Swap-Net: A Memory-Efficient 2.5D Network for Sparse-View 3D Cone Beam CT Reconstruction to ICF Applications

Xiaojian Xu, Marc L. Klasky, Michael T. McCann, *Member, IEEE*, Jason Hu, *Graduate Student Member, IEEE*, and Jeffrey A. Fessler, *Fellow, IEEE*

***Abstract*—Reconstructing 3D cone beam computed tomography (CBCT) images from a limited set of projections is an important inverse problem in many imaging applications from medicine to Inertial Confinement Fusion (ICF). The performance of traditional methods such as filtered back projection (FBP) and model-based regularization is sub-optimal when the number of available projections is limited. In the past decade, deep learning (DL) has gained great popularity for solving CT inverse problems. A typical DL-based method for CBCT image reconstruction is to learn an end-to-end mapping by training a 2D or 3D network. However, 2D networks fail to fully use global information. While 3D networks are desirable, they become impractical as image sizes increase because of the high memory cost. This paper proposes Swap-Net, a memory-efficient 2.5D network for sparse-view 3D CBCT image reconstruction. Swap-Net uses a sequence of novel axes-swapping operations to reconstruct 3D volumes in an end-to-end fashion without using full 3D convolutions. Simulation results on ICF show that Swap-Net consistently outperforms baseline methods both quantitatively and qualitatively in terms of reducing artifacts and preserving details of complex hydrodynamic simulations of relevance to the ICF community.**

***Index Terms*—Deep learning, 3D inverse problem, image reconstruction, denoising, sparse-view, computed tomography, 2.5D network, inertial confinement fusion.**

I. Introduction

THE recovery of high-quality images from limited projection measurements is fundamental in computed tomography (CT) [1]. Cone beam CT (CBCT) is a specialized imaging technique used in fields requiring detailed 3D imaging. In CBCT, an X-ray beam is projected through the 3D object onto a 2D detector. Unlike traditional CT scanners where the X-ray beam is collimated into a narrow fan shape, CBCT systems use a cone-shaped beam, allowing wider coverage of the object in a single rotation. CBCT is a valuable tool in various applications for obtaining detailed structural information [2], [3], [4].

A CBCT scanner captions 2D X-ray projections, also called radiographs, as it rotates around the target object. Computer algorithms process these projections to reconstruct a 3D volumetric image of the object. Developing fast and accurate methods for 3D CBCT image reconstruction is important in many applications [2], [3], [4]. In particular, sparse-view CBCT imaging, which relies on a limited number of projections for image reconstruction, is a common challenge in the physical sciences, including shock physics experiments, high-energy density physics, and national security applications. This challenge arises from the prohibitive cost of acquiring additional projections and the physical constraints that prevent the rotation of the source and/or object in rapidly evolving dynamic systems—a common issue for ICF applications. Additionally, the high computational cost of 3D ICF simulations limits the availability of training datasets, further necessitating the development of new computational imaging algorithms to enhance reconstruction performance. The ability to capture fine features is critical, as these structures significantly impact the burn dynamics of Deuterium Tritium (DT) gas in ICF applications and, consequently, the overall neutron yield. Therefore, advancing algorithms to improve the extraction of these fine details and subsequently optimize topologies through design refinement is of paramount importance to the community.

Filtered back projection (FBP) is a classical algorithm that is computationally efficient and relatively straightforward to implement [5], [6]. However, FBP is sensitive to measurement noise and leads to artifacts when given incomplete or irregularly sampled projection data. Regularized inversion methods view CT imaging as an *inverse problem*, where the unknown object is reconstructed by combining a CT physical model and a hand-crafted regularizer [7], [8], [9], [10], [11], [12], [13], [14], [15]. Recently, *deep learning (DL)* methods have gained popularity in solving CBCT inverse problems [16], [17], [18], [19], [20]. Traditional DL methods are based on training *convolutional neural networks* (CNNs) to map the measurements or low-quality images to the desired high-quality images [21], [22],

Received 6 October 2024; revised 5 March 2025 and 27 April 2025; accepted 7 May 2025. Date of publication 23 May 2025; date of current version 2 July 2025. This work was supported in part by the Laboratory Directed Research and Development program of Los Alamos National Laboratory and in part by the U.S. Department of Energy through Los Alamos National Laboratory (LANL). The associate editor coordinating the review of this article and approving it for publication was Prof. Chao Zuo. *(Corresponding author: Jeffrey A. Fessler.)*

Xiaojian Xu was with the Department of Electrical Engineering & Computer Science, University of Michigan, Ann Arbor, MI 48109 USA. She is now with GE HealthCare, Bellevue, WA 98004 USA.

Jason Hu and Jeffrey A. Fessler are with the Department of Electrical Engineering & Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: fessler@umich.edu).

Marc L. Klasky and Michael T. McCann are with Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545 USA.

Digital Object Identifier 10.1109/TCI.2025.3572699

[23]. *Deep model-based architectures* (DMBAs), such as those based on deep unfolding (DU) [24], [25], have recently extended traditional DL to neural network architectures that combine the CT forward models and CNN regularizers [26], [27], [28], [29], [30], [31], [32], [33], [34].

Despite the rich literature on DL-based methodologies, direct end-to-end 3D CT reconstruction remains a challenging problem due to its high memory and computation cost. Current schemes typically use a 2D approach, where the 3D volume is divided into a series of 2D slices along one or more axes. Then each 2D slice is treated as an independent image, and a 2D neural network is applied to process each slice individually [21], [28], [30], [35], [36]. After processing all 2D slices, the outputs are combined to reconstruct the full 3D volume. Using a 2D network for 3D reconstruction offers several advantages, including computational efficiency, ease of implementation, and compatibility with existing 2D CNN architectures and frameworks. However, it also suffers from drawbacks such as the potential loss of consistency across slices and suboptimal performance in capturing complex 3D structures compared to dedicated 3D reconstruction approaches [22], [23], [37]. Some other works have also explored the feasibility of using a certain number of neighboring slices to reconstruct a slice within a 3D volume [22], [38]. These methods, referred to as neighboring slice-based 2.5D approaches, can be beneficial when memory and data are limited, but they do not fully capture correlations across all dimensions. This paper addresses these issues by presenting a new network—called Swap-Net—for recovering high-quality 3D images from extreme sparse-view measurements. Distinct from the fully 3D volume-based approaches, 2D slice-based approaches, and the neighboring slice-based 2.5D approaches, Swap-Net is developed as a 2.5D CNN where 2D convolution operations are used to extract correlations across all three dimensions of a 3D volume. The key contributions of our work are summarized as follows:

- We present a memory-efficient 2.5D network called Swap-Net to handle end-to-end 3D image reconstruction. The key component in Swap-Net is the new axes-swapping operation that helps combine information along all axes similar to 3D convolution.
- We investigated challenging sparse-view 3D CBCT image reconstruction problems in ICF with as few as 4, 8 and 16 projection views. Moreover, we accounted for non-ideal physics including blur, scatter, and non-white noise. Simulation results demonstrate that the method can restore high-quality 3D volumes across all dimensions, outperforming baseline methods both quantitatively and qualitatively in terms of artifact-reduction and detail-preservation.
- We conducted additional investigations using Swap-Net, e.g., studying the benefits of the axis-swapping and the impact of the swapping order, the advantages of the cascading architecture, the robustness under distribution shifts, the capability of joint reconstruction and artifact correction, the integration with the imaging model within the DU framework. Our results comprehensively demonstrated the effectiveness, robustness, efficiency, and flexibility of Swap-Net.
- We evaluated our method with a greater number of projection views and tested it on a dataset beyond the ICF application, thereby demonstrating the versatility of the method. This experiment further illustrate that Swap-Net is a robust and broadly applicable method for 3D object reconstruction.

The rest of this paper is organized as follows. Section II introduces the background and mathematical formulation of the CBCT imaging problem and discusses related work. Section III presents our proposed approach in detail. Section IV explains our experimental setup, presents the results of our comparisons to other algorithms, and elaborates upon the analysis of the observations. Finally, Section V summarizes our work and discusses potential future directions.

II. Background

A. CT Inverse Problem Formulation

In CT imaging, the relationship between the unknown object $\boldsymbol{x} \in \mathbb{R}^n$ and the (log) projection measurements $\boldsymbol{y} \in \mathbb{R}^m$ is commonly expressed as a linear imaging system

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}, \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ denotes the measurement operator (also known as the forward model or physical model) and $\boldsymbol{e} \in \mathbb{R}^m$ denotes the measurement noise that is sometimes statistically modeled as additive white Gaussian noise (AWGN). The AWGN formulation is a widely used approximation for various imaging systems including CT, magnetic resonance imaging (MRI), etc. [40], [41].

Scatter is another practical corruption that arises in CT imaging due to interactions between X-ray photons and objects. When X-ray photons encounter the object, some of them undergo scattering rather than being absorbed or passing straight through. Since scattered photons have undergone direction changes, they do not provide accurate information about the original object attenuation along the X-ray path. Scattered photons can reach the detector and contribute to errors that reduce the quality of the reconstructed image. Often the post-log scatter-corrupted CBCT projection measurements $\boldsymbol{y}$ are modeled as

$$\boldsymbol{y} = -\log\left(\frac{\Phi(\boldsymbol{I}_0 e^{-\boldsymbol{A}\boldsymbol{x}})}{\boldsymbol{I}_0}\right) \tag{2}$$

where $\boldsymbol{I}_0$ denotes the reference intensity from the source, Φ is a nonlinear function that models the non-ideal physics including blur and scatter and non-white noise corruption (see Section IV-B for details), and log is applied pixelwise. Choices for modeling the scatter component of the function Φ in the literature include kernel convolution with the direct signal followed by Poisson noise [42], [43]. For any noise model, the goal is to reconstruct the image volume $\boldsymbol{x}$ from the projection data $\boldsymbol{y}$.

The motivation for including both AWGN and scatter scenarios is twofold. First, the non-ideal blur and scatter and non-white noise corruption, which we referred to as scatter artifacts for simplicity, are a significant source of distortion in many CT applications such as ICF. Investigating this scattering model is especially beneficial for addressing practical CT-based ICF reconstruction challenges. Second, in CBCT problems it is common to utilize AWGN artifacts as a baseline for investigations. Therefore, we investigate these two scenarios in this work.

B. Related Work for CT Reconstruction

As a 3D-imaging technique, CBCT imaging offers many benefits in clinical, industry, and research. However, due to factors such as computation cost, scatter, noise, limited measurements, and discrepancies in the forward operator model, significant challenges emerge when attempting to efficiently and accurately reconstruct the 3D CT images outlined in (1) and (2) [44], [45], [46]. Classical approaches tackle CT reconstruction by formulating it as a regularized optimization problem

$$\widehat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\{g(\boldsymbol{x}) + r(\boldsymbol{x})\} \tag{3}$$

where g is the data-fidelity term that quantifies the consistency with the measured data $\boldsymbol{y}$, and r is a regularizer that enforces a prior knowledge on the unknown image $\boldsymbol{x}$. For example, two widely-used data-fidelity and regularization terms in imaging are the least-squares and total variation (TV) terms:

$$g(\boldsymbol{x}) = \frac{1}{2}\left\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\right\|_2^2 \quad \text{and} \quad r(\boldsymbol{x}) = \tau\left\|\boldsymbol{D}\boldsymbol{x}\right\|_1 \tag{4}$$

where $\tau > 0$ controls the regularization strength and $\boldsymbol{D}$ denotes the discrete gradient operator [47]. Many handcrafted regularizers similar to TV have also been applied to sparse-view CT reconstruction problems [13], [48], [49], [50], [51], [52]. Beyond handcrafted priors, recent work has also explored the use of learned priors, e.g., [13], [18], [28], [30], [53].

DL has gained great popularity for solving CT inverse problems due to its excellent performance [21], [22], [23], [54], [55], [56]. A widely used supervised DL approach is based on training a CNN to map a corrupted image to its clean target [21], [22], [23], [57]. For example, prior work on DL for CBCT trains a CNN to map FBP reconstructed images to the corresponding ground-truth images. In particular, for CBCT where the target images are 3D, due to the memory limits, the network training is typically done in a slice-by-slice manner, where the 3D volumes are sliced into 2D images along a certain axes and the loss is optimized on the given slices [21], [28], [30], [35], [36]. However, due to the lack of global information, a 2D slice-based approach cannot capture complex 3D structures as well as dedicated 3D reconstruction approaches [23], [58]. An alternative method is to divide the whole volume into small 3D patches, feed the patches to the 3D network, and then fuse the reconstructed patches together [22], [59]. While such 3D patch-based approach can extract and establish features in all dimensions within patches, it cannot model global correlations and the fusion of patches in forming the whole volume requires additional attention to boundary artifacts [22], [60].

