



OPEN

## A residual dense network assisted sparse view reconstruction for breast computed tomography

Zhiyang Fu<sup>1,2</sup>, Hsin Wu Tseng<sup>1</sup>, Srinivasan Vedantham<sup>1,3</sup>, Andrew Karellas<sup>1</sup> & Ali Bilgin<sup>1,2,3</sup>✉

To develop and investigate a deep learning approach that uses sparse-view acquisition in dedicated breast computed tomography for radiation dose reduction, we propose a framework that combines 3D sparse-view cone-beam acquisition with a multi-slice residual dense network (MS-RDN) reconstruction. Projection datasets (300 views, full-scan) from 34 women were reconstructed using the FDK algorithm and served as reference. Sparse-view (100 views, full-scan) projection data were reconstructed using the FDK algorithm. The proposed MS-RDN uses the sparse-view and reference FDK reconstructions as input and label, respectively. Our MS-RDN evaluated with respect to fully sampled FDK reference yields superior performance, quantitatively and visually, compared to conventional compressed sensing methods and state-of-the-art deep learning based methods. The proposed deep learning driven framework can potentially enable low dose breast CT imaging.

Dedicated breast computed tomography (BCT) is an emerging, fully 3D, high-resolution (100–300  $\mu\text{m}$  nearly isotropic voxels) imaging modality that does not employ physical compression of the breast. Compared to digital breast tomosynthesis<sup>1</sup>, BCT almost eliminates tissue superposition and does not suffer from limited-angle acquisition associated artifacts<sup>2</sup> seen in digital breast tomosynthesis. A multi-reader, multi-case receiver operating characteristic (ROC) study employing 18 readers and 235 cases showed improved sensitivity of non-contrast diagnostic BCT over mammography-based diagnostic work-up<sup>3</sup>, leading to its regulatory approval for non-contrast diagnostic use. Non-contrast BCT can have a far greater role if its suitability for breast cancer screening is demonstrated. The radiation dose (mean glandular dose, MGD) from non-contrast diagnostic BCT, while similar to the MGD from mammography-based diagnostic workup, was approximately twice that of 2-view (standard) screening DM<sup>4</sup>. At radiation dose similar to mammography, a prior study using an early prototype showed improved visualization of masses and reduced visualization of microcalcifications with BCT compared to mammography<sup>5</sup>. Hence, the long-term goal is to reduce the radiation dose to be comparable to mammography screening, without loss of detection performance.

Radiation dose reduction in BCT to levels suitable for breast cancer screening can be achieved through improved hardware, acquisition strategies and advanced image reconstruction inclusive of post-processing techniques. In terms of hardware, photon-counting detectors<sup>6,7</sup>, low-noise, high-resolution, complementary metal oxide (CMOS) detectors<sup>8,9</sup> and beam-shaping X-ray filters<sup>10,11</sup> are being investigated. Acquisition strategies being investigated include helical scan<sup>6</sup>, laterally-shifted detector geometry<sup>12,13</sup> short-scan<sup>14</sup>, and sparse-view acquisition<sup>15</sup>. Also, theoretical and empirical optimization of x-ray beam quality for acquiring projection data have been reported<sup>16–19</sup>.

In this study, we describe the potential of advanced image reconstruction employing deep learning techniques that can be used with existing BCT technology. This can lead to lower radiation dose and expedite its translation for breast cancer screening. This study is complementary to ongoing hardware-oriented research. Although, statistical iterative reconstruction<sup>20–22</sup> and denoising techniques<sup>23</sup> have been investigated for BCT, all BCT systems currently use Feldkamp–Davis–Kress (FDK) reconstruction<sup>24</sup>. Deep learning based image reconstruction has not been investigated in the context of BCT or for cone-beam CT; however, it has been explored for conventional multi-detector CT<sup>25–28</sup>. Jin et al.<sup>25</sup> utilized the U-Net with residual learning and demonstrated the feasibility on parallel beam X-ray CT. A similar approach was independently proposed by Chen et al.<sup>26</sup>. The proposed residual encoder-decoder convolutional neural network (RED-CNN)<sup>26</sup> was shown to be quantitatively outperforming the earlier version<sup>29</sup> and the wavelet-domain CNN<sup>30</sup>.

<sup>1</sup>Department of Medical Imaging, University of Arizona, Tucson, AZ, USA. <sup>2</sup>Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, USA. <sup>3</sup>Department of Biomedical Engineering, University of Arizona, Tucson, AZ, USA. ✉email: bilgin@email.arizona.edu

	RED-CNN	P-value	MS-RDN	P-value
<b>(a) Impact of TOI selection</b>				
NMSE	+ 0.090 dB	< 0.0001	+ 0.253 dB	< 0.0001
Bias	$- 0.998 \times 10^{-4} \text{ cm}^{-1}$	< 0.0001	$- 2.099 \times 10^{-4} \text{ cm}^{-1}$	< 0.0001
PSNR	+ 0.090 dB	< 0.0001	+ 0.253 dB	< 0.0001
SSIM	+ 0.0009	< 0.0001	+ 0.0011	< 0.0001
<b>(b) Impact of multi-slice training</b>				
NMSE	+ 0.035 dB	< 0.0001	Not significant	0.211
Bias	$- 0.411 \times 10^{-4} \text{ cm}^{-1}$	< 0.0001	Not significant	0.234
PSNR	+ 0.035 dB	< 0.0001	Not significant	0.211
SSIM	+ 0.0004	< 0.0001	+ 0.0005	< 0.0001

**Table 1.** Statistic analysis of the impact of TOI selection and multi-slice training for RED-CNN and MS-RDN architectures. The evaluation was performed on the entire testing breast dataset using NMSE, Bias, PSNR, and SSIM metrics. (a) shows the performance improvement by including TOI selection on single-slice ( $Z = 1$ ) RED-CNN and MS-RDN. (b) shows the performance difference between multi-slice ( $Z = 5$ ) and single-slice training for RED-CNN and MS-RDN, respectively. Please note that the values corresponding to NMSE and PSNR are identical since these quantities are related as shown in Eqs. (6) and (7).

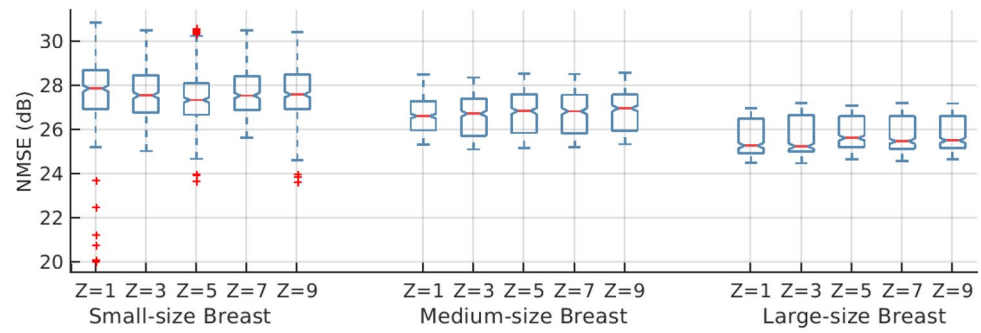
Recently, advanced network architectures using residual blocks<sup>31</sup> or dense blocks<sup>32</sup> have shown improved performance compared to standard convolutional neural networks in computer vision applications<sup>33,34</sup>. In this work, we adopt a derived version of the residual dense network<sup>33</sup> and investigate its potential for low-dose cone-beam BCT image reconstruction.

## Results

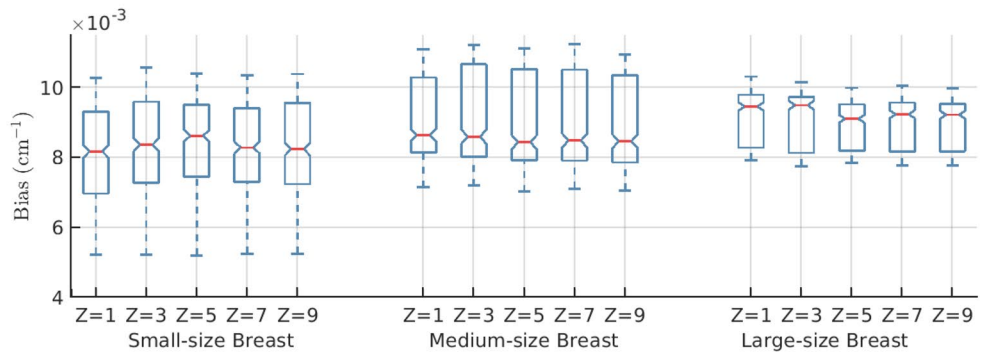
**Breast CT datasets.** This retrospective study was conducted in accordance with relevant guidelines and institutional review-board (IRB) approved protocol (University of Arizona Human Subjects Protection Program, Protocol #1903470973). The study used de-identified projection datasets from 34 women assigned Breast Imaging-Reporting and Data System (BIRADS)<sup>35</sup> diagnostic assessment category 4 or 5, who had previously participated in an IRB approved, Health Insurance Portability and Accountability Act (HIPAA)-compliant research study (ClinicalTrials.gov Identifier: NCT01090687). The study was conducted with informed consent from participants involved. This dataset was used in several prior studies<sup>4,36–41</sup>. All subjects underwent non-contrast dedicated breast CT exam of the ipsilateral breast using a clinical prototype flat-panel cone-beam breast CT system (Koning Corp., West Henrietta, NY). The scan parameters were: 49 kVp, 1.4 mm of Al 1st HVL, 8 ms pulse-width, 300 projection views, 360 degree full-scan acquisition, 12.6 mGy MGD, and 10 s scan time. The 300 view projection datasets were reconstructed using the FDK algorithm with 0.273 mm isotropic voxel pitch and matrix size  $1024 \times 1024$  in the transverse (coronal) plane. Sparse-view (100 views, full scan; 4.2 mGy MGD) projection data were retrospectively undersampled from the 300 view datasets and reconstructed with the FDK algorithm at the same voxel pitch. The longitudinal direction represents the slices. The 34 breast CT datasets were randomly split as follows: 20 for training (total of 8346 2D slices), 5 for validation (total of 1920 slices) and the remaining for testing (total of 4056 slices). The 9 test subjects were evenly divided into groups corresponding to small, medium, and large sized breasts, based on the number of slices in each case. The number of slices for the 9 test subjects were: 250, 315, 390, 426, 450, 462, 523, 600, and 640. The training dataset had diverse lesions (4 soft tissue lesions, 14 calcified lesions, and 2 soft tissue lesions with microcalcifications), BIRADS breast density categories (1, 6, 9, and 3 of categories a through d, respectively), and pathology (5 malignant, 2 hyperplasia, and the remaining benign).

