PRIOR: Prior-Regularized Iterative Optimization Reconstruction For 4D CBCT

Dianlin Hu, Yikun Zhang, Jin Liu, Yi Zhang, Jean-Louis Coatrieux, Yang Chen

PRIOR: Prior-Regularized Iterative Optimization Reconstruction for 4D CBCT

Dianlin Hu, Yikun Zhang, Jin Liu, Yi Zhang, *Senior Member, IEEE*, Jean Louis Coatrieux, *Fellow, IEEE*, Yang Chen, *Senior Member, IEEE*

***Abstract*—4D cone-beam computed tomography (CBCT) is an important imaging modality in image-guided radiation therapy to address the motion-induced artifacts caused by organ movements during the respiratory process. However, due to the extremely sparse projection data for each temporal phase, 4D CBCT reconstructions will suffer from severe streaking artifacts. Therefore, to tackle the streak artifacts and provide high-quality images, we proposed a framework termed Prior-Regularized Iterative Optimization Reconstruction (PRIOR) for 4D CBCT. The PRIOR framework combines the physics-based model and data-driven method simultaneously, with powerful feature extracting capacity, significantly promoting the image quality compared to single model-based or deep learning-based methods. Besides, we designed a specialized deep learning model named PRIOR-Net, which can effectively excavate the static information in the prior image reconstructed from the fully-sampled projections at the encoding stage to improve the reconstruction performance for individual phase-resolved images. Both the simulated and clinical 4D CBCT datasets were performed to evaluate the performance of the PRIOR-Net and the PRIOR framework. Compared with the advanced 4D CBCT reconstruction methods, the proposed methods achieve promising results quantitatively and qualitatively in streak artifact suppression, soft tissue restoration, and tiny detail preservation.**

***Index Terms*—4D CBCT imaging, sparse-view CT reconstruction, prior image, iterative optimization model, deep learning.**

This work was supported in part by the State Key Project of Research and Development Plan under Grant 2022YFC2401600, in part by the National Natural Science Foundation of China under Grant T2225025, in part by the Key Research and development Programs in Jiangsu Province of China under Grant BE2021703 and BE2022768. Yang Chen is with the Jiangsu Provincial Joint International Research Laboratory of Medical Information Processing, Southeast University, Nanjing, China, and also with the Laboratory of Image Science and Technology, the School of Computer Science and Engineering, Southeast University, Nanjing 210096, China (Corresponding author: Yang Chen; Email: chenyang.list.seu.edu.cn).

Dianlin Hu, Yikun Zhang and Yang Chen are with the Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China, and also with the Centre de Recherche en Information Biomedicale Sino-Francais (LIA CRIBs), Rennes, F-3500 France, and also with the Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing 210096, China.

Yi Zhang is with the College of Computer Science, Sichuan University, Chengdu 610065, China (e-mail: yzhang@scu.edu.cn).

Jin Liu is with the College of Computer and Information, Anhui Polytechnic University, Wuhu 241000, China.

Jean-Louis Coatrieux is with Inserm Emeritus Research Director at LTSI (U1099, Rennes, France) and Professor at the Computer Science Department of Southeast University, Nanjing, China).

I. Introduction

ON-board cone-beam computed tomography (CBCT) coupled with a linear accelerator has become a major tool for treatment planning in image-guided radiation therapy (IGRT) [1] by its potential to refine the irradiation plan while minimizing its impact on sensitive organs. However, due to the slow acquisition time (usually 3~4 mins for each scan and 1 min for a rotation) [2, 3], the organ and tissue movements during respiratory generate severe blurring artifacts of the thorax and upper abdomen. Such significantly motion-induced artifacts degrade the CBCT image quality and lead to loss of dose planning and delivery [4].

Later, the 4D CBCT imaging was proposed to address motion-produced artifacts [5]. The collected projection data are split into different groups according to various respiratory phases. Each group projection data can independently reconstruct the time-resolved images using a conventional analytic algorithm [6]. Unfortunately, the reconstructed images of each breathing phase usually still suffer from severe streak artifacts. To tackle this challenging problem and provide high-quality images, numerous algorithms have been proposed for 4D CBCT reconstruction. These methods can be categorized into two main types: motion compensated-based approaches and sparse projection reconstruction methods.

For the compensated-based methods, it is pivotal to compensate for the organ motion of individual phase images by utilizing the deformation vector fields (DVFs) [7-9]. Li *et al.* considered the correlations between images at different phases and integrated the motion model to enhance the 4D CBCT image quality [10]. For more accurate motion modeling, Wang *et al.* incorporated the motion model into the iterative reconstruction method and greatly improved the results [8]. Assisted by the patient-specific planning CT, a hybrid reconstructed method was investigated to make the registration accuracy acceptable in practice [11]. Furthermore, Liu *et al.* proposed a binning-free method based on the 5D model [12], reducing the unknown variables and improving image quality compared to [8, 10]. In general, the performance of the above-mentioned compensation-based methods relies on registration accuracy, which means these methods perform well when respiratory movements conform to the motion-compensation model and vice versa [13].

Conventional sparse-view CT reconstruction methods [14] can be applied to 4D CBCT [15]. These algorithms only

concentrated on single-phase image processing. However, 4D CBCT imaging methods also considered the temporal correlation between reconstructed images of different phases and achieved better results in tissue restoration [16]. Specifically, Jia *et al.* developed an iterative method by introducing the temporal non-local means (TNLM) regularization term for 4D CT, which can reconstruct all the phase images simultaneously and obtain superior performance compared to total variation and tight frames [17]. At the same moment, total variation (TV) was also employed for 4D cardiac imaging by extending spatial TV into the spatial-temporal domain and performed better in artifact suppression [18]. Also, the key to these iterative methods is to design an effective regularization constraint for exactly modeling the tissue features. Therefore, it still has room for improvement because of the limitation of the hand-crafted priors.

The prior image is another significant component to promote the image quality in some 4D CBCT reconstruction methods [16]. It is reconstructed from the fully-sampled projection data and has the static information of the organs and tissues [4, 19]. Noting the above property, Chen *et al.* utilized TV minimization to constrain the differences between the prior image and the time-resolved image to boost the performance. Experiments demonstrated that [19] could provide higher isotropic spatial resolution than TV. Later, to further strengthen the [19], Zhi *et al.* replaced the TV constraint with a spatiotemporal dictionary learning and performed well in detailed structure preservation [20]. Again, prior image-based methods also depend on the design of the regularization term.

Currently, convolutional neural network (CNN) has attracted great attention in medical imaging [15, 21-24], including 4D CBCT [25]. Unlike the mentioned model-based methods, CNN could automatically learn a regularization term from big data and bring promising results over traditional methods [26-30]. For instance, Beaudry *et al.* applied the Unet [31] to interpolate the missing measurements for each phase projection which led to promotion in streak artifact removal and noise reduction [32]. Prior images can also be integrated into the CNN-based methods. Zhi *et al.* not only encoded the degraded image but also encoded the prior image and put them together into the decoding step with considering the relationship between adjacent slices [4]. Experimental results claimed that [4] was robust to different datasets from different vendors. Based on the workflow of [8], Huang *et al.* conducted two deep neural networks to obtain high-quality DVFs and embedded them into [8] to produce final 4D images [33].

In medical imaging, classical model-based methods incorporated the data fidelity with prior knowledge have clear physical meanings, thereby being reliable in various clinical applications. However, they need a careful regularization term design and efficient computer implementation to face their computational loads. In contrast, despite its poor explanation capability, data-driven approaches construct a high-quality image space learned from a large amount of paired samples. Great efforts have been devoted to effectively combining the physics-driven and data-driven methods [34, 35]. Analytic Compressed Iterative Deep (ACID) [36, 37] is a typical hybrid framework initially invented against adversarial attacks, but it is also observed that ACID could improve image quality using a pre-trained neural network for sparse-view CT and fast MRI reconstruction.