C. Our Contribution

This work contributes to the memory-expensive area of efficient 3D CBCT reconstruction using DL methods. We introduce a memory-efficient 2.5D network, called Swap-Net, that refines 3D images reconstructed from artifact-corrupted radiographs. Swap-Net addresses in an end-to-end fashion several common sources of image artifacts, including those due to sparse view sampling, measurement noise, and photon scattering. We extensively test the performance of Swap-Net in the application of ICF, validating that it can be used as an effective end-to-end mapping tool for 3D CBCT image reconstruction.

III. Proposed Method

We propose Swap-Net as an end-to-end mapping network that can handle 3D inverse problems like CBCT reconstruction. Fig. 1 shows the training pipeline (top) and architecture (bottom) of Swap-Net. As illustrated in the top part of Fig. 1, given the corrupted CBCT projections $\boldsymbol{y}$, Swap-Net $\mathsf{R}_{\boldsymbol{\theta}}$ takes the FBP reconstructed images $\boldsymbol{A}^{\dagger}\boldsymbol{y}$ as its input, and maps the whole volume to the desired 3D output $\widehat{\boldsymbol{x}} := \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{A}^{\dagger}\boldsymbol{y})$. Here, $\boldsymbol{\theta}$ represents the parameters of Swap-Net, and $\boldsymbol{A}^{\dagger}$ denotes the FBP reconstruction operation. Swap-Net training seeks to minimize the loss $\mathcal{L}$ between $\widehat{\boldsymbol{x}}$ and the ground truth $\boldsymbol{x}$ over a training set consisting of J samples to obtain the optimized parameters $\boldsymbol{\theta}^*$

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\arg\min}\sum_{j=1}^{J}\mathcal{L}(\boldsymbol{x}_j, \mathsf{R}_{\boldsymbol{\theta}}(\boldsymbol{A}^{\dagger}\boldsymbol{y}_j)) \tag{5}$$

where $\mathcal{L}$ denotes the loss function that measures the discrepancy between the predictions of the Swap-Net and the ground truth.

When the input images $\boldsymbol{A}^{\dagger}\boldsymbol{y}$ are of size $N_x \times N_y \times N_z$, Swap-Net works in a 3D-to-3D manner to produce a whole volume estimate $\widehat{\boldsymbol{x}}$ having the same dimension as its input without slicing and assembling the volume. The efficiency of Swap-Net in facilitating 3D image reconstruction hinges upon the novel and efficient design of its architecture. As illustrated in the bottom part of Fig. 1, Swap-Net is a cascade of three repeating blocks, each includes a sub-network. The sub-network is formed by two convolutional layers (Conv) followed by Rectified Linear Unit (ReLU) activations, one additional convolutional layer, and a residual connection. The convolutional kernels across all layers are uniformly set to a size of 3×3 with a stride of 1. The channel dimensions of the hidden convolutional layers are set to N_x, N_y, and N_z for the first, second, and final sub-networks, respectively, corresponding to the dimensions of the 3D input volume along the x, y, and z axes. A distinctive aspect of Swap-Net lies in its use of the *axes-swapping* operations within each block. This operation cascadingly reorients the channel dimension to the x, y, and z axes, facilitating focused 2D convolutions across the yz, xz, and xy planes, respectively. This strategic approach enables the network to perform artifact reduction axis by axis, thereby ultimately yielding a high-fidelity 3D reconstruction that maintains consistency across all dimensions. Mathematically, the computation of the lth block in Swap-Net can be defined as:

$$\mathsf{B}_{\boldsymbol{\theta}^l}(\boldsymbol{z}^l; a^l) := \mathsf{S}^{-1}(\mathsf{R}_{\boldsymbol{\theta}^l}(\mathsf{S}(\boldsymbol{z}^l; a^l)); a^l)\ , \tag{6}$$

where $\mathsf{R}_{\boldsymbol{\theta}^l}$ represents the lth convolution sub-network of Swap-Net, while S and S^{-1} denote the axes-swapping and inverse axes-swapping operations, respectively. The operation S rearranges the given axis a^l of the 3D input volume $\boldsymbol{z}^l$ onto the convolution dimension, and S^{-1} restores the original orientation after convolution. The inverse axes-swapping operation S^{-1} is applied to ensure that the input and output of each block maintain

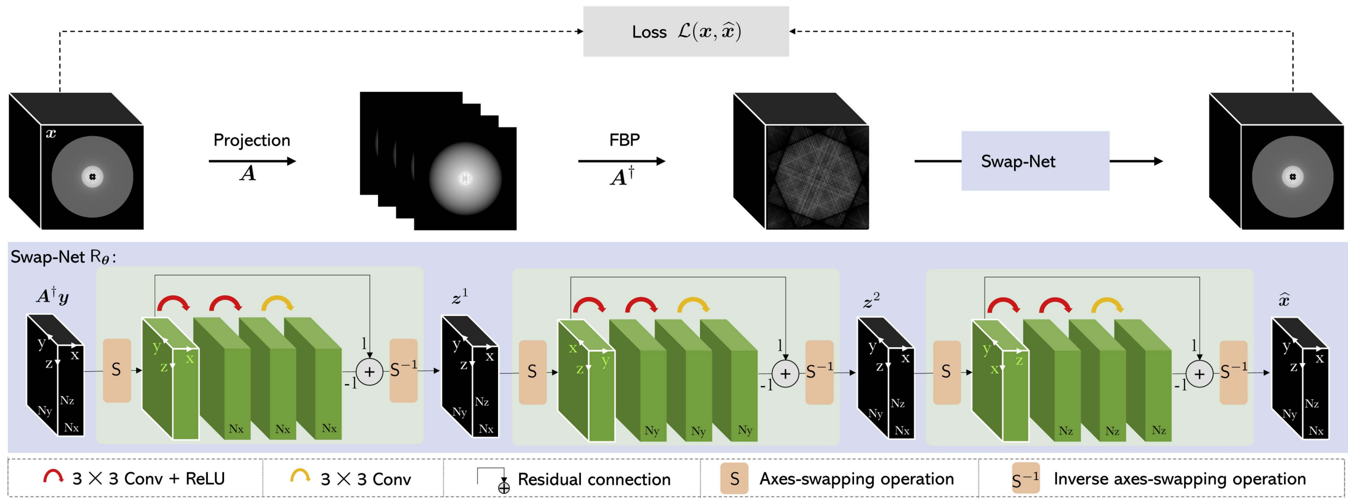


Fig. 1. Overview of the proposed Swap-Net framework for training an end-to-end deep mapping for 3D CBCT image reconstruction using ICF synthetic radiographs. The Swap-Net model $\mathsf{R}_{\boldsymbol{\theta}}$ is implemented as a customized architecture mapping the output of FBP to the desired ground-truth 3D images. The novel axes-swapping operation in Swap-Net allows it to efficiently conduct convolution across all dimensions. The whole network is trained end-to-end in a supervised fashion.

the same spatial alignment. Our implementation of Swap-Net, illustrated in Fig. 1, consists of three cascading blocks for CBCT imaging and can be formulated as:

$$\begin{aligned} \boldsymbol{z}^1 &\leftarrow \mathsf{B}_{\boldsymbol{\theta}^1}(\boldsymbol{A}^{\dagger}\boldsymbol{y}; x) \\ \boldsymbol{z}^2 &\leftarrow \mathsf{B}_{\boldsymbol{\theta}^2}(\boldsymbol{z}^1; y) \\ \widehat{\boldsymbol{x}} &\leftarrow \mathsf{B}_{\boldsymbol{\theta}^3}(\boldsymbol{z}^2; z) . \end{aligned} \tag{7}$$

This design of Swap-Net enables effective feature transformation across different axes of a 3D object, improving reconstruction quality in sparse-view CBCT imaging. We implemented the S and S^{-1} operators using PyTorch's `swapaxes` function. This function facilitates backpropagation by reordering tensor dimensions without modifying data values, thereby preserving the computational graph. PyTorch efficiently tracks this operation, ensuring proper gradient flow during training.

The key novelty of our method is that, to the best of our knowledge, this is the first work presenting such a memory-efficient 2.5D cascade network based on axes-swapping operations. Different from the traditional slice-by-slice mapping methods, Swap-Net instead relies on the axis-by-axis reconstruction, which can be particularly useful when the 3D volume is not uniformly corrupted along each dimension. For example, in CBCT imaging, since the projections are produced by placing sources around a certain axis, e.g., the z axis in our experiments, the insufficient attenuation information along z usually leads to lower-quality images in the xy plane. The cascading axes-swapping operations in Swap-Net allow it to effectively process the reconstruction in the xy plane after accumulating more information in the yz and xz plane, therefore enforcing the global consistency of the reconstructed 3D volume. Moreover, although the output of Swap-Net is the whole 3D volume, it does not involve any computationally expensive 3D convolutions. Instead, it is simply based on 2D convolutions where the convolution is looped over all axes of a 3D volume. Thus, Swap-Net overcomes the suboptimal performance of slice-based 2D CNNs that disregard the information across slices. On the other hand, it also bypasses the expensive computation cost of 3D networks, facilitating solving practical 3D imaging problems.

It also worth highlighting the differences between existing work that uses a certain number of neighboring slices to reconstruct a slice within a 3D volume [22], [38]. Those methods address the issue when training data is limited and the volume is too large to be processed as a whole. In contrast, Swap-Net is designed to process the entire volume at once by performing convolutions across all slices without dividing the volume. This enables deeper correlations along each dimension by processing all slices along multiple axes. We hypothesize that this comprehensive feature extraction contributes to Swap-Net's reconstruction performance especially when trained with limited data. When faced with more severe memory constraints or limited data availability, Swap-Net can also be applied to smaller patched volumes, reconstructing the full volume by combining these patches. Some works have also explored utilizing convolution along different dimensions. For example, [61] processes the 3D volume by slicing it along two different planes and then applies two parallel 2D CNNs to reconstruct the slices, which are subsequently merged to form the final 3D volume. However, this approach does not fully capture correlations across all slices, nor does it adopt the cascading block-by-block optimization architecture used in Swap-Net. These differences highlight Swap-Net's distinct novelty and establish it as a valuable algorithmic advancement in the field of memory-efficient 3D CBCT reconstruction.

IV. Experimental Validation

This section presents numerical results that demonstrate the ability of Swap-Net to provide high-quality 3D reconstructions from sparse-view 2D projections of ICF double-shell capsules as

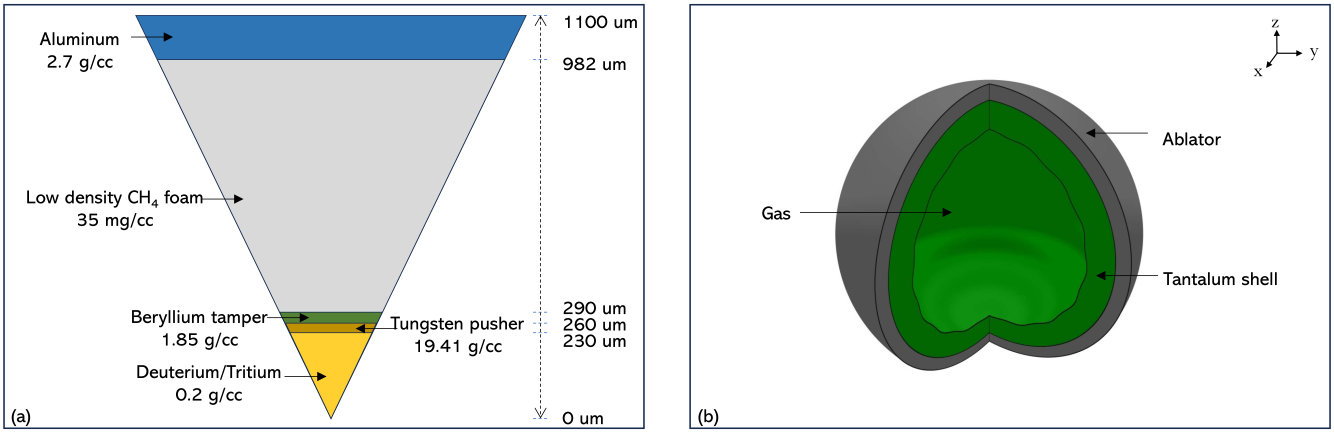


Fig. 2. The ICF models: (a) A typical double-shell ICF capsule containing Deuterium/Tritium, Tungsten pusher, Beryllium tamper, low density CH_4 foam, and Aluminum. (b) A simplified representation of a ICF implosion capsule containing Deuterium/Tritium (gas), Tantalum pusherlator CH_4 foam [39].

depicted in a representative double shell shown in Fig. 2. In particular, we examine Swap-Net under two different practical noise conditions, including AWGN corruption and nonlinear photon scattering corruption, to show its ability to solve challenging CBCT imaging problems.