**Impact of tissue of interest (TOI) selection.** TOI selection was evaluated for the proposed multi-slice residual dense network (MS-RDN) and RED-CNN<sup>26</sup>. Test subject datasets were reconstructed by the single-slice networks with and without TOI selection. FDK reconstructions on the 300-view data (denoted as FDK300) were used as references across all the experiments. The performance was quantitatively evaluated with Normalized Mean Square Error (NMSE), bias, Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Metric (SSIM<sup>42</sup>). All four metrics significantly differed across all reconstructions (Wilks Lambda,  $P < 0.0001$ ). Table 1 panel (a) showed that TOI selection significantly improved all metrics for both RED-CNN and MS-RDN.

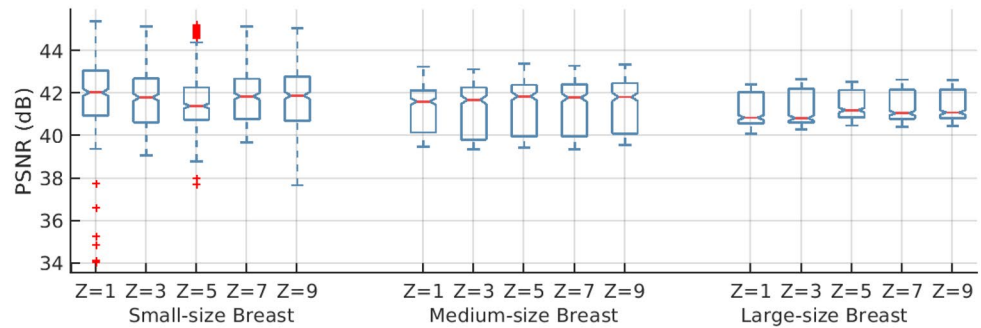
**Impact of multi-slice training.** Over the entire test dataset, MS-RDN with  $Z = 1$  did not differ significantly from MS-RDN with  $Z = 5$  in terms of NMSE ( $P = 0.211$ ), bias ( $P = 0.234$ ), and PSNR ( $P = 0.211$ ) as shown in the panel (b) of Table 1. However, there was a significant improvement with MS-RDN5 compared to MS-RDNZ1 in SSIM ( $P < 0.0001$ ; mean improvement: 0.0005). For RED-CNN, multi-slice training significantly improved all metrics compared to single slice training. The boxplots in Fig. 1 show independent evaluations for small-size, medium-size, and large-size breasts. Figure 1a shows relatively consistent NMSE performance from small-size breasts to large-size breasts. Similar observation of robust performance can be made for the bias, PSNR, and SSIM boxplots shown in Fig. 1b–d, respectively. The quantitative performances of MS-RDN and



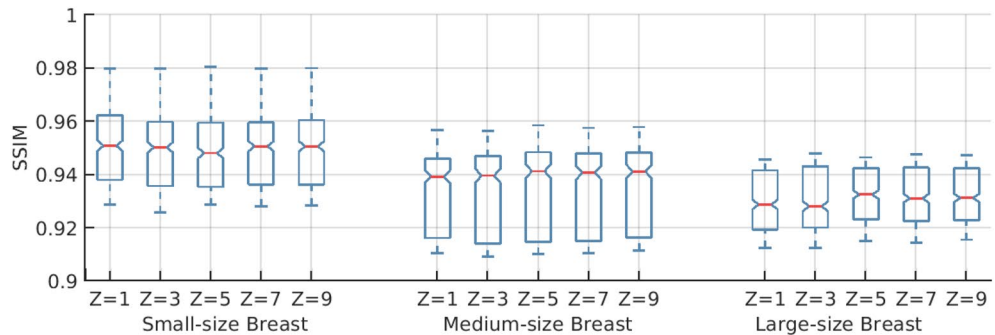
(a)



(b)

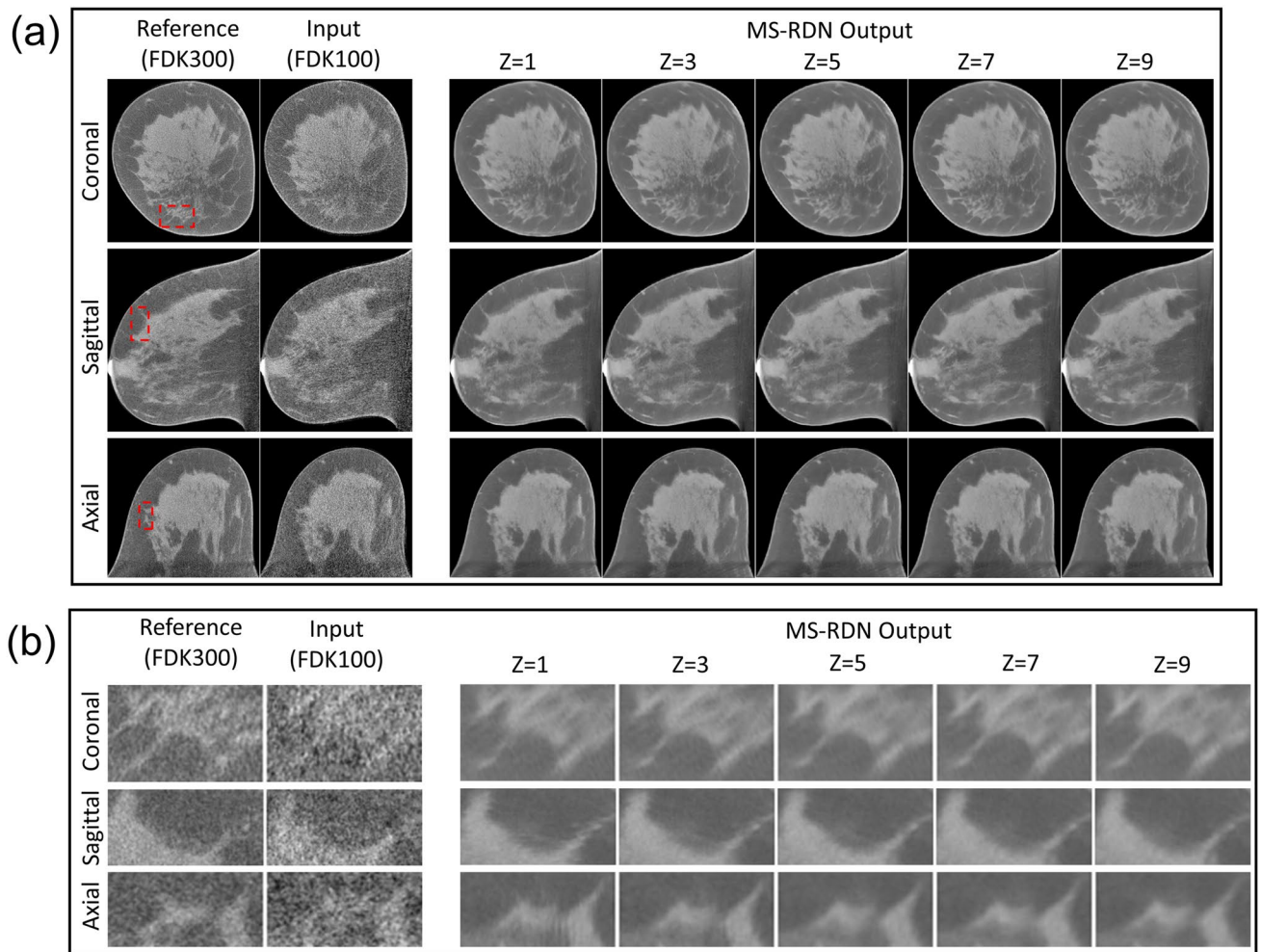


(c)



(d)

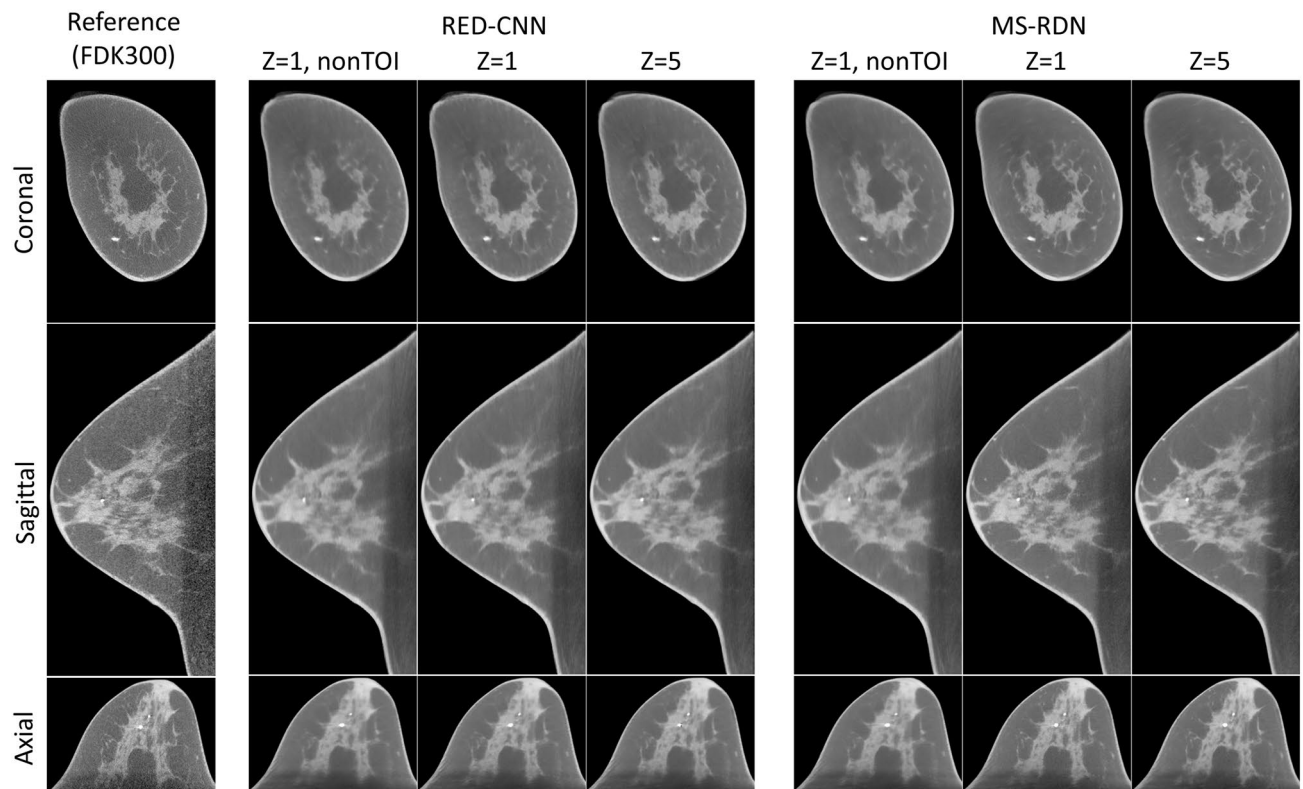
**Figure 1.** MS-RDN reconstructions with different number of adjacent slices ( $Z = 1, 3, 5, 7, 9$ ) are evaluated with (a) NMSE, (b) bias, (c) PSNR, and (d) SSIM for a range of breast sizes. Fully sampled FDK reconstructions are used as reference. These metrics computed along the longitudinal direction are presented using box plots. On each box, the central mark is the median, the top and bottom edges are the 25th and 75th percentiles, respectively. Outliers are denoted as red plus signs.



