Supplementary Material for Lose The Views: Limited Angle CT Reconstruction via Implicit Sinogram Completion

Rushil Anirudh, Hyojin Kim, Jayaraman J. Thiagarajan, K. Aditya Mohan,
Kyle Champley, Timo Bremer

Lawrence Livermore National Laboratory.

1. Generalizability of CT-Net to new domains

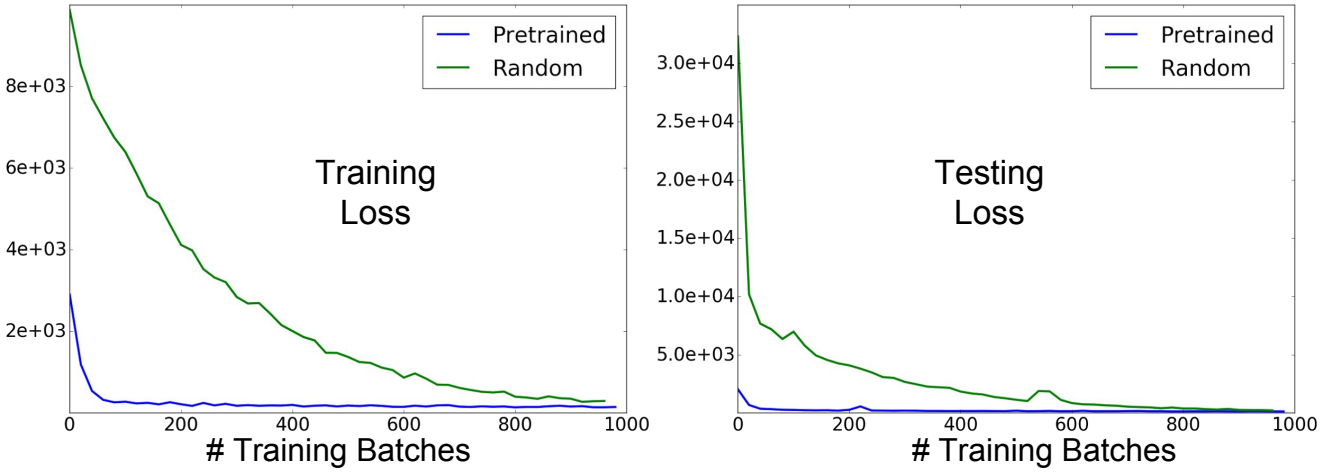
A fundamental aspect of our system is its ability to invert the x-ray transform, which maps 3D objects into its corresponding sinogram representation. Conceptually, the inverse transformation should be applicable to any scenario regardless of the objects being scanned. In this experiment, we validate the generalizability of our system by using the network trained on the transportation luggage dataset[1], to invert the fashion-MNIST [5]. This dataset consists of $60K$ training images of size 28×28 belonging to one of ten classes, consisting of everyday fashion like jeans, dresses, shoes, sandals etc. We resize them to 128×128 using bilinear interpolation. Next, we generate a set of sinograms corresponding to the training data in the fashion-mnist dataset, using the Radon transform which is equivalent to the x-ray transform for 2D images. In particular, we use the implementation of the Radon transform from the scikit-image library [3]. We fine-tune the pre-trained network from the transportation dataset to the target fashion dataset. For training in the partial view scenario, we use only use the top half of the sinogram, corresponding to $0 - 90^{\circ}$. Figure 1aSubfigure 1asubfigure.1.1 shows the training and test losses obtained from training CT-Net on the fashion dataset. As it can be clearly observed, warm-starting the network with weights from the transportation dataset leads to significantly faster convergence within a very few mini-batches, illustrating the effectiveness of the learned inverse x-ray transformation. Furthermore, we illustrate the recovered images in Figure 1bSubfigure 1bsubfigure.1.2.

2. Description of CT-Net

The architectures of the different networks in CT-Net are illustrated in Figure 2cSubfigure 2csubfigure.2.3, and tables 2aSubfigure 2asubfigure.2.1, 2bSubfigure 2bsubfigure.2.2 respectively. We use 5 different filter sizes, with 256 hidden dimensions each, for the 1D CNN, which embeds the sinogram into a $5 \times 256 = 1280$ dimensional latent representation. Next, we decode the CT image using the decoder described in Figure 2cSubfigure 2csubfigure.2.3. Additionally, when using an adversarial loss, we use the discriminator shown in table 2bSubfigure 2bsubfigure.2.2.

