



Sparse2Noise: Low-dose synchrotron X-ray tomography without high-quality reference data



Xiaoman Duan^a, Xiao Fan Ding^a, Naitao Li^a, Fang-Xiang Wu^{a,b,c}, Xiongbiao Chen^{a,c,*}, Ning Zhu^{a,d,e,**}

^a Division of Biomedical Engineering, College of Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

^b Department of Computer Science, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

^c Department of Mechanical Engineering, College of Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

^d Canadian Light Source, Saskatoon, S7N 2V3, SK, Canada

^e Department of Chemical and Biological Engineering, College of Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

ARTICLE INFO

Keywords:

Synchrotron radiation
Computed tomography
Radiation dose
3D reconstruction
Convolutional neural network

ABSTRACT

Background: Synchrotron radiation computed tomography (SR-CT) holds promise for high-resolution *in vivo* imaging. Notably, the reconstruction of SR-CT images necessitates a large set of data to be captured with sufficient photons from multiple angles, resulting in high radiation dose received by the object. Reducing the number of projections and/or photon flux is a straightforward means to lessen the radiation dose, however, compromises data completeness, thus introducing noises and artifacts. Deep learning (DL)-based supervised methods effectively denoise and remove artifacts, but they heavily depend on high-quality paired data acquired at high doses. Although algorithms exist for training without high-quality references, they struggle to effectively eliminate persistent artifacts present in real-world data.

Methods: This work presents a novel low-dose imaging strategy namely Sparse2Noise, which combines the reconstruction data from paired sparse-view CT scan (normal-flux) and full-view CT scan (low-flux) using a convolutional neural network (CNN). Sparse2Noise does not require high-quality reconstructed data as references and allows for fresh training on data with very small size. Sparse2Noise was evaluated by both simulated and experimental data.

Results: Sparse2Noise effectively reduces noise and ring artifacts while maintaining high image quality, outperforming state-of-the-art image denoising methods at same dose levels. Furthermore, Sparse2Noise produces impressive high image quality for *ex vivo* rat hindlimb imaging with the acceptable low radiation dose (i.e., 0.5 Gy with the isotropic voxel size of 26 μm).

Conclusions: This work represents a significant advance towards *in vivo* SR-CT imaging. It is noteworthy that Sparse2Noise can also be used for denoising in conventional CT and/or phase-contrast CT.

1. Introduction

Synchrotron radiation combined with computed tomography (SR-CT) is a nondestructive imaging tool to provide three-dimensional (3D) digital images of internal structure of objects. By exploiting both absorption and refraction information, SR-CT provides excellent image contrasts for low-density objects compared with other imaging methods [1]. In the field of life science, SR-CT presents unique high image quality advantages for soft tissue imaging for various purposes, such as

screening of breast cancer [2], function assessment of brain [3], heart [4], and lung [5], as well as characterization of hydrogel scaffolds for nerve regeneration [6]. Meanwhile, it is noted that the use of high-resolution SR-CT for biomedical imaging may cause the concerns related to the radiation dose received by the objects being imaged as too much radiation possibly damage or hurt the object or its host (such as animals). While it has been reported in the literature [7,8] that SR-CT imaging has reduced radiation dose as compared to conventional (polychromatic X-ray) CT imaging, the demanding for high-resolution *in*

* Corresponding author. Division of Biomedical Engineering, College of Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada.

** Corresponding author. Division of Biomedical Engineering, College of Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada.

E-mail addresses: xbc719@usask.ca (X. Chen), niz504@mail.usask.ca (N. Zhu).

vivo imaging by means of SR-CT still suffer from the issue of radiation dose given the fact that, if doubling resolution, the dose for maintaining the same signal-to-noise ratio can be increased by 16 times [9]. The radiation dose more than 6 Gy is generally considered lethal for small rodents [10]. In longitudinal studies, the object's exposure to radiation has been repeated continuously, so it is essential to keep exposure as low as possible each time of scanning the object. According to typical *in vivo* clinical micro-CT protocols with isotropic voxel size of $\sim 50 \mu\text{m}$, the tolerable radiation dose in live rodents imaging should be less than 0.5 Gy per scan [11–14]. Some strategies have been developed to lower the delivered radiation dose in high-resolution CT scans, such as those by reducing the projection number or lowering the photon flux, but can result in noises and artifacts, thus degrading the image qualities. As such, low-dose imaging is always desired, but challenged by keeping the doses as low as reasonably achievable (ALARA) without sacrificing the image quality.

Low-dose CT imaging has been an emerging research field over the past decades. Many algorithms have been developed based on deep learning (DL) [15–18] so as to accurately denoise and/or remove artifacts in the reconstruction of images while maintaining low computational costs. For example, in recently popular physics-/model-based data-driven methods, one can embed the raw data constraint into the network without the information loss in the process of image reconstruction, while still keeping the advantages of DL such as high processing efficiency, excellent performance, and great clinical potential [54]. Notably, most of these DL-based methods are trained in a supervised mode and their success critically depends on high-quality paired

data (usually acquired with relatively high dose) as the target (as shown in Fig. 1c) in training process. However, obtaining a large number of high-quality paired data can sometimes be challenging or even impossible, such as in longitudinal studies where *in vivo* biomedical imaging is essential, but limited for low-dose scans. Besides, compared with dedicated clinical CT, the imaging objects/features are much more diverse for research purposes with SR-CT imaging, making it difficult to gather a substantial amount of data for training or use as object-specific prior. Limited training data can lead to poor generalization capability of supervised learning algorithms, thus limiting their use in many real-world applications.

For low-dose imaging, it is possible to train models to infer denoised results without high-quality reference data as the target. Deep image prior (DIP, as shown in Fig. 1d) is a popular unsupervised denoising approach that doesn't require additional data [19]; however, if DIP is applied for low-dose CT denoising, it's difficult and unstable to restore detailed features and/or low-contrast objects. Besides, this type of method generates models with many hyper-parameters for 3D inference, which usually become time-consuming tasks for large-scale 3D image reconstruction. CycleGAN is another popular unsupervised algorithm for low-dose CT denoising problem and able to denoise with 4%–20% dose reduction for cardiac CT [20,21]. Nevertheless, it is still hard for CycleGAN to preserve fine structural details in the low-dose CT images for low-density objects imaging. Diffusion models allow for achieving the state-of-the-art performance in the field of image reconstruction and can be trained in unsupervised mode [22,23]. One major drawback associated with training these models is the high computational cost.

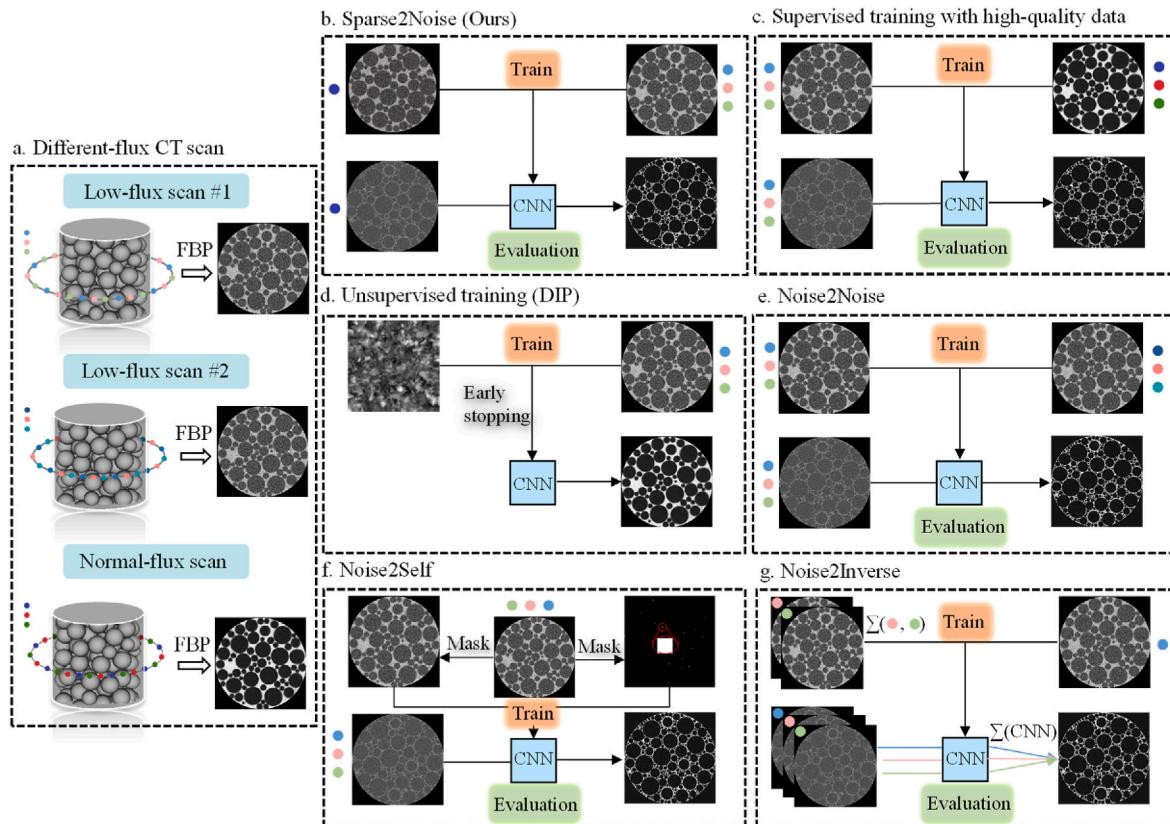


Fig. 1. Illustration of various existing deep learning methods for low-dose CT imaging. **a**, CT scan with varying intensities of photon flux, where different colors of dots represent different flux intensities, and the number of dots indicates the sparsity of the scanning view. For algorithms that use sub-reconstruction, a 1/3 scan is taken as an illustration for simplicity. **b**, Sparse2Noise (developed in this study) utilizes two low-quality images from normal-flux sparse-view CT scan and low-flux full-view CT scan. **c**, Supervised training utilizes paired CT images scanned with low-flux and normal-flux scan. **d**, DIP (one of popular unsupervised trainings) utilizes the random noise image and a low-flux scanned CT image. **e**, Noise2Noise utilizes two noisy images acquired from different low-flux CT scans. **f**, Noise2Self utilizes two masked images acquired from a same noisy image. **g**, Noise2Inverse utilizes two independent noisy images, which are reconstructed from complementary sinograms from a same CT scan.

Additionally, a considerable amount of training data is necessary to achieve an optimal performance, which can be challenging to obtain in certain applications. Another critical concern is the performance degradation of diffusion models if out-of-distribution testing samples or data are to be utilized, which is not uncommon in practical real-world scenarios. In such cases, these diffusion model-based approaches may not be generalized well, leading to suboptimal results.

It is possible to denoise for low-dose CT imaging using paired noisy images (i.e., Noise2Noise as shown in Fig. 1e) [24–27], or denoise using a single noisy image (self-supervised algorithms), such as Noise2Self (as shown in Fig. 1f) [28,29], and Noise2Inverse [30,31] (as shown in Fig. 1g). These algorithms show impressive denoising performance for low-dose CT imaging on simulation data which accord with assumed noise models. However, in many applications, data usually suffer from various artifacts and some image artifacts, which cannot be effectively removed. For SR-CT imaging, ring artifacts are commonly seen on reconstruction due to possible nonlinear and inconsistent responses of detector [32]. These ring artifacts are pronounced for low-dose SR-CT imaging especially for low-density objects/low contrast imaging, which may be destructive for quantitative analysis such as segmentation. Some algorithms reported in the literature [33,34] are able to remove ring artifacts, but have risks of introducing additional artifacts or shadows [35]. In addition, some ring artifacts removal algorithms may cause low-dose noise to be statistically coupled [36], resulting in blurring of denoising results [37]. Furthermore, some of self-supervised algorithms introduce streak artifacts themselves due to the split of projections to generate input and target sub-reconstructions for low-dose imaging.