A. Preparation of 3D Dataset

The emergence of Inertial Confinement Fusion (ICF) as a potential power source has been a major impetus for the continued examination of ICF implosion dynamics. One promising ICF configuration is a double-shell capsule, shown in Fig. 2(a), that employs a high Z metallic shell that is imploded onto a gas-filled cavity via radiation to achieve fusion conditions. Both manufacturing as well as drive asymmetries may lead to hydrodynamic instabilities that can degrade ICF performance. Consequently, quantifying and understanding these instabilities is crucial to the continued success of ICF. To this end, radiography plays an essential role in elucidating the behavior of the metallic shell and quantifying the impact of the asymmetries on ICF performance.

To further simplify the problem, we examine the explosion of a single shell made of tantalum, as this configuration enables the salient features to be captured in the density field, i.e., a complex gas metal interface without needing to increase the simulation complexity. As such we train and test our method with ICF capsules shown in Fig. 2(b) to examine shock propagation and instability growth created using prescribed perturbations on the shell interior surface. All simulations were performed using computational fluid dynamics software. The simplified ICF-like double-shell hydrodynamic simulations have been adopted in multiple prior work [62], [63]. Such ICF capsules, though seemingly simple, encapsulate all the critical and essential structures necessary to illustrate the ICF physical process. In this investigation, the configuration allows for the complex gas/metal interface to form as a consequence of the Richtmyer Meshkoff instability that impacts the ignition and burning of the fuel. Capturing these changes is essential for elucidating the nature of the instabilities that give rise to the complex topologies observed in the simulations. Therefore, evaluating the reconstruction

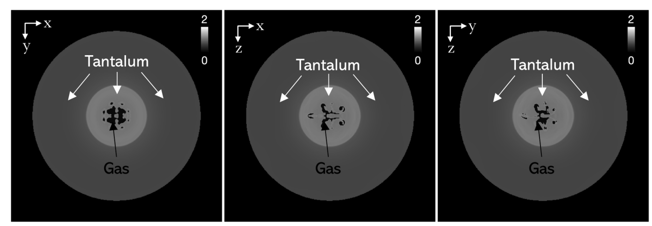


Fig. 3. Central slices along each dimension of an exemplar 3D ICF object generated for an ICF double shell simulation in our dataset. The two materials that form the object, namely gas and metal, are labeled in each image. The images presented here were normalized by the mass attenuation factor to the range of [0, 2] for good visualization (same in the rest of the paper).

performance on such simulated data provides an effective means of assessing a method's general ability to reconstruct experimental ICF data.

We simulated 108 3D objects with different parameters, e.g., initial 3D perturbations, material properties, and/or temporal slice where each case represents a distinct dynamic hydrodynamic configuration. Fig. 3 shows an exemplar object from our datasets with gas and tantalum labeled. In particular, the mass attenuation coefficient of gas is $\xi_{(\text{gas})} = 9.40 \text{ cm}^2/\text{g}$, and tantalum is $\xi_{(\text{tantalum})} = 13.03 \text{ cm}^2/\text{g}$, in the energy range of interest here. Each ICF volume had dimensions of $448 \times 448 \times 448$ with a voxel resolution of $250 \times 250 \times 250\,\mu\text{m}^3$ resulting in a total physical size of $112 \times 112 \times 112\,\text{mm}^3$. To minimize instabilities caused by 3D perturbations, each simulated ICF capsule exhibited mirror symmetry with respect to the yz plane. These 108 objects were split into 90, 18, and 18 for training, validation, and testing, respectively.

B. Generation of Radiographs

Fig. 4 illustrates our CBCT geometry. In our simulation, the ICF capsule undergoes rapid deformation, and due to initial surface asymmetries, it develops a Richtmyer-Meshkov instability following the re-shock of the inner tantalum surface. This instability gives rise to complex topologies in the central region of the images. The source and detectors were uniformly

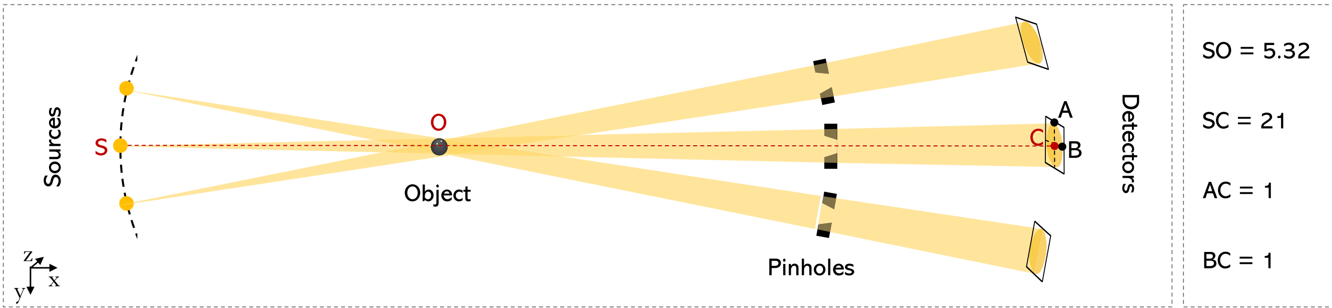


Fig. 4. The CBCT imaging geometry used in our experiments. The configurations on the side are normalized values to show relative distances. The pinhole is a circularly symmetric structure. In our setup, the xy plane is the transaxial plane, with sources and detectors positioned around the z axis, ensuring comprehensive capture of radiographic projections from multiple viewpoints.

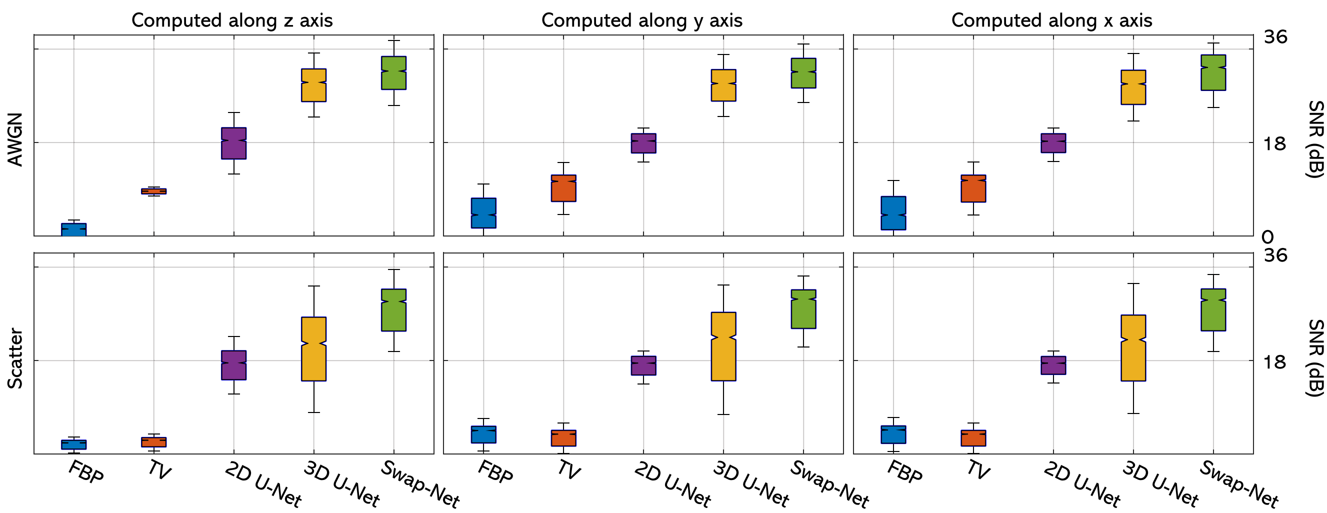


Fig. 5. Statistical summary of SNR values for different reconstruction methods evaluated on 2D slices along each dimension taken from our test set. Plots in the first and second row correspond to the the results with 4 projection views under AWGN and non-ideal physics including blur and scatter and non-white noise corruptions, respectively.

positioned around the capsule so as to provide multiple projections during the dynamic process. The pinhole utilizes the same metal material as in the 3D object with density being $\xi_{(\text{pinhole})} = 0.0404\ \text{cm}^2/\text{g}$. The direct radiographic signals from the area mass were simulated by placing the X-ray source around axis z with base intensity $I_0 = 3.201 \times 10^{-4}$. We tested the performance of Swap-Net on sparse-view CBCT reconstruction with 4, 8 and 16 views. The 2D projections have dimensions of 200×200 with a resolution of $2500 \times 2500\,\mu\text{m}^2$, corresponding to a detector size of $50 \times 50\,\text{cm}^2$. These CBCT views were generated using the ODL package [64], and all views were equally spaced over 180 degrees.

We generated the radiographs under two different corruption scenarios, namely AWGN, as modeled in (1), and non-ideal physics including blur and scatter and non-white noise corruption, as modeled in (2), respectively. For AWGN, the simulated corrupted radiographs included the addition of random AWGN corresponding to an input SNR of 40 dB to the clean ones. For our non-ideal physics investigation, we modeled the total transmission or the noisy radiograph function Φ as the sum of the blurred radiograph, scatter, and noise as follows:

$$\Phi := \boldsymbol{D}_{\text{blur}} + \boldsymbol{D}_{\text{s}} + \boldsymbol{B}_{\text{s}} + \boldsymbol{\eta}. \tag{8}$$

Let $\boldsymbol{D}$denote the uncollided radiation incident on the detector plane. The blurred direct radiation component is given by

$$\boldsymbol{D}_{\text{blur}} = \boldsymbol{D} * \boldsymbol{G}_{\text{blur}}(\sigma_{\text{blur}}) * \boldsymbol{\phi}_{\text{db}}. \tag{9}$$

The source blur $\boldsymbol{G}_{\text{blur}}$ is given by a 2D Gaussian kernel with deviation σ_{blur} chosen randomly between 1 and 3 pixels with an accompanying random orientation between 5 and 26 degrees. This signal was then convolved with a detector blur using another kernel $\boldsymbol{\phi}_{\text{db}}$.

To address the scatter radiation, we included two scatter components. The first was a correlated scatter component given by

$$\boldsymbol{D}_{\text{s}} = \kappa \boldsymbol{D} * \boldsymbol{G}_{\text{scatter}}(\sigma_{\text{scatter}}). \tag{10}$$

Here we convolved the direct radiograph with a 2D Gaussian filter scatter kernel $\boldsymbol{G}_{\text{scatter}}$ having standard deviation σ_{scatter} between 10 and 30 pixels for the kernel, with a scaling factor κ

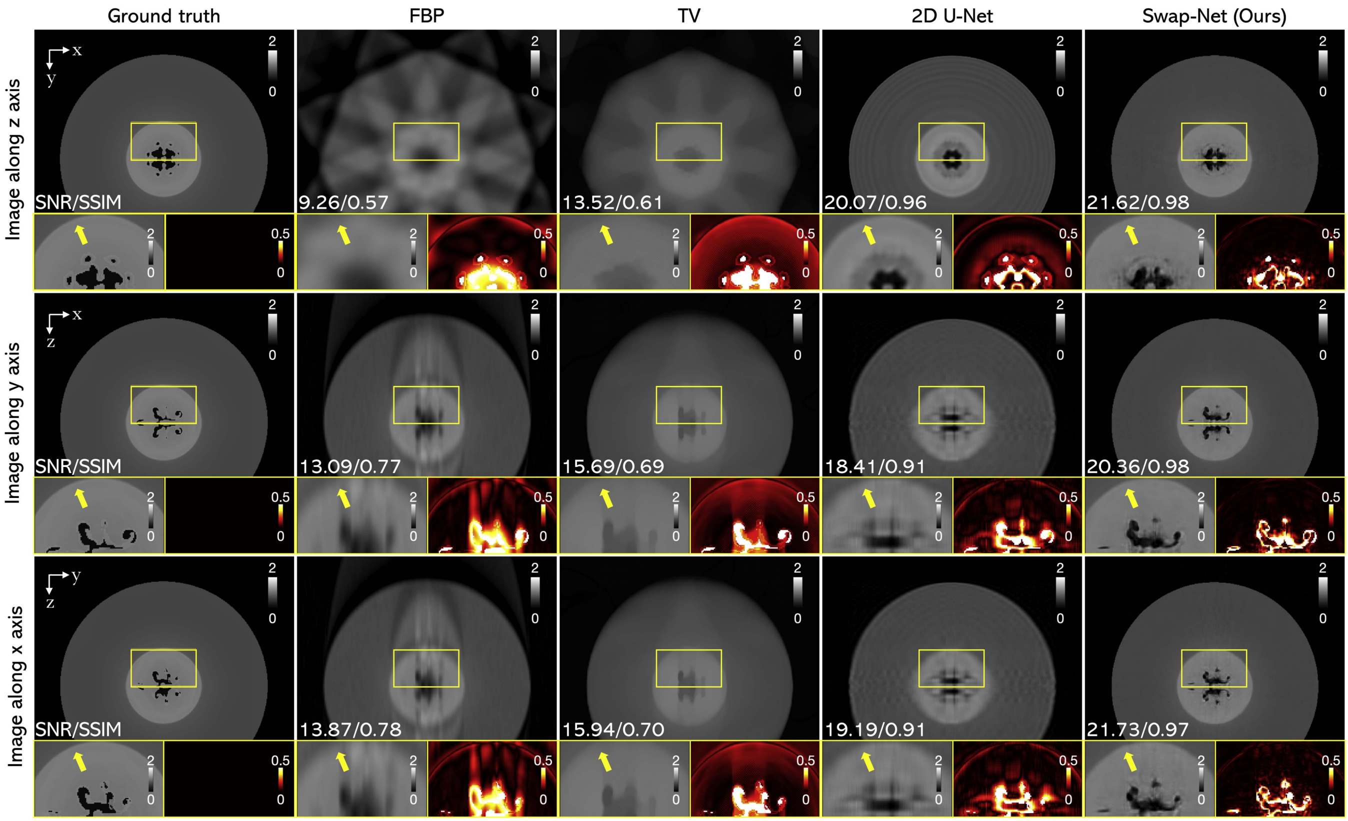


Fig. 6. Visual evaluation of Swap-Net and baseline methods on an exemplar ICF double shell test simulation with 4 projection views under AWGN corruption. Each row shows the middle slice of the whole 3D object along z, y and x axes, respectively. The bottom part of each image provides the SNR and SSIM values and representative $2\times$ zoomed-in regions and their error maps with respect to the ground truth. Arrows in the zoomed-in plots highlight sharp edges that are well reconstructed using Swap-Net. Note the excellent quantitative and qualitative performance of Swap-Net for both artifact correction and detail preservation.

