

Superiorization-inspired unrolled SART algorithm with U-Net generated perturbations for sparse-view and limited-angle CT reconstruction

Yiran Jia, Noah McMichael, Pedro Mokarzel, Brandon Thompson, Dong Si, and Thomas Humphries

School of STEM, University of Washington Bothell, Bothell, WA 98011

E-mail: thumphri@uw.edu

Abstract.

Objective: Unrolled algorithms are a promising approach for reconstruction of CT images in challenging scenarios, such as low-dose, sparse-view and limited-angle imaging. In an unrolled algorithm, a fixed number of iterations of a reconstruction method are unrolled into multiple layers of a neural network, and interspersed with trainable layers. The entire network is then trained end-to-end in a supervised fashion, to learn an appropriate regularizer from training data. In this paper we propose a novel unrolled algorithm, and compare its performance with several other approaches on sparse-view and limited-angle CT.

Approach: The proposed algorithm is inspired by the superiorization methodology, an optimization heuristic in which iterates of a feasibility-seeking method are perturbed between iterations, typically using descent directions of a model-based penalty function. Our algorithm instead uses a modified U-net architecture to introduce the perturbations, allowing a network to learn beneficial perturbations to the image at various stages of the reconstruction, based on the training data.

Main Results: In several numerical experiments modeling sparse-view and limited angle CT scenarios, the algorithm provides excellent results. In particular, it outperforms several competing unrolled methods in limited-angle scenarios, while providing comparable or better performance on sparse-view scenarios.

Significance: This work represents a first step towards exploiting the power of deep learning within the superiorization methodology. Additionally, it studies the effect of network architecture on the performance of unrolled methods, as well as the effectiveness of the unrolled approach on both limited-angle CT, where previous studies have primarily focused on the sparse-view and low-dose cases.

Keywords: computed tomography, deep learning, iterative reconstruction, sparse view, limited angle

Submitted to: *Phys. Med. Biol.*

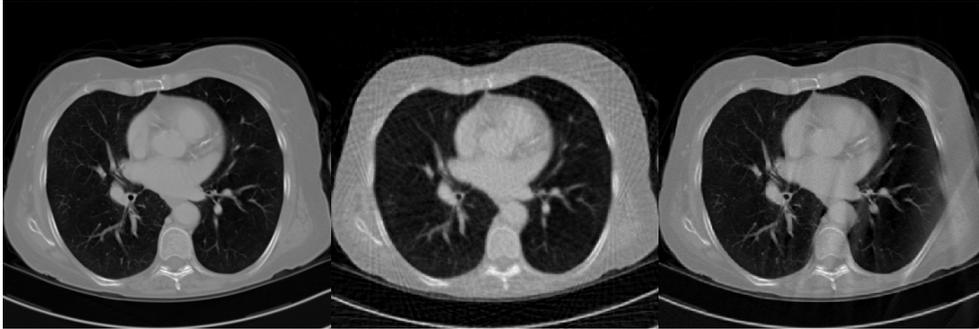


Figure 1. Images reconstructed from simulated sinogram data showing characteristic artifacts in sparse-view (centre) and limited-angle (right) scenarios. Ground truth image shown on left. All three images were reconstructed using the Simultaneous Algebraic Reconstruction Technique (SART)

1. Introduction

Computed Tomography (CT) is widely utilized in clinical, industrial and other applications; for example, it is estimated that roughly 80 million CT scans were performed in the United States in 2017 (OECD 2019). Despite the widespread use of CT in clinics, several longstanding challenges remain in the field. One such scenario is sparse-view CT (Sidky et al. 2006), which uses fewer projection views of the patient than a conventional scan. This approach is motivated by the need to reduce dose to the patient, stemming from rising concerns about the impact of the cumulative dose associated with CT on lifetime risk of cancers (Power et al. 2016). Another problem arising from incomplete data is the limited-angle scenario, wherein the scan can only be performed over a restricted range of angles, due to limited scan flexibility, inspection requirements, or object size (Frikel & Quinto 2013). While the incomplete nature of the data leads to image artifacts in both scenarios, the characteristics of sparse-view and limited-angle artifacts differ considerably, as shown in Figure 1.

Analytic reconstruction techniques such as filtered backprojection (FBP), while computationally efficient, produce severe artifacts when presented with incomplete data. Iterative algorithms, such as the algebraic reconstruction technique (ART) (Gordon et al. 1970) and simultaneous algebraic reconstruction technique (SART) (Andersen & Kak 1984), are more computationally intensive, but offer greater flexibility for dealing with incomplete data. Over the past few decades, researchers have made many efforts on developing regularized iterative reconstruction techniques to solve ill-posed inverse problems and obtain better quality images; in particular, by focusing on the use of prior knowledge of the reconstructed object.

Advances in compressive sensing (CS) showed that an underdetermined linear system can be solved exactly if some prior information of sparsity is used (Foucart & Rauhut 2013), subject to conditions on the measurement model, such as the restricted isometry property. Researchers have proposed various forms of prior knowledge, including dictionaries (Lu et al. 2011, Xu et al. 2012, Xu et al. 2015), wavelet tight frames (Jia et al. 2011, Dong et al. 2013), edge information (Wang et al. 2009, Charbonnier et al. 1997), and others (Zhang, Zeng, Lin, Zhang, Bian, Huang, Gao, Zhang, Zhang, Feng et al. 2017, Zhang, Zeng, Zhang, Wang, Liang & Ma 2017, Zhang, Hu, Yang, Chen, Coatrieux & Luo 2017, Kim et al. 2014). One popular underlying

assumption is the sparsity of the image gradient. By minimizing the ℓ_1 norm of the image gradient (total variation (TV)-based minimization), (Sidky et al. 2006) showed promising results in several incomplete data scenarios. The piecewise-constant assumption implied by gradient sparsity, however, is overly simplistic in many cases. This may lead to blocky artifacts in images reconstructed using TV minimization (Yu & Wang 2009). In addition, the blurred edge artifacts appearing in limited-angle images (cf. Figure 1) suggest that isotropic TV minimization is less well suited for this scenario (Chen et al. 2013). Variants of TV have therefore been proposed to better preserve structural details (Wang et al. 2017, Yu et al. 2017, Shen et al. 2017, Mahmood et al. 2018). In general, however, there is no universal prior knowledge that can fit all types of incomplete data problem, and it is also difficult to adjust the parameters of the regularization terms.

More recently, there has been an explosion in interest in the use of techniques from deep learning (LeCun et al. 2015) to address challenging problems such as low-dose, sparse-view and limited angle CT imaging, as well as other image reconstruction problems. A recent review is provided in (Ongie et al. 2020). For example, (Zhang et al. 2016) suggested a deep learning method to reduce artifacts in limited angle CT images reconstructed with FBP. (Chen, Zhang, Zhang, Liao, Li, Zhou & Wang 2017) developed a three-layer convolutional neural network (CNN) to reduce artifacts in low-dose images. More complex networks were proposed in the same year, such as the wavelet network (Kang et al. 2017) trained in the wavelet domain, and RED-CNN (Chen, Zhang, Kalra, Lin, Chen, Liao, Zhou & Wang 2017), an algorithm using the idea of encoder-decoder with residual learning to denoise low-dose images. (Han et al. 2016) developed a residual learning method using the U-Net based architecture for sparse view image reconstruction. Independently, (Jin et al. 2017) proposed FBPCNN, where the low-dose CT images are first reconstructed with FBP, and then processed by a U-Net based CNN. (Ye et al. 2018) proposed a frame-based deep learning denoising approach that employs an encoder-decoder network. Based on the theory of deep convolutional framelets developed in this paper, (Han & Ye 2018) proposed U-Net variants that meet the frame condition and give better results on sparse-view CT image reconstructions.

The aforementioned approaches can be viewed as postprocessing methods, where an initial reconstruction containing artifacts is fed into the pre-trained CNN. A second type of approach is to use a CNN within an reconstruction algorithm to recover an image from the original sinogram. (Pelt & Batenburg 2013) proposed a neural network filtered backprojection method (NN-FBP), which can be viewed as a weighted combination of FBP method with some learned filters. (Würfl et al. 2016) subsequently expressed filtered back-projection-type algorithms as a deep neural network and demonstrated that learning the weighting and additional filter layers efficiently reduce the reconstruction error. (Adler & Öktem 2017) proposed an approach that partially learns the gradient descent at each stage of iterative optimization by unrolling the entire reconstruction pipeline into a deep CNN. Later, they proposed another approach (Adler & Öktem 2018) that combines the deep learning network with a proximal primal-dual optimization method in the same unrolling manner. (Chen et al. 2018) constructed the LEARN network, which unfolds the iterative procedure into a recurrent residual network with three convolutional layers in each iteration. (Gupta et al. 2018) proposed a guaranteed-convergent iterative image reconstruction algorithm that replace the projector in a projected gradient descent (PGD) with a training of U-Net. Inspired by generative adversarial network (GAN), (Lunz

et al. 2018) proposed a data-driven framework that trains a regularizer through learning to discriminate between the distribution of label images and unregularized reconstructions. (Zheng et al. 2018) proposed a penalized weighted least squares (PWLS) reconstruction incorporating a union of learned transforms, which are pre-learned from image patches extracted from low-dose CT scans.