As motivated, this work aimed to develop a novel low-dose imaging strategy that combines paired sparse-view CT scan (normal-flux) and full-view CT scan (low-flux) based on convolutional neural network (CNN), namely Sparse2Noise (as shown in Fig. 1a). For clarification, the full view of a CT scan in this paper is quantified to meet the requirements for artifact-free inversion, in accordance with the classical Shannon-Nyquist theorem for angular data sampling (i.e., number of projections per 180° for parallel beam: $N_p = \pi R$, where R is the radius of the object in units of voxels). Sparse view therefore indicates the uniform under-sampling ($N_p < \pi R$). The specific number of photons for the normal flux may be varied for different detector systems; nevertheless, a flux that yields the approximate full dynamic range of detector is typically considered as normal flux for SR-CT.

The important features or contributions of this work rest on the following aspects. First, as compared to supervised DL algorithms, the proposed Sparse2Noise does not require high-quality reconstruction as reference during training. Both normal-flux sparse-view scan and low-flux full-view scan involve relatively low radiation dose, which makes Sparse2Noise particularly useful in applications where high-dose scans are not allowed. Second, Sparse2Noise based on CNN is more robust and computationally inexpensive compared to unsupervised deep learning algorithms. Third, Sparse2Noise can effectively obtain results with less noise and reduced ring artifacts, which is highly desired for low-dose SR-CT imaging in real applications. Fourth, this work performs a comprehensive comparison of the image quality of experimental CT data processed using various state-of-the-art denoising algorithms, under varying levels of radiation dose. Notably, many published algorithms only focus on image quality comparisons without the consideration of radiation dose.

This paper is structured as follows. Section 2 (Methods) presents the theoretical framework, proposition, and implementation of proposed Sparse2Noise for low-dose CT imaging; Section 3 (Simulations and Experiments) reports our simulated and experimental studies based on Sparse2Noise, along with Section 4 (Results) as compared to other low-dose imaging strategies; Section 5 are the results for evaluating the influence of hyper-parameters in Sparse2Noise; Section 6 (Discussion) discusses some related, yet important, issues and interpretation; and Section 7 (Conclusion) gives conclusions from the present work.

2. Methods

2.1. Theoretical framework of Sparse2Noise

In SR-CT imaging, let \mathbf{y}^* , \mathbf{y} , and $\tilde{\mathbf{y}}$ (dimension of $Np \times M$, where Np is the number of projections, i.e., the number of scan views and M is the width of the detector) be the clean, normal-flux and low-flux image (i.e., sinogram) collected in transmission measurement domain, respectively. It has been noticed that ring artifacts on reconstructed images originate from angular streak/stripe noise in measurement domain, which can be modeled as an angular correlation noise $\mathbf{y}^* \odot \eta_p$ [36]. Here, \odot denotes the element-wise multiplication, $\eta_p = g \otimes v$ is a $Np \times M$ matrix (\otimes denotes 2D convolution, η_p is the subsection (central part) of convolution which has same size of g , g is an angular dependent matrix with dimension of $Np \times M$ which has a vertically constant line in the middle (value of 1, otherwise value of 0), and $v \sim \mathcal{N}(0, 1)$ is mean-zero normally distributed vector ($1 \times M$). Then the normal-flux image \mathbf{y} can be expressed as,

$$\mathbf{y} = \mathbf{y}^* + \mathbf{y}^* \odot \eta_p. \quad (1)$$

where \odot is element-wise multiplication. On this basis, low-flux image $\tilde{\mathbf{y}}$ collected with filters or reduced exposure time can be modeled by adding i. i.d. mean-zero noise ϵ ,

$$\tilde{\mathbf{y}} = \mathbf{y}^* + \mathbf{y}^* \odot \eta_p + \epsilon. \quad (2)$$

Although the noise is modeled as an additive term ϵ , we note that it also covers non-additive noise Poisson noise [30]. The angular streak/stripe noise η_p on normal-flux image and low-flux image are considered to be independent to each other if complete defects on scintillators and detectors are excluded. According to CT reconstruction algorithm, i.e., filtered back-projection algorithm (FBP), the reconstruction \mathbf{x} and $\tilde{\mathbf{x}}$ ($M \times M$) can be obtained from the measured data \mathbf{y} and $\tilde{\mathbf{y}}$ with linear reconstruction operator \mathbf{R} ($M \times Np$),

$$\mathbf{x} = \mathbf{R}\mathbf{y} = \mathbf{R}(\mathbf{y}^* + \mathbf{y}^* \odot \eta_p), \quad (3)$$

$$\tilde{\mathbf{x}} = \mathbf{R}\tilde{\mathbf{y}} = \mathbf{R}(\mathbf{y}^* + \mathbf{y}^* \odot \eta_p + \epsilon) = \mathbf{x}^* + \mathbf{R}(\mathbf{y}^* \odot \eta_p) + \mathbf{R}\epsilon. \quad (4)$$

The training task of Sparse2Noise is to determine the parameters that best enable the network f_ϕ ,

$$\hat{\phi} = \arg \min_{\phi} \|f_\phi(\mathbf{x}_S) - \tilde{\mathbf{x}}_F\|_2^2 \quad (5)$$

where $\mathbf{x}_S = \mathbf{R}_S \mathbf{y}_S$ ($S \in F$, where $F = \{S_1, S_2, \dots, S_m\}$, and m is the number of uniform subsections) is the reconstruction of normal-flux sparse-view CT scan, and $\tilde{\mathbf{x}}_F = \mathbf{R}_F \tilde{\mathbf{y}}_F$ is the reconstruction of low-flux full-view CT scan.

We denote with μ the joint measure of \mathbf{x} , η_p , and ϵ . In this case, the trained network f_ϕ obtained in Equation (5) approximates the regression function h^* ,

$$h^* = \arg \min_h \mathbb{E}_\mu \|h(\mathbf{x}_S) - \tilde{\mathbf{x}}_F\|^2. \quad (6)$$

The reason we use h^* is grounded in its convenience for analyzing expectations in Section 2.2.

2.2. Proposition in Sparse2Noise

This section will show that Sparse2Noise recovers a clean reconstruction based on a proposition of expected prediction error decomposition.

We introduce \mathbf{x}_F^* in Equation (6) and expand the squared norm,

$$\begin{aligned} \|\mathbf{h}(\mathbf{x}_S) - \tilde{\mathbf{x}}_F\|^2 &= \|\mathbf{h}(\mathbf{x}_S) - \mathbf{x}_F^* + \mathbf{x}_F^* - \tilde{\mathbf{x}}_F\|^2 \\ &= \|\mathbf{h}(\mathbf{x}_S) - \mathbf{x}_F^*\|^2 + \|\mathbf{x}_F^* - \tilde{\mathbf{x}}_F\|^2 + 2[\mathbf{h}(\mathbf{x}_S) - \mathbf{x}_F^*]^T [\mathbf{x}_F^* - \tilde{\mathbf{x}}_F]. \end{aligned} \quad (7)$$

The expectation of the third term is,

$$\mathbb{E}_\mu [h(\mathbf{x}_S) - \mathbf{x}_F^*]^T [\mathbf{x}_F^* - \tilde{\mathbf{x}}_F] = \mathbb{E}_\mu [h(\mathbf{x}_S)^T \mathbf{x}_F^* - \mathbf{x}_F^{*T} \mathbf{x}_F^* - h(\mathbf{x}_S)^T \tilde{\mathbf{x}}_F + \mathbf{x}_F^{*T} \tilde{\mathbf{x}}_F], \quad (8)$$

where $\mathbb{E}_\mu h(\mathbf{x}_S)^T \tilde{\mathbf{x}}_F$ can be expanded as,

$$\begin{aligned} \mathbb{E}_\mu h(\mathbf{x}_S)^T \tilde{\mathbf{x}}_F &= \mathbb{E}_\mu [h(\mathbf{x}_S)^T (\mathbf{x}_F^* + \mathbf{R}_F (\mathbf{y}_F^* \odot \eta_p) + \mathbf{R}_F \epsilon)] \\ &= \mathbb{E}_\mu [h(\mathbf{x}_S)^T \mathbf{x}_F^* + h(\mathbf{x}_S)^T \mathbf{R}_F (\mathbf{y}_F^* \odot \eta_p) + h(\mathbf{x}_S)^T \mathbf{R}_F \epsilon] \\ &= \mathbb{E}_\mu h(\mathbf{x}_S)^T \mathbf{x}_F^* + \mathbb{E}_\mu h(\mathbf{x}_S)^T \mathbf{R}_F \mathbb{E}_\mu \mathbf{y}_F^* \odot \eta_p + \mathbb{E}_\mu h(\mathbf{x}_S)^T \mathbb{E}_\mu \mathbf{R}_F \epsilon \end{aligned} \quad (9)$$

η_p and ϵ are both mean-zero and independent for normal-flux and low-flux scan and \mathbf{y}_F^* is a constant variable free of random noise, therefore,

$$\mathbb{E}_\mu h(\mathbf{x}_S)^T \tilde{\mathbf{x}}_F = \mathbb{E}_\mu h(\mathbf{x}_S)^T \mathbf{x}_F^*. \quad (10)$$

Likewise, we have,

$$\mathbb{E}_\mu \mathbf{x}_F^{*T} \tilde{\mathbf{x}}_F = \mathbb{E}_\mu \mathbf{x}_F^{*T} \mathbf{x}_F^*. \quad (11)$$

Then Equation (8) yields,

$$\mathbb{E}_\mu [h(\mathbf{x}_S^*) - \mathbf{x}_F^*]^T [\mathbf{x}_F^* - \tilde{\mathbf{x}}_F] = 0. \quad (12)$$

As a result, we can obtain the expected prediction error decomposition in Equation (7) as,

$$\mathbb{E}_\mu \|h(\mathbf{x}_S) - \tilde{\mathbf{x}}_F\|^2 = \mathbb{E}_\mu \|h(\mathbf{x}_S) - \mathbf{x}_F^*\|^2 + \mathbb{E}_\mu \|\mathbf{x}_F^* - \tilde{\mathbf{x}}_F\|^2. \quad (13)$$

This equation states the expected prediction error can be decomposed into the supervised prediction error and the variance of the low-flux full-view reconstruction noise error. Note that \mathbf{x}_F^* is a clean reconstruction so the inferred result from the network will be free of noises and ring artifacts. Besides, \mathbf{x}_F^* is the full-view reconstruction, which means Sparse2Noise does not suffer from down-sampling artifacts (i.e., streaking artifacts).

2.3. Implementation of Sparse2Noise

The workflow of Sparse2Noise is depicted in Fig. 2. Initially, two sinogram stacks (one from normal-flux sparse-view acquisition and the other from low-flux full-view scan) are collected and processed using standard flat-dark correction. In the case of SR-CT imaging, phase retrieval (i.e., transport of intensity equation, TIE or Paganin algorithm) [38] can be applied just after flat-dark correction if necessary. FBP reconstruction was realized using *tofu* toolkit [39] with Ram-Lak filter used, as it does not blur the reconstructions [30].

From the theoretical framework of Sparse2Noise, we notice $\|f_\phi(\mathbf{x}_S) - \mathbf{x}_F^*\|^2$ depends on f and this term is a routine supervised sparse-view CT reconstruction in image domain. In this work, the authors tested the U-Net proposed in TomoGAN¹ [40] can produce best removal performance of noise and ring artifacts for Sparse2Noise (as seen later in Section 5.1) while keeping high-frequency details. U-Net network architecture [41] has three down-sampling layers and three up-sampling layers. The main modification of U-Net network is that the input allows for a stack of d adjacent images ($d \geq 1$, $d = 3$ was utilized in this work), and then eight 1×1 convolution kernels are performed. We used the network throughout the simulated study and experimental study in this work. The loss function for generator is L1 loss, which is found to perform well for restoring high-frequency details. Alternative loss, the pixel-wise mean square error (MSE) is also tested. For Sparse2Noise, U-Net was trained using the ADAM algorithm with a mini-batch size of 64 and a learning rate of 10^{-4} . Code of Sparse2Noise is available at: <https://github.com/Xiaoman896/Sparse2Noise>.