between 0.1 and 0.3. We also added a background scatter field $\boldsymbol{B}_{\mathrm{s}}$, which is another essential component of scatter affecting radiographic measurements. Physically, this term represents scatter from our object that is reflected by nearby surrounding objects, e.g., ground and walls, which are particularly difficult to model. This field was modeled with a polynomial of order n given as

$$\boldsymbol{B}_{\mathrm{s}}(x,y) = \sum_{i=0}^{n} a_i x^i + b_i y^i, \tag{11}$$

where x and y denote spatial coordinates and a_i and b_i denote the coefficients of the polynomial. We chose the coefficients of the background scatter field such that the level was randomly between 0.5 and 1.5 times the mean signal level in the center of the image and the tilt was between -10% and 10% .

We modeled gamma and photon noise as Poisson noise denoted by $\boldsymbol{\eta}_g^{\mathrm{Po}}$ and $\boldsymbol{\eta}_p^{\mathrm{Po}}$, respectively. The means of the two distributions were proportional to the total signal $\boldsymbol{D}_{\mathrm{blur}} + \boldsymbol{D}_{\mathrm{s}} + \boldsymbol{B}_{\mathrm{s}}$ (with a scaling for each noise). The noise components were convolved with respective kernels $\boldsymbol{\phi}_g$ and $\boldsymbol{\phi}_p$ to give the total (colored) noise $\boldsymbol{\eta}$ as follows:

$$\boldsymbol{\eta} = \kappa_g(\boldsymbol{\eta}_g^{\mathrm{Po}} * \boldsymbol{\phi}_g) + \kappa_p(\boldsymbol{\eta}_p^{\mathrm{Po}} * \boldsymbol{\phi}_p), \tag{12}$$

where κ_g and κ_p are scaling coefficients for the gamma and photon noise components, respectively. The level of the gamma noise was randomly set in the range of (39,000, 50,000) and the level of the photon noise was randomly set in the range (350, 450). All random parameters were generated independently for each radiograph, so each radiograph was corrupted with different random noise and scatter realizations. In summary, the corruption function Φ serves as a comprehensive model for measurement errors, including scatter, blur (including source, detector, and motion), as will as noise, capturing artifacts present in our imaging system, as shown in Fig. 4. Additionally, it can be flexibly adjusted to simulate different imaging scenarios and configurations.

C. Baseline Methods and Training Settings

We considered several well-known algorithms as baseline methods for CBCT image reconstruction, including *FBP*, *TV* [47], *2D U-Net* [65], and *3D U-Net* [66]. FBP and TV are traditional methods that do not require training, while other methods are all DL methods with publicly available implementations. The FBP method was performed with the Hann filter, and the relative cutoff frequency for the filter was set to 0.3. We used `fminbound` in the `scipy.optimize` toolbox to identify the optimal regularization parameter τ for TV at the inference time. We trained all DL methods on the FBP reconstructed images to handle CBCT reconstruction. For 3D U-Net, we trained the model with 3D patches with patch size set to $112 \times 112 \times 112$.

TABLE I
QUANTITATIVE EVALUATION OF SWAP-NET AND BASELINE METHODS AVERAGED ON THE TEST DATASET FOR DIFFERENT NUMBERS OF PROJECTION VIEWS

Settings	AWGN	AWGN						Scatter					
Metric	PN (Million) / RM (GB) / RT (sec)	SNR (dB)			SSIM			SNR (dB)			SSIM		
Views	4	4	8	16	4	8	16	4	8	16	4	8	16
FBP	— / — / **0.46**	9.04	13.19	15.53	0.63	0.68	0.72	9.44	11.57	12.03	0.64	0.7	0.75
TV	— / — / 722.94	14.01	17.05	18.48	0.57	0.6	0.82	9.79	11.49	12.08	0.73	0.66	0.76
2D U-Net	50.26 / **1.66** / 8.66	20.77	20.93	20.94	0.95	0.95	0.95	17.01	17.77	20.30	0.89	0.91	0.95
3D U-Net	150.75 / 9.25 / 16.22	26.59	27.32	27.62	0.99	0.99	0.99	19.18	19.28	20.42	0.99	0.99	0.99
Swap-Net (Ours)	**16.26** / 5.91 / 0.62	**28.22**	**28.58**	**28.83**	**0.99**	**0.99**	**0.99**	**25.41**	**25.46**	**25.60**	**0.99**	**0.99**	**0.99**

Swap-net contains the fewest network parameters (PN) and uses moderate amount of GPU running memory (RM) yet achieved the highest SNR and SSIM compared with all the baseline methods across different projection settings.

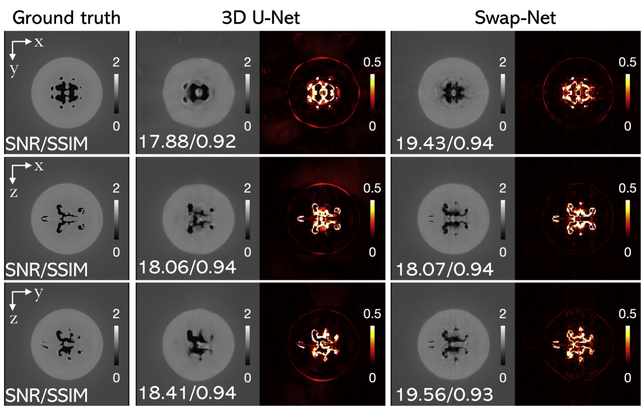


Fig. 7. Visual evaluation of 2.5D Swap-Net and 3D U-Net on an exemplar ICF double shell test simulation with 4 projection views under AWGN corruption. Each row shows the middle slice of the central region of the 3D object and the corresponding error maps with respect to the ground truth along z, y and x axes, respectively. The bottom part of each image provides the SNR and SSIM values. With only about $1/10$ of the parameters of 3D U-Net, Swap-Net still achieves better quantitative and qualitative performance.

The batch size of Swap-Net, 2D U-Net, and 3D U-Net was set to 2, 4, and 1, respectively. We used the ℓ_2 loss function for all training approaches, and set the learning rate to 0.0001 and used Adam [67] as our training optimizer. All models were trained with 2 A40 GPUs until stable convergence was observed. We evaluated reconstruction performance using two widely-adopted metrics: signal to noise ration (SNR) in dB and structural similarity index measure (SSIM) from `skimage.metrics` toolbox. Models that achieved the best performance on our validation dataset were selected for inference.

D. Results and Analysis

This section presents experimental results to demonstrate the efficiency and effectiveness of our proposed method. We organize our presentation as follows. First, we compare the performance of Swap-Net to baseline methods. Next, we provide a comprehensive evaluation of Swap-Net, examining its reconstruction performance across different slices within a complete 3D object volume, as well as its effectiveness in handling dynamic deformation throughout different stages. Unless otherwise specified, all numerical metrics reported in the tables were computed over the full 3D volumes, while the metrics labeled in the visualization figures were calculated specifically for the displayed slices and are provided for reference. All visualized images are axis-aligned and were extracted from planes perpendicular to the x, y, or z axes, as indicated by the axis labels in each figure.

1) Quantitative Performance Comparison With Baseline Methods: We first compared the performance of Swap-Net with baseline methods. Table I summarizes the averaged quantitative evaluation of Swap-Net and baseline methods on our testing dataset with different numbers of projection views. These numerical results were evaluated on the whole 3D volume for both AWGN corruption and non-ideal physics including blur and scatter and photon noise corruption. Swap-Net consistently outperformed the baseline methods, leading to the best SNR and SSIM in different scenarios. As a reference for model complexity, Table I also presents the model size in terms of the number of parameters (PN). Despite obtaining significantly enhanced performance, Swap-Net only uses about $1/3$ as many parameters as 2D U-Net and $1/10$ as many parameters as 3D U-Net. Table I also reports the running memory (RM) usage and the running time (RT) of each algorithm.[1] Note that Swap-Net processes the entire 3D volume during training, whereas the patch-based 3D U-Net processes $1/64$ of the volume, and the slice-based 2D U-Net processes $1/448$ of the volume. Despite these differences in data-related RM demands, Swap-Net's overall RM usage is still lower than that of the 3D U-Net and only slightly higher than the 2D U-Net, making it a memory-efficient solution in practice. To further evaluate the performance of the reconstruction along each of dimension of the 3D object, Fig. 5 summarizes the statistical evaluation for slice-wise reconstruction for both AWGN corruption and non-ideal physics including blur and scatter and photon noise corruptions. Swap-Net achieved consistently good reconstruction performance for 2D image slices along all three dimensions, thanks to the axes-swapping operation in our network design.

2) Visual Performance Comparison With Baseline Methods: Fig. 6 presents visual comparisons from different methods on an exemplar testing data under AWGN corruption with 4 views. We also plotted error maps, which represent the absolute differences from the ground truth in the figure. We specifically highlighted the reconstruction of boundary regions because changes in these areas throughout the dynamic ICF process can serve as

[1] The RM usage is reported as the peak GPU memory consumption. Both RM usage and RT for each method were measured using experiments with a batch size of 1 and 4 projection views

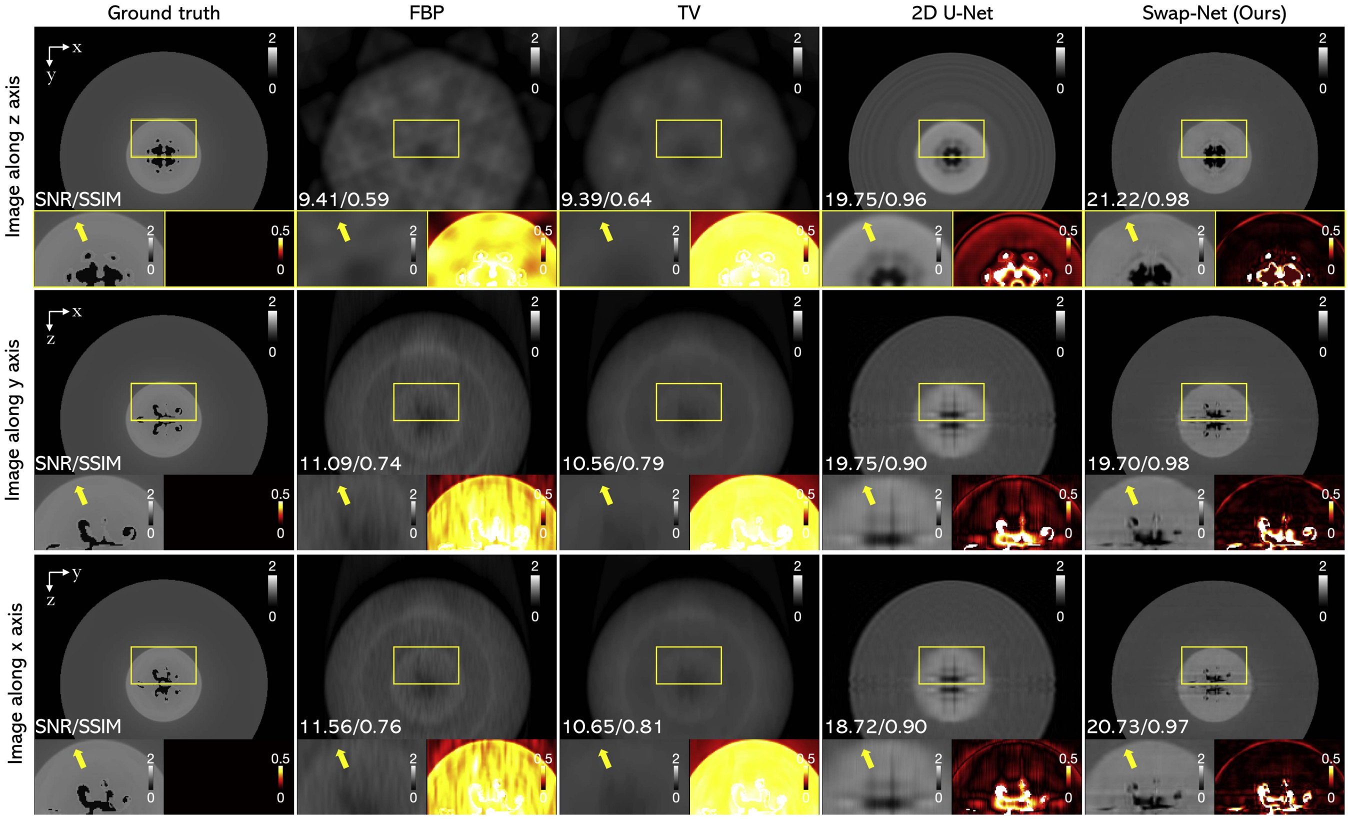


Fig. 8. Visual evaluation of Swap-Net and baseline methods on an exemplar ICF double shell test simulation with 4 projection views under scatter corruption. Each row shows the middle slice of the whole 3D object along z, y and x axes, respectively. The bottom part of each image provides the SNR and SSIM values, and representative $2\times$ zoomed-in regions and their error maps with respect to the ground truth. Arrows in the zoomed-in plots highlight sharp edges that are well reconstructed using Swap-Net. Note the excellent quantitative and qualitative performance of Swap-Net for both artifacts correction and detail preservation.