Building off of these latter works, we propose an unrolled reconstruction algorithm similar to the LEARN algorithm (Chen et al. 2018), which was investigated in the context of sparse-view imaging. The contributions of our paper are the following:

- (i) Rather than regularization, the algorithm is inspired by the superiorization methodology (SM) (Herman et al. 2012), an optimization heuristic which provides theoretical justification for introducing perturbations between iterations of a certain class of algorithms. The simultaneous algebraic reconstruction technique (SART), which we use, is one such algorithm. Given the recent interest in using deep learning for both CT image reconstruction and to enhance optimization algorithms more broadly, it is natural to consider how deep learning can be exploited in the context of the SM as well. This work represents a first step towards doing so.
- (ii) We compare the use of a modified U-net architecture (Ronneberger et al. 2015) to regularize the reconstruction with a simpler 3-layer convolutional architecture used in (Chen et al. 2018). The U-net is well known for its ability to preserve high-resolution image features while also extracting global features of the image using its contracting and expanding paths. We therefore wish to study whether it is also able to improve the performance of an unrolled algorithm.
- (iii) We apply the algorithm to both limited-angle and sparse-view CT. Previous studies using unrolled algorithms (Chen et al. 2018, Adler & Öktem 2017, Adler & Öktem 2018) have typically focused on the sparse-view and low-dose cases. As shown in Figure 1, limited-angle imaging typically presents more severe artifacts than sparse-view imaging, and thus represents a more challenging reconstruction problem. While several papers have used techniques from deep learning to reconstruct CT images from limited angle data, we are not aware of work that has used unrolled algorithms for this purpose.

This paper substantially extends our work first presented in (Jia et al. 2020), by including further experimentation (including the sparse-view scenario), comparison with other approaches, and analysis of the results.

2. Methodology

2.1. Iterative reconstruction

In two dimensional imaging, the CT reconstruction problem is to recover an $n \times n$ pixel transaxial slice of the object from sinogram data acquired from n_v discrete angular views surrounding it, each consisting of n_b measurements (line integrals) through the object. This can be modeled as a linear inverse problem,

$$Ax + \eta = b \tag{1}$$

where $x \in \mathbb{R}^N$ is the unknown image to be recovered ($N = n^2$), $b \in \mathbb{R}^M$ is the sinogram ($M = n_b n_v$), η is the measurement noise, and $A \in \mathbb{R}^{M \times N}$ is the CT system matrix whose (i, j) th element represents the contribution of the j th image pixel to the i th

measurement. To accurately recover x from b without any prior information, some necessary criteria are:

- (i) The magnitude of the noise, η , should not be too large compared to the data, b . This is typically achieved by requiring that the X-ray beam be sufficiently intense. Low-dose imaging (where beam intensity or exposure time is reduced, resulting in much noisier measurements) is a problem that has attracted significant interest in the CT imaging community, though we do not consider it in this paper.
- (ii) The number of angular samples, n_v , must be sufficient to recover the highest-frequency components of the image. Sparse-view CT corresponds to the case when n_v is much smaller than required, resulting in streaking artifacts such as those seen in Figure 1, due to undersampling. Sparse-view imaging is another approach to reducing patient dose, versus low-dose imaging.
- (iii) Data must be acquired over an angular extent that is sufficient to obtain edge information about all structures within the object. In a parallel-beam geometry, this is simply a 180° arc, while in fan-beam geometries the arc must be greater than 180° to account for the fan angle (though a 360° arc is often used for simplicity, as in this paper). Limited-angle CT corresponds to the case where data is acquired over a smaller arc than is required, resulting in blurred edges such as those seen in Figure 1.

The linear inverse problem (1) must be solved iteratively; the size of the problem typically makes direct solution impossible, and the presence of noise in the data makes an exact solution undesirable in any event. To reduce noise and sparse-view or limited angle artifacts, one can incorporate a penalty function $\phi(x) : \mathbb{R}^N \rightarrow \mathbb{R}$ into the iterative algorithm. The value of $\phi(x)$ should be large for undesirable images, and smaller for ones with more desirable properties. The LEARN algorithm (Chen et al. 2018) is motivated by gradient descent applied to the regularized least-squares problem:

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|Ax - b\|^2 + \mu\phi(x), \quad (2)$$

which results in the iteration

$$x^{k+1} = x^k - \alpha_k [A^T (Ax^k - b) + \mu \nabla \phi(x^k)]. \quad (3)$$

Here α_k is the step size (based on a sufficient decrease condition, for example), and μ is parameter that weights the penalty function versus the data fidelity term.

Our approach is inspired by the superiorization methodology (SM) (Herman et al. 2012), whose key concepts we now describe. An iterative algorithm for solving (1) (known as a *basic algorithm* in the context of the SM) can be expressed in the form $x^{k+1} = R(x^k)$ for some function $R : \mathbb{R}^N \rightarrow \mathbb{R}^N$. The *superiorized version* of this basic algorithm R takes the form

$$x^{k+1} = R(x^k + \beta_k v^k), \quad (4)$$

where v^k is chosen to be a nonascending direction of the penalty function ϕ , and β_k is a step size chosen such that $\phi(x^k + \beta_k v^k) \leq \phi(x^k)$. Using the residual of (1) as a measure of proximity to an exact solution, a solution x^* produced by the basic algorithm is said to be ε -compatible if $\|Ax^* - b\| < \varepsilon$.

The basic algorithm R is then said to be *strongly perturbation resilient* if, whenever it produces an ε^* -compatible solution (for some $\varepsilon^* > 0$) from the initial iterate x^0 , the superiorized version (4) is also guaranteed to eventually converge to ε^* -compatible

solution from the same starting point, provided that the directions v^k are bounded in size, and that $\sum_{k=0}^{\infty} \beta_k < \infty$. In other words, the iterate produced by the basic algorithm can be perturbed in every iteration in a nonascending direction of ϕ – as long as the perturbations are summable in size – without jeopardizing convergence to a ε -compatible solution. This solution is also expected to be superior to that produced by the basic algorithm with respect to the function ϕ (i.e., have a smaller function value). Some sufficient conditions for an algorithm to be strongly perturbation resilient are provided in (Herman et al. 2012).

For this work we choose the simultaneous algebraic reconstruction technique (SART) (Andersen & Kak 1984), as the basic algorithm. This algorithm is known to be perturbation resilient (Butnariu et al. 2007), and the superiorized version can be expressed as an alternating iteration (see e.g. (Humphries et al. 2020):

$$x_+^k = x^k - \lambda_k D A^T M (A x^k - b) \quad (5)$$

$$x^{k+1} = x_+^k - \beta_k \nabla \phi(x_+^k) / \|\nabla \phi(x_+^k)\|. \quad (6)$$

Line (5) is the basic SART iteration[‡], where D and M are diagonal scaling matrices whose entries are obtained from column and row sums of A , respectively. The relaxation parameter $\lambda_k \in (0, 2)$ can vary with the iteration number, and is typically chosen to be close to 2 in order to accelerate convergence (Elfving et al. 2012). Line (6) is the superiorization step, where the perturbation v_k is defined to be the normalized negative gradient of ϕ , which is guaranteed to be a nonascending direction of the function.

2.2. Unrolled iterative methods

The iterations (3) and (5),(6) can be run for as many iterations as desired. One way to incorporate a neural network into iterative reconstruction is to fix the total number of iterations at some value K and then “unroll” the sequence of K iterations into a single pipeline. This pipeline may consist of some layers that are not trainable (for example, multiplication by A and A^T) interspersed with layers that employ trainable parameters. The LEARN algorithm (Chen et al. 2018), which inspired our work, mirrors iteration (3), with $\nabla \phi$ replaced by a convolutional neural network (CNN); i.e.

$$x^{k+1} = x^k - \alpha_k A^T (A x^k - b) - \Psi(x^k; \Theta^k). \quad (7)$$

Here $\Psi(x^k; \Theta^k)$ is a CNN with parameters Θ^k , which can be viewed as a trainable, iteration dependent regularizer. The regularization weighting parameter μ is absorbed into Ψ , and the step size α_k is treated as another trainable parameter, instead of performing a line search. The algorithm was applied to sparse-view imaging in (Chen et al. 2018), where it was shown to outperform other state-of-the-art methods.