3. Simulations and Experiments

3.1. Simulations

A simulated cylindrical foam phantom containing randomly-placed bubbles was firstly employed for demonstrating the effectiveness of proposed Sparse2Noise while comparing with other low-dose CT imaging strategies. The phantom generation, forward projection, noise addition, and FBP with Ram-Lak reconstruction with the voxel number of 512^3 were achieved using the open-source *foam_ct_phantom* package [42]. For low-flux full-view imaging in Sparse2Noise, N_p ($804 = 512/2 \times \pi$) projection images were collected with the phantom average absorption of α (20% or 50%), and initial photon count per pixel of I_0 (30). Different combinations of α and I_0 represent varied levels of Poisson noise. As for normal-flux sparse-view imaging, 50 (i.e., 1/16) projection images were acquired with I_0 of 500 and with same α . In order to evaluate the ring artifacts removal effect, streak noises were also added on sinograms and this led to ring artifacts on reconstruction results. We use the parameter of streak noise strength $\|g\|_2 = \text{std}(\eta_p)$ on sinograms to denote the intensity of ring artifacts.

For comparisons, supervised training, DIP, Noise2Inverse, Noise2Noise, and traditional iterative reconstruction (simultaneous iterative reconstruction technique, SIRT) were also assessed. For supervised training, target data were reconstructed from 804 projections with I_0 of 500, while the input data were the normal-flux sparse-view reconstruction collected in Sparse2Noise. For Noise2Inverse, we split the low-flux full-view data following 3:1 training strategy (i.e., the sum of three sub-reconstruction from 201 projections as input while the sub-reconstruction from the remaining projections as target). For Noise2Noise, the low-flux full-view data acquired in Sparse2Noise were taken as one of the noisy data and the other noisy data were generated from same imaging conditions but from a different trial. All above DL-based algorithms were performed using U-Net networks with same structures and training parameters as mentioned in Section 2.3 for avoiding additional influences. 410 slices which are randomly selected from the volume (i.e., 512 slices) are taken as the training data, 77 slices as the test data, while the remaining 25 slices as the validation data. DIP trains on a single slice at one time with original code [19] ($\sigma = 50$) and its output is acquired by early stop with iteration number of 10,000. SIRT was computed using ASTRA toolbox [43], taking the low-flux full-view data as the input with maximum iterations of 500.

3.2. Experiments

3.2.1. Resolution phantom imaging

We evaluated the developed Sparse2Noise on real-world SR-CT data of a high-density resolution phantom (Micro-CT Bar Pattern Phantoms, QRM), which allowed for both quantitative and qualitative evaluation of the image quality. The SR-CT imaging experiments were performed at the 05ID-2 beamline, Canadian Light Source (CLS), Canada. All scans in this work were performed at sample-to-detector distance of 1.5 m (i.e., SR propagation-based imaging CT with great advantage of simple implementation and fast acquisition), and photon energy of 30 keV [6]. The detector with a pixel size of 13 μm and an image depth of 16-bit was used. 900 projections with a scanning angle range of 180° were considered as the full-view scan.

The radiation dose was measured using an ionization chamber (PTW 31010, Freiburg, Germany). The adjustment of photon flux was accomplished by placing neutral density filters (NDFs) with specific thicknesses. For normal-flux imaging, no NDFs were used, resulting in a measured radiation dose rate of 971 mGy/s and an exposure time of 30 ms per projection. The resolution phantom was imaged at three different radiation dose levels: 7.2 Gy, 3.6 Gy, and 0.5 Gy, which corresponded to N_p values of 248, 124, and 18, respectively, as input for Sparse2Noise. The low-flux full-view target for Sparse2Noise was collected with NDFs of 60 mm thickness, resulting in a radiation dose rate of 133 mGy/s and

¹ Code is available at: <https://github.com/lzhengchun/TomoGAN>.

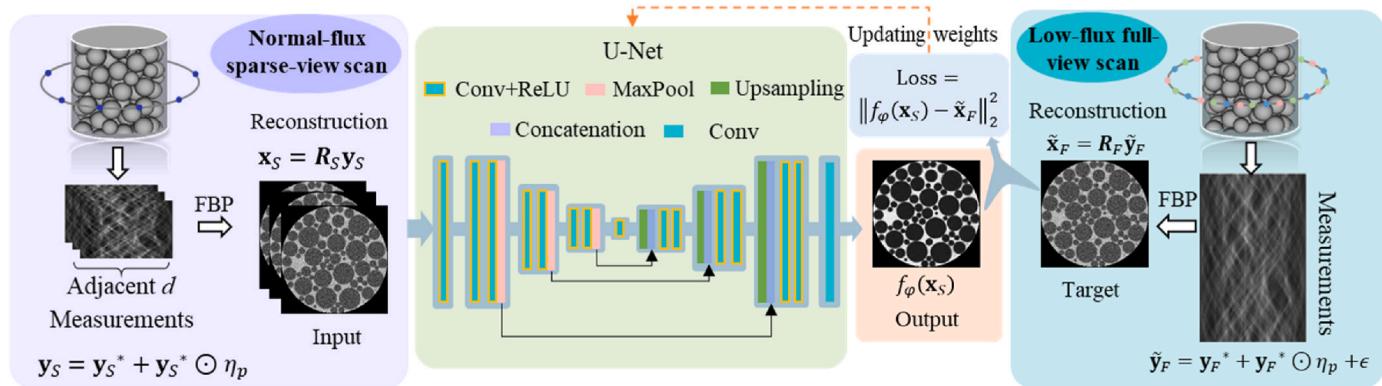


Fig. 2. Workflow of the proposed Sparse2Noise.

a total dose of 3.6 Gy. For Noise2Inverse, the resolution phantom was imaged with doses of 7.2 Gy, 3.6 Gy, and 0.5 Gy by using NDFs with 40 mm, 60 mm, and 120 mm thickness, respectively. For Noise2Noise, the input was the results obtained from scanning with 7.2 Gy, 3.6 Gy, and 0.5 Gy using different thicknesses of NDFs, while the target was the same as that used for Sparse2Noise. SIRT was reconstructed from low-flux full-view results with a maximum of 500 iterations. The dataset consisted of a total of 512 paired slices of 512×512 pixels for each dose level, with 410 slices used as the training dataset, 76 slices as the test dataset, and 26 slices as the validation dataset.

3.2.2. Low-density hydrogel tissue scaffolds imaging

We evaluated Sparse2Noise on experimental SR-CT data of hydrogel tissue scaffolds *in vitro* with same imaging setup and parameters to the previous section. Hydrogel scaffolds made from biomaterials (e.g., alginate, gelatin, chitosan) are used to support and facilitate cell growth and tissue regeneration, and therefore are essential to tissue engineering and regenerative medicine applications [44–47]. Hydrogel scaffolds have low density and SR-CT shows unique advantages for hydrogel scaffolds visualization and/or characterization [1,6,48], especially for combining with phase retrieval. The samples used in this study were prepared by following the procedure developed in our previous study [6, 35]. For DL-based algorithms, the dataset consisting of reconstructions of 4 scaffolds (600 slices with 512×512 pixels (region of interest, ROI) for each scaffold) were generated. Among them, 3 scaffolds were taken as the training data (1740 slices) and validation data (60 slices), while 1 scaffold was the test data.

3.2.3. Ex vivo rat hindlimb imaging

We also assessed Sparse2Noise with low-dose SR-CT imaging data of rat hindlimb (male Sprague–Dawley rat) with hydrogel alginate scaffold implanted (for helping sciatic nerve regeneration). The implantation surgery and euthanasia of rats (2 weeks after the implantation surgery) were performed [6] after the approval of the University of Saskatchewan Committee on Animal Care and Supply. SR-CT imaging of *ex vivo* rat hindlimb was performed with binning 2×2 mode, i.e., the pixel size of 26 μm with image size of 1024×1024 , and the corresponding exposure time per projection was reduced to 5 ms. N_p of 1500 projections with a scanning angle range of 180° were considered as the full-view scan for this case. The *ex vivo* rat hindlimb was imaged at radiation dose of 1.8 Gy and 0.5 Gy, which corresponded to N_p of 371, and 103, respectively, as input for Sparse2Noise. For Noise2Inverse and Noise2Noise, the resolution phantom was imaged with doses of 1.8 Gy and 0.5 Gy by using NDFs with 40 mm and 80 mm thickness, respectively.

4. Results

4.1. Simulated results

Examples of reconstructed results without and with ring artifacts were processed with different algorithms for improving the image quality, as displayed in Fig. 3. The first row in Fig. 3 shows the generated images: the ground truth (i.e., reconstructed from clean projections with N_p of 804) for image quality comparisons, normal-flux sparse-view images without ($\|g\|_2 = 0$) and with ring artifacts ($\|g\|_2 = 0.5$), and low-flux full-view images without and with ring artifacts. The second and third rows in Fig. 3 display the output of six algorithms. Overall, six algorithms can improve the image quality by removing streaking artifacts (i.e., sparse-view artifacts) and/or eliminating noises. Specifically, supervised training algorithm, Noise2Inverse, Noise2Noise, and Sparse2Noise outperform iterative reconstruction SIRT and unsupervised training DIP in terms of removing artifacts/noises. For images without ring artifacts, supervised training and Sparse2Noise can produce the best results which are closest to ground truth while Noise2Inverse and Noise2Noise fail to recover several fine details, as indicated by the yellow circles. Additionally, Noise2Inverse introduces streaking artifacts (i.e., under-sampling artifacts) at the edges due to the split operation. For images with ring artifacts, the supervised training and Sparse2Noise can produce high-quality images that are quite similar to results without ring artifacts. While Noise2Noise can remove most ring artifacts, some artifacts still remain as red arrow indicates. Moreover, the fine details become more difficult to restore compared to results free of ring artifacts. For Noise2Inverse, noticeable ring artifacts remain, and under-sampling artifacts are severe.

The processed results using different algorithms with various imaging conditions of $\|g\|_2$, α , and I_0 , were assessed by peak signal to noise ratio (PSNR) and structural similarity index measurement (SSIM) [49], as displayed in Table 1. From the table, supervised training can always attain the highest metrics for all cases and therefore can be served as references. SIRT and DIP are always inferior to Noise2Inverse, Noise2Noise, and Sparse2Noise for all imaging conditions. Furthermore, it's worth noting that DIP is significantly slower than the other methods (approximately 200 times slower for a CT dataset with 512 slices) due to fresh training on a single image at a time. Noise2Inverse produces close PSNR and SSIM to that of Noise2Noise for results without ring artifacts; however, its performance becomes poor for results with ring artifacts. The image quality of Noise2Noise degrades as artifact levels increase, indicating limited tolerance for ring artifacts. Sparse2Noise shows very close value of both metrics for different ring artifacts levels.