critical indicators of onset of instabilities that are known to degrade ICF performance [63], [68]. Swap-Net outperformed the baseline methods both in terms of removing artifacts and maintaining sharpness. The excellent performance demonstrates that Swap-Net can remove disturbing artifacts while retaining detailed structural information. Such capability is notable for a network having only 9 convolution layers. Because the CBCT projections were simulated by placing the X-ray source around the z axis, it is challenging to reconstruct images along z especially with sparse-view projections (e.g., see the comparatively worse FBP reconstruction in (x, y) plane in Fig. 6). Swap-Net overcomes such asymmetric artifacts by performing cascading convolutions along all axes, resulting in the comparatively consistent reconstruction along all dimensions. Fig. 7 compares Swap-Net and 3D patch-based U-Net methods; to avoid the influence of the edge artifacts, only the central region of the reconstructed object is presented. In Fig. 7, Swap-Net performed better than the 3D U-net. Fig. 8 demonstrates the improved performance of Swap-Net compared with various baseline methods under non-ideal physics including blur and scatter and photon noise corruption. While the baseline methods obviously suffer from scatter corruption, Swap-Net successfully reduced the artifacts, leading to a similar good quantitative and qualitative performance as in the AWGN case. Fig. 9 presents the results of additional investigations with the baseline methods using 8 and 16 views.

3) Effectiveness in Handling 3D ICF Objects: Beyond the results comparison with the baseline method, we further investigated the performance of Swap-Net on the whole dynamic 3D volume. Fig. 10 illustrates the performance of Swap-Net across different slices in a whole 3D object volume. For each slice, we show the side-to-side (top versus bottom) comparison between the results of Swap-Net and the corresponding ground truth. Using only 4 projection views, Swap-Net successfully reconstructed not only the sharp edges but also central details, matching well with the ground truth. The consistent success of Swap-Net on different slices suggests that it can work across the 3D volume, highlighting its effectiveness and adaptability. Fig. 11 additionally shows the visual performance of Swap-Net for different deformation stages of an object. Swap-Net's ability to preserve fine details and maintain sharp edges in our ICF dataset demonstrates its effectiveness in capturing key features of simulated ICF data and, by extension, its potential applicability to experimental ICF data.

E. Additional Study

1) Benefits of Axes-Swapping Operations: To highlight the contribution of Swap-Net's axes-swapping operation, we performed an additional study to examine its influence. First, we investigated a *Non-Swap-Net* network identical to the Swap-Net but without the axes-swapping operations. Comparing to

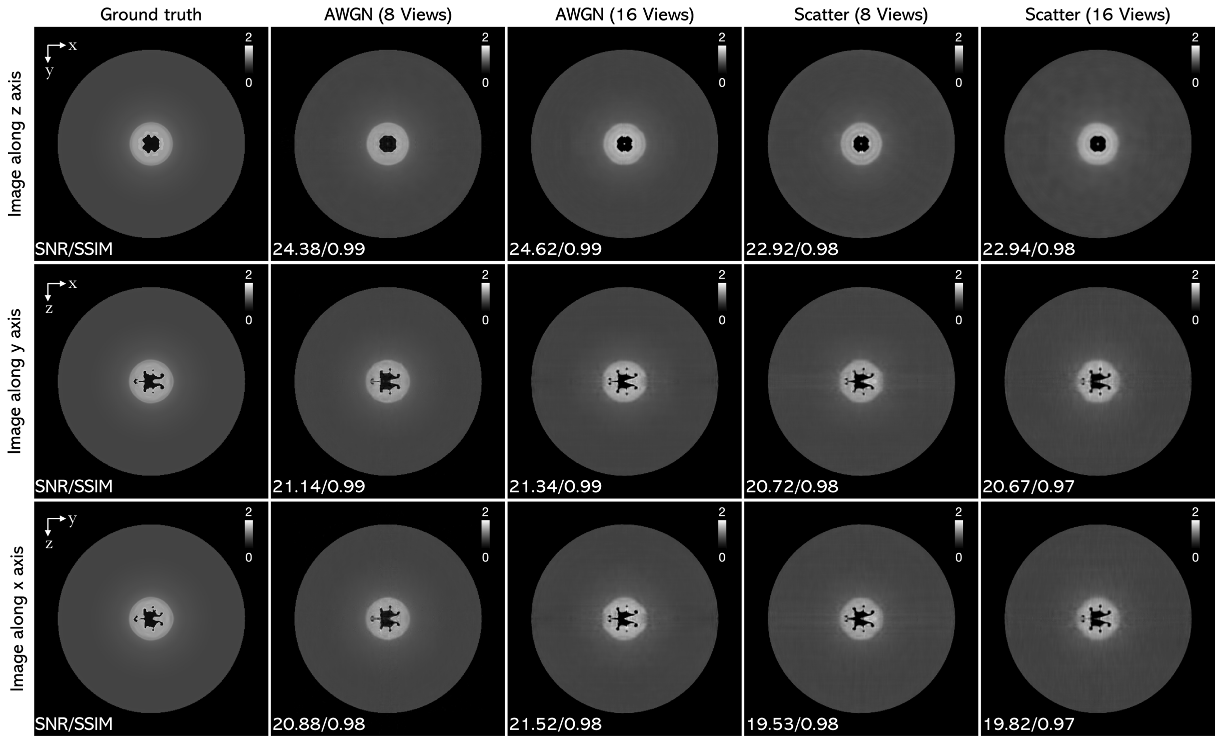


Fig. 9. Visual evaluation of Swap-Net on an exemplar ICF double shell test simulation with 8 and 16 projections views under AWGN and non-ideal physics including blur and scatter and non-white noise (labeled as Scatter) corruptions. Each row shows the middle slice of the whole 3D object along z, y and x axes, respectively. The bottom-left corner of each image provides the SNR and SSIM values with respect to the ground truth. Note the consistently good performance of Swap-Net for different projection views and noise corruptions.

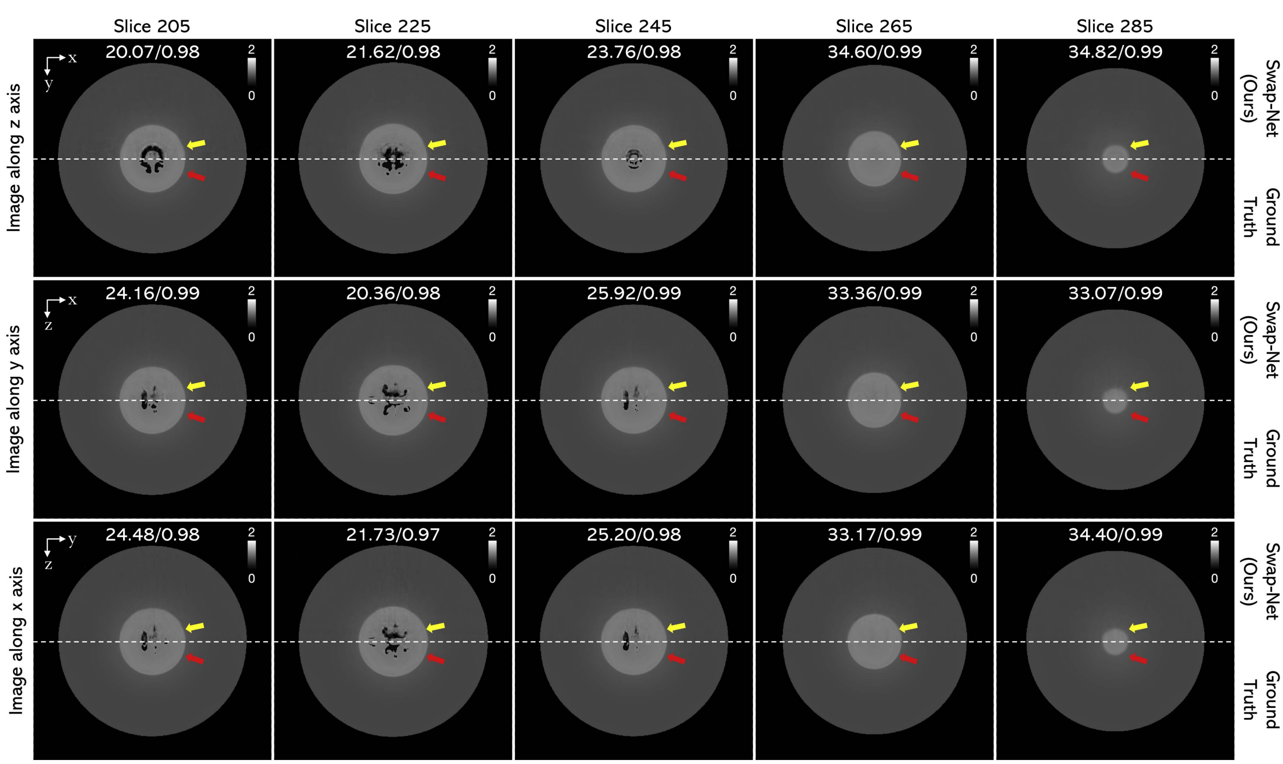


Fig. 10. Visual evaluation of Swap-Net across different slices on an exemplar ICF double shell test simulation with 4 projection views under AWGN corruption. Each row shows different slices of the whole 3D object along z, y and x axes, respectively. In each row, the images to the top of the dashed line are the reconstructed images from Swap-Net, while the images to the bottom are ground truth. The top-middle part of each image provides the SNR and SSIM values with respect to the ground truth. Arrows in the plots highlight sharp edge regions that are well reconstructed using Swap-Net. Note the consistently good performance of Swap-Net across different slices of a 3D object.