In a similar way, we propose modifying the superiorized SART algorithm (5), (6) such that the perturbations are generated by a neural network. In this framework, (6) is replaced by:

$$x^{k+1} = x_+^k + \Psi(x_+^k; \Theta^k), \quad (8)$$

where $\Psi(x_+^k; \Theta^k)$ is again a CNN with parameters Θ^k . Equation (5) is essentially unchanged, although the relaxation parameter λ_k is also made into a trainable parameter. The idea is to exploit the power of deep neural networks to learn data-driven perturbations, with the expectation that this will produce higher-quality images

[‡] Some references call this algorithm SIRT (Simultaneous Iterative Reconstruction Technique).

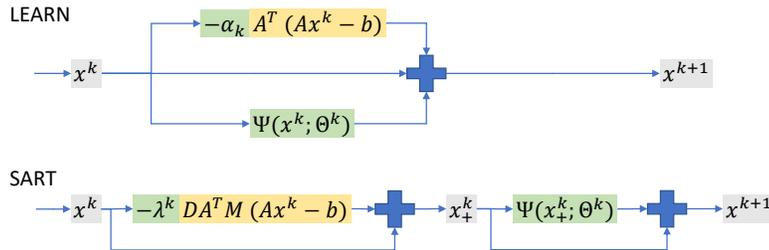


Figure 2. Diagram showing one iteration of each unrolled algorithm. Learned parameters and operations are shown in green, while untrained operations are shown in yellow.

than perturbations derived from nonascending directions of a model-based penalty function. This expectation is based on the flexibility that deep neural networks afford to determine a function appropriate to the distribution of images in training data, as well as the fact that the neural network parameters Θ^k can vary from iteration to iteration, thus allowing for the perturbations to be tailored to the changing properties of the image as it converges.

Visualizing the entire unrolled neural network, the main difference between these two approaches is that the CNN is applied serially with the iterative step in the SART algorithm, as opposed to in parallel as in the LEARN algorithm. We illustrate the difference between the two unrolled architectures in figure 2

2.3. Network architecture

In (Chen et al. 2018), the term $\Psi(x^k; \Theta^k)$ in (7) is implemented as a CNN consisting of three convolutional layers with ReLU activation. We instead consider using a model based on the U-net, a deep neural network originally designed for biomedical segmentation (Ronneberger et al. 2015). We choose the U-net as the base architecture for our CNN model for two reasons. First, the contracting path (left side) of the U-net consists of a sequence of max pooling operations which halve the size of the image, allowing the number of convolutional filters to double without increasing computational cost. The increased number of filters allows for better recognition of image features, while the effective field of view of each filter is also increased, allowing for the extraction of global features of the artifacts. Secondly, the expanding path (right side) of the U-net upsamples the feature map and concatenates with the correspondingly cropped feature map from the contracting path, allowing high-resolution features of the image to be preserved, despite the downsampling applied in the contracting path.

For our algorithm, we made a number of modifications to the architecture in (Ronneberger et al. 2015), as indicated in Figure 3. First, we reduced the size of the four-layer contracting path and expanding path to two layers. This was necessary because the full U-Net includes 31 million trainable parameters; since our approach applies the network between each iteration of SART (with parameters dependent on the iteration), the memory requirements for a full U-net would be prohibitive. Even with the reduced size, we were only able to apply the U-Net based regularization between every four iterations of SART. Second, we replaced the Rectified Linear Unit

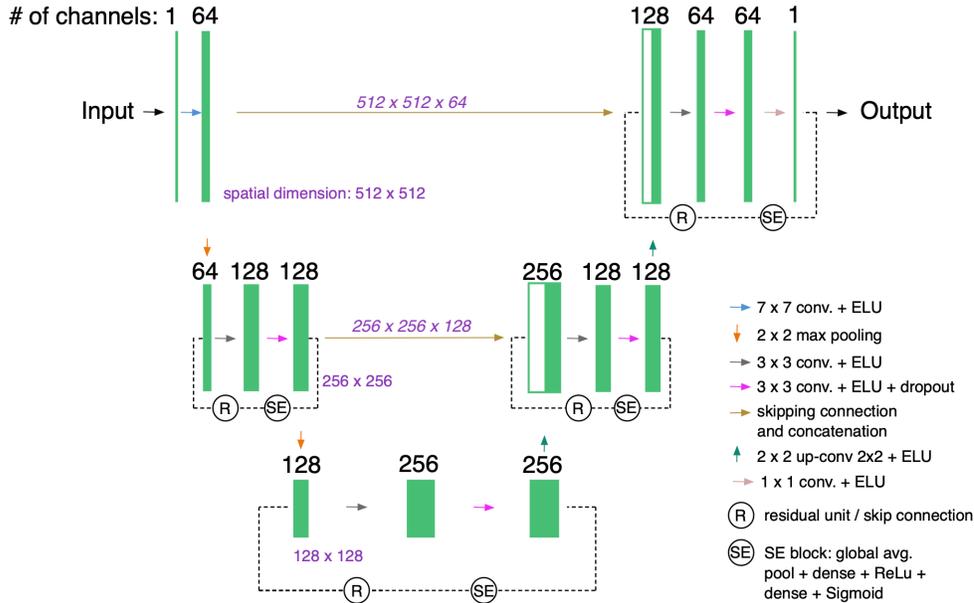


Figure 3. U-net-based CNN implemented in the proposed method. © 2020 IEEE. Reprinted, with permission, from (Jia et al. 2020).

(ReLU) activation functions with the Exponential Linear Unit (ELU) to avoid issues with so-called dying ReLUs (Géron 2017). The dying ReLU issue sometimes occurs due to the flat portion of the ReLU that occurs for negative inputs, which can cause parts of the network to become inactive and impedes training (Lu et al. 2019); ELUs are one way of alleviating this issue. Additionally, inspired by residual networks (He et al. 2016), we added a skip connection across the convolutional layers at every level of the U-Net, to allow the the signal to easily make its way across the whole network within an iteration. We employed Squeeze-and-Excitation (SE) blocks (Hu et al. 2018) for each of the residual units. A SE block analyzes the output of the unit to which it is attached, and learns which features are usually most active together. It then uses this information to recalibrate the feature maps. Finally, we replaced the last convolutional and sigmoid layer with one convolutional layer which reduces the 64 channels to a single channel array with the same size as the initial input. The He initialization method (He et al. 2015) was used to initialize all the convolutional layers, to aid in convergence.

2.4. Connection to the superiorization methodology

While our unrolled algorithm is inspired by the SM, there are several ways in which it does not meet the requirements for a conventionally superiorized algorithm, as described in (Herman et al. 2012) and elsewhere. First, since the perturbations are generated by a neural network, they are not descent directions of any prescribed secondary criterion $\phi(x)$, as in the conventional case (6). This is a feature of the approach, as it allows the network to learn a beneficial sequence of data-driven perturbations without being constrained by a model-based secondary criterion such as

TV. Nonetheless, in the absence of such a criterion, it is reasonable to ask in what sense the images produced by the unrolled algorithm are “superior” to those produced by the basic algorithm. In general, we expect that this will be dependent on the criteria used to train the network; for example, the qualitative and quantitative results presented in the next section show that these images are significantly improved versus those obtained from the basic algorithm, as a consequence of training the unrolled network to minimize the ℓ_2 loss between reconstructed and true images in the paired training data.

Secondly, a main result in the SM is that the superiorized version of a bounded perturbation resilient algorithm is guaranteed to eventually converge to an ε -compatible solution to the linear inverse problem, provided that the perturbations are summable. This guarantee requires that the algorithm can be iterated indefinitely, however, while an unrolled algorithm requires specifying a finite number of iterations. Thus, the solution produced by an unrolled algorithm is not guaranteed to be ε -compatible, even if it superior in quality to the image produced by the basic algorithm. We discuss a potential way to address this in Section 4.

Finally, in order to guarantee convergence to a ε -compatible solution, the perturbations generated by the superiorized algorithm must be summable in the limit as $k \rightarrow \infty$ (k being the iteration number), which is typically achieved by forcing the perturbation size to decrease geometrically. While it is possible to bound the size of the perturbations generated by the U-Net in the same way, we have not done so in our algorithm; since the unrolled algorithm includes only finitely many perturbations, these perturbations will be summable in any event, provided that they are bounded in size. We discuss this further together with the issue of ε -compatibility in Section 4.