Inspired by [36, 37], this paper proposes a Prior-Regularized Iterative Optimization Reconstruction (PRIOR) framework for 4D CBCT. PRIOR utilizes a well-trained neural network as the regularization constraint to improve the reconstruction image via an effective iterative strategy. The main contributions of PRIOR are two-fold. First, to excavate the static structural information of the prior image, we develop a PRIOR-Net. Even though CycN-Net [4] generated superior results than some advanced methods, its usage of prior images is inefficient. Unlike [4], the proposed PRIOR-Net utilizes an effective prior feature fusion (PFF) module, which encodes the degraded images and prior images and fuses them at the encoding step to enhance the feature extraction from the prior image. PIROR-Net could bring more promising reconstruction images compared to CycN-Net. Second, ACID [36, 37] is a simple but practical framework to be applied to inverse problems, which only need a pre-trained neural network. To employ ACID to 4D CBCT imaging and further improve its performance, we extend and modify it to formulate a more concise and strong method (PRIOR). Compared to [36, 37], the advantage of the PRIOR framework is mainly in the utilization of prior images. Specifically, PRIOR-Net can provide more effective regularization on the reconstructed image than the network used in ACID [36, 37]. Further, by introducing prior images into the iterative process, PRIOR leads to promotion in detail preservation without generating fake structures. Experimental results demonstrate that the PRIOR framework performs well in artifact removal, detail preservation, and tissue restoration.

The rest of this paper is organized as follows. Section II gives the background related to the 4D CBCT and ACID scheme. Section III describes architecture details of the PRIOR-Net and the implementation of the PRIOR framework. In section IV, simulated and clinical experiments are performed, and the ablation study is also conducted to explore the effectiveness of different modules in the PRIOR framework. In section V, we will discuss some issues and make conclusions.

II. Background

A. FDK Reconstruction

FDK (Feldkamp-Davis-Kress) is a widely used method for cone-beam CT reconstruction [6] based on circular orbit scanning. Its reconstruction process can be written as:

$$f_{FDK}(x,y,z) = \int_0^{2\pi} \frac{R^2}{(R+x\cos\beta+y\sin\beta)^2}\,\tilde{p}\big(\beta, a(x,y,\beta), b(x,y,z,\beta)\big)d\beta \quad (1)$$

with:

$$\tilde{p}(\beta,a,b) = \left(\frac{R}{\sqrt{R^2+a^2+b^2}}\cdot p(\beta,a,b)\right) * h(a) \quad (2)$$

$$a(x,y,\beta) = R\frac{-x\sin\beta + y\cos\beta}{R + x\cos\beta + y\sin\beta} \quad (3)$$

$$b(x,y,z,\beta) = \frac{zR}{R + x\cos\beta + y\sin\beta} \quad (4)$$

where $p(\beta,a,b)$ is the projection data with Cartesian coordinates a and b, β stands the scanning angular, R presents the distance between x-ray source to rotation center, $h(\cdot)$ is the ramp-filter and $*$ denotes the 1-D convolution and $f_{FDK}(x,y,z)$ is the reconstructed image.

FDK can provide high-quality images fast when the projection data is noiseless and complete.

B. Prior Image and Time-resolved Image Reconstruction

For 4D CBCT, the static 3D prior image (PI) $\in R^{M\times N\times S}$ can be reconstructed by the FDK algorithm [6] as below.

$$\mathrm{PI} = \mathrm{FDK}(y_F) \quad (5)$$

where y_F is the fully-sampled projection data that contains all breathing processes. Therefore, PI has motion blurring artifacts caused by organ and tissue movements and lacks dynamic changes, but with most structural features [4]. For time-resolved image reconstruction, the fully-sampled projections y_F are first grouped into different K phases according to the collected respiratory signals. Then each phase image can be independently reconstructed using the corresponding projection set as follows.

$$x_k = \mathrm{FDK}(y_k) \quad (6)$$

where y_k is the sub-projections belonging to k^{th} respiratory phase, and x_k presents the corresponding 3D reconstructed image, which has the same size as PI. The workflow of prior image and time-resolved image reconstruction is illustrated in Fig. 1.

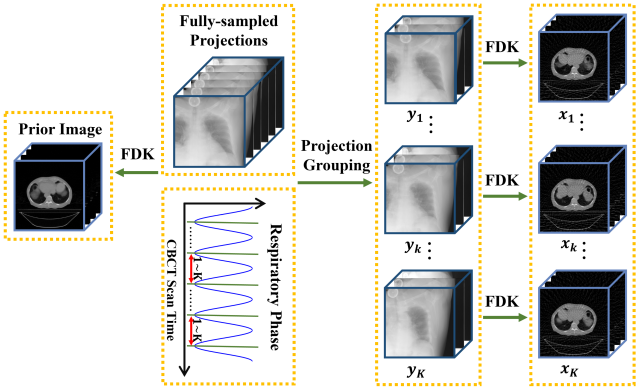

Fig. 1 The workflow of prior image and time-resolved image reconstruction.

C. ACID Scheme

The PRIOR framework is derived from the ACID framework [36, 37], whose aim is to overcome the instabilities of deep learning in medical imaging. In CT reconstruction, its objective function is given as follows.

$$\underset{p,x}{\operatorname{argmin}} \frac{1}{2}\|\Phi(Ax+p)-x\|_2^2 + \frac{\lambda}{2}\|y-Ax-p\|_2^2 + \frac{\mu}{2}\|p\|_2^2 + \zeta\|Hx\|_1 \quad (7)$$

where $\lambda > 0$, $\mu \geq 0$, and $\zeta > 0$ are hyperparameters to balance different constraint terms, x is the reconstructed image, y presents the collected projection, p stands for the error between x and y in sinogram domain, A is the system matrix, Φ indicates the well-trained neural network and it can map the projection data to an image, H is the discrete gradient transform.

Referring to [36, 37], the solution of p and x can be obtained as:

$$p^{n+1} = \frac{\lambda(y - Ax^n)}{1+\lambda} \quad (8)$$

$$x^{n+1} = H^{-1}S_{\frac{\zeta}{\lambda}}H\left(x^n + \frac{1}{\lambda}(\Phi(\mathrm{FDK}(p^{n+1})))\right) \quad (9)$$

H^{-1} is the inverse transform of H and $S_{\frac{\zeta}{\lambda}}$ is the soft-thresholding operator (the detailed mathematical derivation is referred to [36, 37]).

Unlike the traditional data fidelity term, ACID innovatively introduces a residual error p as the correction mechanism to improve the detail preservation and leads to impressive results in sparse-view CT reconstruction. Specifically, the Φ is a pre-trained neural network to remove streak artifacts (e.g. DDNet [38]). However, Φ usually encounters over-smoothing and detail missing. By incorporating Φ into the ACID scheme, the drawbacks of Φ can be overcome to some degree. Particularly, Eq. (8) can measure the distance between the reconstructed image and y in the sinogram domain and then be mapped into the error image by FDK algorithm. Next, the pre-trained Φ is used to boost the error image and get a more accurate approximation to the ground truth. Last, the soft-thresholding operation is to eliminate some noise or artifacts. Because p mainly indicates the high-frequency errors, ACID can bring improvements in detail preservation compared to Φ.

III. Proposed Method

A. PRIOR-Net

In this work, we propose a PRIOR-Net (as illustrated in Fig. 2) as the deep regularization constraint to improve the image quality for each respiratory phase reconstruction. The PRIOR-Net adopts U-Net-based architecture as the backbone, which has been widely used in image restoration [4, 34, 35]. Besides, to enhance the feature extraction, accelerate the information flow and the gradient backpropagation, skip connection, and residual learning [39] are also utilized in the PRIOR-Net. Moreover, PRIOR-Net employs a global residual connection to decrease the learning difficulty, promote convergence and improve generalization. Noting that there are no additional operations on the prior image and single-phase image before being input into the PRIOR-Net.