3. More Reconstruction Results

We show failure cases and successful reconstruction examples from unseen test data in Figures 3**Failure cases:** Images with high frequency content are much harder to recover, and the proposed method does not provide a significant improvement over existing approaches like FBP or WLS. In addition, if the objects are not visible in the views available, they will naturally appear invisible in the current setup, as seen for the object in the middle row here.figure.caption.3 and 4**Successful Cases:** Reconstruction results, with PSNR listed on top of each reconstructed image slice. Apart from the baselines shown in the main paper, we also include CT-Net -mse+FBP,CT-Net -mse+WLS,CT-Net -adv+FBP here for comparison.figure.caption.4 respectively. In Figure 4**Successful Cases:** Reconstruction results, with PSNR listed on top of each reconstructed image slice. Apart from the baselines shown in the main paper, we also include CT-Net -mse+FBP,CT-Net -mse+WLS,CT-Net -adv+FBP here for comparison.figure.caption.4, along with the baselines shown in the main paper, we also include CT-Net -mse+FBP,CT-Net -mse+WLS,CT-Net -adv+FBP here for comparison. In general, we observe that sinogram completion works much better than any other approach. Further, sinogram completion with WLS works better than with FBP. Finally, we find CT-Net-mse and CT-Net-adv perform very similar for completion and do not differ greatly in the final reconstruction. This is because the adversarial loss is measured in the image space. An end-to-end system with an adversarial loss in sinogram completed space is expected to work better, but that is left as future work. From Figure



(a) Comparing training and testing losses for the network warm-started with weights vs randomly initialized. We see that the warm-started model converges within just 50 batches of training, with batch size=100.

(b) Reconstructed samples from the test set

Figure 1: **Transferring knowledge to new domains:** We "warm start" CT-Net, with weights trained on the luggage dataset [1], and fine-tune for the fashion-MNIST dataset. It can be seen that within just a few batches of training, our network learns to recover the images with just half the views.

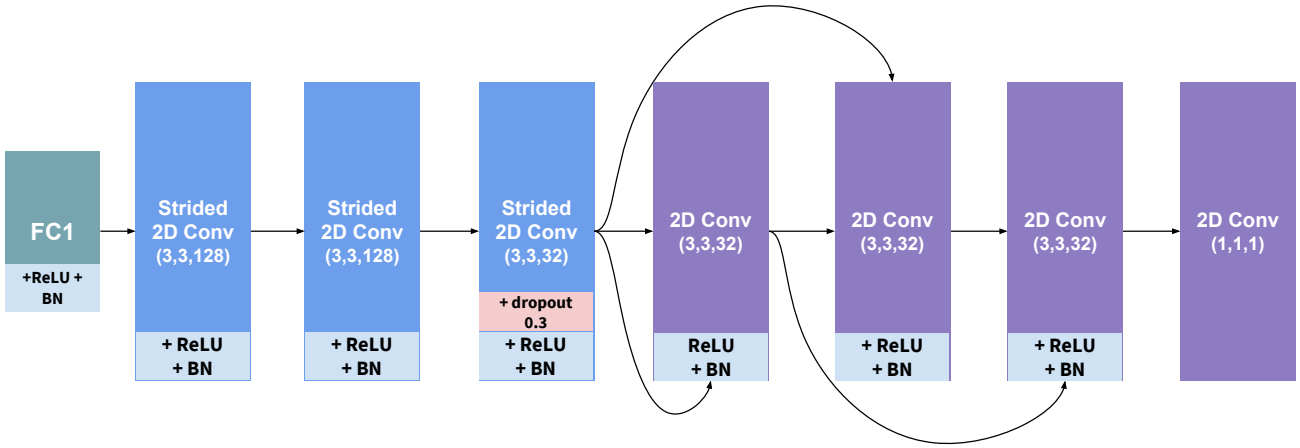
3**Failure cases:** Images with high frequency content are much harder to recover, and the proposed method does not provide a significant improvement over existing approaches like FBP or WLS. In addition, if the objects are not visible in the views available, they will naturally appear invisible in the current setup, as seen for the object in the middle row here.figure.caption.3, it can be observed that test images with lot of high-frequency content are not well recovered with CT-Net. Further, when the objects are not visible in any of the views available, the network has not informa-

Input	Filters
Sinogram 360×720	$(f, 128, 1, 256)$, $f = [1, 2, 3, 4, 5]$

(a) 1D CNN

Input	Filters
I:128×128	Conv 2D $(7, 7, 1, 64) \rightarrow$ BN $\rightarrow$ LReLU
h1:$64 \times 64 \times 64$	Conv 2D $(7, 7, 64, 32) \rightarrow$ BN $\rightarrow$ LReLU
h2:$32, 768$	FC $32, 768 \times 16 \rightarrow$ BN $\rightarrow$ LReLU
h3: 16	FC $16 \times 1 \rightarrow$ BN $\rightarrow$ LReLU

(b) Discriminator



(c) Generator

Figure 2: Architecture details of the different networks in CT-Net.