We note as well that since the U-Net based perturbation is only applied every fourth iteration of SART in our unrolled algorithm, there are two ways of interpreting this in the context of the SM; either the basic algorithm is four iterations of SART, with a perturbation applied between each iteration of the basic algorithm; or one iteration of SART still represents the basic algorithm, and the perturbation size in three out of four iterations is zero. Both are compatible with the SM as originally described.

3. Numerical Experiments

3.1. Description

The training and test datasets were generated using 512×512 pixel slices from a lung CT study obtained from the Cancer Imaging Archive (Clark et al. 2013, Goldgof et al. 2015). A total of 3629 slices obtained from forty-two patients were used to generate the training set, while 325 slices obtained from five different patients were used to generate the validation set. The dataset included images ranging from the lower abdomen to upper thorax.

Our algorithm is implemented in Python using Tensorflow and the recently developed Pyro-NN library (Syben et al. 2019), which allows forward and backward projection operations to be readily integrated within a neural network. Simulated sinogram data of each image were generated using the forward projection operators in Pyro-NN. All sinograms were generated assuming a fan-beam geometry and pixel size of 0.65 mm. The source-isocenter and source-detector distance were 1000 mm and 2000 mm, respectively, and the detector consisted of 729 pixels with width 1.73

Table 1. Scan parameters for the five imaging scenarios that were considered. $\Delta\theta$ refers to the spacing between consecutive views of the object.

Scenario	Angular range ($^{\circ}$)	Number of views	$\Delta\theta$ ($^{\circ}$)
sv180	360	180	2
sv90	360	90	4
sv60	360	60	6
1a162	162	405	0.2
1a144	144	360	0.2

mm, corresponding to a fan angle of 35° . The Poisson noise was added subsequently, proportional to an initial intensity of 10^6 counts per line. This represents a relatively small amount of noise, but ensured that there was some inconsistency between the forward projection of the ground truth and the measured image, as in a true scan. Five different scenarios were simulated: three sparse-view scenarios with 180, 90, and 60 views (denoted as **sv180**, **sv90**, and **sv60**, respectively) and two limited-angle scenarios with angular ranges of 162° and 144° (**1a162** and **1a144**). The parameters for these scenarios are summarized in Table 1.

To compare the effects of the different network architectures, we implemented a total of six reconstruction approaches in Pyro-NN. These approaches fell into three categories:

- (i) Basic algorithms (**learn_basic** and **sart_basic**): in these approaches the CNN component of the algorithm was omitted, resulting in an unregularized reconstruction. Thus, **learn_basic** essentially implements gradient descent of the data fidelity term only, and **sart_basic** implements a standard SART iteration. For ease of comparison with the other approaches, however, the line search parameter α_k in (7) and relaxation parameter λ_k in (5) were still learned during training.
- (ii) LEARN algorithms (**learn_3L** and **learn_unet**): these approaches implemented the LEARN algorithm, as described in (7). **learn_3L** used the 3-layer CNN as described in the original paper (Chen et al. 2018), while **learn_unet** uses our modified U-net architecture, as described in Section 2.3.
- (iii) SART algorithms (**sart_3L** and **sart_unet**): these approaches implemented the SART algorithm with learned regularizer, again using either a 3-layer CNN (**sart_3L**) or modified U-net (**sart_unet**).

All algorithms were unrolled using a fixed number of $K = 40$ iterations. Due to the much larger size of the modified U-Net versus the three-layer CNN, the regularization term could only be included between every fourth iteration of the **learn_unet** and **sart_unet** approaches, while **learn_3L** and **sart_3L** included regularization between every iteration. All networks were trained to minimize the ℓ_2 loss between the ground truth image and the images reconstructed from the sinograms. Training and testing of the networks was performed on a dedicated server with two NVIDIA Quadro RTX 5000 GPUs, each having 16 GB of SDRAM.

For the sake of comparison, reconstructions using filtered back projection (FBP) were also computed in all experiments; while FBP typically performs more poorly than iterative methods for ill-posed problems, it is significantly less computationally expensive, and can produce good results in a sufficiently well-posed setting. We also considered two competing algorithms designed for ill-posed reconstruction problems.

The first is a superiorized version of SART using total variation (TV):

$$\phi(x) = \sum_{m,n} \sqrt{(x_{m+1,n} - x_{m,n})^2 + (x_{m,n+1} - x_{m,n})^2 + \epsilon^2}. \quad (9)$$

Here x is represented as a two-dimensional image, and $\epsilon = 10^{-6}$ is a small parameter to ensure differentiability. TV-superiorized SART was used for sparse-view CT in one of the earliest papers introducing the SM (Herman et al. 2012), as well as for sparse-view and limited-angle CT in (Humphries et al. 2017). Typically, the basic framework of (5),(6) is modified to allow for $N \geq 1$ nonascending steps (6) of the penalty function between every iteration of the basic algorithm, to allow for further improvement with respect to ϕ ; additionally, the parameter β_k is chosen to decrease geometrically by some factor $0 < \gamma < 1$ to ensure that the steps are summable. Finally, a sufficient decrease condition is included to ensure that $\phi(x)$ decreases in each iteration; for more details, see (Herman et al. 2012). In our experiments, we denoted the algorithm as `sart.TV` and implemented it in Python using the ASTRA toolbox (Palenstijn et al. 2011, Van Aarle et al. 2016); implementation in ASTRA was more straightforward than PYRO-NN, as the approach does not include any learned parameters. For the three sparse-view experiments, we used $\gamma = 0.995$ and $N = 10$, while for the limited-angle experiments, we used $\gamma = 0.9995$ and $N = 60$, which yielded better results. We also used ordered subsets and a relaxation parameter value of $\lambda_k = 1.9$ to accelerate convergence of the algorithm.

The second competing approach was the learned gradient descent (LGD) method (Adler & Öktem 2017). This is another unrolled method, whose basic iterative formula is given by:

$$(s^{k+1}, \Delta x^{k+1}) = \Psi(s^k, x^k, \nabla \mathcal{D}(x^k), \nabla \phi(x^k); \Theta^k) \quad (10)$$

$$x^{k+1} = x^k + \Delta x^{k+1} \quad (11)$$

As before, $x^k \in \mathbb{R}^N$ represents the k th iterate of the image being reconstructed, ϕ is the penalty function (total variation, in our experiments), and Ψ represents a CNN with parameters Θ^k . A discrepancy functional is specified by $\mathcal{D} : \mathbb{R}^N \rightarrow \mathbb{R}$; we used $\mathcal{D}(x) = \frac{1}{2} \|Ax - b\|^2$, and so $\nabla \mathcal{D}(x) = A^T(Ax - b)$. Finally, $s^k \in \mathbb{R}^{N \times M}$ represents a persistent memory term that allows the algorithm to use information from earlier iterates, inspired by quasi-Newton methods. This algorithm is a natural point of comparison with LEARN (7) and our unrolled SART approach, as all are based on a gradient descent-like iteration with a penalty term. The key difference is that LGD freely chooses a search direction based on information about the gradients of \mathcal{D} and a suitable regularizer ϕ ; LEARN and our approach fix the descent direction with respect to \mathcal{D} and choose the regularization freely.

We used the implementation of the LGD method provided by the authors§ and modified it to allow for training on our dataset with the five described scenarios. The network Ψ in this implementation is a simple three-layer CNN, similar to that used by the standard LEARN algorithm, and the CT projection operations are implemented using the Operator Discretization Library (ODL)||. As with the other CNN-based approaches, the LGD algorithm was trained to minimize the ℓ_2 loss between the ground truth image and the images reconstructed from the sinograms. The sinogram data for both the `sart.TV` and LGD experiments were generated using the same parameters as in Table 1, but with the CT system matrices native to the respective software

§ https://github.com/adler-j/learned_primal_dual

|| <https://github.com/odlgroup/odl>

Table 2. Summary of different reconstruction approaches used in the numerical experiments

Algorithm	Type	Implementation	Regularization	Parameters
FBP	Analytic	ASTRA	None	0
learn_basic	Iterative	Pyro-NN	None	40
sart_basic	Iterative	Pyro-NN	None	40
sart_TV	Iterative	ASTRA	TV	3
LGD	Unrolled	ODL	TV-guided	13,318
learn_3L	Unrolled	Pyro-NN	Learned	1,525,800
learn_unet	Unrolled	Pyro-NN	Learned	18,627,890
sart_3L	Unrolled	Pyro-NN	Learned	1,525,800
sart_unet	Unrolled	Pyro-NN	Learned	18,627,890

packages, rather than those used by Pyro-NN. This ensured a fair comparison across all approaches, as there was no mismatch between the system matrices used to generate the data, and those used during reconstruction.