B. Prior Feature Fusion Module

The prior image contains static structural information but has motion-induced artifacts. Conversely, the time-resolved image introduces severe streak artifacts yet indicates dynamic tissue changes. Therefore, taking the degraded and prior images together as input for the CNNs is a popular strategy [4, 40]. For instance, CycN-Net independently encodes the prior

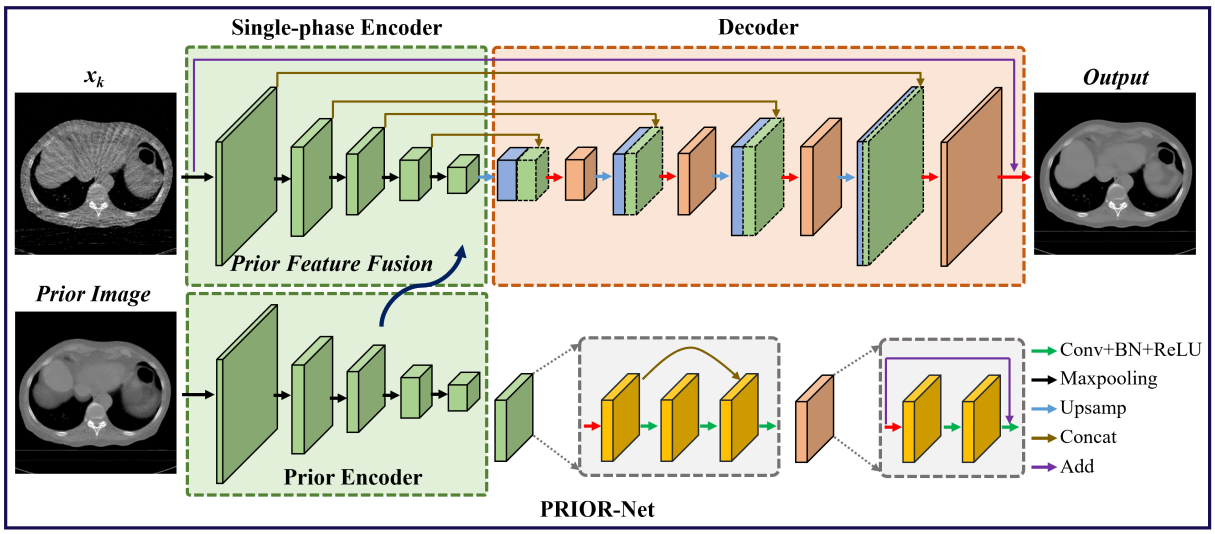

Fig. 2. The architecture of the proposed PRIOR-Net.

and single-phase images and decodes those two feature maps jointly to get the final high-quality images. Nevertheless, this scheme still has room for improvement that the prior and degraded images can be further uniformly processed at the encoding stage. Thus, our PRIOR-Net designs a specialized prior feature fusion (PFF) module with a more efficient static structural detail extraction from prior images. Specifically, the PFF module (as depicted in Fig.3) has two encoders, a single-phase encoder for time-resolved images and a prior encoder for prior images, respectively. Both two encoders own four downsampling operations to generate feature maps with different spatial resolutions. The outputted feature maps of the prior encoder are concatenated to the exported feature maps of the single-phase encoder at the same scale and then these fused features are input into the next block. That means inputted feature maps of each block in the single-phase encoder share both dynamic information and static structures. Meanwhile, aided by skip connection, the merged feature maps are delivered to the decoder step. Consequently, the PFF module could significantly improve the usage of prior images and achieve better reconstruction results than CycN-Net in artifact reduction and small detail restoration.

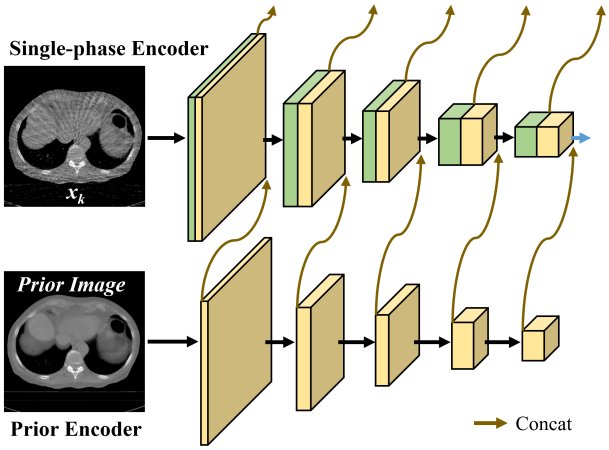

Fig. 3 Illustration of prior feature fusion module.

C. Loss Function for PRIOR-Net

The cost function of PRIOR-Net is defined as:

$$L = \sum_{k=1}^{K} \|I_k - \Psi(x_k, \mathrm{PI})\|_F^2 \tag{10}$$

where I_k is the reference image at the k^{th} respiratory phase, Ψ stands for the PRIOR-Net. In this work, only one PRIOR-Net model is trained to improve image quality for all phases.

D. PRIOR Framework

Although ACID can leads to improvements in small structure restoration, it will introduce fake textures when the projection data is extremely sparse. Therefore, based on Eqs. (8)-(9), we specifically made two modifications to construct the PRIOR framework (as demonstrated in Fig. 4) for 4D CBCT. The first one is to simplify the update procedure. Observing that the hyper-parameter λ mainly plays the role of controlling the update step, the weighted operation for projection error calculation in Eq. (8) can be eliminated. The second one is to expand Eq. (9) to fit the 4D CBCT imaging. That means the prior image is integrated with the iterative process as well. The iterative process of the proposed PRIOR framework is expressed as follows.

For the 1^{st} iteration, the intermediate reconstructed image can be obtained directly by:

$$x_k^1 = \Psi(g_k^1, m_k^1) = \Psi(x_k^0, \mathrm{PI}) \tag{11}$$

where x_k^0 is reconstructed by FDK of the k^{th} respiratory phase.

For the n^{th} iteration, we first calculate the error image with the following expression:

$$g_k^n = \mathrm{FDK}(y_k - \mathrm{A_k} x_k^{n-1}) \tag{12}$$

where $\mathrm{A_k}$ is the projection matrix for the k^{th} respiratory phase. g_k^n indicates the distance between the x_k^{n-1} and the reference image.

Then, the other error image is computed to compensate for the missing static structural details.

$$m_k^n = x_k^{n-1} - \mathrm{PI} \tag{13}$$

Next, the g_k^n and m_k^n are fed into the PRIOR-Net to get a more accurate difference map than g_k^n. In practice, the value

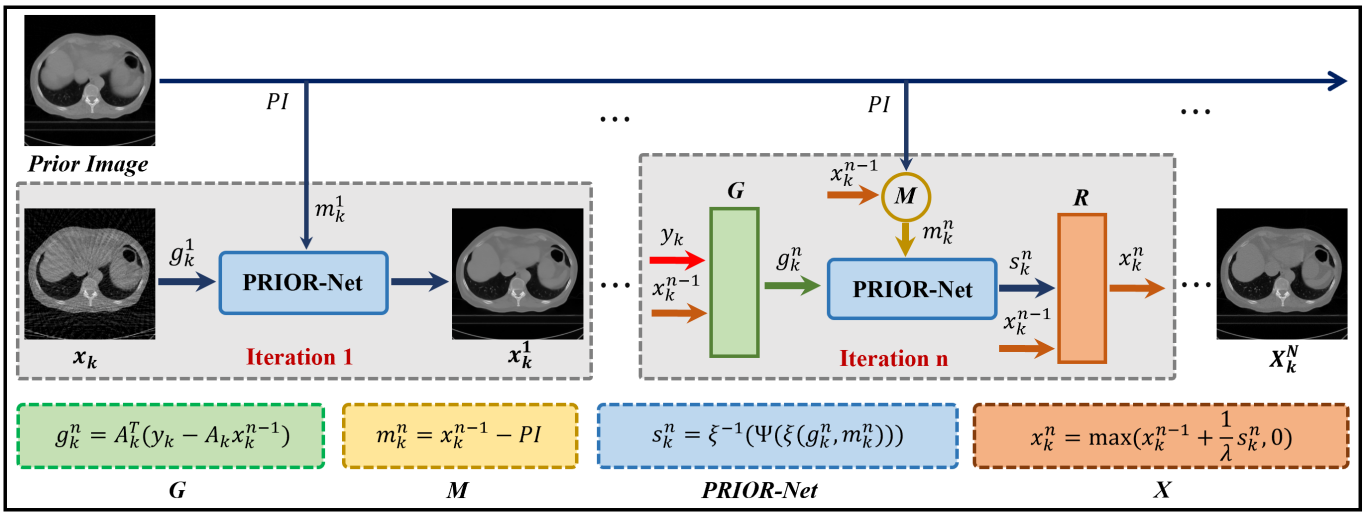

Fig. 4. The flowchart of the proposed PRIOR framework.