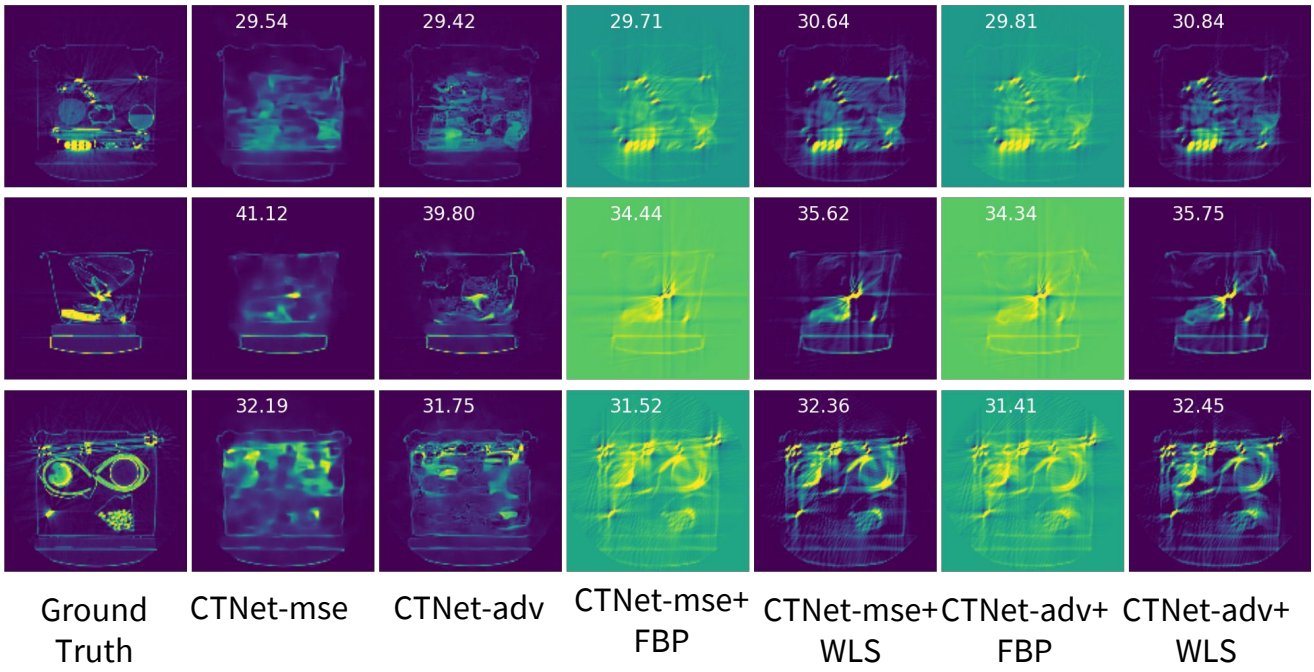


Figure 3: **Failure cases:** Images with high frequency content are much harder to recover, and the proposed method does not provide a significant improvement over existing approaches like FBP or WLS. In addition, if the objects are not visible in the views available, they will naturally appear invisible in the current setup, as seen for the object in the middle row here.

tion to recover that object.

4. 3D Segmentation Results

In this experiment, we used a popular region growing segmentation similar to the method used in [2]. It is a simplified version of the method in [4], with a randomly chosen starting position and a fixed kernel size. The purpose of this experiment is to understand how reconstruction quality affects object segmentation. The luggage dataset contains segmentation labels of objects of interest, and the evaluation focused on how well each segmentation extracts the labeled object. We reconstructed all slices of each bag through the proposed method and combined them into a single bag in 3D. Then we run the region growing in 3D at multiple, hand-tuned parameter settings (intensity threshold ranging from 0.005 to 0.02), and reported the results from the best-performing setting. This is done as some reconstruction results are poor and very sensitive to the threshold (especially when using the partial-view FBP and iterative method). Examples for the 3D segmentation on reconstructions obtained using CT-Net-adv+WLS, using region growing method are shown in figures 5aSubfigure 5asubfigure.5.1,5bSubfigure 5bsubfigure.5.2, and 5Test bag 2 with 274 reconstructed slices.figure.caption.6. It is easy to see that the proposed reconstruction segments the objects of interest very similar to the ground truth images, than compared to using WLS for reconstruction.