A summary of all the reconstruction algorithms that were tested is provided in Table 2. In addition to the basic type of algorithm, the library used to implement it, and the type of regularization used (if any), we also specify the number of parameters required or learned by each algorithm. For **learn_basic** and **sart_basic**, these are the line search or relaxation parameter, respectively, in every iteration, while for **SART_TV**, these are the values of the relaxation parameter, N , and γ mentioned above. The LGD method has relatively few network parameters to learn, compared with the **learn_3L** and **sart_3L** approaches (38,145 trainable parameters per iteration, times 40 iterations) and especially the **learn_unet** and **sart_unet** approaches; each copy of our reduced U-Net model consist of 1,862,789 trainable parameters, times 10 iterations. The **learn_3L** algorithm is our own implementation in Pyro-NN of the algorithm presented in (Chen et al. 2018), while the final three algorithms listed are our new contributions.

The quality of the images reconstructed using the various approaches was compared using two standard image quality metrics: peak signal-to-noise ratio (PSNR) and Structural SIMilarity index (SSIM)(Wang et al. 2004). The PSNR of a reconstructed image $x \in \mathbb{R}^N$, compared with the ground truth y , is calculated as

$$PSNR(x, y) = 10 \log_{10} \left(\frac{y_{max}}{MSE(x, y)} \right), \text{ where} \quad (12)$$

$$MSE(x, y) = \frac{1}{N} \sum_j (x_j - y_j)^2, \text{ and } y_{max} = \max_j \{y_j\}. \quad (13)$$

The better the agreement between the images, the higher the value of the PSNR; identical images produce infinite PSNR.

The SSIM is calculated as

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (14)$$

where μ_x, μ_y are the means of the two images, and σ_x, σ_y and σ_{xy} the variances and covariance, respectively. The small constants C_1 and C_2 are included to avoid instability when the denominator is near zero. The SSIM produces a score between 0 and 1, with 1 corresponding to identical images. Both metrics were computed for the entire 512×512 pixel images.

Table 3. Average PSNR and SSIM values for the five scenarios and eight reconstruction methods. Best overall results are highlighted in bold. Multiple results are bolded in cases where ANOVA indicated insignificant differences.

Method	sv180		sv90		sv60		1a162		1a144	
	PSNR	SSIM								
FBP	34.740	0.762	29.226	0.614	26.558	0.534	26.975	0.792	24.325	0.758
learn_basic	36.093	0.867	32.052	0.785	29.872	0.723	31.547	0.886	28.179	0.829
sart_basic	33.275	0.874	31.414	0.794	29.870	0.745	30.124	0.850	27.522	0.810
sart_TV	41.687	0.973	37.042	0.937	33.777	0.891	36.098	0.953	32.720	0.932
LGD	43.807	0.972	41.066	0.964	38.030	0.943	37.933	0.960	34.880	0.945
learn_3L	43.773	0.978	42.856	0.973	40.068	0.956	40.772	0.967	38.019	0.954
learn_unet	44.257	0.979	41.717	0.964	39.040	0.948	40.614	0.956	37.456	0.935
sart_3L	41.795	0.970	38.674	0.947	37.160	0.935	37.717	0.958	35.391	0.938
sart_unet	44.265	0.979	41.236	0.966	39.734	0.957	42.841	0.979	38.956	0.967

3.2. Results

Table 3 shows the PSNR and SSIM values, averaged over the 325 test images, for every scenario and reconstruction method. One-way analysis of variance (ANOVA) and a multiple comparison test were run to determine if differences in image quality between the reconstruction methods were statistically significant or not. It is clear that FBP performs most poorly, followed by the two basic (unregularized) approaches, and that all approaches incorporating regularization resulted in significant improvement. The **sart_unet** method gave the best results for both of the limited-angle scenarios (**1a162** and **1a144**). For the sparse-view scenarios, it was tied for best performance with the **learn_3L** and **learn_unet** methods for the **sv180** and **sv60** datasets, but gave somewhat worse results than **learn_3L** for the **sv90** dataset. The **sart_3L** method, on the other hand, performed more poorly than these three approaches.

The **LGD** method was generally competitive with the **LEARN** and **SART**-based methods for sparse-view reconstruction, though its performance tended to be slightly worse than the best-performing methods. For the the limited-angle experiments, there was a greater discrepancy in performance. Meanwhile, the **sart_TV** approach was the least competitive of all the regularized approaches, especially for more challenging scenarios such as **sv60** and **1a144**.

Figure 4 shows two slices from the **sv60** test set, reconstructed using all 9 methods shown in Table 3. The images reconstructed using FBP (b) and the two basic approaches (c and d) exhibit the streaking artifacts characteristic of sparse-view reconstruction, which are especially severe for FBP. While the **sart_TV** method (e) smooths these artifacts out to some extent, they are still visible, and the texture of the resulting images is blocky. The **LGD** (f) and **sart_3L** (i) images are somewhat better, but some blockiness is still apparent. Visually, the best results appear to be from the **learn_unet** (h) and **sart_unet** (j) approaches, particularly with fine details such as the tip of the shoulderblade in the top image, and the small air bubbles in the bottom image (highlighted with the green arrows in Figure 4).

The same two slices reconstructed from the **1a162** dataset are shown in Figure 5. As a result of the missing segment of views in this experiment, the edges of various structures in the bottom right part of the images are severely blurred when using FBP and the basic algorithms (b–d), and streaking is apparent in other parts of the images as well. We observe that the **sart_TV** (e) and **LGD** (f) methods offer some improvement, but are less effective at recovering details, such as the edges of the ribs and shoulder