ranges of g_k^n and m_k^n are varied at different iterations, while the PRIOR-Net is sensitive to the pixel values. Hence, before inputting them into the PRIOR-Net, the scaling transform ξ is conducted on the input data to make the PRIOR-Net work effectively. So, the corrected difference image s_k^n can be gained after the inverse transform ξ^{-1} as shown below.

$$s_n^k = \xi^{-1}(\Psi(\xi(g_k^n, m_k^n))) \tag{14}$$

Last, we adopt a positive constrain on the weighted sum of x_k^{n-1} and s_k^n to acquire the reconstructed image at the n^{th} iteration for the k^{th} breathing phase:

$$x_k^n = \max\left(x_k^{n-1} + \frac{1}{\lambda} s_k^n, 0\right) \tag{15}$$

Through Eqs. (11)-(15), the reconstructed image can be gradually improved.

IV. Experimental Results

A. Setup

1) Dataset Acquisition

In this study, two types of datasets were performed to validate the performance of different methods, including simulated data and clinical data.

Simulated data came from the 4D-Lung Cancer Imaging Archive (https://wiki.cancerimagingarchive.net/display/Public/4D-Lung). It had twenty high-quality 4D lung images and each of which had ten volumetric images corresponding to ten breathing phases. By setting the respiratory signal for high-quality 4D CT images, the corresponding single-phase degraded images and prior images can be simulated. Last, the high-quality 4D images were treated as the reference images, and degraded and prior images were treated as the input data to train the PRIOR-Net. We selected seventeen patients (16870 images) for training and three patients (3060 images) for testing. Referred to the respiratory signal setting in [4], we also split one breathing cycle [0% 100%] into ten phases. The geometry configuration of the simulated cone-beam scan was as follows. The distance from the source to detector and isocenter was 1500 mm and 1000 mm, respectively. The detector had a size of 400×900, and each of them covered an area of 1.5×1.5 mm^2. The reconstructed volume had a spatial resolution of $0.9 \times 0.9 \times 0.9$ mm^3. For each scan, 600 projections were collected through 360°, which represented that a specific respiratory phase had 60 views. All the images in simulated data were generated by the full-scan mode.

Clinical data was provided by the SPARE Challenge (https://image-x.sydney.edu.au/spare-challenge/) [2]. And it contains the Elekta dataset and the Varian dataset. (a) For the Elekta dataset, one case was chosen to test the various methods. This case was acquired from an Elekta system using full-fan geometry through 199.4° gantry rotation. Specifically, the Elekta dataset gave two types of projection data, over-sampled, and down-sampled projection sets. Since there are no reference images, the over-sampled projection (1010 views) data were reconstructed as ground truth. And the down-sampled projection data (340 views) was performed for performance evaluation of comparisons. The related parameter settings were listed as follows. The respiratory phase was also set to ten. The source-to-detector and source-to-isocenter distances were 1536 mm and 1000 mm, respectively. The detector size was 512×512, and each pixel represented an area of 0.8×0.8 mm^2. The size of the reconstructed images was 256×256×200, and each voxel stood for 1×1×1 mm^3. (b) For the Varian dataset, one patient was selected to test the different methods. This data was provided by the Varian system using half-fan scan mode and the detector shifted 148 mm to increase the field of view. The detector had a size of 1024×728 and each pixel covered an area of 0.388×0.388 mm^2. The x-ray source to the rotation center and detector were 1000 mm and 1500 mm, respectively. 600 projections were collected via 360°. The size of the reconstructed image was $512 \times 512 \times 210$ and each voxel was $1 \times 1 \times 1$ mm^3. The breathing phase was ten. It was worth noting that clinical datasets were tested by the PRIOR-Net trained on the simulated images because clinical datasets have no ground truth images.

2) Comparison Methods & Evaluation Metrics

To validate the projection PRIOR-Net and PRIOR framework, PICCS [19], DDNet [38], and CycN-Net [4] were treated as comparisons. All the methods were performed on a

Table I
QUANTITATIVE EVALUATIONS OF DIFFERENT METHODS FROM SELECTED PHASES FOR THE SIMULATED DATASET (RMSE: HU)

	Metric	Gated-FDK	PICCS	DDNet	CycN-Net	PRIOR-Net	PRIOR
Phase1	RMSE	111.38	39.11	33.41	30.06	19.33	**15.49**
	PSNR	26.48	36.92	37.15	38.06	41.88	**43.84**
	SSIM	0.4542	0.8633	0.9435	0.9476	0.9703	**0.9726**
Phase3	RMSE	122.84	36.63	35.89	32.06	16.75	**12.66**
	PSNR	25.76	37.50	36.66	37.65	43.17	**45.59**
	SSIM	0.4535	0.8747	0.9475	0.9517	0.9782	**0.9806**
Phase5	RMSE	125.29	38.09	33.75	32.24	17.54	**12.19**
	PSNR	25.57	36.93	37.05	37.73	42.71	**45.82**
	SSIM	0.4541	0.8258	0.9495	0.9212	0.9771	**0.9800**
Phase7	RMSE	118.36	37.29	31.84	29.78	16.82	**11.56**
	PSNR	26.05	37.15	37.57	38.19	43.04	**46.25**
	SSIM	0.4605	0.8292	0.9527	0.9551	0.9781	**0.9820**
Phase9	RMSE	118.62	35.88	33.12	30.52	16.74	**12.08**
	PSNR	25.96	37.59	37.23	37.98	43.11	**45.94**
	SSIM	0.4595	0.8787	0.9505	0.9531	0.9782	**0.9817**