4.2. Experimental results

4.2.1. Resolution phantom imaging

The imaging results of the resolution phantom with different levels of

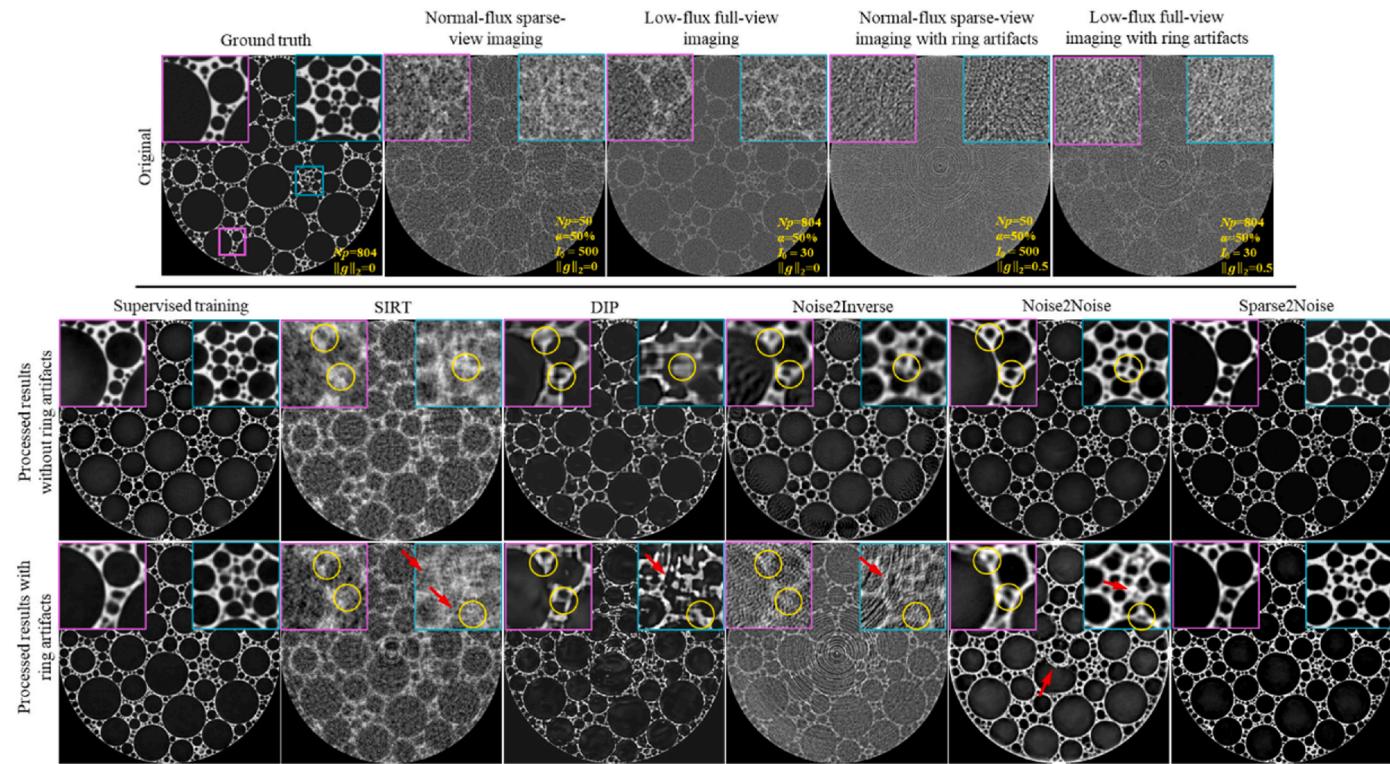


Fig. 3. Generated reconstruction images on simulated foam phantoms without and with ring artifacts and corresponding outputs of supervised training, SIRT, DIP, Noise2Inverse, Noise2Noise, and proposed Sparse2Noise. Regions in yellow circles show the details comparison and red arrows indicate the remaining ring artifacts.

Table 1

Comparison of PSNR and SSIM metrics for supervised training, SIRT, DIP, Noise2Inverse, Noise2Noise, and the proposed Sparse2Noise on the simulated foam phantom (test dataset) with different imaging conditions of $\|g\|_2$, α , and I_0 . Font in bold is used to emphasize the best metrics excluding supervised training.

$\ g\ _2$	α	I_0	Metrics	Supervised training (avg. \pm std.)	SIRT (avg. \pm std.)	DIP (avg. \pm std.)	Noise2Inverse (avg. \pm std.)	Noise2Noise (avg. \pm std.)	Sparse2Noise (avg. \pm std.)
0	20%	30	PSNR	23.27(\pm 0.30)	16.80(\pm 0.21)	17.45(\pm 0.13)	19.17(\pm 0.50)	19.63 (\pm 0.49)	23.21(\pm0.48)
			SSIM	0.88(\pm 0.04)	0.40(\pm 0.08)	0.73(\pm 0.41)	0.79(\pm 0.17)	0.79 (\pm 0.15)	0.86(\pm0.12)
—	50%	—	PSNR	25.66(\pm 0.21)	16.89(\pm0.23)	18.75(\pm 1.87)	19.13(\pm 0.48)	20.33 (\pm 0.45)	25.60(\pm0.41)
			SSIM	0.91(\pm 0.02)	0.46(\pm 0.08)	0.75(\pm 0.38)	0.80(\pm 0.18)	0.83 (\pm 0.11)	0.89(\pm0.07)
0.1	50%	30	PSNR	22.19(\pm 0.23)	14.28(\pm 0.45)	13.06(\pm 2.27)	14.51(\pm 0.49)	18.22 (\pm 0.48)	21.06(\pm0.41)
			SSIM	0.83(\pm 0.03)	0.39(\pm 0.09)	0.49(\pm 0.32)	0.46(\pm 0.27)	0.76 (\pm 0.15)	0.83(\pm0.07)
0.3			PSNR	22.20(\pm 0.21)	13.02(\pm 0.43)	12.77(\pm 2.30)	13.95(\pm 0.51)	18.09 (\pm 0.47)	21.35(\pm0.42)
			SSIM	0.84(\pm 0.03)	0.36(\pm 0.09)	0.48(\pm 0.29)	0.42(\pm 0.25)	0.75 (\pm 0.14)	0.84(\pm0.07)
0.5			PSNR	22.22(\pm 0.25)	12.19(\pm 0.48)	11.96(\pm 2.24)	13.34(\pm 0.53)	17.18 (\pm 0.47)	21.29(\pm0.41)
			SSIM	0.84(\pm 0.03)	0.31(\pm 0.09)	0.46(\pm 0.24)	0.38(\pm 0.24)	0.74 (\pm 0.13)	0.83(\pm0.07)

radiation dose and the corresponding processed outputs with different algorithms are presented in Fig. 4. It is observed that low-flux full-view images exhibit ring artifacts and noise. Sparse-view artifacts are visible in normal-flux sparse-view images, leading to distortion of the resolution, particularly for a dose of 0.5 Gy. SIRT shows low noise but also causes blurring. Noise2Inverse significantly reduces noise without blurring but still suffers from noticeable ring artifacts. Noise2Noise can remove ring artifacts to some extent for doses of 7.2 Gy and 3.6 Gy, but the effect becomes limited for a dose of 0.5 Gy. In contrast, no significant ring artifacts or sparse-view artifacts are observed in the results obtained with Sparse2Noise for all three radiation dose levels. Profiles along the red lines in the 0.5 Gy images show that the results obtained using Sparse2Noise match well with the reference (red dotted lines), which was reconstructed from normal-flux full-view imaging results of ~ 26 Gy.

Results obtained with dose levels lower than 0.5 Gy and the results processed by Sparse2Noise are given in Supplementary Material Fig. S-1.

Fig. 5 shows the modulation transfer function (MTF) [50] values for

the spatial resolution analysis, calculated based on the contrast difference of resolutions. Sparse2Noise outperformed other imaging conditions for all three different radiation dose levels, with the highest MTF amplitudes with respect to frequency (Lp/mm).

4.2.2. Low-density hydrogel tissue scaffolds imaging

Processed results before phase retrieval with different algorithms and the corresponding grey value profiles (only displayed for results of 0.5 Gy) are given in Fig. 6. As shown in Fig. 6, reducing the delivered radiation dose by lowering the photon flux will reduce the intensity of scaffolds and increase the noise and ring artifacts. For the radiation dose lower than 3.6 Gy, the ring artifacts are serious for low-flux full-view imaging and disrupt the continuity of the strand greyscale values in the image and thus impair the capacity to accurately visualize/analyze strand properties. Ring artifacts are less prominent on reconstruction images using normal-flux sparse-view imaging strategy; however, the images suffer from under-sampling artifacts. It can be observed that the outcomes obtained using SIRT and Noise2Inverse exhibit a low level of

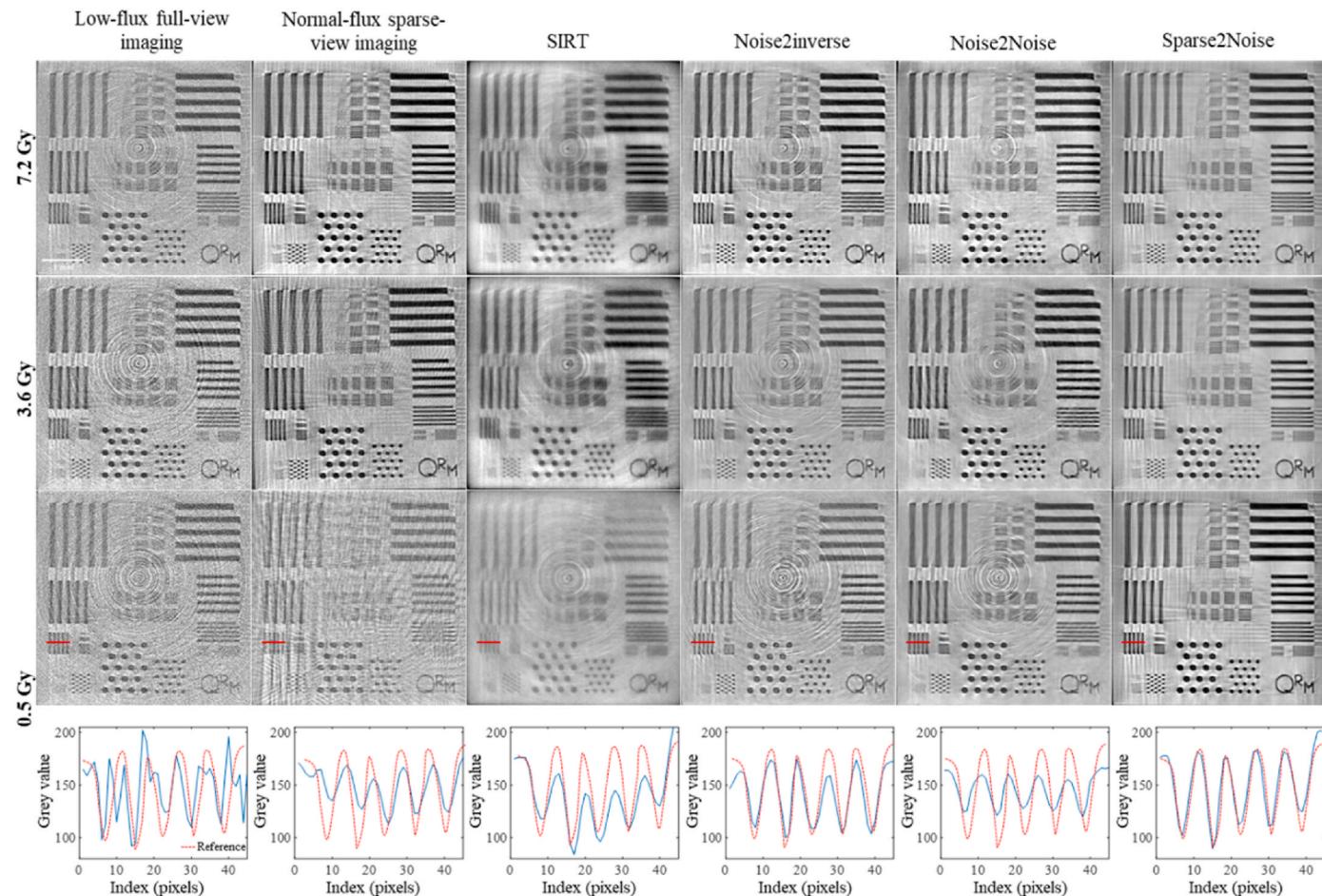


Fig. 4. Imaging results of resolution phantom using SR-CT with different levels of radiation dose and corresponding processed results with various algorithms for improving the image quality. The bottom row displays the grey value profiles of different results measured at the red line positions for dose of 0.5 Gy. Red dotted line indicates the profiles of reference images that was obtained with normal-flux full-view imaging at dose of 26 Gy. Scale bar: 1 mm.