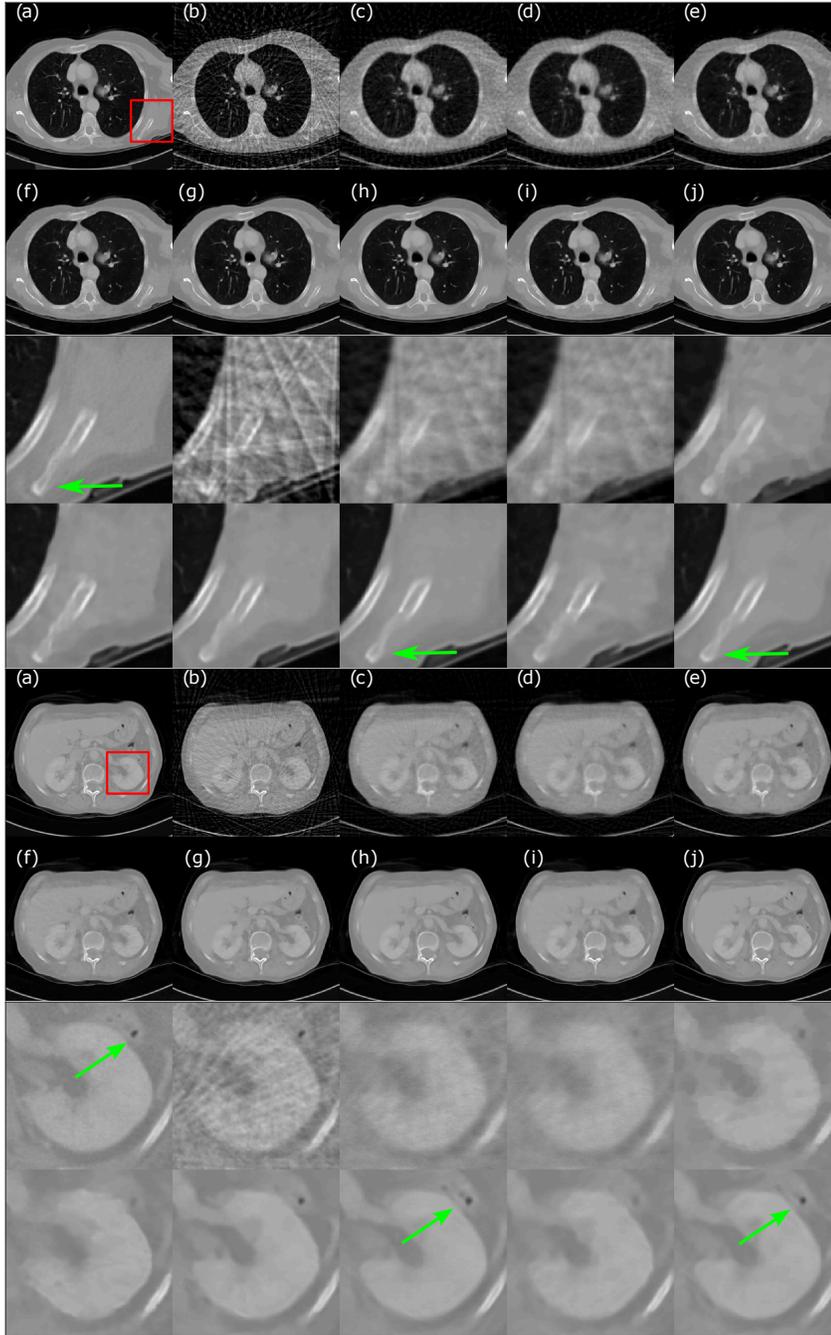


Figure 4. Reconstructed images of thoracic (top three rows) and abdominal (bottom three rows) CT slices from the test dataset for the *sv60* experiment. The labels are: (a) true image, (b) FBP (c) *learn_basic*, (d) *sart_basic*, (e) *sart_TV*, (f) LGD, (g) *learn_3L*, (h) *learn_unet*, (i) *sart_3L*, (j) *sart_unet*. Rightmost three columns show the zoomed in 128×128 pixel region indicated by the red box for each slice.

blade in the top image, and some small soft tissue contrast on the right side of the kidney in the bottom image. As in the `sv60` case, the two methods using a U-net (h and j) appear to give the best results visually.

4. Discussion

We can make several observations based on the results presented in Table 3 and the representative images in Figures 4 and 5. First, it is clear that `sart_TV` was the worst-performing of any of the regularized approaches; the PSNR and SSIM scores in Table 3 were significantly lower in almost every experiment, and the qualitative appearance of the images is worse as well. This is not surprising, as total variation minimization is entirely model-based (under the assumption that the image is roughly piecewise constant) with no ability to adapt to the data, and is known to sometimes result in blocky texture artifacts. There is some ability to improve the performance of the method by fine-tuning the parameters γ and N described in section 3.1, but this process is tedious, and the optimal choice of parameters may change depending on the imaging scenario.

While the LGD method was more competitive – especially in the three sparse-view test cases – its performance was still worse than the top-performing algorithms in all five test cases. As mentioned earlier, a major difference between the LGD approach and the other four unrolled methods that we considered is that LGD determine an overall, holistic search direction by incorporating information about the gradients of the data discrepancy functional and the regularizer. The LEARN and SART-based approaches are prescribed the descent direction with respect to the data discrepancy, and only focus on learning a regularization step. It is possible that this second problem is easier to solve than the first, which would account for the somewhat better performance of these approaches.

Additionally, while LGD picks the search direction freely, it seems plausible that providing it with information about the gradient of a specific regularizer (TV, in this case) may direct it towards providing TV-like regularization. This may be appropriate in sparse-view imaging, but less so in limited-angle imaging, where TV regularization is less appropriate. The LEARN and SART based approaches, on the other hand, are not given any information about the TV or its gradient at any point, and determine a regularization strategy based only on minimizing the error of the images in the training set. This may also account for the improved performance of those methods in the limited-angle experiments.

It is difficult to make any general claims on the merits of the U-Net versus three-layer CNN for regularization, or of the SART algorithms versus the LEARN algorithms. The most obvious result is that the `SART_3L` method gave significantly poorer results than the three related approaches; this is apparent both from the blurrier appearance of the images in Figure 4 and Figure 5, as well as the quantitative metrics computed in Table 3. Despite some additional experimentation, we were not able to determine what accounted for the poorer performance, which seems anomalous given the success of the other three algorithms.

Aside from this, the quantitative performance of the other three approaches (`sart_unet`, `learn_3L`, `learn_unet`) in most of the experiments is comparable, which does not clearly support a claim that the U-net architecture we used is preferable to the 3-layer CNN, or that the SART architecture is clearly to the LEARN architecture. The most salient result is that the `sart_unet` method did provide significantly better

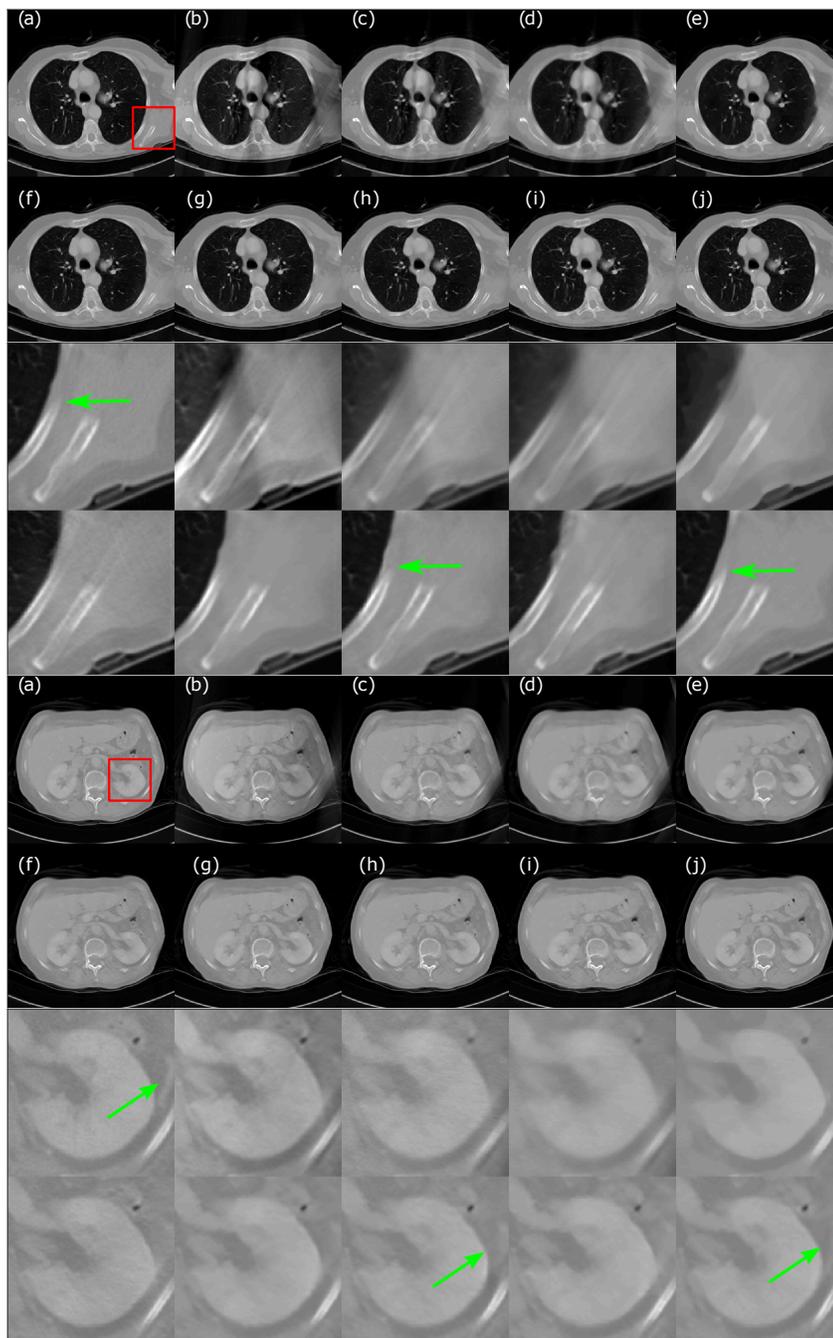


Figure 5. Reconstructed images of thoracic (top three rows) and abdominal (bottom three rows) CT slices from the test dataset for the 1a162 experiment. See Figure 4 caption for legend.

PSNR and SSIM results (Table 3) than the other approaches in the two limited-angle experiments. Given that the performance of `learn_3L` and `learn_unet` was quantitatively comparable in these experiments, however, it is difficult to attribute this to the use of the U-net, or to the SART architecture, since `sart_3L` performed poorly. That being said, qualitatively, the zoomed-in images in Figure 4 do appear to have sharper contrast for the `sart_unet` and `learn_unet` reconstructions than for the `learn_3L` reconstruction, especially in the regions highlighted with green arrows. We note that both (Chen et al. 2018) and (Adler & Öktem 2018) mentioned some loss of contrast in their sparse-view experiments, suggesting that the reason was the use of the ℓ_2 loss function when training the network; since we also used this loss function, it seems possible that the U-Net is more effective at preserving contrast than a simpler CNN.