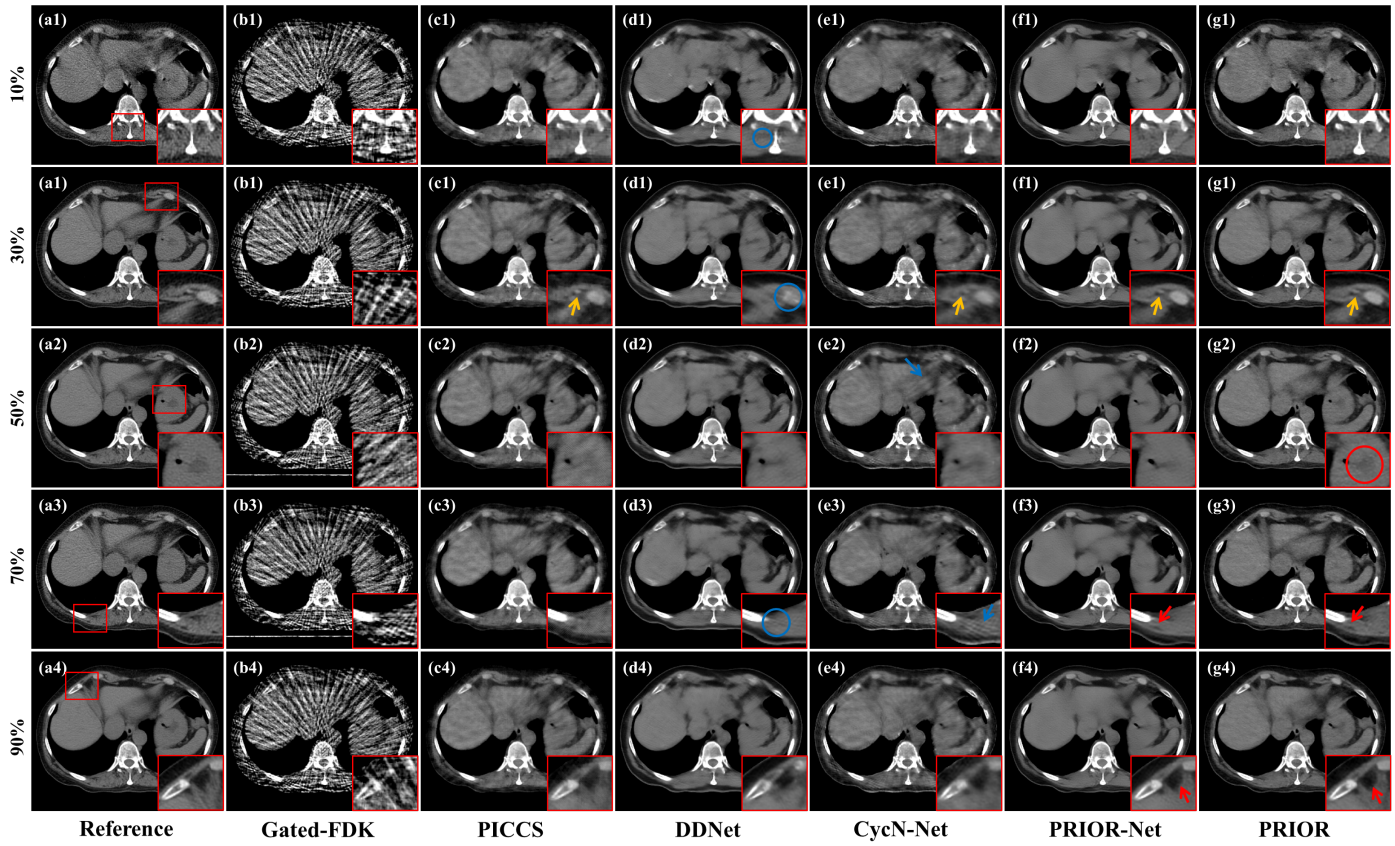

Fig. 5. Reconstructed results from four selected phases (30%, 50%, 70%, and 90%) at the same axial view of the simulated 4D CBCT dataset for different methods, including reconstructed images and magnified regions-of-interest (ROIs). (a1)-(a4) Reference images from 4D CT, (b1)-(b4) Images reconstructed by FDK algorithm, (c1)-(c4) Images reconstructed by PICCS method, (d1)-(d4) Images processed by DDNet, (e1)-(e4) Images processed by CycN-Net, (f1)-(f4) Images processed by PRIOR-Net, (g1)-(g4) Images processed by PRIOR framework. The display window of reconstructed results is [-160, 240] HU.

PC(CPU was Inter(R) Xeon(R) E5-2683, 2 GHz, GPU was NVIDIA GTX TITAN with 12G memory).

In our experiments, the PRIOR-Net was optimized using the Adam [41] algorithm with setting β_1=0.9 and β_2=0.999. The learning rate was initially set to 10^{-3} and slowly reduced to 10^{-6} with 50 epochs. The batch size was 2. It spent about 30 hours training the PRIOR-Net. Noting that all the reconstructed 3D results were processed by the PRIOR-Net slice by slice and then restored to the original 3D shape for the next step. For the PRIOR framework, the hyper-parameter λ was 15, and the iteration number N was 100. The codes of PRIOR-Net and PRIOR are available at https://github.com/lonelyatu/ PRIOR-Net.

We adopted the root mean square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) to evaluate different methods.

B. Simulated Data Results

Table I lists the average quantitative evaluations of volumetric images reconstructed by different methods for five

selected respiratory phases. It can be seen that FDK unavoidably gets the worst indicators in all cases. Considering the domain knowledge that the difference maps between time-resolved and prior images are sparse, PICCS brings significant improvements compared to FDK in terms of RMSE, PNSR, and SSIM via TV minimization. DDNet is a classical CNN-based sparse-view CT reconstruction method, and it can also promote the reconstructed results because of the powerful feature extraction capacity. Intuitively, the data-driven method should be superior to the traditional iterative approach [15, 26, 38]. However, comparing DDNet with PICCS, it can be found that these two methods have similar performance in some evaluations. That means the model-based method is competitive enough when the regularization term is designed suitably. This observation also claims the significance of prior images for the 4D CBCT reconstruction. Therefore, more encouraging results could be obtained when combing the prior image and CNN model. Specifically, CycN-Net outperforms PICCS and DDNet since it benefits from the feature extraction of CNNs as well as shares static structures from prior images. Furthermore, the proposed PRIOR-Net provides more promising assessments than CycN-Net. Two reasons can be attributed to that. First, PRIOR-Net adopts the PFF module, which utilizes the prior image with a more effective scheme than CycN-Net. Second, PRIOR-Net employs advanced techniques, such as batch normalization [42] and global residual learning, all of which have been validated in medical imaging. Based on the PRIOR-Net, the PRIOR framework still boosts the quantitative metrics, which proves the effectiveness of the iterative strategy in this study for 4D CBCT.

Fig. 5 depicts the qualitative results of different methods for the simulated 4D CBCT dataset at 30%, 50%, 70%, and 90% phases with the same axial view. Similar to the conclusion in Table I, FDK algorithm leads to severe streak artifacts due to the extremely sparse projection data. From Fig. 5(c1)-(c4), it can be observed that PICCS greatly improves the visual results in artifact removal and tissue restoration. Also, DDNet generates more smoothing results and fewer artifacts than PICCS. However, it loses some minor details (as marked by

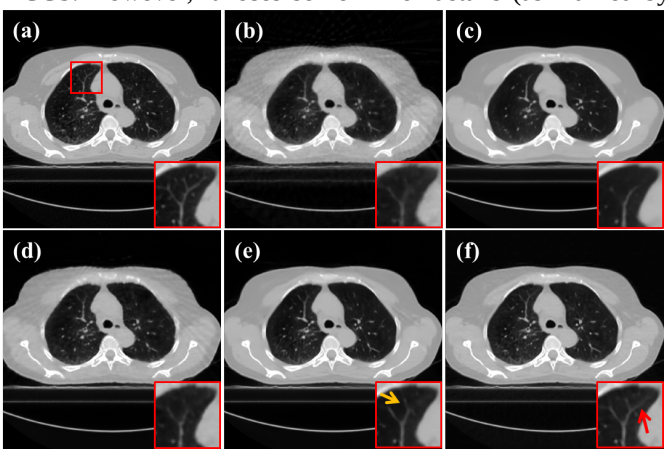

Fig. 6. Reconstructed results of different methods from the simulated 4D CBCT dataset at the 50% breathing phase. (a)-(f) Reference image and images processed by PICCS, DDNet, CycN-Net, PRIOR-Net and PRIOR algorithms. The display window of reconstructed results is [-900, 600] HU.