References

[1] D. Center of Excellence at Northeastern University. ALERT TO4 datasets for automated threat recognition. available at http://www.northeastern.edu/.

[2] H. Kim, J. J. Thiagarajan, and P.-T. Bremer. A randomized ensemble approach to industrial ct segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1707–1715, 2015.

[3] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors. scikit-image: image processing in Python. *PeerJ*, 2:e453, 6 2014.

[4] D. F. Wiley, D. Ghosh, and C. Woodhouse. Automatic segmentation of ct scans of checked baggage. In *Proceedings of the 2nd International Meeting on Image Formation in X-ray CT*, pages 310–313, 2012.

[5] H. Xiao, K. Rasul, and R. Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms, 2017.

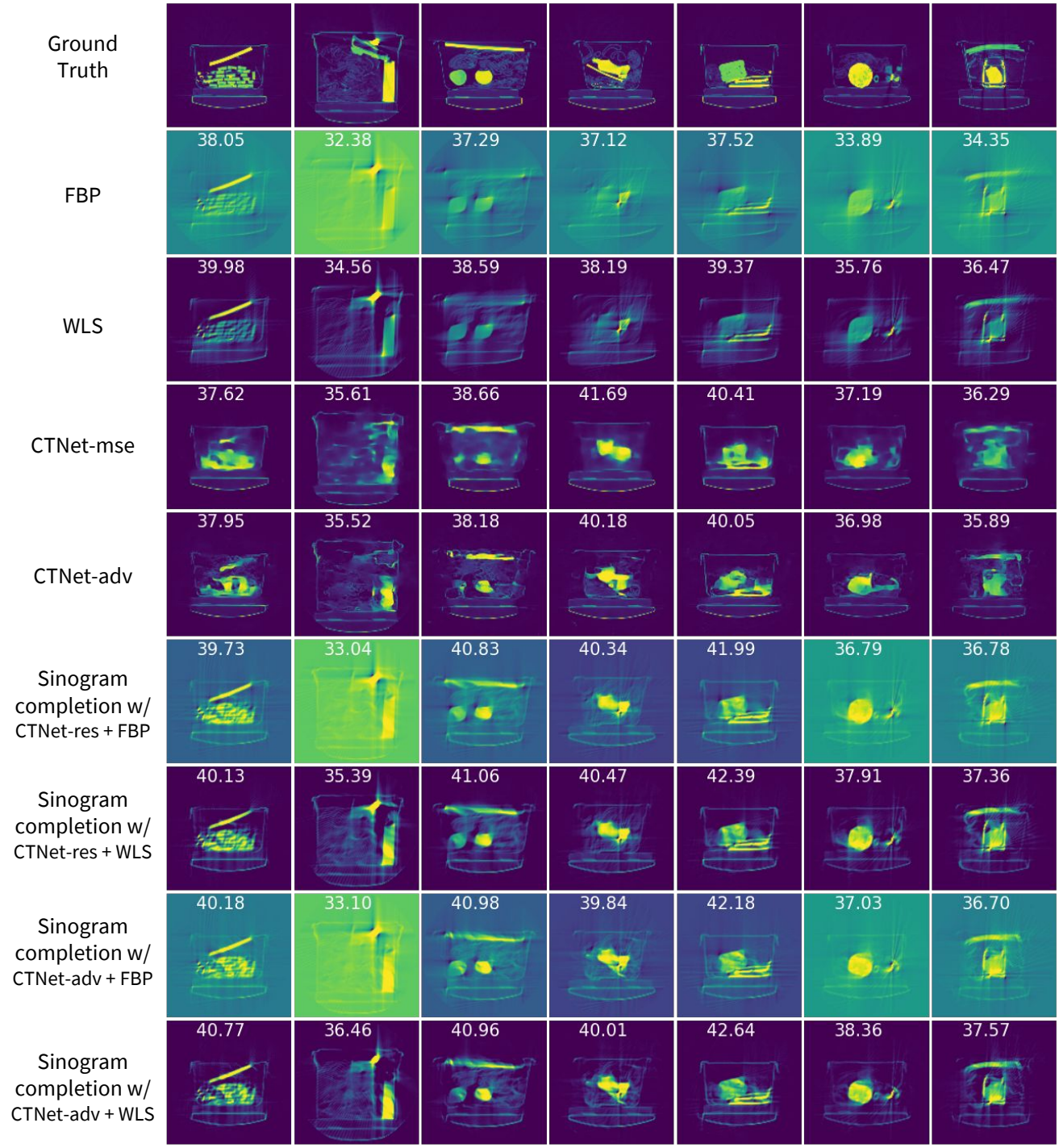
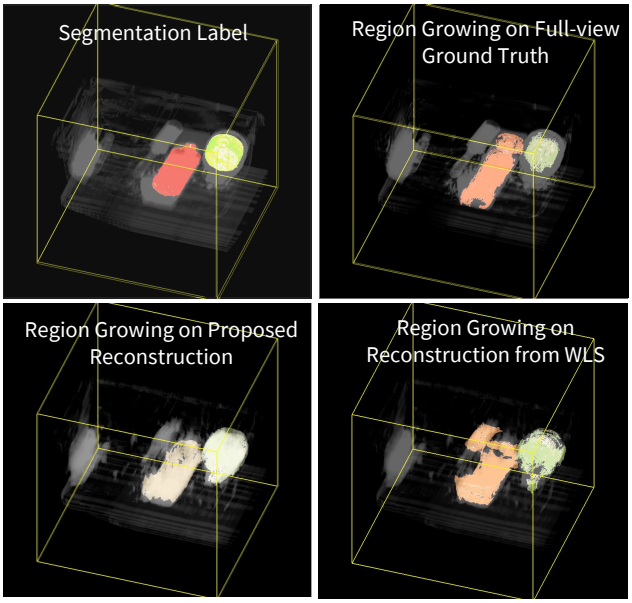
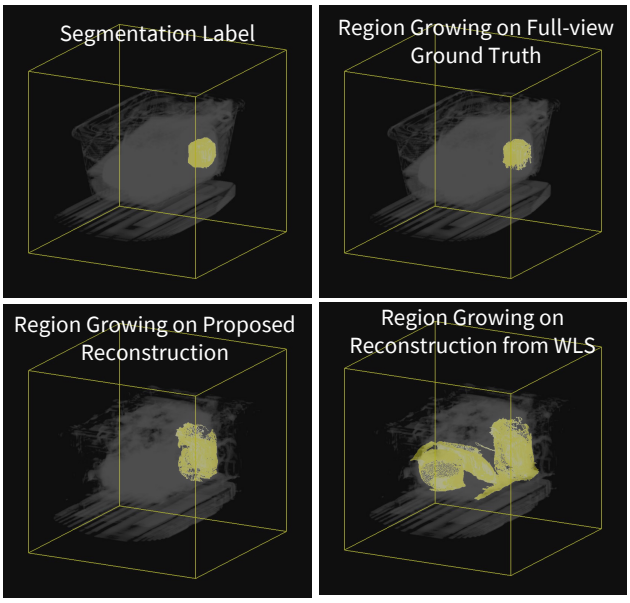