While the U-net does provide a more sophisticated network architecture than the 3-layer CNN, its use in the unrolled algorithm is hampered by the fact that it cannot be applied in every iteration of the algorithm, due to the prohibitive number of parameters that must be trained. There is likely some tradeoff that can be explored between reducing the complexity of the network and allowing it to be applied in more iterations, but this is a computationally expensive problem to investigate. One solution may be to adapt the SART component to use ordered subsets of the data (as in (Humphries et al. 2017)), which would accelerate convergence, and therefore allow for the use of fewer SART iterations.

It is also worth noting that the total number of learned parameters varies dramatically across the unrolled algorithms that we studied, as indicated in table 2. The LGD algorithm uses on the order of 10^4 network parameters, while the 3-layer algorithms require use on the order of 10^6 , and the U-net based algorithms on the order of 10^7 . Therefore it is reasonable to ask whether the difference in performance noted above can simply be attributed to the number of parameters used in each model. In (Adler & Öktem 2017) the authors of the LGD algorithm note that they deliberately use a small number of parameters to reduce over-fitting; this refers to the phenomenon wherein a network that is over-parameterized may learn to fit the training data too well, resulting in poorer performance on the test data set. Indeed, in the original LEARN paper, the authors tested their algorithm by varying the number of network layers per iteration from two up to six (cf. Table VIII in (Chen et al. 2018)), and settled on three layers because the performance of the algorithm did not improve significantly after that point (and in fact, worsened when five or six layers were used). So while the U-net based approaches do feature the most parameters, it is not likely that this alone accounts for the observed improvement in performance of the `sart_unet` algorithm in the limited angle scenarios, and we do not see evidence of overfitting, based on the performance on the test dataset. We note as well that while the total number of parameters in the `learn_unet` and `sart_unet` algorithms is large, it is not out of line with other deep learning architectures, such as the original U-Net (Ronneberger et al. 2015), which has over 30 million parameters.

As mentioned in Section 2.4, the unrolled algorithm is not guaranteed to produce an ε -compatible solution, since it requires specifying a finite number of iterations to run. If a ε -compatible solution is desired, the output of the unrolled algorithm could be used as the initial iterate in a second run of the basic (or conventionally superiorized) algorithm, which could be run until ε -compatibility is attained. Ideally this would only require a small number of additional iterations, due to the high quality of the initial iterate, as running a large number of iterations could result

in artifacts being reintroduced into the image. Such an approach could be viewed, on the whole, as a superiorized version of the basic algorithm, in which a finite number of perturbations are generated by the U-Net, followed by perturbations of size zero (if the basic algorithm is run subsequently) or with geometrically decreasing size (if a conventionally superiorized is run subsequently). In both cases, the perturbation size as a whole would be summable, as required.

In this paper, all deep-learning methods we investigated are unrolled algorithms, which are fully supervised methods where the forward operator (A) is known during both training and testing (Ongie et al. 2020). A complementary class of methods are “plug-and-play” type approaches, where a network is trained without knowledge of the forward operator, using only ground truth images (i.e., not sinograms) to learn the distribution of natural images; see (Chang et al. 2017) and (Ongie et al. 2020), §4.2.1 and references therein. Such approaches have the advantage of not requiring paired data, and of being more generalizable; for example, a single network trained on a suitably large set of “natural” CT images could be plugged into an iterative algorithm for reconstructing sparse-view or limited-angle images at various sampling rates or geometries. An unrolled algorithm, on the other hand, is typically trained for a specific problem, e.g. sparse-view reconstruction for a given acquisition and sampling rate, and thus may not perform well if the acquisition parameters change. So, while we expect unrolled algorithms to give better results on specific sparse-view or limited-angle reconstruction tasks than a plug-and-play approach, it would be useful to perform a comparison to assess the relative performances of the two methods. A plug-and-play approach might also lend itself more naturally to an implementation within the SM, as it allows the algorithm to be run indefinitely until ε -compatibility is attained.

5. Conclusions

In this paper we present a novel unrolled iterative algorithm for reconstruction of sparse-view and limited-angle CT data. The algorithm is inspired by the LEARN algorithm (Chen et al. 2018) and the superiorization methodology, using a modified U-net to introduce perturbations in each iteration, rather than descent directions of a model-based penalty function. In numerical experiments, the approach is found to provide significantly better results than a total variation superiorization approach, as well as a competing unrolled iterative algorithm (Adler & Öktem 2017). It also outperforms the original LEARN algorithm in limited-angle imaging scenarios, while providing comparable performance in the sparse-view case.

6. Acknowledgments and Contributions

YJ, NM, PM and BT were all undergraduate students at the time of working on the project. NM and PM provided the original implementation of the LEARN method in Pyro-NN, as well as preliminary experiments. YJ first implemented the modified U-net approach, performed all numerical experiments involving the SART and LEARN-based approaches, and co-authored the manuscript. BT modified code provided by the authors of (Adler & Öktem 2017) to run on our five test datasets. TH and DS co-supervised the students, and TH performed additional numerical experimentation and co-wrote the manuscript. This work was supported by the Royalty Research Fund 68-2304 of the University of Washington. Initial code for our method was adapted from

code provided by the authors of (Zhang, Zuo, Chen, Meng & Zhang 2017). We thank the two anonymous referees for their helpful and constructive comments on the original manuscript.

References

- Adler, J. & Öktem, O. (2017). Solving ill-posed inverse problems using iterative deep neural networks, *Inverse Problems* **33**(12): 124007.
- Adler, J. & Öktem, O. (2018). Learned primal-dual reconstruction, *IEEE transactions on medical imaging* **37**(6): 1322–1332.
- Andersen, A. H. & Kak, A. C. (1984). Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm, *Ultrasonic imaging* **6**(1): 81–94.
- Butnariu, D., Davidi, R., Herman, G. T. & Kazantsev, I. G. (2007). Stable convergence behavior under summable perturbations of a class of projection methods for convex feasibility and optimization problems, *IEEE Journal of Selected Topics in Signal Processing* **1**(4): 540–547.
- Chang, J.-H. R., Li, C.-L., Poczos, B., Kumar, B. V. & Sankaranarayanan, A. C. (2017). One network to solve them all – solving linear inverse problems using deep projection models., *ICCV*, pp. 5889–5898.
- Charbonnier, P., Blanc-Féraud, L., Aubert, G. & Barlaud, M. (1997). Deterministic edge-preserving regularization in computed imaging, *IEEE Transactions on image processing* **6**(2): 298–311.
- Chen, H., Zhang, Y., Chen, Y., Zhang, J., Zhang, W., Sun, H., Lv, Y., Liao, P., Zhou, J. & Wang, G. (2018). LEARN: learned experts’ assessment-based reconstruction network for sparse-data CT, *IEEE transactions on medical imaging* **37**(6): 1333–1347.
- Chen, H., Zhang, Y., Kalra, M. K., Lin, F., Chen, Y., Liao, P., Zhou, J. & Wang, G. (2017). Low-dose CT with a residual encoder-decoder convolutional neural network, *IEEE Trans. Med. Imag.* **36**(12): 2524–2535.
- Chen, H., Zhang, Y., Zhang, W., Liao, P., Li, K., Zhou, J. & Wang, G. (2017). Low-dose ct via convolutional neural network, *Biomedical optics express* **8**(2): 679–694.
- Chen, Z., Jin, X., Li, L. & Wang, G. (2013). A limited-angle ct reconstruction method based on anisotropic tv minimization, *Physics in Medicine & Biology* **58**(7): 2119.
- Clark, K. et al. (2013). The cancer imaging archive (TCIA): maintaining and operating a public information repository, *Journal of digital imaging* **26**(6): 1045–1057.
- Dong, B., Li, J. & Shen, Z. (2013). X-ray ct image reconstruction via wavelet frame based regularization and radon domain inpainting, *Journal of Scientific Computing* **54**(2): 333–349.
- Elfving, T., Hansen, P. C. & Nikazad, T. (2012). Semiconvergence and relaxation parameters for projected SIRT algorithms, *SIAM Journal on Scientific Computing* **34**(4): A2000–A2017.
- Foucart, S. & Rauhut, H. (2013). An invitation to compressive sensing, *A mathematical introduction to compressive sensing*, Springer, pp. 1–39.
- Frikel, J. & Quinto, E. T. (2013). Characterization and reduction of artifacts in limited angle tomography, *Inverse Problems* **29**(12): 125007.
- Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*, ” O’Reilly Media, Inc.”.
- Goldgof, D., Hall, L., Hawkins, S., Schabath, M., Stringfield, O., Garcia, A., Gillies, R. et al. (2015). Data from QIN_LUNG_CT. The Cancer Imaging Archive.
URL: <http://doi.org/10.7937/K9/TCIA.2015.NPGZYBZ>
- Gordon, R., Bender, R. & Herman, G. T. (1970). Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography, *Journal of theoretical Biology* **29**(3): 471–481.
- Gupta, H., Jin, K. H., Nguyen, H. Q., McCann, M. T. & Unser, M. (2018). CNN-based projected gradient descent for consistent CT image reconstruction, *IEEE Trans. Med. Imag.* **37**(6): 1440–1453.
- Han, Y. S., Yoo, J. & Ye, J. C. (2016). Deep residual learning for compressed sensing ct reconstruction via persistent homology analysis, *arXiv preprint arXiv:1611.06391* .
- Han, Y. & Ye, J. C. (2018). Framing u-net via deep convolutional framelets: Application to sparse-view CT, *IEEE Trans. Med. Imag.* **37**(6): 1418–1429.
- He, K., Zhang, X., Ren, S. & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034.