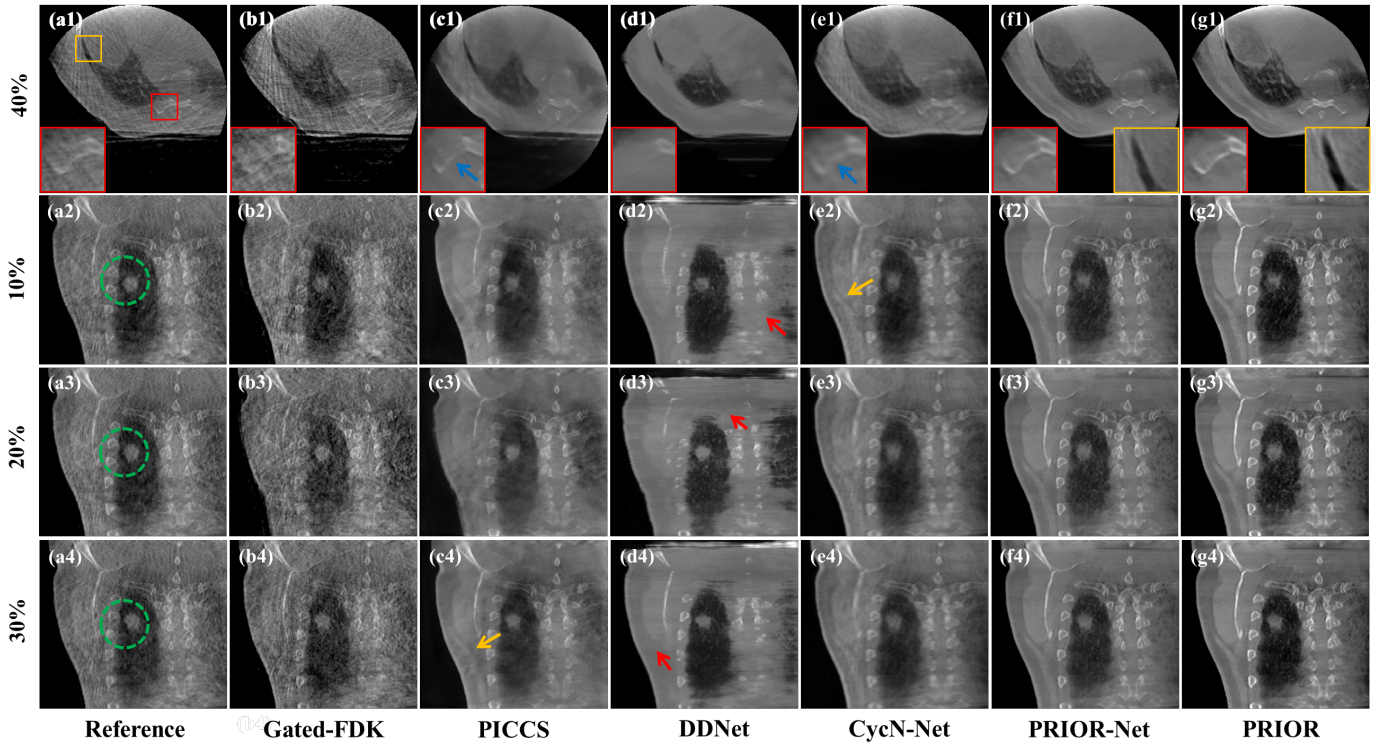

Fig. 7. Reconstructed axial and coronal results from four selected phases (10%, 20%, 30%, and 40%) of the clinical Elketa dataset for different methods, including reconstructed images and magnified ROIs. (a1)-(a4) Reference images reconstructed from over-sampled projections, (b1)-(b4) Images reconstructed by FDK algorithm, (c1)-(c4) Images reconstructed by PICCS, (d1)-(d4) Images processed by DDNet, (e1)-(e4) Images processed by CycN-Net, (f1)-(f4) Images processed by PRIOR-Net, (g1)-(g4) Images processed by PRIOR framework. The display window of reconstructed results is [0, 0.025] mm^{-1}.

blue circles in Fig. 5(d1)(d3)). As expected, all the prior image-based methods can recover tiny tissues (as pointed by yellow arrows in Fig. 5(c1)(e1)-(g1)), including PICCS, which states the necessity of prior images again. Unlike the remarkable performance in Table I, CycN-Net still reserves little streak artifacts (as indicated by blue arrows in Fig. (e2)(e3)). Compared to DDNet and CycN-Net, the proposed PRIOR-Net reduces all the streak artifacts and generates clearer tissue edges (as seen by the red arrow in Fig. 5(f4)). This implies that the prior image and powerful CNN models are mutually beneficial. Last, the PRIOR framework successfully preserves some unobvious lesions (as marked by the red circle in Fig. 5(g2)) using the well-trained PRIOR-Net.

Fig. 6 shows the reconstructed images of different methods from the simulated 4D CBCT dataset at 50% breathing phase. The PRIOR-Net gives more lung vessel structures (as observed by the yellow arrow in Fig. 6(e)). Further, the PRIOR framework provides the best vessel restoration (as pointed by the red arrow in Fig. 6(f)).

C. Clinical Data Results

Fig. 7 demonstrates the axial and coronal results of different methods for the clinical Elekta dataset at 10%, 20%, 30%, and 40% phases. In particular, the reference image is not the golden standard in this section, and it is the auxiliary guidance for the results provided by different methods. Again, DDNet works well in streak artifacts reduction compared to FDK (as shown in Fig. 7(d1)-(d4)). However, it misses many soft tissues and bone structures and results in over-smoothing regions (as illustrated by red arrows in Fig. (d2)(d4)). The reason is that the DDNet recognizes the normal structures as artifacts and removes them wrongly. PICCS remains stable in balancing the trade-off between artifact removal and tissue restoration in either simulated or clinical datasets, which indicates that the model-based method is robust to different imaging conditions. Taking the prior image as additional input, CycN-Net can preserve more bones than DDNet (as observed in the blue arrow in Fig. 7(e1)). However, PICCS and CycN-Net still generate over-smoothing regions (as demonstrated by yellow arrows in Fig. 7(c4)(e2)). The proposed PRIOR-Net and PRIOR framework generate clear edges and accurate tissue details. Moreover, Fig. 7(a2)-(a4) mark the tumor with green circles to observe its motion trajectory. It can be found that all the methods can reconstruct the tumor movements successfully. Specifically, the tumor provided by DDNet has high contrast than PICCS and CycN-Net. The PRIOR-Net can also generate high-quality tumor and the PRIOR method further improve the sharpness of tumor contour.

To probe the performance of proposed methods on different datasets, the experiments based on the Varian dataset were conducted. Fig. 8 illustrates the reconstructed results of different methods at 50%, 70%, and 90% phases. Compared to FDK, all the methods can significantly promote image quality.