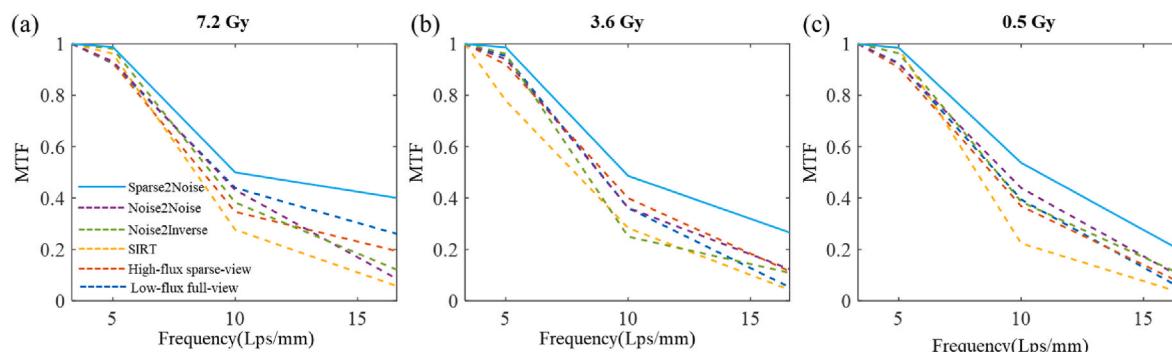


Fig. 5. MTF amplitudes with respect to frequency of low-dose imaging results obtained under different conditions.

noise but are still plagued by prominent ring artifacts. In contrast, Noise2Noise demonstrates significant noise and ring artifact reduction; however, the efficacy of Noise2Noise diminishes at a dose of 0.5 Gy as compared to 7.2 Gy and 3.6 Gy. The image characteristics, such as the contours of the scaffold microstructure, obtained using Sparse2Noise are more easily distinguishable, and the reconstruction remains free of noise and ring artifacts for all three dose levels.

Although the processed results of scaffolds above demonstrate reduced noise and fewer artifacts for low-dose SR-CT imaging, the strands of the scaffolds still exhibit similar grey values to that of the background. As a result, it remains challenging to perform quantitative

analysis, such as segmentation. In this regard, phase retrieval (i.e., TIE) can convert the phase contrast effects to equivalent absorption contrast effects, therefore allow for a distinction between strands region and background. Results after phase retrieval with different algorithms and the segmentation (only displayed for results of 0.5 Gy) by using Biomedisa [51] are given in Fig. 7. As the radiation dose decreases, the noise and ring artifacts become more noticeable and adversely impact quantitative analysis. The results indicate that while SIRT and Noise2Inverse can eliminate low-dose noise, they also introduce blurring. Additionally, these algorithms are unable to effectively eliminate ring artifacts and, in some cases, can even worsen them, particularly at a dose

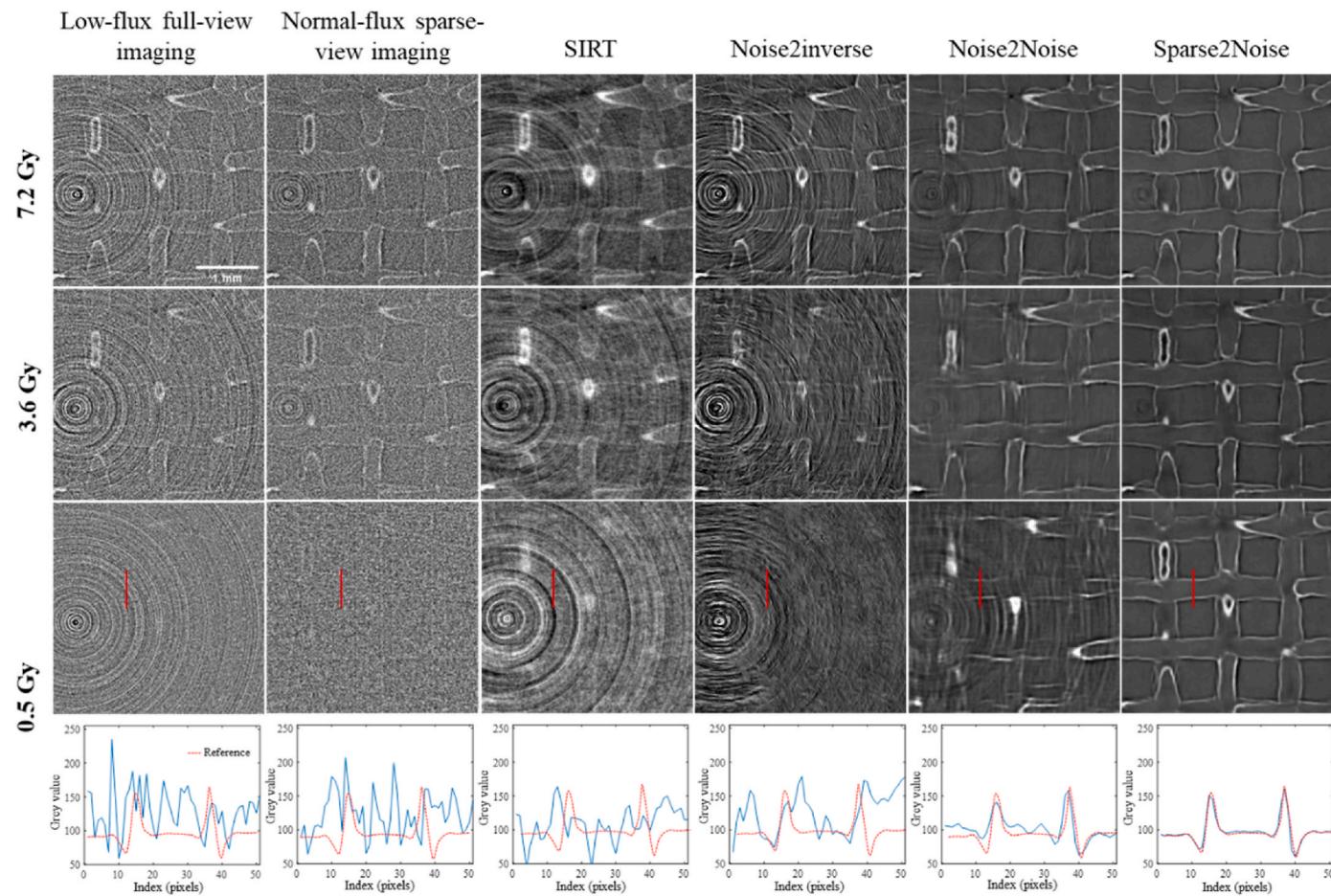


Fig. 6. *In vitro* imaging results of hydrogel scaffolds before phase retrieval using SR-CT with different levels of radiation doses and corresponding processed results using different algorithms. The bottom row displays the grey value profiles of different results measured at the red line positions for dose of 0.5 Gy. Red dotted line indicates the profiles of reference images (26 Gy).

of 0.5 Gy. Although Noise2Noise reduces both ring artifacts and noise, it demonstrates limited improvement at a dose of 0.5 Gy. On the other hand, Sparse2Noise exhibits a substantially improved image quality with reduced noise and ring artifacts across all three dose levels.

For quantitative analysis, Dice score [52], one of most popular performance metrics in medical image segmentation, was measured,

$$Dice(A, B) = \frac{2|A \bullet B|}{|A| + |B|} \quad (14)$$

The binary vectors A and B are defined as having values of 1 for elements belonging to the scaffolds' strands and 0 otherwise. In this study, A is the segmented reference obtained using a radiation dose of 26 Gy (i.e., normal-flux full-view scan), while B represents the measured vectors obtained from low-dose results or the corresponding processed results.

4.2.3. Ex vivo rat hindlimb imaging

Fig. 8 illustrates the results of *ex vivo* rat hindlimb imaging and corresponding corrected images obtained using various methods. For a dose of 1.8 Gy, both low-flux full-view imaging and normal-flux sparse-view imaging demonstrate limited visualization of muscle tissue and the implanted hydrogel scaffold with noise and ring artifacts. While SIRT, Noise2Inverse, and Noise2Noise significantly improve the image quality, they still suffer from ring artifacts. In contrast, Sparse2Noise exhibits minimal ring artifacts, and the implanted hydrogel scaffold is more evident. For a dose of 0.5 Gy, muscle tissue and the implanted hydrogel scaffold are entirely disrupted by noise and ring artifacts for both low-flux full-view imaging and normal-flux sparse-view imaging. They are

still challenging to identify even after processing using SIRT and Noise2Inverse. While Noise2Noise shows significant improvement in muscle tissue, the implanted hydrogel scaffold is still disrupted by ring artifacts. The image processed by Sparse2Noise shows the highest image quality and is comparable to the processed results of the dose of 1.8 Gy.

5. Hyper-parameters evaluations

We examined the impact of various factors, including the neural network and loss function, the size of the dataset for fresh training, and sparsity strategy for a given radiation dose level, on the effectiveness of Sparse2Noise. To conduct these tests, we employed the same foam phantom as in Section 3.1, and acquired noisy projection data with N_p of 804, α of 50% and I_0 of 30.

5.1. Neural networks and loss function

Mixed-scale dense (MS-D) network [30], DD-Net [53], and U-Net used in work [40] have been tested and evaluated the image quality with respect to epochs. The MS-D network has 100 single-channel intermediate layers with total 45,652 trainable network parameters. DD-Net² is one of popular networks tailored to sparse-view CT reconstruction, which employs DenseNet to increase the size of neural network for more feature extraction maps and applies deconvolution for better image

² Code is available at: https://github.com/zzc623/DD_Net.

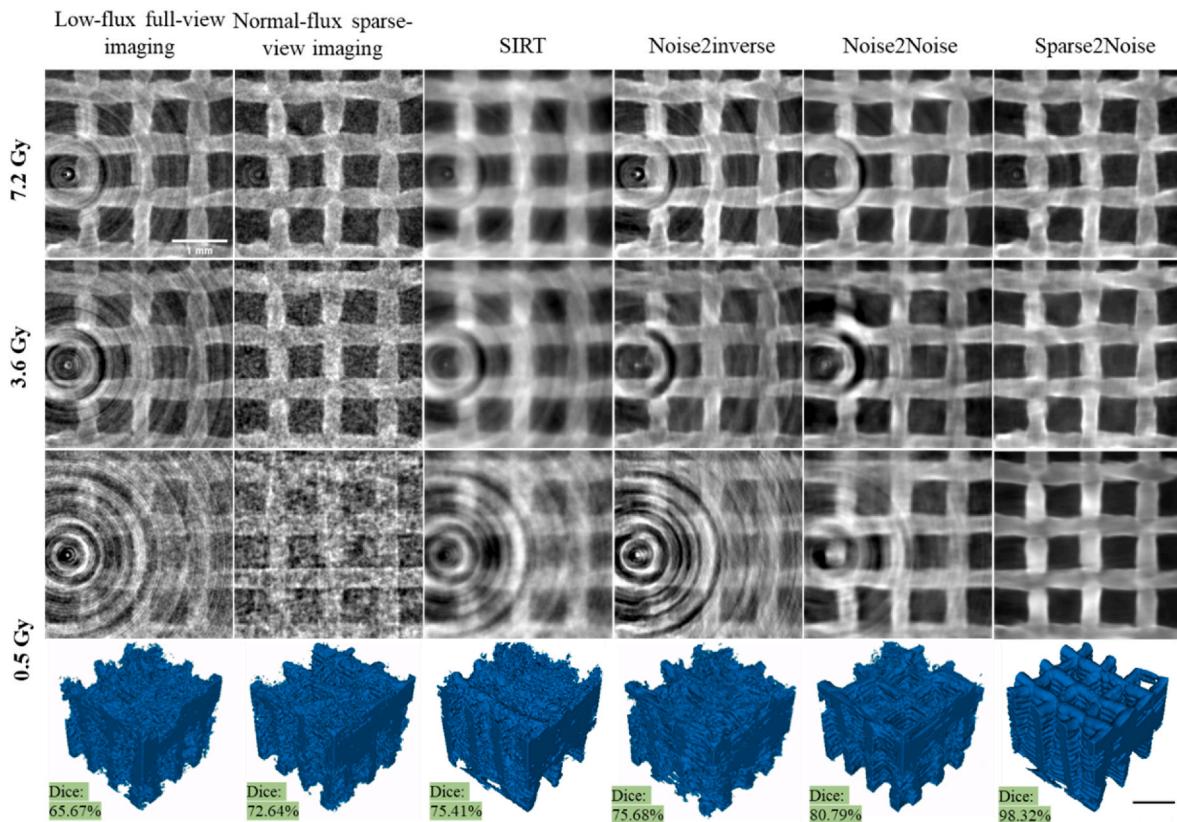


Fig. 7. *In vitro* imaging results of hydrogel scaffolds after phase retrieval using SR-CT with different levels of radiation doses and corresponding processed results using different algorithms. The bottom row displays the segmentation results for dose of 0.5 Gy. Scale bar: 1 mm.