- He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.
- Herman, G. T., Garduño, E., Davidi, R. & Censor, Y. (2012). Superiorization: An optimization heuristic for medical physics, *Med. Phys.* **39**: 5532–5546.
- Hu, J., Shen, L. & Sun, G. (2018). Squeeze-and-excitation networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132–7141.
- Humphries, T., Loreto, M., Halter, B., O’Keeffe, W. & Ramirez, L. (2020). Comparison of regularized and superiorized methods for tomographic image reconstruction, *J. Appl. Numer. Optim* **2**(1): 77–99.
- Humphries, T., Winn, J. & Faridani, A. (2017). Superiorized algorithm for reconstruction of CT images from sparse-view and limited-angle polyenergetic data, *Phys. Med. Biol.* **62**(16): 6762. **URL:** <http://stacks.iop.org/0031-9155/62/i=16/a=6762>
- Jia, X., Dong, B., Lou, Y. & Jiang, S. B. (2011). Gpu-based iterative cone-beam ct reconstruction using tight frame regularization, *Physics in Medicine & Biology* **56**(13): 3787.
- Jia, Y., McMichael, N., Mokarzel, P., Si, D. & Humphries, T. (2020). Algorithm for limited angle ct reconstruction with u-net based regularization, *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–4.
- 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.
- Kang, E., Min, J. & Ye, J. C. (2017). A deep convolutional neural network using directional wavelets for low-dose x-ray CT reconstruction, *Medical physics* **44**(10).
- Kim, K., Ye, J. C., Worstell, W., Ouyang, J., Rakvongthai, Y., El Fakhri, G. & Li, Q. (2014). Sparse-view spectral ct reconstruction using spectral patch-based low-rank penalty, *IEEE transactions on medical imaging* **34**(3): 748–760.
- LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning, *Nature* **521**(7553): 436.
- Lu, L., Shin, Y., Su, Y. & Karniadakis, G. E. (2019). Dying relu and initialization: Theory and numerical examples, *arXiv preprint arXiv:1903.06733*.
- Lu, Y., Zhao, J. & Wang, G. (2011). Few-view image reconstruction with dual dictionaries, *Physics in Medicine & Biology* **57**(1): 173.
- Lunz, S., Öktem, O. & Schönlieb, C.-B. (2018). Adversarial regularizers in inverse problems, *Advances in Neural Information Processing Systems*, pp. 8507–8516.
- Mahmood, F., Shahid, N., Skoglund, U. & Vanderghenst, P. (2018). Adaptive graph-based total variation for tomographic reconstructions, *IEEE Signal Processing Letters* **25**(5): 700–704.
- OECD (2019). *Health at a Glance 2019*. **URL:** <https://www.oecd-ilibrary.org/content/publication/4dd50c09-en>
- Ongie, G., Jalal, A., Metzler, C. A., Baraniuk, R. G., Dimakis, A. G. & Willett, R. (2020). Deep learning techniques for inverse problems in imaging, *IEEE Journal on Selected Areas in Information Theory* **1**(1): 39–56.
- Palenstijn, W. J., Batenburg, K. J. & Sijbers, J. (2011). Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs), *Journal of structural biology* **176**(2): 250–253.
- Pelt, D. M. & Batenburg, K. J. (2013). Fast tomographic reconstruction from limited data using artificial neural networks, *IEEE Transactions on Image Processing* **22**(12): 5238–5251.
- Power, S. P., Moloney, F., Twomey, M., James, K., O’Connor, O. J. & Maher, M. M. (2016). Computed tomography and patient risk: facts, perceptions and uncertainties, *World journal of radiology* **8**(12): 902.
- Ronneberger, O., Fischer, P. & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation, *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp. 234–241.
- Shen, Y., Liu, Q., Lou, S. & Hou, Y.-L. (2017). Wavelet-based total variation and nonlocal similarity model for image denoising, *IEEE Signal Processing Letters* **24**(6): 877–881.
- 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.
- Syben, C., Michen, M., Stimpel, B., Seitz, S., Ploner, S. & Maier, A. K. (2019). PYRO-NN: Python reconstruction operators in neural networks, *Medical physics* **46**(11): 5110–5115.
- Van Aarle, W., Palenstijn, W. J., Cant, J., Janssens, E., Bleichrodt, F., Dabrovolski, A., De Beenhouwer, J., Batenburg, K. J. & Sijbers, J. (2016). Fast and flexible x-ray tomography using the ASTRA toolbox, *Optics express* **24**(22): 25129–25147.
- Wang, J., Li, T. & Xing, L. (2009). Iterative image reconstruction for cbct using edge-preserving prior, *Medical physics* **36**(1): 252–260.

- 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, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity, *IEEE transactions on image processing* **13**(4): 600–612.
- Würfl, T., Ghesu, F. C., Christlein, V. & Maier, A. (2016). Deep learning computed tomography, *International conference on medical image computing and computer-assisted intervention*, Springer, pp. 432–440.
- Xu, Q., Yu, H., Mou, X., Zhang, L., Hsieh, J. & Wang, G. (2012). Low-dose x-ray ct reconstruction via dictionary learning, *IEEE transactions on medical imaging* **31**(9): 1682–1697.
- Xu, Q., Yu, H., Wang, G. & Mou, X. (2015). Dictionary learning based low-dose x-ray ct reconstruction, *Frontiers of Medical Imaging*, World Scientific, pp. 99–119.
- Ye, J. C., Han, Y. & Cha, E. (2018). Deep convolutional framelets: A general deep learning framework for inverse problems, *SIAM Journal on Imaging Sciences* **11**(2): 991–1048.
- Yu, H. & Wang, G. (2009). Compressed sensing based interior tomography, *Physics in medicine & biology* **54**(9): 2791.
- Yu, W., Wang, C. & Huang, M. (2017). Edge-preserving reconstruction from sparse projections of limited-angle computed tomography using ℓ_0 -regularized gradient prior, *Review of Scientific Instruments* **88**(4): 043703.
- Zhang, H., Li, L., Qiao, K., Wang, L., Yan, B., Li, L. & Hu, G. (2016). Image prediction for limited-angle tomography via deep learning with convolutional neural network, *arXiv preprint arXiv:1607.08707*.
- Zhang, H., Zeng, D., Lin, J., Zhang, H., Bian, Z., Huang, J., Gao, Y., Zhang, S., Zhang, H., Feng, Q. et al. (2017). Iterative reconstruction for dual energy ct with an average image-induced nonlocal means regularization, *Physics in Medicine & Biology* **62**(13): 5556.
- Zhang, H., Zeng, D., Zhang, H., Wang, J., Liang, Z. & Ma, J. (2017). Applications of nonlocal means algorithm in low-dose x-ray ct image processing and reconstruction: A review, *Medical physics* **44**(3): 1168–1185.
- Zhang, J., Hu, Y., Yang, J., Chen, Y., Coatrieux, J.-L. & Luo, L. (2017). Sparse-view x-ray ct reconstruction with gamma regularization, *Neurocomputing* **230**: 251–269.
- Zhang, K., Zuo, W., Chen, Y., Meng, D. & Zhang, L. (2017). Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising, *IEEE transactions on image processing* **26**(7): 3142–3155.
- Zheng, X., Ravishankar, S., Long, Y. & Fessler, J. A. (2018). PWLS-ULTRA: An efficient clustering and learning-based approach for low-dose 3d CT image reconstruction, *IEEE Trans. Med. Imag.* **37**(6): 1498–1510.