**Figure 2.** (a) A comparison of breast images reconstructed by MS-RDNs with different slice depth ( $Z = 1, 3, 5, 7, 9$ ) on retrospectively undersampled 100-view cone-beam data. The network inputs are obtained using FDK on the 100-view breast data, denoted as FDK100, and the references are obtained using FDK on the 300-view breast data, denoted as FDK300. The bounding boxes on the reference images indicate the ROIs enlarged in (b). Note that the sagittal and axial ROIs were rotated 90 degrees clockwise for presentation. The display window is  $[0.15, 0.35] \text{ cm}^{-1}$ .

RED-CNN with multi-slice training were breast size dependent with smaller improvements, or degradation, for smaller breasts than for medium and large breasts. For the medium-size and large-size breasts, MS-RDN with  $Z = 5$  (MS-RDNZ5) achieved the best performance for all metrics. For small-size breasts, the single-slice MS-RDN (MS-RDNZ1) appeared to perform better than multi-slice networks. The lower cone-angle of small-size breasts could reduce longitudinal correlation for the multi-slice networks to exploit, and the under-representation of small-size breasts (approximately 16% of slices) in the training dataset may be contributing factors to the above observation. Studies into these aspects will be pursued in future with the availability of larger datasets. Figure 2a shows the (medium-size) breast images reconstructed by FDK and MS-RDNs with varying slice depths on the retrospectively undersampled 100-view data together with the reference image obtained using FDK on the 300-view data. Figure 2b shows the zoomed-in views corresponding to the red bounding boxes indicated in Fig. 2a. Note that the sagittal and axial ROIs were rotated 90 degrees clockwise for display. Compared to the reference images, all MS-RDN outputs appear less noisy. It is worth noting that the Venetian blind artifacts appear in the longitudinal reconstructions of MS-RDN with single slice training. As the slice depth increases, these artifacts are suppressed but the glandular tissues become blurred gradually. Importantly, multi-slice training eliminates longitudinal artifacts and enhances the reconstructions as well. On the other hand, MS-RDN with large slice depths increases computational complexity in training and testing without gaining substantial performance. Hence, we opted to train MS-RDN with 5 adjacent slices in the following experiments as a balance between performance and complexity.

**Comparison with RED-CNN.** Our MS-RDN was compared with RED-CNN in three sets of network configurations: single slice training without TOI selection ( $Z = 1$ , nonTOI), single slice training ( $Z = 1$ ), and multi-slice training ( $Z = 5$ ). Figure 3 shows the breast images (small-size) reconstructed by RED-CNN and MS-RDN



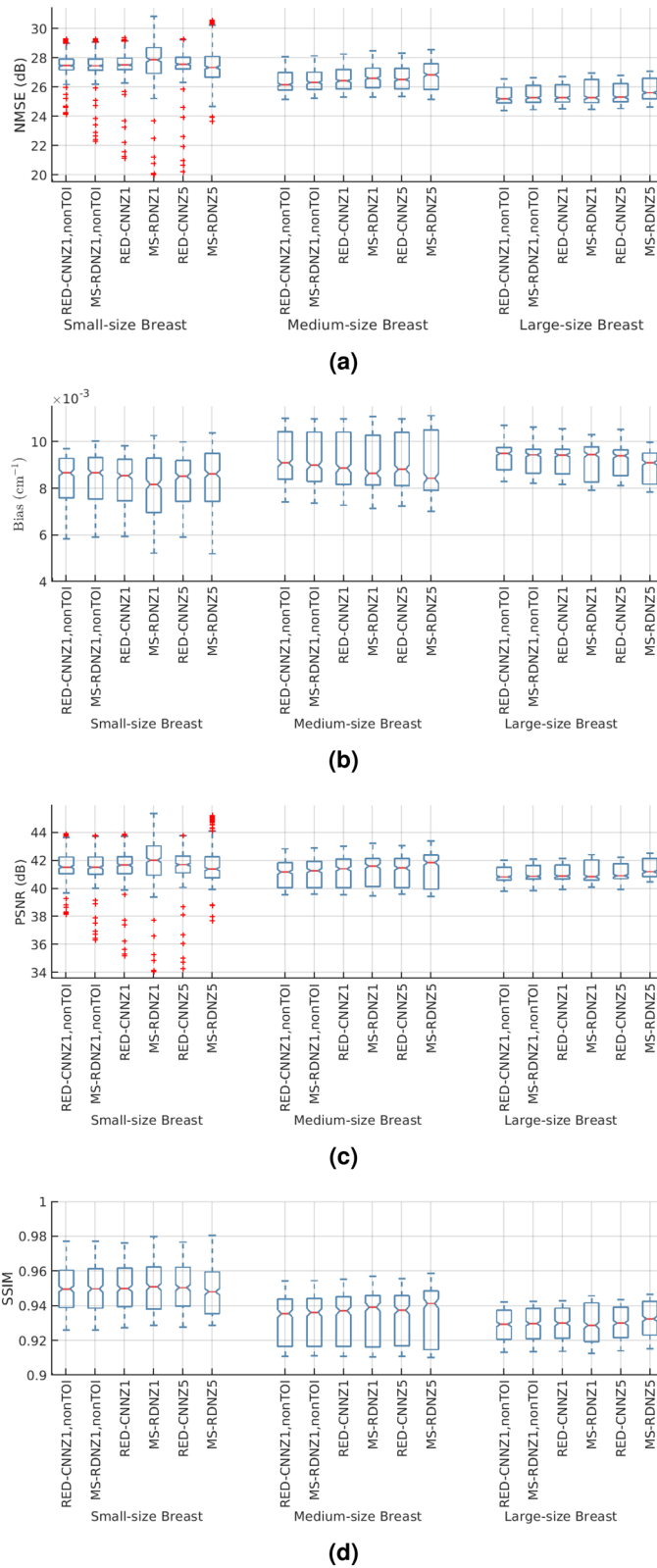
**Figure 3.** Comparisons to the residual encoder–decoder convolutional neural network (RED-CNN). The proposed MS-RDN was compared with RED-CNN in three sets of configurations: single slice training without TOI oriented patch extraction ( $Z = 1$ , nonTOI), single slice training ( $Z = 1$ ), and multi-slice training ( $Z = 5$ ). Breast images of the test subject were reconstructed by these RED-CNNs and MS-RDNs using the retrospectively undersampled 100-view data. The reference images were obtained using FDK on the 300-view data. The display window is  $[0.15, 0.35] \text{ cm}^{-1}$ .

on the retrospectively undersampled 100-view data together with the reference image obtained using FDK on the 300-view data. Overall, MS-RDNs preserved high-frequency features such as edges and textures better than their RED-CNN counterparts. In addition, the aforementioned Venetian blind artifacts are also presented in the non-transverse images obtained using RED-CNN with single slice training. Figure 4 shows the boxplots of (a) NMSE, (b) bias, (c) PSNR, and (d) SSIM for the RED-CNN and MS-RDN reconstructions of various-size breasts. For small-size breasts, MS-RDN with single slice training ( $Z = 1$ ) attained the best NMSE and bias performance. For medium-size and large-size breasts, it can also be observed that TOI selection and multi-slice training improve performance of MS-RDN independently. Table 2 shows that MS-RDN outperforms RED-CNN significantly in all configurations.

#### Comparison with the fast, iterative, TV-regularized, statistical reconstruction technique (FIRST<sup>22</sup>).

Figure 5 illustrates the breast (large-size) reference images reconstructed by FDK and FIRST using the 300-view data as well as the reconstructions obtained using FIRST and MS-RDNZ5 on the 100-view data. Compared to the 300-view FDK reconstructions (FDK300), the 300-view FIRST reconstructions (FIRST300) suppress the noise and preserve breast tissue structures in fine scale. However, the FIRST reconstructions with the 100-view data (FIRST100) exhibits blurred structures/textures and increased streak artifacts. In contrast, MS-RDNZ5 with 100-view data is able to remove the streaks as well as suppress the noise. In Table 3, the performance of FIRST and MS-RDNZ5 are evaluated with NMSE, bias, PSNR, and SSIM using 300-view FDK and 300-view FIRST reconstructions as references, respectively. For all these metrics, MS-RDNZ5 outperforms FIRST considerably. It is noteworthy that these metrics are improved by a large margin (roughly 5–8 dB NMSE increase,  $4\text{--}6 \times 10^{-3} \text{ cm}^{-1}$  bias decrease, 5–8 dB PSNR increase, and 0.04–0.07 SSIM increase) when FIRST300 images rather than FDK300 reconstructions are used as references.