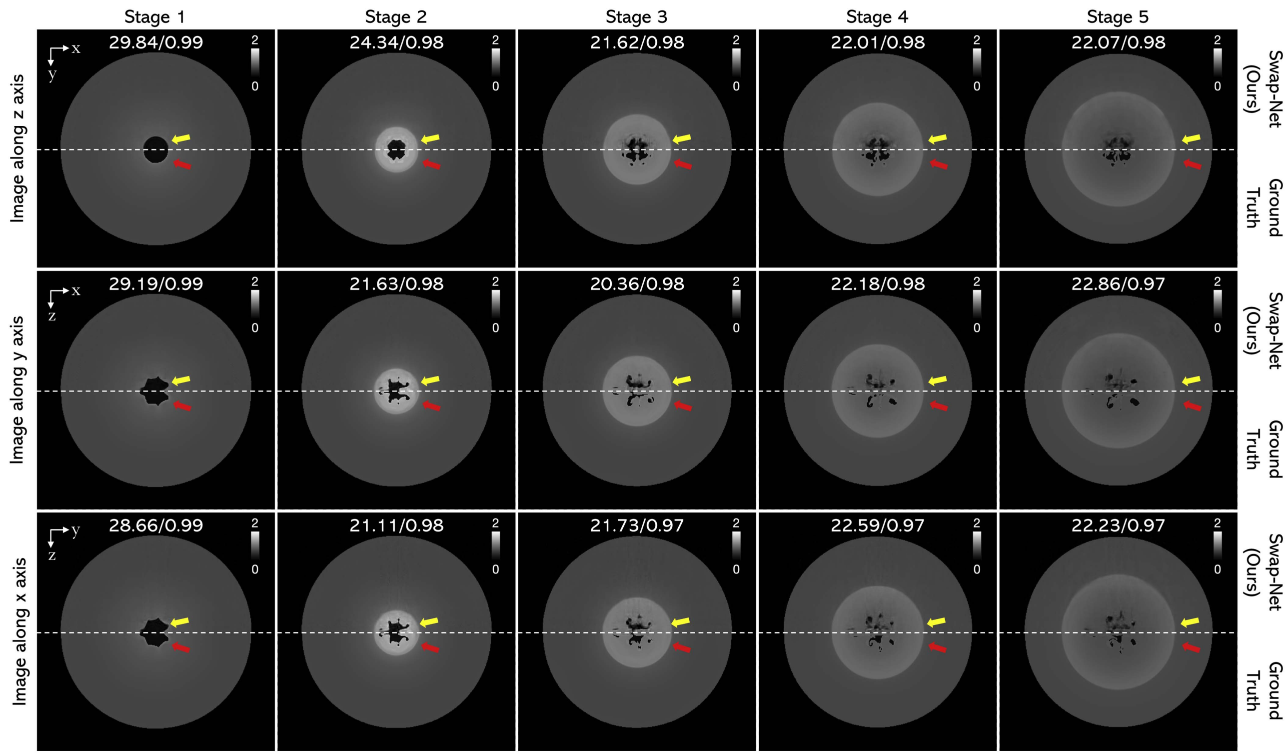


Fig. 11. Visual evaluation of Swap-Net on different exemplar ICF double shell test simulations with 4 projection views under AWGN corruption. Each row shows different stage of an exploding 3D object along z, y and x axes, respectively. In each row, the images above the dashed line are the images reconstructed from Swap-Net, while the images below are the ground truth. The top-middle part of each image provides the SNR and SSIM values with respect to the ground truth. Arrows in the plots highlight sharp edge regions that are well reconstructed using Swap-Net. Swap-Net had consistently good performance across different deformation stages of dynamic 3D ICF objects.

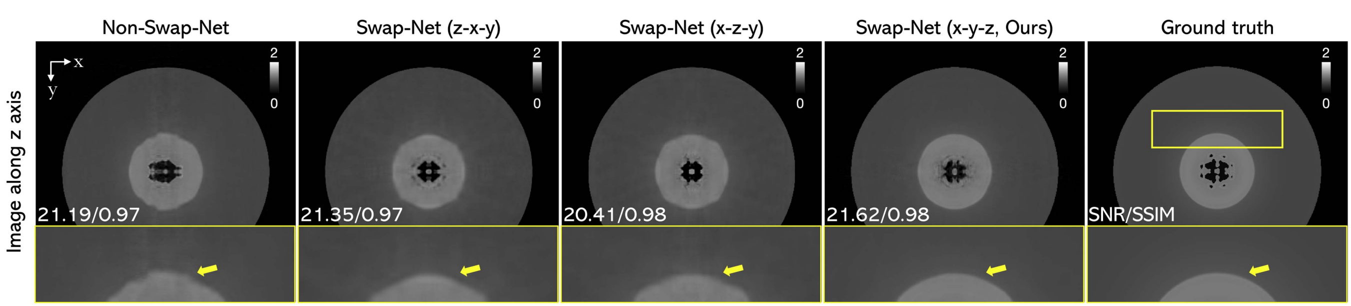


Fig. 12. Quantitative and visual evaluation of Swap-Net variants with different axes-swapping settings on an exemplar ICF double shell test simulation with 4 projection views under AWGN corruption. The middle slice of the whole 3D object along z axis is plotted. The bottom-left corner of each image provides the SNR and SSIM values, and 2× zoomed-in region. Arrows in the zoomed-in plots highlight sharp edges that are well reconstructed using Swap-Net with swapping order x-y-z. Note the the improvement from non-Swap-Net to Swap-Net variants, and the the influence of the order of axes-swapping operations in the reconstruction.

Non-Swap-Net helps to illustrate improvements due to axes-swapping operations. We also tested Swap-Net with different axes-swapping orders. Given that z is the CBCT axis around which the X-ray sources and detectors were placed, we checked the the following axes-swapping orders: (a) z-x-y, (b) x-z-y, and (c) x-y-z, namely putting the convolution in the xy plane in the beginning, middle, and end of the Swap-Net pipeline. Order (c) x-y-z is the strategy adopted in our paper. Fig. 12 presents the reconstruction performance of those Swap-Net variants; it shows z-axis slices (similar results were observed for images along x and y axes and therefore were omitted here). Clearly, Non-Swap-Net gave the worst results with obvious artifacts, and Swap-Net with axes-swapping orders of z-x-y and x-z-y did not perform as well as the x-y-z order. We hypothesize this is because the relative worse FBP reconstruction along the z axis makes the learning in xy plane more challenging, so putting the convolution in xy plane at the end of the Swap-Net pipeline allows it to exploit the intermediate object reconstruction.

2) Benefits of the Cascading Architecture: As defined in (7), Swap-Net employs a cascading architecture where the output from each reconstruction stage serves as the input for the next stage. This cascading structure enables the optimization of the final output in a block-by-block manner. To highlight the advantages of this cascading design, we investigated a parallel

TABLE II
COMPARISON EVALUATION OF SWAP-NET WITH PARALLEL AND CASCADING ARCHITECTURE AVERAGED OVER THE TEST DATASET FOR 4 PROJECTION VIEWS UNDER AWGN CORRUPTION

Architectures	Parallel-Net	Swap-Net (Ours)
SNR (dB)	24.08	**28.22**
SSIM	0.94	**0.99**

TABLE III
QUANTITATIVE EVALUATION OF SWAP-NET ON THE WALNUT TEST DATASET FOR DIFFERENT PROJECTION VIEWS

Metric	SNR (dB)		
Views	32	48	64
FBP	10.44	11.06	11.16
Swap-Net (Ours)	13.96	15.29	15.61

workflow, defined as:

$$\widehat{\boldsymbol{x}} \leftarrow (\mathsf{B}_{\boldsymbol{\theta}^1}(\boldsymbol{A}^\dagger\boldsymbol{y};x) + \mathsf{B}_{\boldsymbol{\theta}^2}(\boldsymbol{A}^\dagger\boldsymbol{y};y) + \mathsf{B}_{\boldsymbol{\theta}^3}(\boldsymbol{A}^\dagger\boldsymbol{y};z))/3\,. \tag{13}$$

Here, three blocks, $\mathsf{B}_{\boldsymbol{\theta}^1}$, $\mathsf{B}_{\boldsymbol{\theta}^2}$, and $\mathsf{B}_{\boldsymbol{\theta}^3}$, share the same architecture as the blocks in Swap-Net but take the same input $\boldsymbol{A}^\dagger\boldsymbol{y}$ and operate convolutions for each axis independently in a parallel manner. The final reconstruction, $\widehat{\boldsymbol{x}}$, is obtained by averaging the outputs from each block. We refer to this architecture as Parallel-Net and summarized its performance comparison with Swap-Net in Table II. The results show that Swap-Net significantly outperformed this parallel variant, demonstrating the benefits of its cascading architecture.

3) Additional Evaluation on the Walnut Dataset: To demonstrate the practical applicability of our method beyond the limited-view ICF application, we incorporated additional views and conducted experiments on the public Walnut dataset [69]. We scaled the 42 walnut volumes to a size of $256 \times 256 \times 256$ and split them into 90% for training and 10% for testing. We experimented with the CBCT geometry setup used in our ICF experiments, where the 2D projections had dimensions of 200×200 with a resolution of $2500 \times 2500\,\mu\mathrm{m}^2$, corresponding to a detector size of $50 \times 50\,\mathrm{cm}^2$. We performed CBCT reconstruction using 32, 48, and 64 views. The testing results are summarized in Table III and an exemplar result of 64 views is shown in Fig. 13. These experiment further validate that Swap-Net is a robust and broadly applicable method for 3D object reconstruction.

4) Additional Comparison With Other Network Architectures: Given the wide range of 2D network variants available for CT reconstruction [70], [71], we selected the classical U-Net as a representative 2D network due to its popularity and well-documented strong performance. We further adopted the RED-CNN architecture from the related CT work [71] as an 2D baseline for additional comparison. The results, summarized in Table IV, show that RED-CNN fell short compared to our method.

5) Robustness Under Distribution Shifts: We have also conducted experiments to evaluate the robustness of our method under distribution shifts. Specifically, we trained our model on

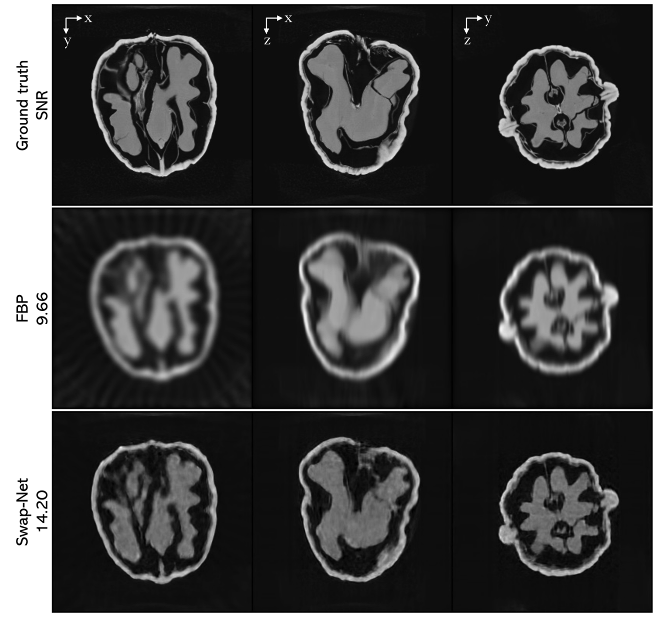


Fig. 13. An exemplar reconstruction from the Walnut dataset to demonstrate the practical applicability of our method beyond the sparse-view ICF application. SNR values computed over the full 3D walnut volumes are reported for each method.

TABLE IV
QUANTITATIVE EVALUATION OF SWAP-NET AND BASELINE METHODS RED-CNN AVERAGED OVER THE TEST DATASET FOR 4 PROJECTION VIEWS UNDER AWGN CORRUPTION

Methods	RED-CNN	Swap-Net (Ours)
SNR (dB)	25.37	**28.22**
SSIM	0.95	**0.99**

TABLE V
THE ROBUST PERFORMANCE OF SWAP-NET UNDER DISTRIBUTION SHIFTS BETWEEN TRAINING AND TESTING DATA

Scenarios	AWGN					
Metric	SNR (dB)			SSIM		
Level	35	40	45	35	40	45
Swap-Net	26.83	27.11	27.19	0.99	0.99	0.99
Scenarios	Scatter					
Metric	SNR (dB)			SSIM		
Level	high	medium	low	high	medium	low
Swap-Net	24.61	25.11	24.88	0.99	0.99	0.99

AWGN corruption at 40 dB and tested it with AWGN at 35 dB and 45 dB, simulating scenarios where the testing conditions are better or worse than the training ones. Similarly, we trained our model with medium scatter and tested it on low and high scatter levels (where the low and high scatter levels correspond to 0.8 and 1.2 times the training non-ideal blur and scatter and photon noise corruption level, respectively). The results, summarized in Table V, show that although the performance of our method fluctuated when the testing conditions differed