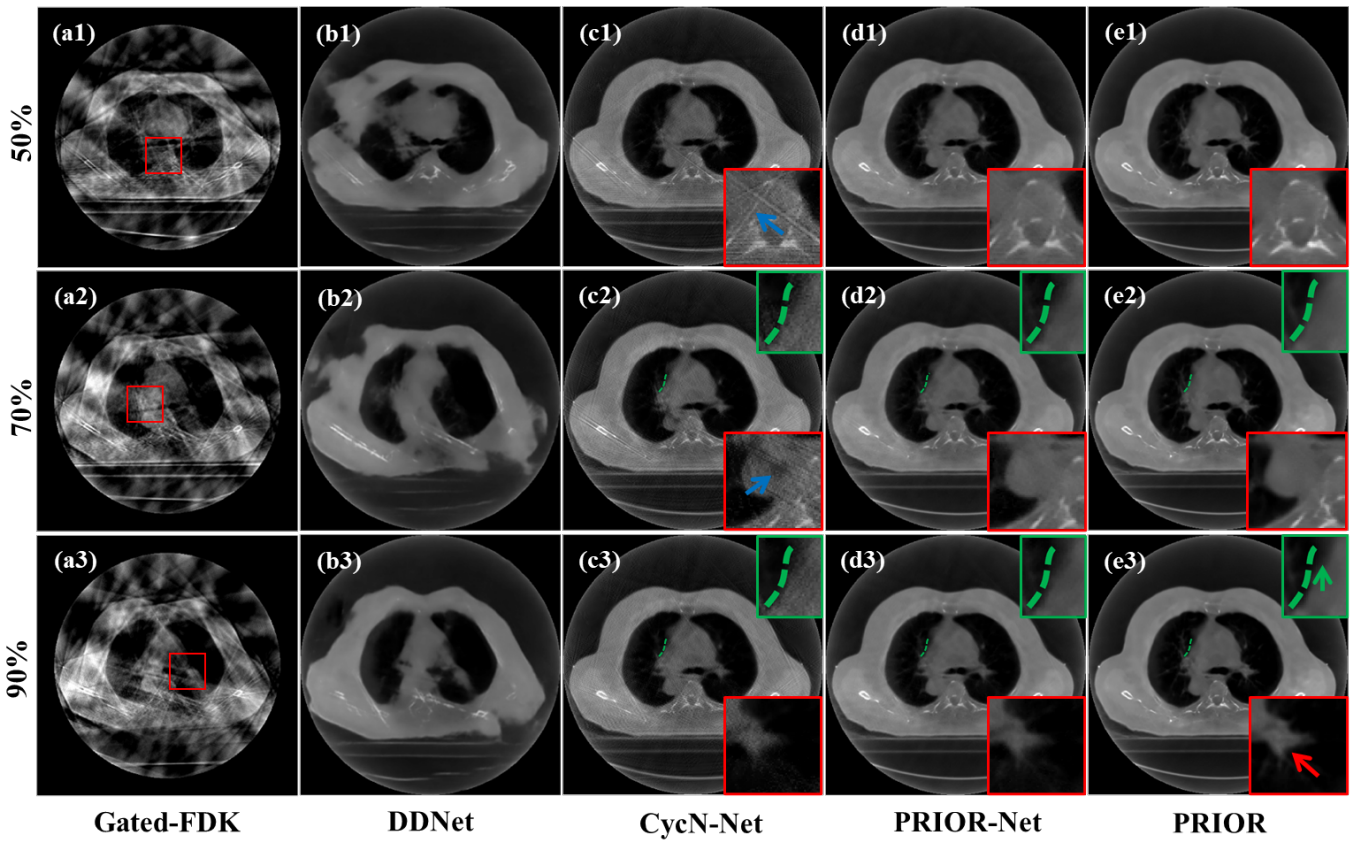

Fig. 8. Reconstructed results from three selected phases (50%, 70%, and 90%) of the clinical Varian dataset for different methods, including reconstructed images and magnified ROIs. (a1)-(a3) Images reconstructed by FDK algorithm, (b1)-(b3) Images reconstructed by DDNet, (c1)-(c3) Images processed by CycN-Net, (d1)-(d3) Images processed by PRIOR-Net, (e1)-(e3) Images processed by PRIOR framework. The display window of reconstructed results is [0, 0.03] mm^{-1}.

Nevertheless, from Fig. 8(b1)-(b3), it can be observed the results estimated by DDNet are degraded with residual artifacts and some organs are deformed severely. In contrast, prior-image-based methods not only eliminate the streaking artifacts but also accurately restore most of the tissue details. However, some artifacts and fake structures can be found in some regions processed by the CycN-Net (as pointed by blue arrows in Fig. 8(c1)(c2)). Because of the iterative process, the PRIOR method provides the clearest tissue edges (as shown by the red arrow in Fig. 8(e3)). To demonstrate the performance of different methods in the cardiac movement track, the green dash line is marked at the same position in different phases. Because the green line is fixed, therefore, the relative distance between the line and the cardiac indicates the cardiac movements. From Fig. 8(c2)(c3), the distance between the green line and cardiac has no obvious changes, which means CycN-Net fails to reconstruct the dynamic information. And the PRIOR-Net has the same defect. Comparing the zoomed ROIs in Fig. 8(e2)(e3), the cardiac motion can be observed noted by the green arrow in Fig. 8(e3), because the PRIOR framework involves data fidelity term optimization and brings accurate evaluation in organ motions.

D. Ablation Study

In this section, an ablation study was conducted to investigate the influence of different modules in the PRIOR framework on the reconstruction results based on the simulated 4D CBCT dataset. As for the ablation analysis, a progressive verification strategy was adopted.

The neural network without the prior encoder module based on the PRIOR-Net was taken as the baseline model. Then, the ACID scheme was adopted using the well-trained baseline model as the first comparison model. Next, the PFF module was added to the baseline model to establish the second comparison model (PRIOR-Net). Last, the PRIOR-Net was combined with the iterative optimization to construct the third comparison model, i.e., PRIOR.

Table II
QUANTITATIVE EVALUATIONS FOR THE PROGRESSIVE STUDY BASED ON THE SIMULATED DATASET (RMSE: HU)

	Metric	Baseline	ACID	PRIOR-Net	PRIOR
Phase1	RMSE	29.82	26.74	18.77	**15.15**
	PSNR	37.78	38.74	41.82	**43.85**
	SSIM	0.9463	0.9332	0.9736	**0.9760**
Phase4	RMSE	32.73	25.80	16.35	**11.45**
	PSNR	37.33	39.40	43.36	**46.45**
	SSIM	0.9492	0.9453	0.9800	**0.9818**
Phase7	RMSE	32.69	24.85	16.82	**11.69**
	PSNR	37.23	39.62	43.04	**46.20**
	SSIM	0.9490	0.9489	0.9786	**0.9810**

Table II gives the quantitative results of the progressive ablation study. We can see that the baseline model is close to the DDNet in Table I because they have similar architecture and techniques. Based on the baseline model, ACID performs better in RMSE and PSNR but compromises SSIM in some cases. This implies that the ACID strategy can indeed promote the accuracy of CT values. However, it may decrease the perceptual quality in structure preservation as the baseline model is ineffective and negatively affects it. Considering the prior image information, PRIOR-Net outperforms the first two models, stating the positive effect of prior images for 4D CBCT imaging convincingly. So, assisted with the powerful PRIOR-Net, the proposed PRIOR framework obtains the best performance in all metric evaluations.

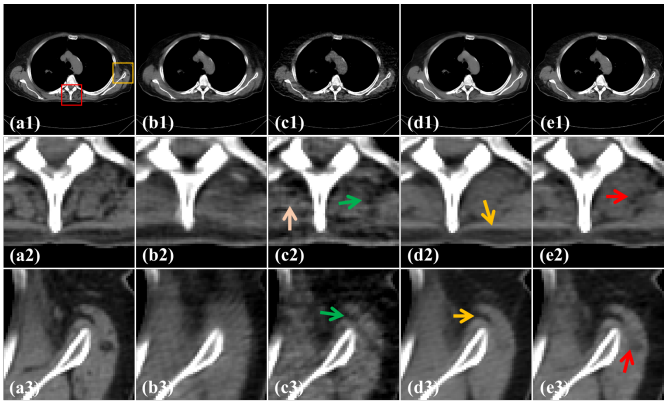

Fig. 9. Reconstruction results and zoomed ROIs from the simulated 4D CBCT dataset at 50% breathing phase. (a1)-(a2) Reference images, (b1)-(b3) Images processed by baseline model, (c1)-(c3) Images processed by ACID, (d1)-(d3) Images processed by PRIOR-Net, (e1)-(e3) Images processed by PRIOR. The display window is [-160, 240] HU.

Fig. 9 illustrates the reconstruction results of different comparison models based on the simulated 4D CBCT dataset at 50% phase. Like DDNet, the baseline model achieves the target of artifact removal but at the cost of missing detail. ACID preserves more tiny details through the forward and backward operation (as shown in green arrows in Fig. 9(c2)(c3)). Nevertheless, some fake structures maybe occur (as pointed by the pink arrow in Fig. 9(c2)), which is consistent with the observation in Table II. The yellow arrows in Fig. 9 show that PRIOR-Net generates more clear edges than ACID. Moreover, PRIOR restores small details (as pointed by red arrows in Fig. 9). The ablation study indicates that the modifications and extensions based on the ACID are potential for 4D CBCT.