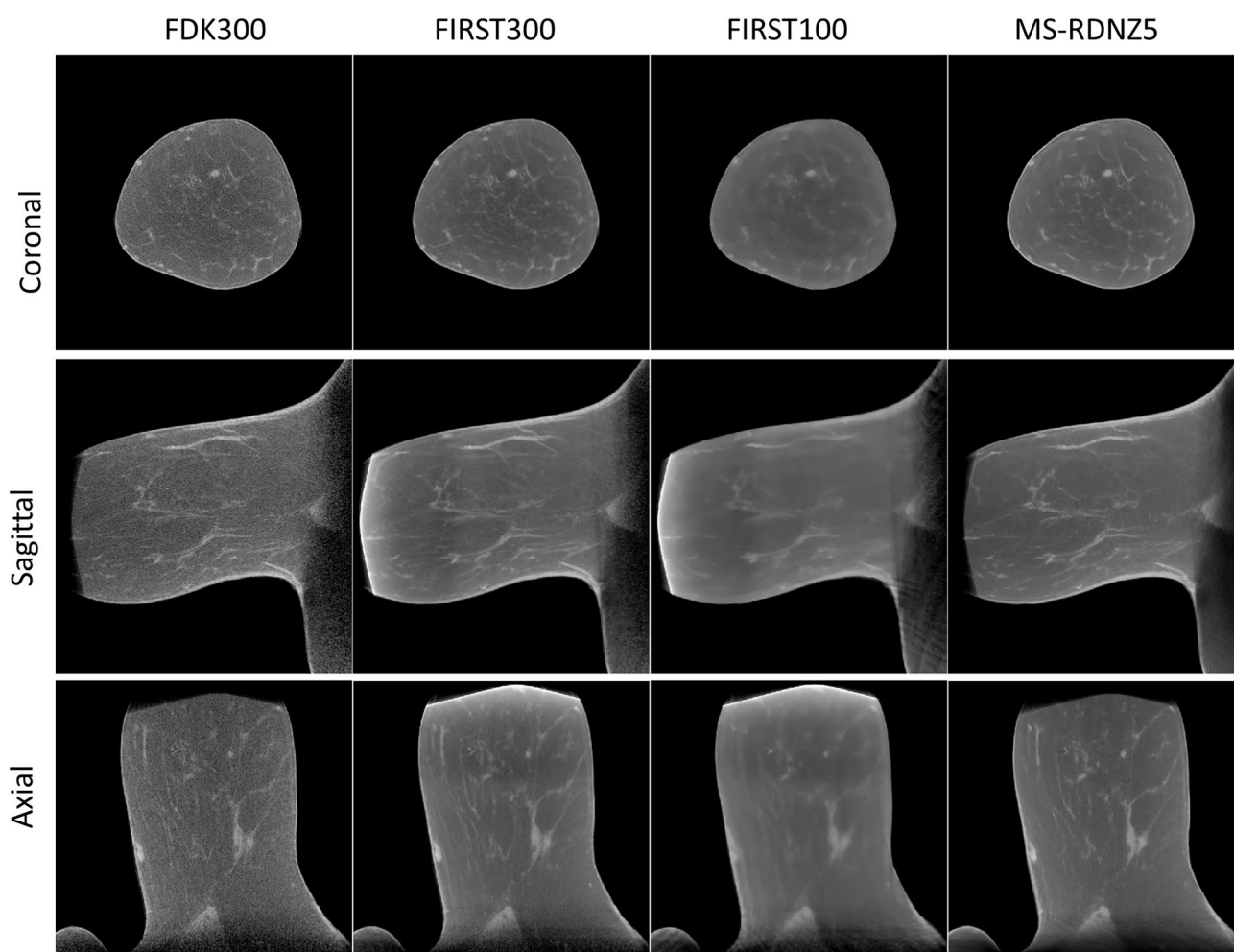
**Outlier inspection.** The slice with the worst NMSE for MS-RDNZ5 was identified in Fig. 4. This slice was from a small heterogeneously dense breast (BI-RADS density category c). Figure 6 shows the reconstructions obtained using the investigated methods for this slice. A hyper-intense signal, corresponding to a calcification, is located near the center of the breast, which was biopsied subsequent to breast CT. Pathology indicated a benign finding—fibrosis with calcification. It is interesting to note that this calcification is not reconstructed well by any of the deep-learning techniques in terms of the shape, whereas the iterative reconstruction captures the shape



**Figure 4.** The boxplots of (a) NMSE, (b) bias, (c) PSNR, and (d) SSIM for the reconstructions obtained using RED-CNN and MS-RDN with the following configurations: single slice training without TOI oriented patch extraction ( $Z = 1$ , nonTOI), single slice training ( $Z = 1$ ), and multi-slice training ( $Z = 5$ ). For example, “MS-RDNZ1” represents MS-RDN with single slice training. On each box, the central mark is the median, the top and bottom are the 25th and 75th percentiles respectively. Outliers are denoted as red plus signs. Note that, in each breast-size group, MS-RDN and RED-CNN with the same configurations are placed next to each other for comparison.

	NMSE (dB)	Bias (cm <sup>-1</sup> )	PSNR (dB)	SSIM
Single slice training, non-TOI	+0.034	-0.399 × 10 <sup>-4</sup>	+0.034	+0.0003
Single slice training	+0.197	-1.501 × 10 <sup>-4</sup>	+0.197	+0.0006
Multi-slice training (Z = 5)	+0.144	-0.975 × 10 <sup>-4</sup>	+0.144	+0.0007

**Table 2.** Statistical analysis of MS-RDN and RED-CNN reconstructions using generalized linear models. The table reports the performance gained by MS-RDN over RED-CNN for three different configurations and four quantitative metrics. All improvements are significant with  $P < 0.0001$ . Please note that the values corresponding to NMSE and PSNR are identical since these quantities are related as shown in Eqs. (6) and (7).



**Figure 5.** A comparison to the FIRST algorithm. Breast reference images, FDK300 and FIRST300, are obtained using FDK and FIRST algorithms on the 300-view data respectively. Similarly, FIRST100 represents FIRST reconstructions on the retrospectively undersampled 100-view data. On the same undersampled data, breast images were reconstructed using MS-RDN with multi-slice training ( $Z = 5$ ), indicated as MS-RDNZ5. The display window is [0.15, 0.35] cm<sup>-1</sup>.

better. However, there is loss of detail and texture in other regions, such as the edges between adipose and fibroglandular tissues, with the iterative reconstruction.

## Discussion

In this study, we presented a deep learning (DL) based reconstruction framework for 3D sparse-view breast CT. In reference to full view FDK reconstructions, the proposed framework yields image quality superior to compressed sensing techniques such as FIRST while requiring comparable reconstruction times. In this study, the reconstructed FOV was relatively large (280 mm × 280 mm or 1024 pixel × 1024 pixel) to accommodate breasts with large diameter at the chest-wall<sup>36</sup>, which leads to large fraction of background in some of the datasets. Thus,

Metrics	FDK300 reference		FIRST300 reference	
	FIRST100	MS-RDNZ5	FIRST100	MS-RDNZ5
<b>NMSE (dB)</b>				
S	27.19 (1.03)	<b>27.50</b> (1.03)	32.67 (1.01)	<b>32.99</b> (1.17)
M	24.85 (0.19)	<b>25.67</b> (0.31)	29.15 (0.36)	<b>31.45</b> (1.53)
L	26.11 (0.38)	<b>26.78</b> (0.29)	34.93 (0.52)	<b>35.68</b> (0.71)
<b>Bias (<math>10^{-3} \text{ cm}^{-1}</math>)</b>				
S	9.08 (1.24)	<b>8.60</b> (1.65)	4.55 (0.71)	<b>4.12</b> (0.81)
M	11.80 (0.21)	<b>10.69</b> (0.29)	6.98 (0.26)	<b>5.19</b> (0.79)
L	8.68 (0.34)	<b>8.07</b> (0.19)	2.92 (0.19)	<b>2.75</b> (0.26)
<b>PSNR (dB)</b>				
S	41.17 (1.17)	<b>41.42</b> (1.37)	46.77 (1.09)	<b>47.00</b> (1.28)
M	38.95 (0.16)	<b>39.80</b> (0.25)	43.23 (0.31)	<b>45.56</b> (1.43)
L	41.62 (0.32)	<b>42.28</b> (0.21)	50.45 (0.53)	<b>51.22</b> (0.73)
<b>SSIM</b>				
S	0.941 (0.020)	<b>0.946</b> (0.021)	0.988 (0.004)	<b>0.989</b> (0.003)
M	0.893 (0.004)	<b>0.914</b> (0.003)	0.964 (0.003)	<b>0.985</b> (0.002)
L	0.938 (0.003)	<b>0.944</b> (0.002)	<b>0.994</b> (0.001)	<b>0.994</b> (0.001)

**Table 3.** Quantitative analysis of the proposed method (MS-RDNZ5) and the FIRST algorithm. One small-size breast (S), one medium-size breast (M), and one large-size breast (L) were selected for testing, respectively. The suffixes “100” and “300” denote the number of projections in the data. The MS-RDNZ5 network was always trained using FDK100 as input and FDK300 as label. However, either FDK300 or FIRST300 were used as the reference when computing the quality metrics, as indicated by the column labels “FDK300 Reference” and “FIRST300 Reference”, respectively. Median and interquartile range in the bracket are shown. Bolded values indicate better performance in pairwise comparison.

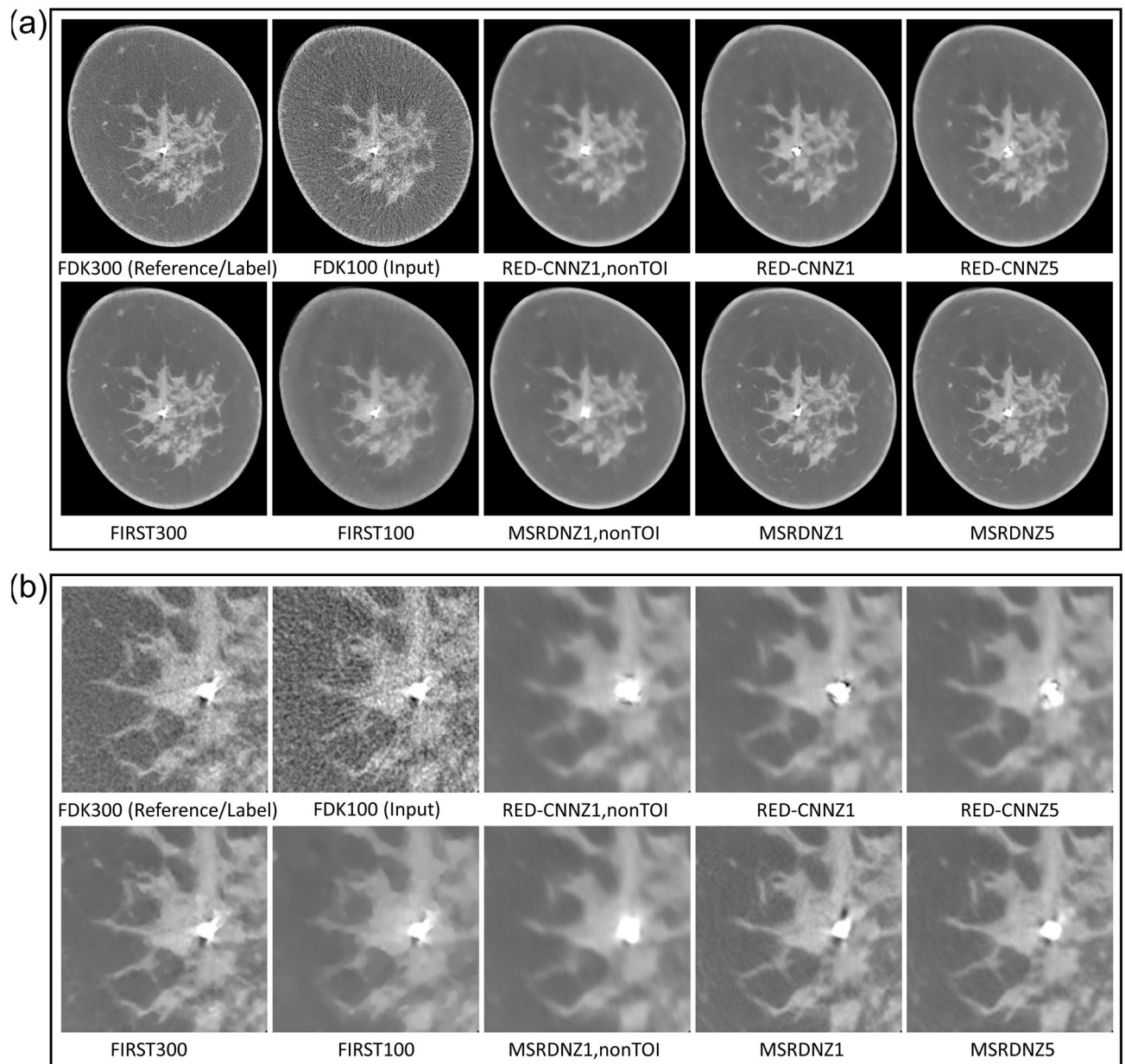
we adopted a tissue of interest oriented patch extraction strategy, termed TOI selection, during the network training to enforce learning on the breast tissue region rather than the irrelevant background regions. Importantly, patches that contain less than 50% background pixels were also included in training to ensure the recovery of breast anatomy boundary. This TOI selection alone enhanced the sharpness of breast textures and achieved improved NMSE and bias compared to random patch extraction.