restoration simultaneously. This work utilized 45 layers in total, including 4 Dense blocks (4 convolution layers for each block). This architecture resulted in 26,832,112 trainable network parameters in total. For MS-D and DD-Net, same batch size, patch size, optimization algorithm, and learning rate that were used as U-Net for comparisons. U-Net utilized in this work has 944,097 trainable network parameters in total.

The previous experiment was repeated with Sparse2Noise on the dataset containing input reconstructed from 50 projections with α of 50% and I_0 of 500. The networks were trained for 2000 epochs. The results of metrics of PSNR are displayed in Fig. 9. The figure shows that the U-Net network with L1 loss can achieve overall highest performance in improving image quality. In addition, the figure shows that the PSNR of MS-D with L1 loss approaches that of MS-D with MSE loss, and both become steady after 1000 epochs. Besides, DD-Net with L1 loss produces worse PSNR after 1000 epochs, likely due to the DD-Net starting to fit the noises. By contrast, the PSNR of the U-Net network continued to increase.

5.2. Sparsity strategy

The performance of the Sparse2Noise was also evaluated quantitatively with respect to the number of subsections for input data (i.e., sparsity, m) while keeping same radiation dose in each group (i.e., same value of $N_p \times I_0$). Experiments were presented with m of 2, 4, 8, 16, 32, 64, and 128 with three radiation dose levels. Results were also compared with Noise2Inverse and Noise2Noise with same radiation dose for each subsection. The corresponding PSNR metrics are displayed in Fig. 10.

In general, when the same dose of input is used, a smaller m value results in a higher peak signal-to-noise ratio (PSNR). Specifically, m values smaller than 32 are found to yield considerably better results than both Noise2Noise and Noise2Inverse for normal-flux imaging. For low-

flux and ultra-low-flux imaging with the same radiation dose, m values smaller than 16 are recommended for optimal results. The figure demonstrates that despite the reduced input flux for low- and ultra-low-flux imaging, Sparse2Noise can still produce higher image quality than Noise2Noise and Noise2Inverse, as demonstrated in Fig. 10b. However, for ultra-low-flux inputs, the improvement is limited, as shown Fig. 10c.

5.3. Sizes of dataset for fresh training

Fresh training with Sparse2Noise is similar to the convenience of conventional reconstruction methods, as it is much more flexible and can be used for processing unseen (for training model) samples. The performance of Sparse2Noise in terms of PSNR metrics was tested with respect to epochs for different sizes of training datasets (i.e., 10, 20, 30, and 40 slices of foam data), as shown in Fig. 11. The results indicate that when the dataset size is smaller than 30 slices, the PSNR reaches its peak value around 600 epochs, but then rapidly decreases as the number of epochs increases. However, when the dataset exceeds 30 slices, the PSNR becomes steady after 600 epochs for 30 slices and continues to increase for 40 slices. Therefore, Sparse2Noise allows for fresh training on a limited dataset size to improve image quality in low-dose scans, which is an efficient and effective approach.

6. Discussion

In this work, we developed a novel low-dose imaging strategy named Sparse2Noise to achieve high image quality and low radiation dose. Sparse2Noise combines paired normal-flux sparse-view CT scan and low-flux full-view CT scan (both are low-dose data) based on CNN. Although paired data are also required in Sparse2Noise, it's relatively accessible to collect two paired low-dose data rather than high-quality references for SR-CT. By taking the normal-flux sparse-view CT

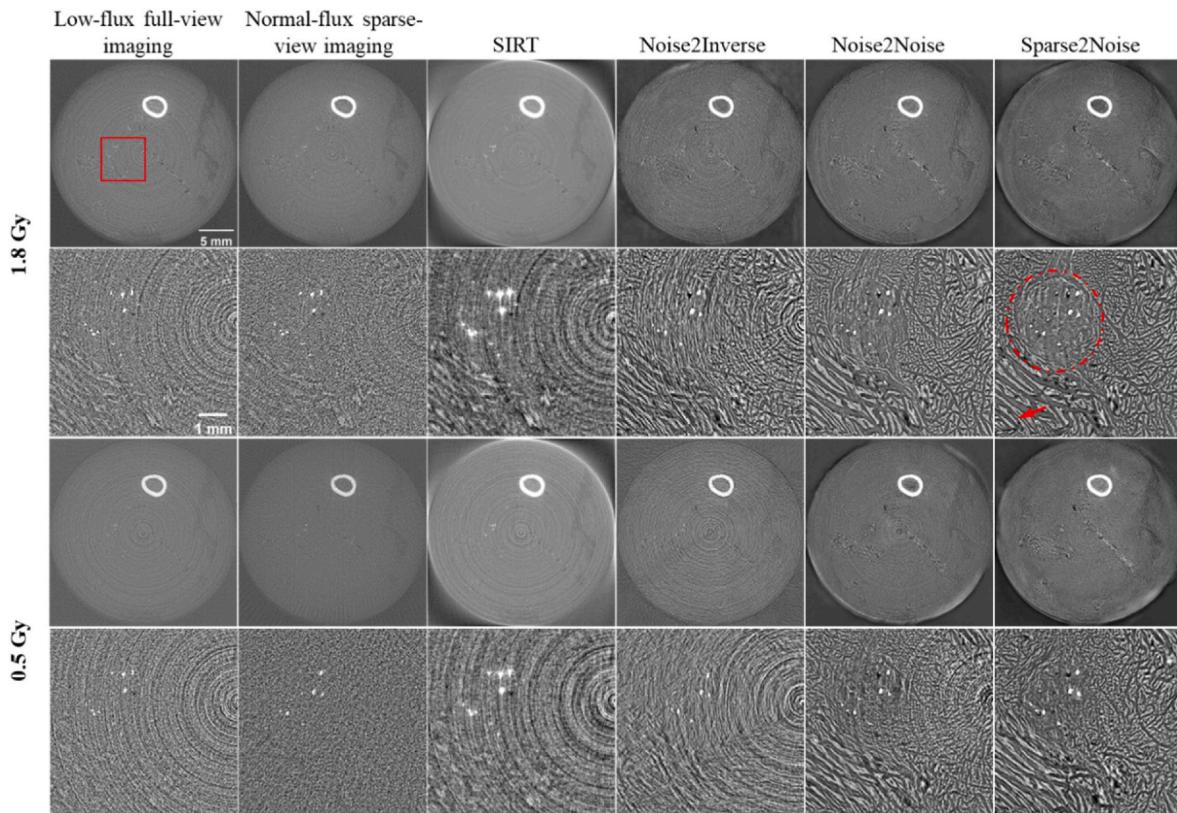


Fig. 8. *Ex vivo* imaging results of rat hindlimb using SR-CT with different levels of radiation doses and corresponding processed results using different algorithms. The images in second and fourth rows are the enlarged ROIs at position of red rectangle for each case. The region indicated by red dotted circle is the implanted hydrogel tissue scaffold and the area marked by red arrow is the muscle tissue.

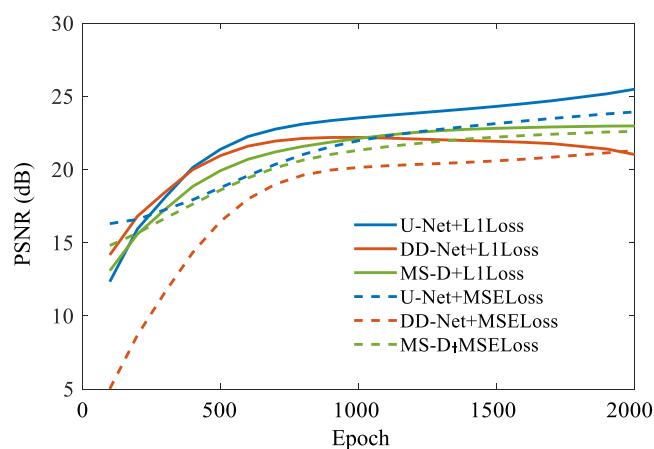


Fig. 9. Comparison of the PSNR metric with respect to epoch using Sparse2Noise for three different networks (U-Net, DD-Net, and MS-D) and two loss functions (L1 loss and MSE loss).

reconstruction as the input while low-flux full-view CT reconstruction as the target for training, we are utilizing the physical information, i.e., measured noise is element-wise independent and the streak noise on projections resulted in ring artifacts are mutually independent for both normal-flux and low-flux measurements. This approach enables us to achieve high-quality CT results by exploiting the inherent physics of the same imaging objects under varying imaging conditions. Sparse2Noise holds significant relevance for practical applications in SR-CT, particularly when dealing with new scenarios and/or new data. It should be noticed that the SR-CT (phase-contrast) data is quite different from the

clinical CT data (absorption contrast) in terms of image resolution, noise, and artifacts distribution. For example, in SR-CT ring artifacts are more common and severe, as compared with clinical CT, due to high resolution imaging that may enlarge the influence of inconsistency response of the detector. Therefore, the techniques or methods used for clinical CT data may not directly apply to SR-CT data. While it is feasible to construct a large model using an extensive dataset of retrospective clinical data and subsequently apply transfer learning (both feature-based and model-based) to actual SR-CT data, this approach can prove to be complex and may not always yield successful outcomes when realistic data is involved. Most importantly, for stable and accurate output, it still necessitates limited high-quality reconstructed data as labels during transfer learning. Obtaining such high-quality data during actual imaging is sometimes challenging or even impossible since it involves high-dose scanning when employing SR-CT for high-resolution imaging purposes. The method presented in this paper not only achieves denoising but also effectively removes ring artifacts that are severe for low-dose SR-CT. Although a denoising model could potentially be trained using retrospective data with low-dose images and high-quality reconstructed data, there is a lack of corresponding retrospective data with realistic ring artifacts. Most supervised learning methods for artifact removal rely on simulated ring artifacts, which may not accurately capture the unique characteristics of ring artifacts in low-dose SR-CT, particularly in imaging low-density objects. Sparse2Noise can be trained with experimental data that includes both ring artifacts and noise as the input and target. This ensures that Sparse2Noise is specifically tailored to address the challenges associated with low-dose SR-CT imaging. This approach allows the model to learn from real-world scenarios, making it more robust and capable of producing high-quality reconstructions that closely resemble the original undistorted images. Sparse2Noise fills the under-sampling artifacts (process to learn signals on images) in input but does not learn the low-dose

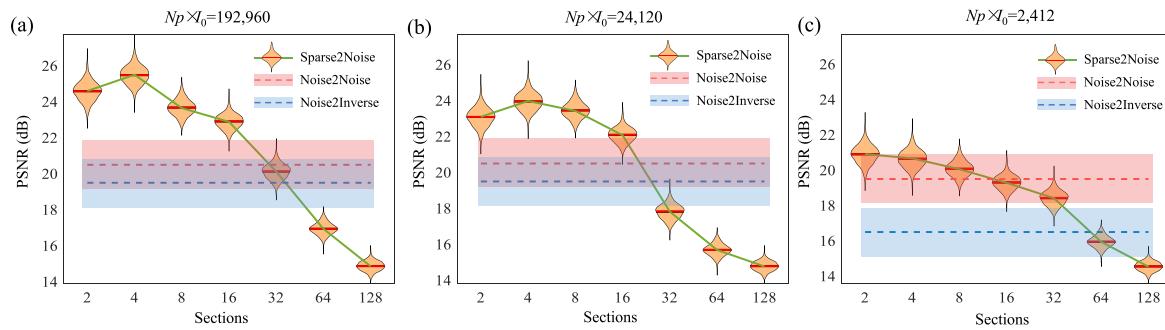


Fig. 10. Comparison of the PSNR metric with respect to the number of subsections using Sparse2Noise for three different dose levels and comparisons with Noise2Noise and Noise2Inverse. **a** Normal-flux with total initial photon count per pixel of 192,960. **b** low-flux with total initial photon count per pixel of 24,120. **c** Ultra-low-flux with total initial photon count per pixel of 2412.