Figure 4: **Successful Cases:** Reconstruction results, with PSNR listed on top of each reconstructed image slice. Apart from the baselines shown in the main paper, we also include CT-Net -mse+FBP,CT-Net -mse+WLS,CT-Net -adv+FBP here for comparison.



(a) Test Bag 3 with 270 reconstructed slices.



(b) Test Bag 4 with 250 reconstructed slices.

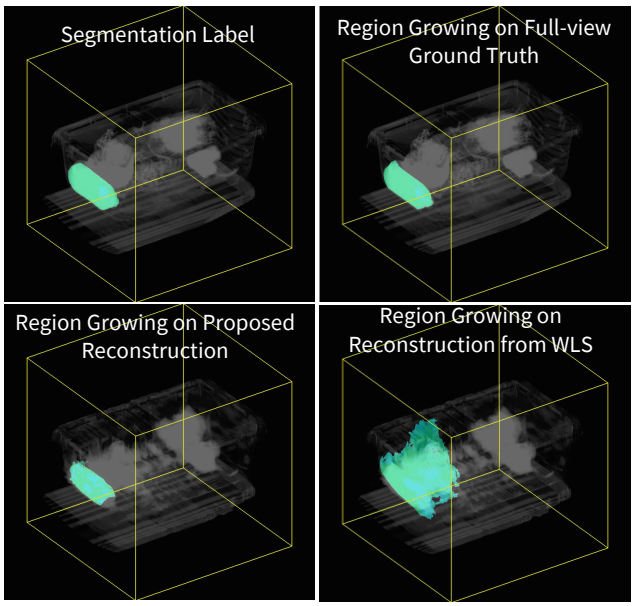


Figure 5: Test bag 2 with 274 reconstructed slices.