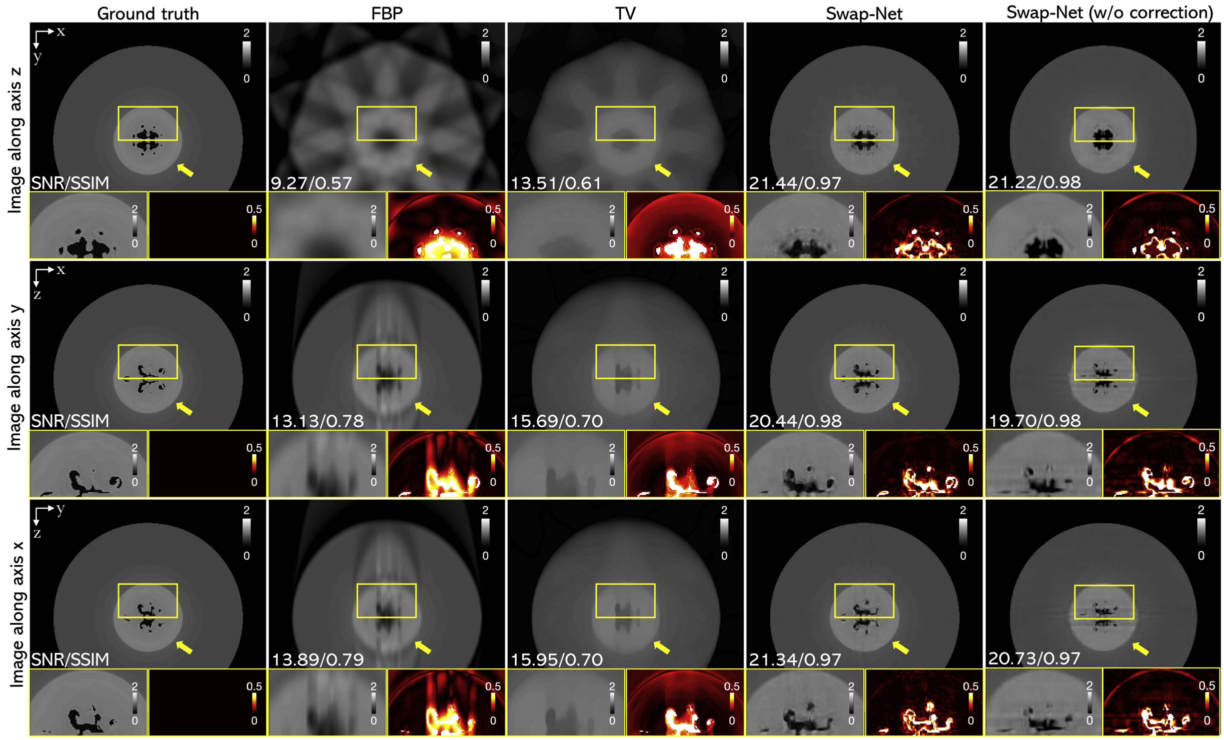


Fig. 14. Visual evaluation of Swap-Net on an exemplar ICF double shell test simulation with 4 projection views when a "perfect" scatter correction is applied to the baseline methods. Swap-Net with and without scatter correction are shown in column 4 and 5, respectively. Each row shows the middle slice of the whole 3D object along z, y and x axes, respectively. The bottom part of each image provides the SNR and SSIM values and representative $2\times$ zoomed-in regions and their error maps with respect to the ground truth. Arrows in the zoomed-in plots highlight sharp edges that are well reconstructed using Swap-Net. Note the superior quantitative and qualitative performance of Swap-Net even when baseline methods benefit from "perfect" motion correction.

from the training conditions, it remained within a reasonable range. This demonstrates that our method is capable of handling potential mismatches between training and testing deployments. Given that it is common practice to train and deploy models under similar data distributions, we expect Swap-Net to deliver robust performance in the presence of reasonable mismatches between training and testing distributions.

6) Capability of Reconstruction With Artifact Correction: To further validate the performance of our method in reconstruction with artifact correction, we conducted experiments without the non-ideal blur and scatter and photon noise, simulating a "perfect" artifact correction process before applying the baseline methods, FBP and TV. The results, presented in Fig. 14, show that FBP and TV (columns 2 and 3) incorporated correction for the non-ideal blur and scatter and photon noise, while our Swap-Net (column 5) did not. As a comparison, Swap-Net with the artifact correction was also plotted in the figure (column 4). Notably, Swap-Net still outperformed these baseline methods, even when they benefit from "perfect" artifact correction. This result demonstrates the joint reconstruction and artifact correction performance of our approach.

7) Integration With the Model-Based Reconstruction: As a model-free method, Swap-Net can be directly applied without requiring knowledge of the imaging model. Benefiting from its memory-efficient design, Swap-Net can also be used in deep model-based reconstruction frameworks when the information of the forward model $\boldsymbol{A}$ is provided. For example, deep unfolding (DU) is such a DMBA paradigm that was widely adopted in solving inverse problems due to its ability to provide a systematic connection between iterative algorithms and deep neural network architectures [72], [73], [74], [75]. We consider the following DU algorithm

$$\boldsymbol{x}^k \leftarrow \mathsf{R}_{\boldsymbol{\theta}}\left(\boldsymbol{x}^{k-1} - \gamma \nabla g(\boldsymbol{x}^{k-1})\right) , \tag{14}$$

where ∇g is the gradient of the data-fidelity term in (3), $\gamma > 0$ is the step size, and $\mathsf{R}_{\boldsymbol{\theta}}$ is a deep network module. This DU frameworks unroll the play-and-play priors (PnP) algorithm [76] and can be trained with fixed number of iterations in a supervised fashion. By using Swap-Net as $\mathsf{R}_{\boldsymbol{\theta}}$ and jointly training it with the measurement model, DU leads to an 3D model-based reconstruction method for the given CBCT inverse problem. To evaluate this capability, we ran the DU framework described in (14) with Swap-Net by setting the unfolding iteration $k = 4$ and stepsize $\gamma = 0.005$. We have also conducted the comparison with the PnP-BM3D approach, a method that uses BM3D [77] as a denoiser within the PnP framework. Our results showed that PnP-BM3D performed similarly to the TV methods within 1 dB difference in terms of SNR, while Swap-Net and its DU variant led to superior performance. Fig. 15 shows an exemplar testing result of this Swap-Net-empowered DU model. We notice that though improved, incorporating the imaging model did not significantly enhance reconstruction performance. We hypothesize that this is due to the limitations imposed by ultra-sparse measurements as well as the imperfect imaging

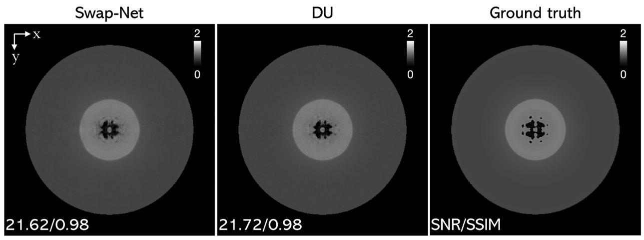


Fig. 15. Visual evaluation of Swap-Net and DU on an exemplar ICF double shell test simulation with 4 projection views under AWGN corruption. The middle slice of 3D object along z axis is plotted as an example. The bottom part of each image provides the SNR and SSIM values. The performance of Swap-Net was slightly improved when combined with the imaging forward model in the DU framework.

model, which restrict the contribution of the data consistency. This highlights its effectiveness as a standalone deep learning-based solution, particularly valuable for memory-efficient and model-agnostic applications in computational imaging. Consequently, Swap-Net is poised to make a significant impact across various physics domains, including material science investigations at the Advanced Photon Source and numerous national security applications within the Department of Energy (DOE) complex.

V. Conclusion

This paper presents a memory-efficient 2.5D network, namely Swap-Net, for handling 3D image reconstruction problems like sparse-view CBCT. The major challenge in this problem is to reconstruct high-quality 3D images efficiently and accurately when only a limited number of projections and training data are available, and when complicated corruptions are presented. Swap-Net uses a novel axes-swapping operation that allows for sequential convolution along all three dimension of a 3D object. We optimized the network weights by minimizing the loss between the output of the Swap-Net and the ground-truth 3D images on the training dataset using FBP reconstruction as inputs. We demonstrated the enhanced performance of our method on sparse-view 3D CBCT image reconstruction relative to model-based regularization (such as TV), 2D, and 3D CNNs under both AWGN and non-ideal physics including blur and scatter and photon noise corruptions. Our extensive validation elaborated the potential of Swap-Net on producing high-quality images from artifact-corrupted measurements. As demonstrated by our comprehensive evaluation, although this paper focuses on CBCT reconstruction and ICF application, our network can be extended to other 3D imaging applications.

In conclusion, our method exploits the lower computational cost of 2D convolution while bridging the gap to 3D convolution via axes-swapping operations, thereby offering a computationally efficient strategy for handling memory-demanding 3D reconstructions. Its simple yet effective design makes it practically appealing. Although Swap-Net generally improves reconstruction quality for the ICF capsule—particularly in capturing sharp edges—we acknowledge its limitations in recovering fine details within the central region, underscoring the challenges posed by extreme sparse-view measurements and limited data. Further improvement may be possible by increasing the depth of each convolution block in Swap-Net. This work kept the channel dimension to be the same value for all Swap-Net blocks for simplicity; future work could optimize the feature dimensions. It could also be interesting to explore connections between Swap-Net and tensor decomposition methods, e.g., [78]. Applying Swap-Net to other imaging tasks is planned in the future.

Acknowledgment

The authors acknowledge Jennifer Schei Disterhaupt from LANL for contributing to the code for the noise and scatter model, and Robert Reinovsky (LANL) for his support of the project.

Data availability statement: Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Disclosures statement: The authors declare no conflicts of interest.

References

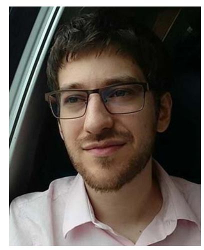
[1] T. Selig, T. Marz, M. Storath, and A. Weinmann, “Low-dose CT image reconstruction by fine-tuning a UNet pretrained for gaussian denoising for the downstream task of image enhancement,” 2024, *arXiv:2403.03551*.
[2] H. M. Alamri, M. Sadrameli, M. A. Alshalhoob, and M. A. Alshehri, “Applications of CBCT in dental practice: A review of the literature,” *Gen. Dent.*, vol. 60, no. 5, pp. 390–400, 2012.
[3] K. Horner, L. O’Malley, K. Taylor, and A. -M. Glenny, “Guidelines for clinical use of CBCT: A review,” *Dentomaxillofacial Radiol.*, vol. 44, no. 1, 2015, Art. no. 20140225.
[4] J. Casselman et al., “Cone beam CT: Non-dental applications,” *J. Belg. Soc. Radiol.*, vol. 96, no. 6, pp. 333–353, 2013.
[5] L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone beam algorithm,” *J. Opt. Soc. America A*, vol. 1, no. 6, pp. 612–619, Jun. 1984.
[6] X. Pan, E. Y. Sidky, and M. Vannier, “Why do commercial ct scanners still employ traditional, filtered back-projection for image reconstruction?,” *Inverse Problems*, vol. 25, no. 12, 2009, Art. no. 123009.
[7] K. Sauer and C. Bouman, “A local update strategy for iterative reconstruction from projections,” *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 534–548, Feb. 1993.
[8] K. Kim et al., “Sparse-view spectral CT reconstruction using spectral patch-based low-rank penalty,” *IEEE Trans. Med. Imag.*, vol. 34, no. 3, pp. 748–760, Mar. 2015.
[9] J. A. Fessler, M. Sonka, and J. M. Fitzpatrick, “Statistical image reconstruction methods for transmission tomography,” *Handbook Med. Imag.*, vol. 2, pp. 1–70, 2000.
[10] I. A. Elbakri and J. A. Fessler, “Statistical image reconstruction for polyenergetic X-ray computed tomography,” *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, Feb. 2002.
[11] J. -B. Thibault, C. A. Bouman, K. D. Sauer, and J. Hsieh, “A recursive filter for noise reduction in statistical iterative tomographic imaging,” *Proc. SPIE*, vol. 6065, 2006, Art. no. 60650X.
[12] M. Beister, D. Kolditz, and W. A. Kalender, “Iterative reconstruction methods in X-ray CT,” *Physica Medica*, vol. 28, no. 2, pp. 94–108, 2012.
[13] C. Zhang, T. Zhang, M. Li, C. Peng, Z. Liu, and J. Zheng, “Low-dose ct reconstruction via l1 dictionary learning regularization using iteratively reweighted least-squares,” *Biomed. Eng. Online*, vol. 15, pp. 1–21, 2016.
[14] W. Yu, C. Wang, X. Nie, M. Huang, and L. Wu, “Image reconstruction for few-view computed tomography based on l0 sparse regularization,” *Procedia Comput. Sci.*, vol. 107, pp. 808–813, 2017.
[15] C. Xu, B. Yang, F. Guo, W. Zheng, and P. Poignet, “Sparse-view CBCT reconstruction via weighted Schatten p-norm minimization,” *Opt. Exp.*, vol. 28, no. 24, pp. 35469–35482, 2020.
[16] R. Anirudh, H. Kim, J. J. Thiagarajan, K. A. Mohan, K. Champley, and T. Bremer, “Lose the views: Limited angle CT reconstruction via implicit sinogram completion,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 6343–6352.