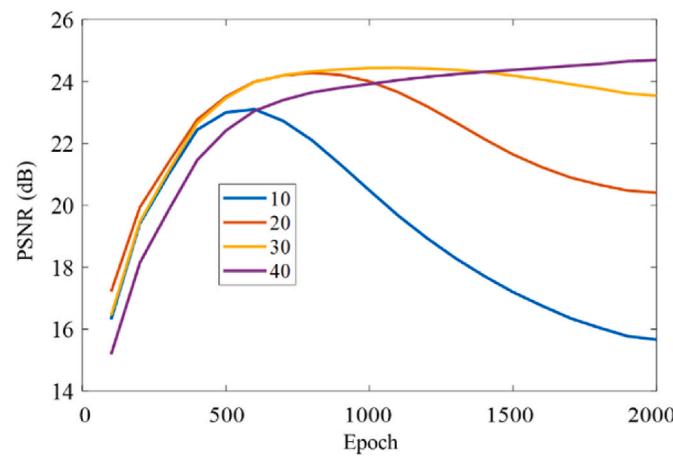


Fig. 11. Comparison of the PSNR metric with respect to epochs using Sparse2Noise for different sizes of training dataset.

noise, which employs similar idea to DIP, which utilizes network parameters with high noise impedance and low signal impedance. Sparse2Noise demonstrates greater robustness and faster convergence than DIP by using under-sampling reconstruction as input instead of random noise, which is particularly important for real-world experimental applications. Additionally, compared with unsupervised training algorithms, Sparse2Noise utilizes a CNN architecture (i.e., U-Net) that is efficient and computationally inexpensive, making it a practical and accessible solution for low-dose CT imaging. In addition to DIP, a popular unsupervised method, we also compared the results of Sparse2Noise with a very recent unsupervised method based on the diffusion model, as discussed in Supplementary Material Fig. S-3.

Our theoretical rationale illustrates the Sparse2Noise will converge to the expectation of full-view clean reconstruction that is free of noise and ring artifacts, therefore no additional artifacts will be introduced. The U-Net proposed in TomoGAN with L1 loss can achieve overall highest performance in terms of image quality with fast training speed (Fig. 9). In our study, we specifically chose $d = 3$ as adjacent reconstructed slices as the input for training because it yielded the best image quality, as depicted in Supplementary Material Fig. S-4. It is important to note that TomoGAN may not be suitable for the task at hand, as the adversarial networks used in the model can be interpreted as a trainable loss function. Since GANs attempt to predict noise or artifact-like textures and structures, they may not perform well in this particular scenario. If the Sparse2Noise conditions are not met, e.g., the flux of input is not sufficient, then the output of the trained model can be sub-optimal but still effective for reducing the noise and produce higher image quality compared with Noise2Inverse and Noise2Noise with same

radiation dose (Fig. 10) with m smaller than 16. It's important to note that these results only indicate the trend in image quality with respect to section number. For real-world applications, one should choose an appropriate m value based on the specific imaging object being processed. This could involve selecting a value either smaller or larger than 16. Sparse2Noise allows for fresh training on a limited dataset (larger than 30 slices) for improving the image quality with low-dose scan (Fig. 11). Fresh training can be highly beneficial for processing unseen samples as it is much more flexible and efficient.

For validating the feasibility of Sparse2Noise, both simulated and experimental CT data were employed. For simulation studies, although high-quality data were not used in Sparse2Noise, it was observed that Sparse2Noise produced high-quality results that are comparable to traditional supervised DL and outperform to other popular low-dose imaging algorithms. Noise2Inverse cannot remove ring artifacts since the artifacts arising from the sub-sampling data are correlated. Low-dose CT imaging applications require the total projections as few as possible, which is not suitable for the use of the Noise2Inverse algorithm due to the even less sub-sampling data. For experimental studies, we notice Noise2Noise has limited ring removal effect for low-dose imaging condition at 0.5 Gy in real-world applications. In low dose condition, streak noise on two paired noisy measurements (input and target) may not be statistically independent. Also, the application of phase retrieval on the Noise2Noise data can enlarge those streak noises on projections, therefore progressively corrupting statistically independency. However, all those effects don't appear in the results of Sparse2Noise since Sparse2Noise takes the advantage of the ring artifacts are not pronounced on normal-flux reconstruction, despite sparse-view imaging.

Apart from sparse-view scan mode, limited-angle CT scan mode (spanning angle less than 180°) is another way to reduce the radiation dose. Limited-angle imaging results in streaking artifacts (i.e., missing wedge artifacts) on reconstruction images. The author extended the feasibility of the similar idea with Sparse2Noise but taking the normal-flux limited-angle scan projections as the input while the low-flux full-view scan is still the target (namely Limited2Noise). Results show Limited2Noise (Fig. S-2. in Supplementary Material) can restore the image quality for limited-angle scan with 22.5° and 45° by a large margin. Limited2Noise provides greater flexibility for acquiring low-dose projections. For instance, in cases where there are limitations on imaging angles, projections can be obtained without following the uniform sampling principle required by Sparse2Noise.

It is worthwhile to note that for some applications, Sparse2Noise may have limitations. For Sparse2Noise, it is assumed that the measured noise is element-wise independent and mean-zero and that the streak noise on projections resulted in ring artifacts are mutually independent for both normal-flux and low-flux measurements. Based on this we remark that, similar to Noise2Inverse and Noise2Noise, Sparse2Noise is able to eliminate the ring artifacts from inconsistent response of detector, but not able to completely remove the ring artifacts that result from

the defects on scintillators and detectors since these artifacts exist on both normal-flux and low-flux measurements statically. Besides, in comparison to recently popular physics-/model-based data-driven methods [54,55] that integrate raw measured data into the training process, Sparse2Noise only employs the reconstructed data, which may loss some details from raw measured data. These details are often challenging to be reproduced solely through the utilization of a CNN without raw data constraint into the network.

For improvement, integration of a physics-based strategy into Sparse2Noise would be promising. By integrating physics-based constraints (e.g., taking the CT raw data as input and revising the architecture of network accordingly), Sparse2Noise can effectively incorporate prior knowledge into the reconstruction process and help to preserve important image features, thus enhancing the accuracy of the reconstructed images and reducing the potential loss of information that may occur in purely data-driven approaches. Another promising avenue for future research is the exploration of reinforcement learning techniques within the framework of Sparse2Noise to further enhance low-dose SR-CT imaging. Reinforcement learning offers the potential to optimize the denoising and artifact removal process by learning from interactions with the imaging environment. By training a reinforcement learning agent to make intelligent decisions on how to denoise and remove artifacts, Sparse2Noise can adaptively improve its performance over time.

7. Conclusions

SR-CT can be of a nondestructive imaging tool with excellent image contrasts and shows great potential for *in vivo* imaging. The main concern for *in vivo* imaging using SR-CT is the radiation dose verse image quality trade-off, as simply reducing the radiation dose can lead to noises and various artifacts, thus degrading the image quality. This paper presents the development of Sparse2Noise for improving low-dose image quality of SR-CT without the need for high-quality reference data based on CNN. Sparse2Noise utilizes the reconstruction data from paired sparse-view CT scan (normal-flux) and full-view CT scan (low-flux), both of which are low-dose imaging results. Sparse2Noise can fill the undersampling artifacts in input but does not learn the low-dose noise, which utilizes network parameters in CNN with high noise impedance and low signal impedance. Our simulated results based on a foam phantom demonstrate that Sparse2Noise outperforms the traditional iterative reconstruction (SIRT), unsupervised learning method (DIP), self-supervised method (i.e., Noise2Inverse), as well as Noise2Noise in terms of PSNR and SSIM metrics, with or without ring artifacts present in the images. Our experimental results on a resolution phantom, hydrogel tissue scaffolds, and *ex vivo* rat hindlimb, show that Sparse2Noise can be effectively applied to low-dose SR-CT imaging of varying dose levels, with results or data that have less noise and fewer ring artifacts (even for dose of 0.5 Gy with isotropic voxel size of 13 μm and 26 μm) than other state-of-the-art algorithms. This would represent a significance advance for *in vivo* studies using SR-CT. Sparse2Noise can also be applied for traditional CT and common synchrotron radiation CT imaging techniques for improving the image quality by removing noise and ring artifacts.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC, No. RGPIN 06007-2019 and No. RGPIN 06396-2019). Experimental research described in this paper were performed at the Canadian Light Source, which is a national

research facility supported by the Canada Foundation for Innovation (CFI), NSERC, National Research Council (NRC), Canadian Institutes of Health Research (CIHR), Government of Saskatchewan, and the University of Saskatchewan.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.combiomed.2023.107473>.

