9), 2236–40 (1971)
19. Ren, D., Zuo, W., Hu, Q., Zhu, P., Meng, D.: Progressive image deraining networks: a better and simpler baseline. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
20. Sidky, E.Y., Pan, X.: Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Phys. Med. Biol. **53**, 4777 (2008)
21. Tao, X., Gao, H., Wang, Y., Shen, X., Wang, J., Jia, J.: Scale-recurrent network for deep image deblurring. CoRR abs/1802.01770 (2018)
22. Wang, W., et al.: An end-to-end deep network for reconstructing CT images directly from sparse sinograms. IEEE Trans. Comput. Imaging **6**, 1548–1560 (2020)
23. Wecksung, G.W., Kruger, R.P., Morris, R.A.: Fan-to parallel-beam conversion in cat by rubber sheet transformation. Appl. Digital Image Process. **III**(0207), 76–83 (1979)
24. Xie, Q., et al.: Robust low-dose CT sinogram preprocessing via exploiting noise-generating mechanism. IEEE Trans. Med. Imaging **36**(12), 2487–2498 (2017)
25. Xu, L., Ren, J.S., Liu, C., Jia, J.: Deep convolutional neural network for image deconvolution. Adv. Neural Inf. Process. Syst. **27**, 1790–1798 (2014)
26. Xu, Q., Yu, H., Mou, X., Zhang, L., Hsieh, J., Wang, G.: Low-dose x-ray ct reconstruction via dictionary learning. IEEE Trans. Med. Imaging **31**, 1682–1697 (2012)
27. Yan, K., Wang, X., Lu, L., Summers, R.M.: Deeplesion: automated deep mining, categorization and detection of significant radiology image findings using large-scale clinical lesion annotations. CoRR abs/1710.01766 (2017)
28. Yang, Q., et al.: Low-dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss. IEEE Trans. Med. Imaging **37**(6), 1348–1357 (2018)
29. Yu, W., Wang, C., Nie, X., Huang, M., Wu, L.: Image reconstruction for few-view computed tomography based on l0 sparse regularization. Procedia Comput. Sci. **107**, 808–813 (2017)
30. Zhang, H., Sindagi, V., Patel, V.M.: Joint transmission map estimation and dehazing using deep networks. IEEE Trans. Circ. Syst. Video Technol. **30**(7), 1975–1986 (2019)
31. Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans. Image Process. **26**(7), 3142–3155 (2017)
32. Zhang, Y., Zhang, W.H., Chen, H., Yang, M., Li, T.Y., Zhou, J.L.: Few-view image reconstruction combining total variation and a high-order norm. Int. J. Imaging Syst. Technol. **23**, 249–255 (2013)
33. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. CoRR abs/1807.02758 (2018)
34. Zhu, Y., Zhao, M., Zhao, Y., Li, H., Zhang, P.: Noise reduction with low dose CT data based on a modified ROF model. Opt. Express **20**(16), 17987–18004 (2012)