This work used multi-slice training as a compromise between 2D and 3D network training. We demonstrated that multi-slice training is effective in exploiting the correlations between adjacent slices. Most importantly, it eliminated the Venetian blind artifacts in images obtained using single slice training. However, we also noticed that the performance gained with increased slice depth of MS-RDN saturates at small number of slice depth. This suggests the longitudinal correlation is largely local. One future extension to the current work could be assembling three networks trained in the axial, coronal, and sagittal planes respectively. The ensemble of three 3D networks explores local similarities along all three orientations similar to what a 3D network does but it would still require much less GPU memory and training data.

Our DL-based framework uses residual dense blocks<sup>33,43</sup> as the backbone of the network. It has been shown that such combination of residual connections<sup>31</sup> and densely connected structures<sup>32</sup> improved network parameter efficiency and reconstruction accuracy in single image super resolution problems<sup>33,43</sup>. Our MS-RDN was comprehensively compared with the residual learning based RED-CNN and showed superior reconstruction quality of breast CT images. While this study demonstrated promise in the task of sparse-view breast CT reconstruction, it has several limitations. The reference FDK reconstruction exhibits higher noise than multi-detector CT used for imaging other organs, due to the hardware limitations and radiation dose constraints. Our MS-RDN reconstructions looked (perceptually) more similar to the FIRST approaches in terms of signal-to-noise ratio. Recent studies<sup>44–46</sup> suggest that pixel-wise losses, such as  $\ell_1$  or  $\ell_2$  loss, are prone to overly smoothing image structures. In contrast, adversarial training<sup>47,48</sup>, perceptual loss<sup>49</sup>, as well as texture matching loss<sup>50</sup> are proven to preserve high frequency image content and improve the perceptual quality. However, it should be noted that these techniques may hallucinate high frequency textures<sup>44</sup>, which makes them less appealing for medical applications. In breast CT imaging, hallucinated high frequency texture may mimic microcalcifications. Nevertheless, the impact of alternative loss functions in dedicated breast CT needs to be investigated and can be an extension of the current work.

We also investigated the possible failure cases for the proposed deep learning technique. For the example shown in Fig. 6, we found out that both MS-RDN and RED-CNN (irrespective of their configurations) produced poor reconstructions of the shape of a calcification. Note that the calcification is a minor class compared to the fibroglandular or adipose tissues in the training dataset. Unlike the iterative compressed sensing method, which includes data consistency and model based priors, the proposed method learns from training samples. Hence, the network may not learn the characteristics of tissues that are scarcely represented in the training data. It would be interesting to develop deep learning techniques that can yield improved reconstructions of such calcifications in future works.





**Figure 6.** Reconstructions of the slice that yields the worst NMSE performance for MS-RDNZ5 in Fig. 4a. Reconstructions from all investigated methods are shown in (a). The zoomed regions of the central part of the breast tissue with a calcification are shown in (b). The display window is  $[0.15, 0.35] \text{ cm}^{-1}$ .

## Methods

**Projection acquisition and three-dimensional image reconstruction.** In 3D cone-beam BCT, multi-projection data  $\mathbf{P} \in \mathbb{R}^{N_d \times N_p}$  were acquired in a complete circular trajectory composed of  $N_p$  projections using a two-dimensional (2D) X-ray area detector consisting of  $N_d$  pixels. From the cone-beam projections  $\mathbf{P}$ , an estimate of the underlying image volume  $\mathbf{V} \in \mathbb{R}^{N_x \times N_y \times N_z}$  was reconstructed using the conventional analytical FDK algorithm<sup>24</sup>. The reconstruction process can be expressed using the following equation

$$\mathbf{V} = \mathbf{F}(\mathbf{P}), \quad (1)$$

where  $\mathbf{F}$  denotes the FDK reconstruction operator interpolated by voxel-driven approach<sup>51,52</sup>. Reconstructed volumes are assumed to have isotropic voxel resolution as the voxel sizes are principally determined by size of the imaging detectors. However, the spatial resolution can be location-dependent and anisotropic due to reduced sampling at the periphery of the field of view within a transverse slice and due to geometric distortions arising from cone-beam geometry (commonly referred to as cone-beam artifacts) as the acquisition does not satisfy data-completeness requirement<sup>53,54</sup> with the exception of the central transverse slices.

To reduce radiation dose, a common way is to uniformly reduce the number of projections without compromising the full angular coverage<sup>55–57</sup>. This sparse-view projection data was obtained by retrospectively under-sampling the full-view projection data  $\mathbf{P}$  using

$$\mathbf{P}_u = \mathbf{P}[1 : 1 : N_d, 1 : u : N_p], \quad (2)$$

where  $\mathbf{P}_u \in \mathbb{R}^{N_d \times \lfloor \frac{N_p}{u} \rfloor}$  represents the sparse-view projection data,  $u$  denotes the undersampling factor, and the notation  $i : j : k$  in Eq. (1) denotes regularly spaced sampling between indices  $i$  and  $k$  using  $j$  as the increment. Similarly, an estimate of the image volume  $\mathbf{V}_u$  was reconstructed from the sparse-view data  $\mathbf{P}_u$  using the FDK algorithm, that is

$$\mathbf{V}_u = \mathbf{F}(\mathbf{P}_u). \quad (3)$$

It should be noted that the reconstructed image volume  $\mathbf{V}_u$  typically exhibits streaking artifacts due to undersampling.

**Deep neural network reconstruction.** Earlier studies on abdominal contrast-enhanced CT<sup>58</sup> and optoacoustic tomography<sup>59</sup> showed promising performance of deep neural network reconstruction with sparse data. The goal of this work is to combine sparse-view data acquisition with deep neural network reconstruction to reduce undersampling artifacts. A deep neural network  $\mathbf{D}(\mathbf{w}, \cdot)$  can be utilized to recover  $\mathbf{V}$  from  $\mathbf{V}_u$ , where  $\mathbf{w}$  are the weights of  $\mathbf{D}$ . In supervised learning,  $\mathbf{w}$  are optimized by minimizing a pre-defined loss function  $\mathbf{L}(\cdot)$ , namely,

$$\hat{\mathbf{w}} = \arg \max_{\mathbf{w}} \mathbf{L}(\mathbf{D}(\mathbf{w}, \mathbf{V}_u), \mathbf{V}) \quad (4)$$

over a training dataset.

Our proposed framework uses supervised training where the inputs and targets of the network are obtained using Eqs. (1) and (3), respectively. While it may be ideal to process the entire volume using a 3D neural network, there are practical constraints associated with 3D networks<sup>60–65</sup>. Conventional denoising methods for 3D CT images based on non-local means<sup>66</sup> or block matching filter<sup>67</sup> showed that a multi-slice approach is able to leverage inter-slice spatial dependencies with small growth in computational complexity. Hence, we jointly reconstruct  $Z \in \mathbb{Z}^+$  adjacent slices as a compromise between 2D and 3D processing.

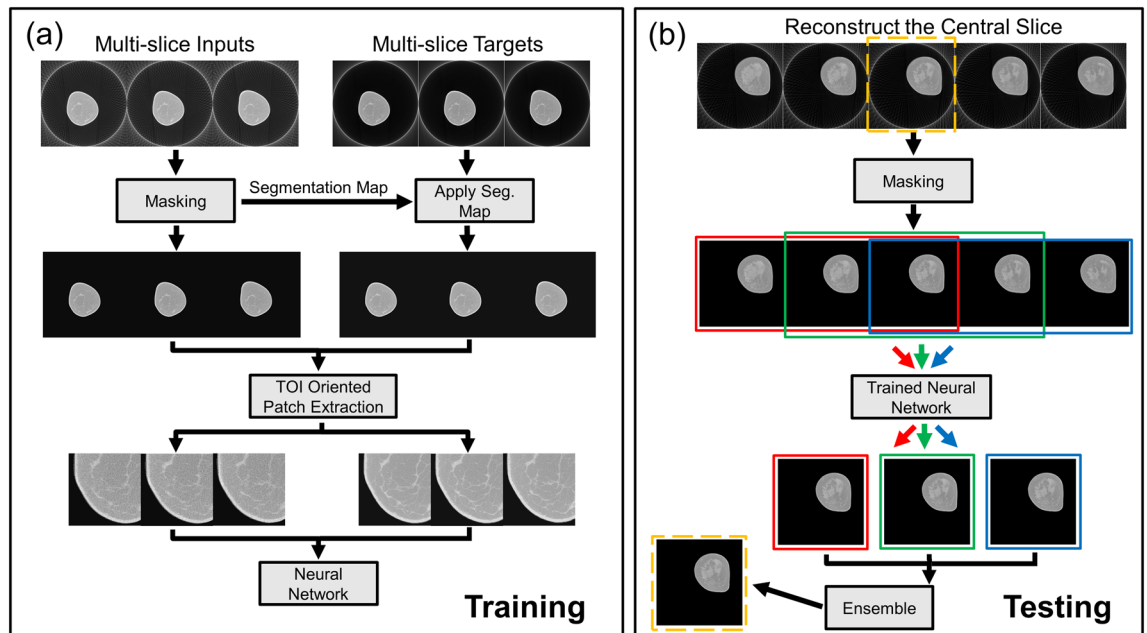
Figure 7a illustrates the proposed training procedure for  $Z = 3$ . The first step in processing is a masking procedure to remove the background regions in each slice. Figure 8 illustrates this masking process for an individual image slice. In this process, masking was performed to remove the artifacts outside of the circular Field of View (FOV). The image data within the circular FOV across all slices were used to create a histogram of linear attenuation coefficients for the entire volume. Based on the observation that the background noise and undersampling artifacts (streaks) are well separated from the breast tissue in this histogram, we selected the bin center with the lowest bin count as the hard threshold and created segmentation maps that identify the breast tissue in each slice. We further dilated the segmentation maps using a flat disk-shaped structuring element with a radius of 2 pixels. Segmentation maps created from the input slices were shared with the corresponding target slices as shown in Fig. 7a. Training is performed using patch pairs extracted from the input and target volumes. Selection of training samples is a well-studied area in machine learning literature and numerous methods have been proposed to reduce bias through training sample selection<sup>68–70</sup>. Inspired by these techniques, patches that contain more than 50% foreground pixels were selected as training samples. This patch extraction process is referred to as tissue-of-interest (TOI) selection.