References

- [1] X. Duan, N. Li, X. Chen, N. Zhu, Characterization of tissue scaffolds using synchrotron radiation microcomputed tomography imaging, *Tissue Eng. C Methods* 27 (2021) 573–588.
- [2] L. Brombal, F. Arfelli, P. Delogu, S. Donato, G. Mettivier, K. Michielsen, P. Oliva, A. Taibi, I. Sechopoulos, R. Longo, Image quality comparison between a phase-contrast synchrotron radiation breast CT and a clinical breast CT: a phantom based study, *Sci. Rep.* 9 (2019), 17778.
- [3] M. Töpperwien, A. Markus, F. Alves, T. Salditt, Contrast enhancement for visualizing neuronal cytoarchitecture by propagation-based x-ray phase-contrast tomography, *Neuroimage* 199 (2019) 70–80.
- [4] A. Gonzalez-Tendero, C. Zhang, V. Balicevic, R. Cárdenas, S. Loncaric, C. Butakoff, B. Paun, A. Bonnin, P. García-Cañadilla, Emma Muñoz-Moreno, Eduard Gratacós, Fatima Crispí, Bart Bijnens, Whole heart detailed and quantitative anatomy, myofibre structure and vasculature from X-ray phase-contrast synchrotron radiation-based micro computed tomography 18 (2017) 732–741.
- [5] S. Bayat, L. Porra, P. Suortti, W. Thomlinson, Functional lung imaging with synchrotron radiation: methods and preclinical applications, *Phys. Med.* 79 (2020) 22–35.
- [6] L. Ning, N. Zhu, A. Smith, A. Rajaram, H. Hou, S. Srinivasan, F. Mohabatpour, L. He, A. McLennan, V. Serpooshan, P. Papagerakis, X. Chen, Noninvasive three-dimensional *in situ* and *in vivo* characterization of bioprinted hydrogel scaffolds using the X-ray propagation-based imaging technique, *ACS Appl. Mater. Interfaces* 13 (2021) 25611–25623.
- [7] S. Tavakoli Taba, P. Baran, Y.I. Nesterets, S. Pacile, S. Wienbeck, C. Dullin, K. Pavlov, A. Maksimenko, D. Lockie, S.C. Mayo, Comparison of propagation-based CT using synchrotron radiation and conventional cone-beam CT for breast imaging, *Eur. Radiol.* 30 (2020) 2740–2750.
- [8] H. Labriet, C. Nemoz, M. Renier, P. Berkvens, T. Brochard, R. Cassagne, H. Elleaume, F. Estève, C. Verry, J. Balosso, Significant dose reduction using synchrotron radiation computed tomography: first clinical case and application to high resolution CT exams, *Sci. Rep.* 8 (2018) 1–7.
- [9] K.D. Harrison, E. Sales, B.D. Hiebert, A. Panahifar, N. Zhu, T. Arnason, K.J. Sweekla, P. Pivonka, L.D. Chapman, D.M. Cooper, Direct assessment of rabbit cortical bone basin multicellular unit longitudinal erosion rate: a 4D synchrotron-based approach, *J. Bone Miner. Res.* 37 (2022) 2244–2258.
- [10] E.L. Ritman, Micro-computed tomography—current status and developments, *Annu. Rev. Biomed. Eng.* 6 (2004) 185–208.
- [11] M.L. Bouxsein, S.K. Boyd, B.A. Christiansen, R.E. Guldberg, K.J. Jepsen, R. Müller, Guidelines for assessment of bone microstructure in rodents using micro-computed tomography, *J. Bone Miner. Res.* 25 (2010) 1468–1486.
- [12] N. Miyahara, T. Kokubo, Y. Hara, A. Yamada, T. Koike, Y. Arai, Evaluation of X-ray doses and their corresponding biological effects on experimental animals in cone-beam micro-CT scans (R-mCT2), *Radiol. Phys. Technol.* 9 (2016) 60–68.
- [13] K. Laperre, M. Depypere, N. Van Gastel, S. Torrekens, K. Moermans, R. Bogaerts, F. Maes, G. Carmeliet, Development of micro-CT protocols for *in vivo* follow-up of mouse bone architecture without major radiation side effects, *Bone* 49 (2011) 613–622.
- [14] I. Willekens, N. Buls, T. Lahoutte, L. Baeyens, C. Vanhove, V. Caveliers, R. Deklerck, A. Bossuyt, J. de Mey, Evaluation of the radiation dose in micro-CT with optimization of the scan protocol, *Contrast Media Mol. Imaging* 5 (2010) 201–207.
- [15] E. Kang, J. Min, J.C. Ye, A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction, *Med. Phys.* 44 (2017) e360–e375.
- [16] G. Wang, J.C. Ye, B. De Man, Deep learning for tomographic image reconstruction, *Nat. Mach. Intell.* 2 (2020) 737–748.
- [17] H. Shan, A. Padole, F. Homayounieh, U. Kruger, R.D. Khera, C. Nitiwarangkul, M. K. Kalra, G. Wang, Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction, *Nat. Mach. Intell.* 1 (2019) 269–276.
- [18] H. Chen, Y. Zhang, M.K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, G. Wang, Low-dose CT with a residual encoder-decoder convolutional neural network, *IEEE Trans. Med. Imag.* 36 (2017) 2524–2535.
- [19] D. Ulyanov, A. Vedaldi, V. Lempitsky, Deep image prior, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 9446–9454.
- [20] E. Kang, H.J. Koo, D.H. Yang, J.B. Seo, J.C. Ye, Cycle-consistent adversarial denoising network for multiphase coronary CT angiography, *Med. Phys.* 46 (2019) 550–562.

- [21] J. Gu, T.S. Yang, J.C. Ye, D.J. Yang, CycleGAN denoising of extreme low-dose cardiac CT using wavelet-assisted noise disentanglement, *Med. Image Anal.* 74 (2021), 102209.
- [22] Y. Song, L. Shen, L. Xing, S. Ermon, Solving Inverse Problems in Medical Imaging with Score-Based Generative Models, 2021 arXiv preprint arXiv:08005.
- [23] J. Liu, R. Anirudh, J.J. Thiagarajan, S. He, K.A. Mohan, U.S. Kamilov, H. Kim, DOLCE: A Model-Based Probabilistic Diffusion Framework for Limited-Angle Ct Reconstruction, 2022 arXiv preprint arXiv:12340.
- [24] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, T. Aila, Noise2Noise: Learning Image Restoration without Clean Data, 2018, 04189 arXiv preprint arXiv.
- [25] D. Wu, K. Gong, K. Kim, X. Li, Q. Li, Consensus neural network for medical imaging denoising with only noisy training samples, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 741–749.
- [26] N. Yuan, J. Zhou, J. Qi, Half2Half: deep neural network based CT image denoising without independent reference data, *Phys. Med. Biol.* 65 (2020), 215020.
- [27] A.M. Hasan, M.R. Mohebbian, K.A. Wahid, P. Babyn, Hybrid-collaborative Noise2Noise denoiser for low-dose CT images, *IEEE Transact. Radiat. Plasma Med. Sci.* 5 (2020) 235–244.
- [28] J. Batson, L. Royer, Noise2self: blind denoising by self-supervision, in: International Conference on Machine Learning, PMLR, 2019, pp. 524–533.
- [29] K. Choi, J.S. Lim, S. Kim, Self-supervised inter-and intra-slice correlation learning for low-dose CT image restoration without ground truth, *Expert Syst. Appl.* 209 (2022), 118072.
- [30] A.A. Hendriksen, D.M. Pelt, K.J. Batenburg, Noise2inverse: self-supervised deep convolutional denoising for tomography, *IEEE Transact. Computat. Imag.* 6 (2020) 1320–1335.
- [31] A.A. Hendriksen, M. Bührer, L. Leone, M. Merlini, N. Vigano, D.M. Pelt, F. Marone, M. Di Michiel, K.J. Batenburg, Deep denoising for multi-dimensional synchrotron X-ray tomography without high-quality reference data, *Sci. Rep.* 11 (2021) 1–13.
- [32] Y. Li, S. Han, Y. Zhao, F. Li, D. Ji, X. Zhao, D. Liu, J. Jian, C. Hu, Synchrotron microtomography image restoration via regularization representation and deep CNN prior, *Comput. Methods Progr. Biomed.* 226 (2022), 107181.
- [33] B. Münch, P. Trtik, F. Marone, M. Stampanoni, Stripe and ring artifact removal with combined wavelet–Fourier filtering, *Opt Express* 17 (2009) 8567–8591.
- [34] N.T. Vo, R.C. Atwood, M. Drakopoulos, Superior techniques for eliminating ring artifacts in X-ray micro-tomography, *Opt Express* 26 (2018) 28396–28412.
- [35] X. Duan, N. Li, D.M. Cooper, X.F. Ding, X. Chen, N. Zhu, Low-density tissue scaffold imaging by synchrotron radiation propagation-based imaging computed tomography with helical acquisition mode, *J. Synchrotron Radiat.* 30 (2023) 417–429.
- [36] Y. Mäkinen, S. Marchesini, A. Foi, Ring artifact reduction via multiscale nonlocal collaborative filtering of spatially correlated noise, *J. Synchrotron Radiat.* 28 (2021) 876–888.
- [37] M. Salehjahromi, Q. Wang, Y. Zhang, L.A. Gjestebø, D. Harrison, G. Wang, P. M. Edic, H.J.M.p. Yu, A new iterative algorithm for ring artifact reduction in CT using ring total variation 46 (2019) 4803–4815.
- [38] D. Paganin, S.C. Mayo, T.E. Gureyev, P.R. Miller, S.W. Wilkins, Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object, *J. Microsc.* 206 (2002) 33–40.
- [39] T. Faragó, S. Gasilov, I. Emslie, M. Zuber, L. Helfen, M. Vogelgesang, T. Baumbach, Tofu: a fast, versatile and user-friendly image processing toolkit for computed tomography, *J. Synchrotron Radiat.* 29 (2022) 916–927.
- [40] Z. Liu, T. Bicer, R. Kettimuthu, D. Gursoy, F. De Carlo, I. Foster, TomoGAN: low-dose synchrotron x-ray tomography with generative adversarial networks: discussion, *JOSA A* 37 (2020) 422–434.
- [41] H. Shan, Y. Zhang, Q. Yang, U. Kruger, M.K. Kalra, L. Sun, W. Cong, G. Wang, 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network, *IEEE Trans. Med. Imag.* 37 (2018) 1522–1534.
- [42] D.M. Pelt, A.A. Hendriksen, K.J. Batenburg, Foam-like phantoms for comparing tomography algorithms, *J. Synchrotron Radiat.* 29 (2022) 254–265.
- [43] W. Van Aarle, W.J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabavolski, J. De Beenhouwer, K.J. Batenburg, J. Sijbers, Fast and flexible X-ray tomography using the ASTRA toolbox, *Opt. Express* 24 (2016) 25129–25147.
- [44] F. Mohabatpour, X. Duan, Z. Yazdanpanah, X.L. Tabil, L. Lobanova, N. Zhu, S. Papagerakis, X. Chen, P. Papagerakis, Bioprinting of alginate-carboxymethyl chitosan scaffolds for enamel tissue engineering in vitro, *Biofabrication* 15 (2022), 015022.
- [45] D.X. Chen, *Extrusion Bioprinting of Scaffolds*, Springer, Cham, 2019.
- [46] F. You, X. Wu, M. Kelly, X. Chen, Bioprinting and in vitro characterization of alginate dialdehyde-gelatin hydrogel bio-ink, *Bio-Design. Manufact.* 3 (2020) 48–59.
- [47] X. Chen, A.F. Anvari-Yazdi, X. Duan, A. Zimmerling, R. Gharraei, N. Sharma, S. Sweilem, L. Ning, Biomaterials/bioinks and extrusion bioprinting, *Bioact. Mater.* 28 (2023) 511–536.
- [48] Z. Izadifar, A. Honaramooz, S. Wiebe, G. Belev, X. Chen, D. Chapman, Low-dose phase-based X-ray imaging techniques for in situ soft tissue engineering assessments, *Biomaterials* 82 (2016) 151–167.
- [49] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.* 13 (2004) 600–612.
- [50] H. Fujita, D.-Y. Tsai, T. Itoh, K. Doi, J. Morishita, K. Ueda, A. Ohtsuka, A simple method for determining the modulation transfer function in digital radiography, *IEEE Trans. Med. Imag.* 11 (1992) 34–39.
- [51] P.D. Lösel, T. van de Kamp, A. Jayme, A. Ershov, T. Faragó, O. Pichler, N. Tan Jerome, N. Aadepur, S. Bremer, S.A.J.N.C. Chilingaryan, Introducing Biomedisa as an open-source online platform for biomedical image segmentation 11 (2020) 1–14.
- [52] T. Eelbode, J. Bertels, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, M. B. Blaschko, Optimization for medical image segmentation: theory and practice when evaluating with dice score or jaccard index, *IEEE Trans. Med. Imag.* 39 (2020) 3679–3690.
- [53] Z. Zhang, X. Liang, X. Dong, Y. Xie, G. Cao, A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution, *IEEE Trans. Med. Imag.* 37 (2018) 1407–1417.
- [54] W. Xia, H. Shan, G. Wang, Y. Zhang, Physics-/model-based and data-driven methods for low-dose computed tomography: a survey, *IEEE Signal Process. Mag.* 40 (2023) 89–100.
- [55] B. Zhu, J.Z. Liu, S.F. Cauley, B.R. Rosen, M.S.J.N. Rosen, Image reconstruction by domain-transform manifold learning, *Nature* 555 (2018) 487–492.