[17] S. Ravishankar, J. C. Ye, and J. A. Fessler, "Image reconstruction: From sparsity to data-adaptive methods and machine learning," *Proc. IEEE*, vol. 108, no. 1, pp. 86–109, Jan. 2020.
[18] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessler, "PWLS-ULTRA: An efficient clustering and learning-based approach for low-dose 3D CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1498–1510, Jun. 2018.
[19] H. Chen et al., "LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1333–1347, Jun. 2018.
[20] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1440–1453, Jun. 2018.
[21] Z. Jiang, Y. Chen, Y. Zhang, Y. Ge, F. -F. Yin, and L. Ren, "Augmentation of CBCT reconstructed from under-sampled projections using deep learning," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2705–2715, Nov. 2019.
[22] A. Ziabari, D. H. Ye, S. Srivastava, K. D. Sauer, J. -B. Thibault, and C. A. Bouman, "2.5D deep learning for CT image reconstruction using a multi-GPU implementation," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, 2018, pp. 2044–2049.
[23] A. A. Yunker, B. R. Kettimuthu, and C. J. C. Roeske, "Low dose CBCT denoising using a 3D U-Net," in *Proc. 2024 IEEE Int. Conf. Acoust., Speech, Signal Process. Workshops*, 2024, pp. 85–86.
[24] J. R. Hershey, J. L. Roux, and F. Weninger, "Deep unfolding: Model-based inspiration of novel deep architectures," 2014, *arXiv:1409.2574*.
[25] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 18–44, Mar. 2021.
[26] A. Hauptmann et al., "Model-based learning for accelerated, limited-view 3-D photoacoustic tomography," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1382–1393, Jun. 2018.
[27] J. Adler and O. Öktem, "Learned primal-dual reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, Jun. 2018.
[28] J. Liu, Y. Sun, W. Gan, X. Xu, B. Wohlberg, and U. S. Kamilov, "SGD-Net: Efficient model-based deep learning with theoretical guarantees," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 598–610, 2021.
[29] S. Mukherjee, M. Carioni, O. Öktem, and C. -B. Schönlieb, "End-to-end reconstruction meets data-driven regularization for inverse problems," in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, vol. 34, pp. 21413–21425.
[30] J. Liu et al., "Online deep equilibrium learning for regularization by denoising," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, vol. 35, pp. 25363–25376.
[31] D. Wu, K. Kim, G. El Fakhri, and Q. Li, "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2479–2486, Dec. 2017.
[32] I. Y. Chun, Z. Huang, H. Lim, and J. A. Fessler, "Momentum-Net: Fast and convergent iterative neural network for inverse problems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4915–4931, Apr. 2023.
[33] B. Zhou, S. K. Zhou, J. S. Duncan, and C. Liu, "Limited view tomographic reconstruction using a cascaded residual dense spatial-channel attention network with projection data fidelity layer," *IEEE Trans. Med. Imag.*, vol. 40, no. 7, pp. 1792–1804, Jul. 2021.
[34] Y. Huang, A. Preuhs, G. Lauritsch, M. Manhart, X. Huang, and A. Maier, "Data consistent artifact reduction for limited angle tomography with deep learning prior," in *Proc. Int. Workshop Mach. Learn. Med. Image Reconstruction*, 2019, pp. 101–112.
[35] Y. Han, J. Kang, and J. C. Ye, "Deep learning reconstruction for 9-view dual energy ct baggage scanner," 2018, *arXiv:1801.01258*.
[36] S. Guan, A. A. Khan, S. Sikdar, and P. V. Chitnis, "Limited-view and sparse photoacoustic tomography for neuroimaging with deep learning," *Sci. Reports*, vol. 10, no. 1, 2020, Art. no. 8510.
[37] S. Lee, H. Chung, M. Park, J. Park, W. -S. Ryu, and J. C. Ye, "Improving 3D imaging with pre-trained perpendicular 2D diffusion models," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2023, pp. 10710–10720.
[38] A. Ziabari et al., "Enabling rapid X-ray CT characterisation for additive manufacturing using cad models and deep learning-based reconstruction," *npj Comput. Materials*, vol. 9, no. 1, 2023, Art. no. 91.
[39] E. C. Merritt et al., "Experimental study of energy transfer in double shell implosions," *Phys. Plasmas*, vol. 26, no. 5, 2019, Art. no. 052702.
[40] J. -B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Med. Phys.*, vol. 34, no. 11, pp. 4526–4544, 2007.
[41] J. A. Fessler, "Model-based image reconstruction for MRI," *IEEE Signal Process. Mag.*, vol. 27, no. 4, pp. 81–89, Jul. 2010.
[42] M. Sun and J. M. Star-Lack, "Improved scatter correction using adaptive scatter kernel superposition," *Phys. Med. Biol.*, vol. 55, no. 22, pp. 6695–6720, Oct. 2010.
[43] M. T. McCann, M. L. Klasky, J. L. Schei, and S. Ravishankar, "Local models for scatter estimation and descattering in polyenergetic X-ray tomography," *Opt. Exp.*, vol. 29, no. 18, pp. 29423–29438, 2021.
[44] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: An application to limited-angle tomography," *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 204–221, Feb. 1998.
[45] D. F. Yu and J. A. Fessler, "Edge-preserving tomographic reconstruction with nonlocal regularization," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 159–173, Feb. 2002.
[46] M. Pasha, A. K. Saibaba, S. Gazzola, M. I. Español, and E. de Sturler, "A computational framework for edge-preserving regularization in dynamic inverse problems," *Electron. Trans. Numer. Anal.*, vol. 58, pp. 486–516, 2023.
[47] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 14, pp. 259–268, Nov. 1992.
[48] E. Y. Sidky, C. -M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," J. X-ray *Sci. Technol.*, vol. 14, no. 2, pp. 119–139, 2006.
[49] G. T. Herman and R. Davidi, "Image reconstruction from a small number of projections," *Inverse Problems*, vol. 24, no. 4, 2008, Art. no. 045011.
[50] G. -H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, no. 2, pp. 660–663, 2008.
[51] J. Bian et al., "Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT," *Phys. Med. Biol.*, vol. 55, no. 22, 2010, Art. no. 6575.
[52] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 677–688, Mar. 2012.
[53] L. Pfister and Y. Bresler, "Model-based iterative tomographic reconstruction with adaptive sparsifying transforms," *Proc. SPIE*, vol. 9020, 2014, Art. no. 90200H.
[54] J. Dong, J. Fu, and Z. He, "A deep learning reconstruction framework for X-ray computed tomography with incomplete data," *PLoS One*, vol. 14, no. 11, 2019, Art. no. e0224426.
[55] H. Kim, R. Anirudh, K. A. Mohan, and K. Champley, "Extreme few-view CT reconstruction using deep inference," 2019, *arXiv:1910.05375*.
[56] J. C. Montoya, C. Zhang, Y. Li, K. Li, and G. -H. Chen, "Reconstruction of three-dimensional tomographic patient models for radiation dose modulation in CT from two scout views using deep learning," *Med. Phys.*, vol. 49, no. 2, pp. 901–916, 2022.
[57] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.
[58] X. Xu et al., "Learning-based motion artifact removal networks for quantitative r2∗ mapping," *Magn. Reson. Med.*, vol. 88, no. 1, pp. 106–119, 2022.
[59] D. Karimi and R. K. Ward, "Patch-based models and algorithms for image processing: A review of the basic principles and methods, and their application in computed tomography," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 11, pp. 1765–1777, 2016.
[60] S. Majee, T. Balke, C. A. J. Kemp, G. T. Buzzard, and C. A. Bouman, "4D X-Ray CT reconstruction using multi-slice fusion," in *Proc. IEEE Int. Conf. Comput. Photography*, 2019, pp. 1–8.
[61] A. Kofler, M. Haltmeier, T. Schaeffter, and C. Kolbitsch, "An end-to-end-trainable iterative network architecture for accelerated radial multi-coil 2D cine MR image reconstruction," *Med. Phys.*, vol. 48, no. 5, pp. 2412–2425, 2021.
[62] A. Lahiri, G. Maliakal, M. L. Klasky, J. A. Fessler, and S. Ravishankar, "Sparse-view cone beam CT reconstruction using data-consistent supervised and adversarial learning from scarce training data," *IEEE Trans. Comput. Imag.*, vol. 9, pp. 13–28, 2023.
[63] X. Xu, J. A. Fessler, M. Klasky, G. Sidharth, J. L. Schei, and M. T. McCann, "An end-to-end learning approach for subpixel feature extraction," presented at the Imag. Syst. Appl., Boston, MA, USA, 2023, Paper. JW2A.1.
[64] J. Adler, H. Kohr, and O. Öktem, "Operator discretization library (odl)," Jan. 2017. [Online]. Available: https://doi.org/10.5281/zenodo.249479
[65] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.- Assist. Intervention*, Munich, Germany, Oct. 2015, pp. 234–241.

[66] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Proc. Med. Image Comput. Comput.-Assist. Interv.–MICCAI: 19th Int. Conf.*, Athens, Greece, Springer, 2016, pp. 424–432.
[67] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations*, San Diego, CA, USA, May 2015, pp. 1–13.
[68] D. A. Serino, M. L. Klasky, B. T. Nadiga, X. Xu, and T. Wilcox, "Reconstructing Richtmyer–Meshkov instabilities from noisy radiographs using low dimensional features and attention-based neural networks," *Opt. Exp.*, vol. 32, no. 24, pp. 43366–43386, 2024.
[69] H. Der Sarkissian, F. Lucka, M. van Eijnatten, G. Colacicco, S. B. Coban, and K. J. Batenburg, "A cone-beam X-ray computed tomography data collection designed for machine learning," *Sci. Data*, vol. 6, no. 1, 2019, Art. no. 215.
[70] C.-M. Fan, T.-J. Liu, and K.-H. Liu, "SUNet: Swin transformer UNet for image denoising," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2022, pp. 2333–2337.
[71] H. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.
[72] J. Sun et al., "Deep ADMM-Net for compressive sensing MRI," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, vol. 29, pp. 10–18.
[73] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019.
[74] J. Zhang and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 1828–1837.
[75] J. Rudzusika, B. Bajić, T. Koehler, and O. Öktem, "3D helical CT reconstruction with a memory efficient learned primal-dual architecture," *IEEE Trans. Comput. Imag.*, vol. 10, pp. 1414–1424, 2024.
[76] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. IEEE Glob. Conf. Signal Process. Inf. Process.*, Austin, TX, USA, Dec. 2013, pp. 945–948.
[77] A. Danielyan, V. Katkovnik, and K. Egiazarian, "BM3D frames and variational image deblurring," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1715–1728, Apr. 2012.
[78] M. E. Kilmer, L. Horesh, H. Avron, and E. Newman, "Tensor-tensor algebra for optimal representation and compression of multiway data," *Proc. Nat. Acad. Sci.*, vol. 118, no. 28, 2021, Art. no. e2015851118.

Xiaojian Xu received her B.E. degree in communication engineering in 2014 and M.E. degree in communication and information systems in 2017 from the University of Electronic Science and Technology of China, Chengdu, China. She earned her Ph.D. in Computer Science from Washington University in St. Louis, MO, USA, in 2022, and subsequently joined the University of Michigan, Ann Arbor, MI, USA, as a Postdoctoral Research Fellow in the Department of Electrical Engineering and Computer Science. She is currently an AI Scientist with GE HealthCare in Bellevue, WA, USA. Her research interests include computational imaging, deep learning, optimization, inverse problems, computer vision, and signal processing.

Marc L. Klasky is currently a Senior Scientist with Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA. He is developing new algorithms in tomographic analysis for the past 15 years to support analysis of images acquired with the Dual Axis Hydrodynamic Radiographic Test Facility. He has made major contributions in a number of diverse areas in computational physics including the development of inversion methods for spectral analysis of compton spectrometers, scatter treatment in radiographic analysis, resistive magneto-hydrodynamics, corrosion science, computational chemistry, and dynamic tomography. His research interests include the development of machine learning algorithms to enable parameter estimation in 3-D hydrodynamics simulation codes using sparse tomographic projections and the development of reduced order models to describe the evolution of hydrodynamic features.



Michael T. McCann (Member, IEEE) received the B.S.E. degree in biomedical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2010 and the Ph.D. degree in biomedical engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2015. He is currently a Staff Scientist with Applied Mathematics and Plasma Physics Group (T-5), Los Alamos National Laboratory, Los Alamos, New Mexico. His research focuses on using signal and imaging processing tools for scientific image reconstruction and analysis. He is an Associate Editor for the IEEE Transactions on Computational Imaging.

Jason Hu (Graduate Student Member, IEEE) is currently working toward the Ph.D. degree in electrical and computer engineering with the University of Michigan, Ann Arbor, MI, USA. His research interests include developing generative AI and ML based algorithms for image processing, using score-based diffusion models coupled with deep learning and optimization techniques to solve computational imaging problems with applications in medicine, physics, and computer vision, theoretical and foundational analysis of imaging.

Jeffrey A. Fessler (Fellow, IEEE) received the Ph.D. degree in EE from Stanford University, Stanford, CA, USA, in 1990. He is currently the William L. Root Distinguished University Professor of EECS with the University of Michigan, Ann Arbor, MI, USA. He is the Interim Chair of ECE. He joined UM as a Postdoctoral Fellow. He is a Professor in EECS, BME and Radiology. He is a Fellow of the IEEE, for his contributions to image reconstruction. He was the recipient of the Edward Hoffman Medical Imaging Scientist Award, IEEE EMBS Technical Achievement Award and the 2023 Steven S. Attwood Award, the highest honor for faculty in the UM College of Engineering. He is a Senior Associate Editor for the IEEE Transactions on Computational Imaging.