The network testing phase is illustrated in Fig. 7b. Since the proposed network reconstructs multiple slices simultaneously, a target slice (indicated by dotted yellow bounding box) is reconstructed multiple times in different slice contexts (indicated by red, green, and blue bounding boxes). In this illustration, 5 adjacent slices were first preprocessed using the same masking procedure as the training phase. Using a sliding window of size 3 and stride of size 1, the target slice is processed three times by the network. The three reconstructions are then combined using an ensemble strategy. In summary, for any trained network  $\mathbf{D}_Z(\hat{\mathbf{w}}, \cdot)$  with slice depth  $Z$ , the ensemble strategy to obtaining the target slice reconstruction  $\hat{\mathbf{S}}_t$  can be formulated as

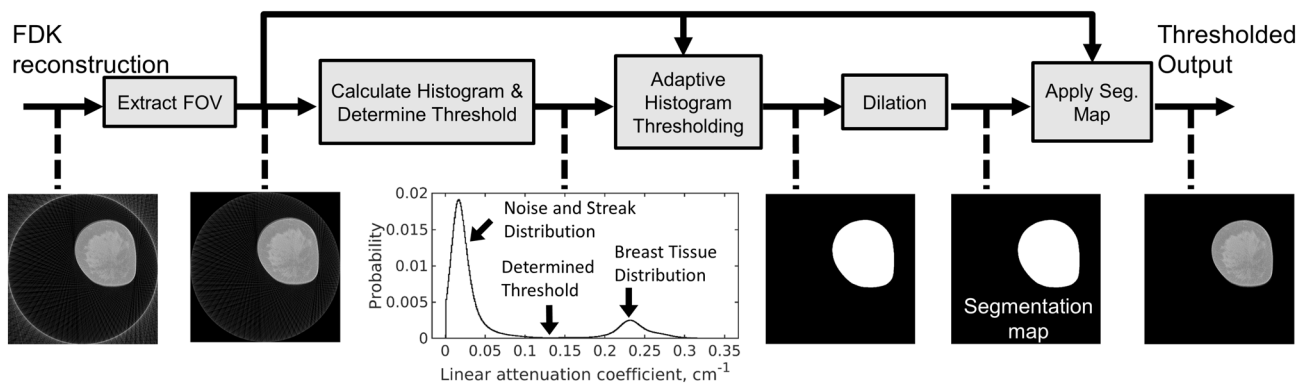
$$\begin{aligned} \hat{\mathbf{S}}_t = f(g_t(\mathbf{D}_Z(\hat{\mathbf{w}}, \mathbf{S}_{t-Z+1}, \mathbf{S}_{t-Z+2}, \dots, \mathbf{S}_t)), \\ g_t(\mathbf{D}_Z(\hat{\mathbf{w}}, \mathbf{S}_{t-Z+2}, \mathbf{S}_{t-Z+3}, \dots, \mathbf{S}_{t+1})), \\ \dots, \\ g_t(\mathbf{D}_Z(\hat{\mathbf{w}}, \mathbf{S}_t, \mathbf{S}_{t+1}, \dots, \mathbf{S}_{t+Z-1}))) \end{aligned} \quad (5)$$

where  $f$  denotes the ensemble function,  $g_t$  only retains the reconstruction of the target slice  $t$ , and  $\mathbf{S}_i$  denotes the slice  $i$  of the input. In our experiment, we found evenly averaging is a simple yet effective ensemble approach. We replicate border slices to handle slices at edges.

**Network architecture.** The proposed MS-RDN architecture is shown in Fig. 9a. Multi-slice inputs are first processed by a shared 2D convolutional layer. The resulting 3D spatial features are then consecutively propagated through the high resolution and low resolution feature branches. Learned high resolution and low resolution features are summed using a trainable weighting factor. In the end, the output convolutional layer reconstructs multi-slice outputs from the fused feature maps. Inspired by Ledig et al.<sup>44</sup>, our feature branch is sequentially composed of multiple dense compression units (DCUs)<sup>33</sup>, a  $3 \times 3$  convolutional layer and a skip connection. As

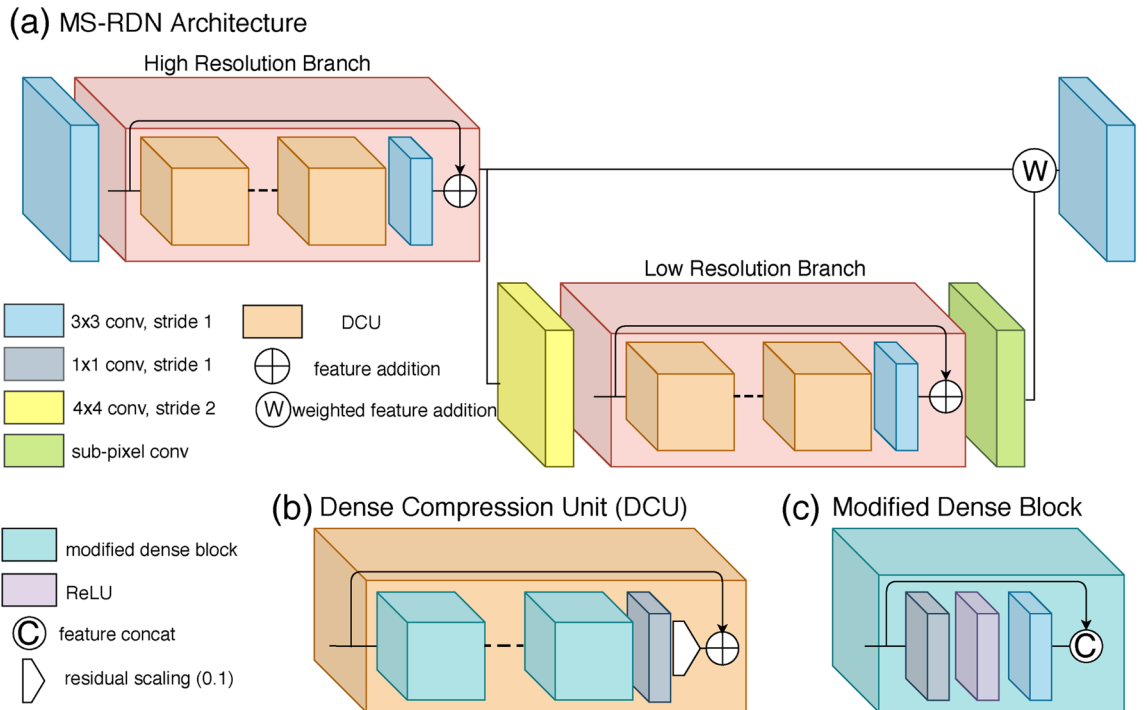


**Figure 7.** Network multi-slice (a) training and (b) testing framework. Training with three slices is shown as an example. (a) Multi-slice inputs reconstructed from sparse projection data is processed with the masking procedure described in Fig. 8. The generated segmentation maps are shared with multi-slice targets reconstructed from full projection data. Patches are extracted as training samples only when they contain more than 50% foreground pixels based on the generated masks, termed tissue of interest (TOI) oriented. (b) Five consecutive testing slices are used to reconstruct the central slice, indicated by the yellow bounding box. Three sets of multi-slice inputs, where the target slice has different slice context, are independently processed by the same trained network. Only the target slices are retained and aggregated to obtain the final reconstruction of the target slice.



**Figure 8.** The masking procedure. Circular Field of View (FOV) of the FDK reconstruction is extracted to remove out-of-FOV artifacts. Typically, streaks and breast tissue are well separated in the histogram of linear attenuation coefficients. Based on the histogram, an adaptive thresholding algorithm that selects the bin center with lowest bin counts as the hard threshold is used to generate the segmentation map and the thresholded output. The images and plots linked by dashed line show the intermediate outputs of the entire processing pipeline.

shown in Fig. 9b, the DCU consists of stacked densely connected blocks, a  $1 \times 1$  convolutional layer, a residual scaling (0.1) and a local skip connection. The  $1 \times 1$  convolutional layer compresses accumulated features to the same number of input features, which enables the residual connection within the dense structure. The constant scaling stabilizes network training, when the number of filters is high<sup>34,71</sup>. The DCU structure efficiently merges local feature information and periodically breaks dense connections to improve back projection of gradients<sup>33</sup>. Figure 9c details the layout of modified dense block, where all batch normalization layers are removed compared to the original DenseNet configuration<sup>32</sup>.



**Figure 9.** The architecture of multi-slice residual and dense network (MS-RDN). **(a)** Overall layouts; **(b)** the detailed layouts of dense compression unit (DCU); **(c)** the detailed layouts of modified dense block.

**Network evaluation.** To demonstrate the superiority of multi-slice training, we first trained multiple MS-RDNs with the same configurations except for the number of adjacent slices, i.e.,  $Z = 1, 3, 5, 7, 9$ , respectively. Note that when  $Z = 1$ , MS-RDN reduces to the single slice network, i.e. 2D network.

Our MS-RDN was compared with the residual encoder–decoder convolutional neural network (RED-CNN)<sup>26</sup> designed for low dose CT image reconstruction. We followed the implementation of RED-CNN from [https://github.com/SSinyu/RED\\_CNN](https://github.com/SSinyu/RED_CNN) and adopted the suggested network parameters (for example, convolutional kernel size is set to 5). Note that unlike our proposed deep learning reconstruction framework, RED-CNN<sup>26</sup> was trained with *randomly extracted single-slice patches*. We therefore applied the TOI selection and multi-slice training scheme to the RED-CNN architecture for comparison.

Nine randomly selected test subjects were evenly grouped by the size of breast. To reduce the impact of breast size or slice location, we always select a constant number of measurement samples within the breast for quantitative analysis. The network reconstructions were evaluated with normalized mean square error (NMSE), bias, peak signal-to-noise ratio (PSNR), and Structural Similarity Index Metric (SSIM<sup>42</sup>). The NMSE metric was computed as the ratio of mean square error to mean square of the reference image and converted into decibel (dB), that is

$$\text{NMSE}(\mathbf{x}, \mathbf{x}_{ref}) = -10 \times \log_{10} \left( \frac{\|\mathbf{x} - \mathbf{x}_{ref}\|_2^2}{\|\mathbf{x}_{ref}\|_2^2} \right). \tag{6}$$

The bias metric was computed as the mean absolute error. The PSNR metric was computed as the ratio of the maximum pixel intensity ( $I_{max}$ ) squared to mean square error as

$$\text{PSNR}(\mathbf{x}, \mathbf{x}_{ref}) = 10 \times \log_{10} \left( \frac{I_{max}^2}{\|\mathbf{x} - \mathbf{x}_{ref}\|_2^2} \right). \tag{7}$$

The SSIM index was computed using the default hyper-parameters except that the dynamic range of pixel values was set to the maximum pixel intensity within the entire dataset. All metrics were calculated in the longitudinal direction as the representation.

The fast, iterative, tv-regularized, statistical reconstruction technique (FIRST<sup>22</sup>) was also used for sparse-view image reconstruction. This algorithm is an ultra-fast variant of the adaptive steepest descent-projection on to convex sets (ASD-POCS<sup>72</sup>) and has been shown to suppress additional artifacts on the periphery of the object. The performance of FIRST was compared to MS-RDN using one small-size breast, one medium-size breast, and one large-size breast.

**Implementation.** We construct our MS-RDN with a high resolution branch and a low resolution branch, where each branch consists of 9 DCUs and each DCU is composed of 8 modified dense blocks. The initial number of features is set to 64 with a growth-rate of 32. To evaluate the impact of network depth on RED-CNN

performance, we implemented RED-CNN with 10, 22, and 42 convolutional layers. Note that the 10-layer architecture corresponds to what was proposed in the RED-CNN paper<sup>26</sup> and the 42-layer RED-CNN with  $Z = 5$  has roughly the same number of trainable parameters (9,243,941) as our MS-RDN with  $Z = 5$  (9,237,126). In line with observations made in earlier studies<sup>26,73</sup>, we determined that deeper RED-CNNs perform roughly the same as the 10-layer RED-CNN in our application (see Supplementary Fig. S1). Thus, we used the 10-layer RED-CNN for its computational simplicity.

All models were optimized using ADAM with its standard settings ( $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ , and  $\epsilon = 10^{-8}$ ) for 100 epochs. Each mini-batch consists of 8 training samples with patch size  $128 \times 128 \times Z$ , and was normalized by the mean and standard deviation of the entire training data. All networks were trained with  $\ell_1$  loss. The learning rate was initially set to  $1 \times 10^{-4}$  and halved every  $2 \times 10^5$  mini-batch updates. The single slice network was trained from scratch and used as a pre-trained model for other multi-slice networks. To fine-tune on the pre-trained single slice network, we replicated the single channel weights along the channel dimension at the input and output convolutional layers, respectively<sup>74</sup>. Pre-training, as an approach to initializing network weights, has been shown to improve training stability of larger networks<sup>27,74</sup>. In contrast, we found that further training of the single-slice network does not lead to considerable improvements (see Supplementary Fig. S2). The model with the best validation loss was evaluated at inference time.

Our MS-RDN was implemented in PyTorch<sup>75</sup> with CUDA backend and CUDNN support, and trained on a NVIDIA Quadro P6000 GPU. The network took about 60 hours on average for 100 epochs training. The FDK and FIRST algorithms were implemented in MATLAB with GPU acceleration. Ram-Lak filter was used for the FDK algorithm and FDK reconstructions were used as the initialization of the FIRST algorithm. Other standard hyperparameters of FIRST were:  $\beta = 1$ ,  $\beta_{\text{residual}} = 0.995$ ,  $\alpha = 0.001$ ,  $\alpha_{\text{residual}} = 0.95$ ,  $r_{\text{max}} = 0.95$ , 100 total iterations, and 30 Total Variation iterations. On average, MS-RDN, RED-CNN, FDK, and FIRST require about 2.3 s, 1.2 s, 0.01 s and 3.1 s per slice ( $1024 \times 1024$  matrix size), respectively, on a single NVIDIA Quadro P6000 GPU. Note that MS-RDN and RED-CNN are able to reconstruct breast images in a slice-by-slice manner, whereas FDK and FIRST reconstruct the entire breast volume simultaneously. MS-RDN, RED-CNN, FDK, and FIRST require about 9.0 GB, 2.4 GB, 2.5 GB, and 6.3 GB GPU memory, respectively.

**Statistical analysis.** Generalized linear models (repeated measures analysis of variance) were used to test if the metric (NMSE, bias, PSNR, and SSIM) differed between the reconstructions, as the same set of test cases were reconstructed using different methods. Effects associated with  $P < 0.05$  were considered statistically significant. If the generalized linear model showed significant difference, then follow-up paired t-tests were performed to determine (i) if the metric differed between TOI and non-TOI strategies for MS-RDN and RED-CNN; (ii) if the metric differed between  $Z = 1$  and  $Z = 5$  for MS-RDN and RED-CNN; and (iii) if MS-RDN differed from RED-CNN for the TOI strategy when  $Z = 1$  and  $Z = 5$ . For each metric, this results in a total of 7 comparisons. Hence, Bonferroni-adjusted alpha of 0.007 was considered statistically significant for these pairwise comparisons. The data analysis for this paper was generated using SAS software, Version 9.4 of the SAS System for Windows.

Received: 28 January 2020; Accepted: 18 November 2020

Published online: 03 December 2020

## References

- Vedantham, S., Karellas, A., Vijayaraghavan, G. R. & Kopans, D. B. Digital breast tomosynthesis: state of the art. *Radiology* **277**, 663–684. <https://doi.org/10.1148/radiol.2015141303> (2015).
- Sujlana, P. S. *et al.* Digital breast tomosynthesis: image acquisition principles and artifacts. *Clin. Imaging* **55**, 188–195. <https://doi.org/10.1016/j.clinimag.2018.07.013> (2019).
- Cole, E. B., Campbell, A. S., Vedantham, S., Pisano, E. D. & Karellas, A. Clinical performance of dedicated breast computed tomography in comparison to diagnostic digital mammography. In *Radiological Society of North America 2018 Scientific Assembly and Annual Meeting* (Chicago IL, 2015).
- Vedantham, S., Shi, L., Karellas, A., Oonnell, A. M. & Conover, D. L. Personalized estimates of radiation dose from dedicated breast CT in a diagnostic population and comparison with diagnostic mammography. *Phys. Med. Biol.* **58**, 7921–7936. <https://doi.org/10.1088/0031-9155/58/22/7921> (2013).
- Lindfors, K. K. *et al.* Dedicated breast CT: initial clinical experience. *Radiology* **246**, 725–733. <https://doi.org/10.1148/radiol.2463070410> (2008).
- Rößler, A.-C. *et al.* Performance of photon-counting breast computed tomography, digital mammography, and digital breast tomosynthesis in evaluating breast specimens. *Acad. Radiol.* **24**, 184–190. <https://doi.org/10.1016/j.acra.2016.09.017> (2017).
- Kalluri, K. S., Mahd, M. & Glick, S. J. Investigation of energy weighting using an energy discriminating photon counting detector for breast CT: investigation of energy weighting with CZT for breast CT. *Med. Phys.* **40**, 081923. <https://doi.org/10.1118/1.4813901> (2013).
- Gazi, P. M. *et al.* Evolution of spatial resolution in breast CT at UC Davis. *Med. Phys.* **42**, 1973–1981 (2015).
- Vedantham, S., Shi, L. & Karellas, A. Scintillator performance considerations for dedicated breast computed tomography. In *Radiation Detectors in Medicine, Industry, and National Security XVIII* (eds Grim, G. P. *et al.*) (SPIE, San Diego, 2017).
- Lück, F., Kolditz, D., Hupfer, M. & Kalender, W. A. Effect of shaped filter design on dose and image quality in breast CT. *Phys. Med. Biol.* **58**, 4205–4223. <https://doi.org/10.1088/0031-9155/58/12/4205> (2013).
- Vedantham, S. & Karellas, A. Dedicated Breast CT: Numerical Evaluation of Improvement in X-Ray Fluence Uniformity Using 3d Beam-Shaping X-Ray Filter. In *59th Annual Meeting of the American Association of Physicists in Medicine* (American Association of Physicists in Medicine, Denver, 2017).
- Mettivier, G., Russo, P., Lanconelli, N. & Meo, S. L. Cone-beam breast computed tomography with a displaced flat panel detector array: CBBCT with a displaced flat panel detector array. *Med. Phys.* **39**, 2805–2819. <https://doi.org/10.1118/1.4704641> (2012).

13. Vedantham, S., Tseng, H.-W., Konate, S., Shi, L. & Karellas, A. Dedicated cone-beam breast CT using laterally-shifted detector geometry: quantitative analysis of feasibility for clinical translation. *J. X-ray Sci. Technol.* **28**, 405 (2020).
14. Tseng, H., Vedantham, S. & Karellas, A. Dedicated breast CT: impact of short-scan source trajectory and sparse-view acquisition on image quality. In *60th Annual Meeting of the American Association of Physicists in Medicine (AAPM)* (Med Phys, Nashville, TN, 2018).
15. Tseng, H., Vedantham, S. & Karellas, A. Impact of GPU-accelerated sparse-view reconstruction for radiation dose reduction in dedicated breast CT. In *60th Annual Meeting of the American Association of Physicists in Medicine (AAPM)* (Med Phys, Nashville, TN, 2018).
16. Glick, S. J., Vedantham, S. & Karellas, A. Investigation of optimal kVp settings for CT mammography using a flat-panel imager. In *Medical Imaging 2002* (eds Antonuk, L. E. & Yaffe, M. J.) 392–402 (SPIE, San Diego, 2002).
17. Prionas, N. D., Huang, S.-Y. & Boone, J. M. Experimentally determined spectral optimization for dedicated breast computed tomography: experimentally determined spectral optimization for dedicated breast CT. *Med. Phys.* **38**, 646–655. <https://doi.org/10.1118/1.3537077> (2011).
18. Crotty, D. J., McKinley, R. L. & Tornai, M. P. Experimental spectral measurements of heavy K-edge filtered beams for x-ray computed mammotomography. *Phys. Med. Biol.* **52**, 603–616. <https://doi.org/10.1088/0031-9155/52/3/005> (2007).
19. McKinley, R. L., Tornai, M. P., Samei, E. & Bradshaw, M. L. Simulation study of a quasi-monochromatic beam for x-ray computed mammotomography. *Med. Phys.* **31**, 800–813. <https://doi.org/10.1118/1.1668371> (2004).
20. Makeev, A. & Glick, S. J. Investigation of statistical iterative reconstruction for dedicated breast CT: statistical iterative reconstruction for dedicated breast CT. *Med. Phys.* **40**, 081904. <https://doi.org/10.1118/1.4811328> (2013).
21. Bian, J. *et al.* Investigation of iterative image reconstruction in low-dose breast CT. *Phys. Med. Biol.* **59**, 2659–2685. <https://doi.org/10.1088/0031-9155/59/11/2659> (2014).
22. Tseng, H. W., Vedantham, S. & Karellas, A. Cone-beam breast computed tomography using ultra-fast image reconstruction with constrained, total-variation minimization for suppression of artifacts. *Phys. Med.* **73**, 117–124. <https://doi.org/10.1016/j.ejmp.2020.04.020> (2020).
23. Xia, J. Q., Lo, J. Y., Yang, K., Floyd, C. E. & Boone, J. M. Dedicated breast computed tomography: volume image denoising via a partial-diffusion equation based technique: PDE based volume image denoising in breast CT. *Med. Phys.* **35**, 1950–1958. <https://doi.org/10.1118/1.2903436> (2008).
24. Feldkamp, L. A., Davis, L. C. & Kress, J. W. Practical cone-beam algorithm. *J. Opt. Soc. Am. A* **1**, 612. <https://doi.org/10.1364/JOSAA.1.000612> (1984).
25. Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).
26. Chen, H. *et al.* Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans. Med. Imaging* **36**, 2524–2535. <https://doi.org/10.1109/TMI.2017.2715284> (2017).
27. Shan, H. *et al.* 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network. *IEEE Trans. Med. Imaging* **37**, 1522–1534. <https://doi.org/10.1109/TMI.2018.2832217> (2018).
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**, 1348–1357. <https://doi.org/10.1109/TMI.2018.2827462> (2018).
29. Chen, H. *et al.* Low-dose CT via convolutional neural network. *Biomed. Opt. Express* **8**, 679. <https://doi.org/10.1364/BOE.8.000679> (2017).
30. Kang, E., Chang, W., Yoo, J. & Ye, J. C. Deep convolutional framelet denoising for low-dose CT via wavelet residual network. *IEEE Trans. Med. Imaging* **37**, 1358–1369. <https://doi.org/10.1109/TMI.2018.2823756> (2018).
31. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778 (IEEE, Las Vegas, NV, USA, 2016). <https://doi.org/10.1109/CVPR.2016.90>.
32. Huang, G., Liu, Z., Maaten, L. v. d. & Weinberger, K. Q. Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2261–2269 (IEEE, Honolulu, HI, 2017). <https://doi.org/10.1109/CVPR.2017.243>.
33. Wang, Y. *et al.* A fully progressive approach to single-image super-resolution. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 977–97709 (IEEE, Salt Lake City, 2018). <https://doi.org/10.1109/CVPRW.2018.00131>.
34. Lim, B., Son, S., Kim, H., Nah, S. & Lee, K. M. Enhanced deep residual networks for single image super-resolution. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1132–1140 (IEEE, Honolulu, HI, USA, 2017). <https://doi.org/10.1109/CVPRW.2017.151>.
35. D'Orsi, C. J. *et al.* *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System* 39–48 (American College of Radiology, Reston, 2013).
36. Vedantham, S., Shi, L., Karellas, A. & OConnell, A. M. Dedicated breast CT: fibroglandular volume measurements in a diagnostic population. *Med. Phys.* **39**, 7317–7328. <https://doi.org/10.1118/1.4765050> (2012).
37. Vedantham, S., Shi, L., Karellas, A., OConnell, A. M. & Conover, D. Dedicated Breast CT: Anatomic Power Spectrum. In *2nd International Conference on Image Formation in X-ray Computed Tomography* (Salt Lake City, UT, 2012).
38. Shi, L., Vedantham, S., Karellas, A. & OConnell, A. M. Technical note: skin thickness measurements using high-resolution flat-panel cone-beam dedicated breast CT. *Med. Phys.* **40**, 031913. <https://doi.org/10.1118/1.4793257> (2013).
39. Shi, L., Vedantham, S., Karellas, A. & Zhu, L. Library based x-ray scatter correction for dedicated cone beam breast CT: library-based scatter correction. *Med. Phys.* **43**, 4529–4544. <https://doi.org/10.1118/1.4955121> (2016).
40. Shi, L., Vedantham, S., Karellas, A. & Zhu, L. X-ray scatter correction for dedicated cone beam breast CT using a forward-projection model. *Med. Phys.* **44**, 2312–2320. <https://doi.org/10.1002/mp.12213> (2017).
41. Shi, L., Vedantham, S., Karellas, A. & Zhu, L. The role of off-focus radiation in scatter correction for dedicated cone beam breast CT. *Med. Phys.* **45**, 191–201. <https://doi.org/10.1002/mp.12686> (2018).
42. Wang, Z., Bovik, A., Sheikh, H. & Simoncelli, E. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612. <https://doi.org/10.1109/TIP.2003.819861> (2004).
43. Zhang, Y., Tian, Y., Kong, Y., Zhong, B. & Fu, Y. Residual dense network for image super-resolution. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2472–2481 (IEEE, Salt Lake City, UT, 2018). <https://doi.org/10.1109/CVPR.2018.00262>.
44. Ledig, C. *et al.* Photo-realistic single image super-resolution using a generative adversarial network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 105–114 (IEEE, Honolulu, HI, 2017). <https://doi.org/10.1109/CVPR.2017.19>.
45. Sajjadi, M. S. M., Scholkopf, B. & Hirsch, M. EnhanceNet: single image super-resolution through automated texture synthesis. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 4501–4510 (IEEE, Venice, 2017). <https://doi.org/10.1109/ICCV.2017.481>.
46. Wolterink, J. M., Leiner, T., Viergever, M. A. & Išgum, I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Trans. Med. Imaging* **36**, 2536–2545. <https://doi.org/10.1109/TMI.2017.2708987> (2017).
47. Goodfellow, I. *et al.* Generative adversarial nets. In *Advances in Neural Information Processing Systems 27* (eds Ghahramani, Z. *et al.*) 2672–2680 (Curran Associates Inc., New York, 2014).
48. 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**, 1348–1357 (2018).
49. Johnson, J., Alahi, A. & Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In *Computer Vision—ECCV 2016* (eds Leibe, B. *et al.*) 694–711 (Springer, Cham, 2016).

50. Gatys, L. A., Ecker, A. S. & Bethge, M. Image style transfer using convolutional neural networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2414–2423 (IEEE, Las Vegas, NV, USA, 2016). <https://doi.org/10.1109/CVPR.2016.265>.
51. Peters, T. M. Algorithms for fast back- and re-projection in computed tomography. *IEEE Trans. Nucl. Sci.* **28**, 3641–3647. <https://doi.org/10.1109/TNS.1981.4331812> (1981).
52. Zhuang, W., Gopal, S. & Hebert, T. Numerical evaluation of methods for computing tomographic projections. *IEEE Trans. Nucl. Sci.* **41**, 1660–1665. <https://doi.org/10.1109/23.322963> (1994).
53. Tuy, H. K. An inversion formula for cone-beam reconstruction. *SIAM J. Appl. Math.* **43**, 546–552. <https://doi.org/10.1137/0143035> (1983).
54. Smith, B. D. Image reconstruction from cone-beam projections: necessary and sufficient conditions and reconstruction methods. *IEEE Trans. Med. Imaging* **4**, 14–25. <https://doi.org/10.1109/TMI.1985.4307689> (1985).
55. Hsieh, J., Wei, Y. & Wang, G. Fractional scan algorithms for low-dose perfusion CT. *Med. Phys.* **31**, 1254–1257. <https://doi.org/10.1118/1.1708653> (2004).
56. Abbas, S., Lee, T., Shin, S., Lee, R. & Cho, S. Effects of sparse sampling schemes on image quality in low-dose CT: sparse sampling schemes. *Med. Phys.* **40**, 111915. <https://doi.org/10.1118/1.4825096> (2013).
57. Bian, J. *et al.* Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. *Phys. Med. Biol.* **55**, 6575–6599. <https://doi.org/10.1088/0031-9155/55/22/001> (2010).
58. Urase, Y. *et al.* Simulation study of low-dose sparse-sampling ct with deep learning-based reconstruction: usefulness for evaluation of ovarian cancer metastasis. *Appl. Sci.* **10**, 4446 (2020).
59. Davoudi, N., Deán-Ben, X. L. & Razansky, D. Deep learning optoacoustic tomography with sparse data. *Nat. Mach. Intell.* **1**, 453–460 (2019).
60. Roth, H. R. *et al.* A new 2.5 D representation for lymph node detection using random sets of deep convolutional neural network observations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 520–527 (Springer, 2014).
61. Setio, A. A. A. *et al.* Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans. Med. Imaging* **35**, 1160–1169 (2016).
62. Tran, D., Ray, J., Shou, Z., Chang, S.-F. & Paluri, M. ConvNet Architecture Search for Spatiotemporal Feature Learning. <http://arxiv.org/abs/1708.05038> (2017).
63. Shan, H. *et al.* 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network. *IEEE Trans. Med. Imaging* **37**, 1522–1534 (2018).
64. Tran, D. *et al.* A closer look at spatiotemporal convolutions for action recognition. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6450–6459 (IEEE, Salt Lake City, UT, 2018). <https://doi.org/10.1109/CVPR.2018.00675>.
65. Ziabari, A. *et al.* 2.5D deep learning for CT image reconstruction using a multi-GPU implementation. In *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, 2044–2049 (IEEE, Pacific Grove, CA, USA, 2018). <https://doi.org/10.1109/ACSSC.2018.8645364>.
66. Storozhilova, M., Lukin, A., Yurin, D. & Sinityn, V. 2.5 D extension of neighborhood filters for noise reduction in 3D medical CT images. In *Transactions on Computational Science XIX*, 1–16 (Springer, 2013).
67. Hasan, A. M., Melli, A., Wahid, K. A. & Babyn, P. Denoising low-dose CT images using multiframe blind source separation and block matching filter. *IEEE Trans. Radiat. Plasma Med. Sci.* **2**, 279–287 (2018).
68. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
69. Chang, H.-S., Learned-Miller, E. & McCallum, A. Active bias: training more accurate neural networks by emphasizing high variance samples. in *Advances in Neural Information Processing Systems*, 1002–1012 (2017).
70. Ren, M., Zeng, W., Yang, B. & Urtasun, R. Learning to reweight examples for robust deep learning. In *ICML* (2018).
71. Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-First AAAI Conference on Artificial Intelligence* (2017).
72. Sidky, E. Y. & Pan, X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys. Med. Biol.* **53**, 4777–4807. <https://doi.org/10.1088/0031-9155/53/17/021> (2008).
73. Dong, C., Loy, C. C., He, K. & Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 295–307 (2015).
74. Wang, L. *et al.* Temporal segment networks: towards good practices for deep action recognition. In *Computer Vision—ECCV 2016* (eds Leibe, B. *et al.*) 20–36 (Springer, Cham, 2016).
75. Paszke, A. *et al.* Automatic differentiation in PyTorch. In *NIPS-W* (2017).

## Acknowledgements

This work was supported in part by the Technology and Research Initiative Fund (TRIF) Improving Health Initiative. This work was supported in part by the National Cancer Institute (NCI) of the National Institutes of Health (NIH) Grants R21 CA134128, R01 CA195512 and R01 CA199044. The contents are solely the responsibility of the authors and do not represent the official views of the NIH or the NCI.

## Author contributions

A.B., S.V. and A.K. conceived the presented study. Z.F. carried out the experiment and developed the multi-slice residual dense network. H.T. implemented the FDK and FIRST algorithms for the work. S.V. performed the statistical analysis. S.V. wrote the Introduction and Z.F. and A.B. wrote the other sections. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-77923-0>.

**Correspondence** and requests for materials should be addressed to A.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020