E. Comparison of ACID and PRIOR

The previous section demonstrates that both ACID and PRIOR can improve the detail restoration based on a pre-trained model. This part aims to deeply explore their mechanisms. Fig. 10 presents the zoomed ROIs (as shown in Fig. 9) of different methods at different iterations. According to Eq. (8), p can evaluate the residual error between the reconstructed image and the ground truth in the sinogram domain. After mapping p into the error image and processing it by the pre-trained neural network, the small structures will be enhanced during the iterative process (as indicated by the red arrow in Fig. 10(e1)). However, ACID will introduce some fake textures (as demonstrated by yellow arrows in Fig. 10(e1)(d1)) when dealing with the error image because the pre-trained neural network was trained on the image domain and cannot generalize well in the error image domain. With the assistance of static information in prior images, PRIOR

also performs well in the error image space. Therefore, PRIOR can overcome the drawback of ACID and gradually improves the image quality without negative effects (as observed in Fig. 10(a2)-(e2)).

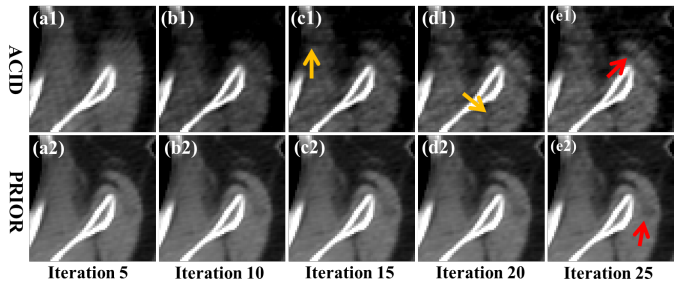

Fig. 10. Zoomed ROIs as marked in yellow square in Fig. 9(a1) of different methods at different iterations. The display window is [-160, 240] HU.

F. XCAT Phantom Results

To compare the CycN-Net and the proposed method fairly and convincingly, the experiments on the datasets utilized in CycN-Net were performed. Specifically, for simulated datasets, 15 patients (10210 images) for training, and 2 patients (1190 images) for validation. And 4D extended cardiac-torso (XCAT) phantom (1600 images) was used for the test (more description about datasets can be found in [4]). Because CycN-Net did not provide the projection data, we only provided the results of PRIOR-Net in this section.

Table III
QUANTITATIVE EVALUATIONS FOR DIFFERENT METHODS ON THE TRAINING DATASET

	Phase1		**Phase3**		**Phase5**	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
CycN-Net	27.44	0.8840	27.53	0.8896	27.23	0.8798
PRIOR-Net	**31.23**	**0.9105**	**31.62**	**0.9129**	**30.83**	**0.9065**

Table IV
QUANTITATIVE EVALUATIONS FOR DIFFERENT METHODS ON THE VALIDATION DATASET

	Phase1		**Phase3**		**Phase5**	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
CycN-Net	25.51	0.8629	25.72	0.8665	25.44	0.8671
PRIOR-Net	**28.17**	**0.8933**	**28.42**	**0.8979**	**27.87**	**0.8906**

Table V
QUANTITATIVE EVALUATIONS FOR DIFFERENT METHODS ON THE XCAT PHANTOM DATASET

	Phase1		**Phase3**		**Phase5**	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
CycN-Net	20.37	0.8419	20.41	0.8456	20.29	0.8401
PRIOR-Net	**22.40**	**0.8654**	**22.54**	**0.8702**	**22.61**	**0.8616**

Tables III, VI, and V list the quantitative results of different datasets for three selected respiratory phases. It can be found that the performance of all methods on the XCAT phantom dataset decrease greatly because XCAT phantom images are quite different from the simulated datasets. With more effective utilization of prior images, PRIOR-Net achieves better evaluation scores than CycN-Net in all datasets.

G. Convergence Analysis

The convergence analysis of the PRIOR framework is probed based on the simulated 4D CBCT dataset. Fig. 11 plots the RMSE with 100 iterations of four selected phases. We can see that all the RMSE values decrease quickly at initial iterations and tend to converge after 80 iterations. Moreover, PRIOR can significantly boost the results with 25%~31% improvements in RMSE value based on the PRIOR-Net results. These verify the practicability of the iterative scheme of PRIOR.

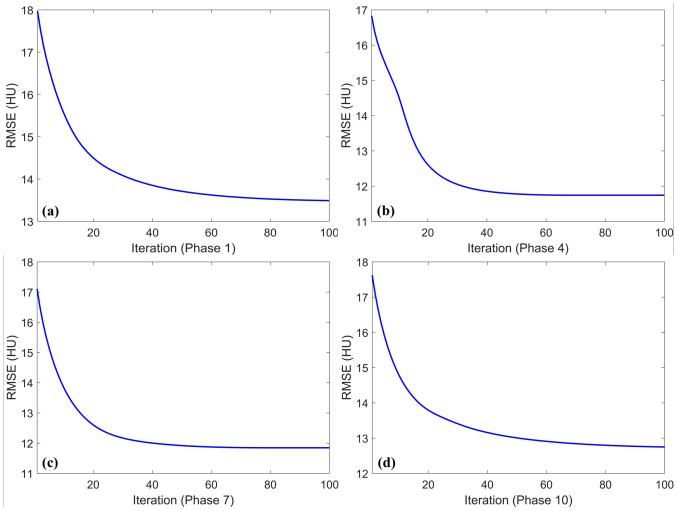

Fig. 11. Convergence analysis of RMSE values vs. iteration number of four selected phases based on the simulated 4D CBCT dataset.

H. Computational Cost

Table VI provides the computational cost of different deep learning-based methods on the Varian dataset for single-phase image reconstruction. All the image-domain-based methods can infer the results within a short time. Because of the iterative process, PRIOR takes much longer time than other methods. However, it can be reduced by adjusting the update size λ.

Table VI
COMPUTATIONAL COST OF DIFFERENT DEEP LEARNING METHODS (UINT: SECOND)

Method	DDNet	CycN-Net	PRIOR-Net	PRIOR
Time	10.64	14.22	12.49	2250.00

V. CONCLUSION AND DISCUSSION

Model-based methods have explainable physical meanings and rigorous mathematical forward-backward update operations. Hence, these methods are reliable in clinical applications. However, the hand-crafted regularization and empirical hyper-parameter selection may limit their performance. Recently, deep learning has been widely used in medical imaging [4, 15, 21, 26, 38]. Because of the powerful non-linear representation capacity, deep learning-based methods usually perform better than traditional methods [27, 40]. But owing to the black box property, these methods are hampered in the real scenario. Combing the advantages of physics- and data-driven methods, hybrid approaches have been developed and produced promising results [34, 35, 40].

Therefore, to improve the interpretability and performance of deep learning, we develop a PRIOR framework based on ACID [36] for 4D CBCT. The proposed PRIOR framework has two main components, i.e., PRIOR-Net and iterative strategy. Like to CycN-Net, PRIOR-Net also takes the prior image as input to extract the static structural features. Dissimilarly, PRIOR-Net designs a specialized PFF module to boost the utilization of prior images, leading to better results in tiny detail restoration and edge preservation. Then, the PRIOR framework adopts an iterative scheme inspired by [36, 37], which uses the pre-trained PRIOR-Net as the deep regularization to improve image quality gradually. Both simulated and clinical datasets are performed to prove the effectiveness of the PRIOR-Net and PRIOR framework in artifact removal and structure recovery.

Although the PRIOR framework shows improved results in 4D CBCT, some issues should be noticed. First, the performance of the PRIOR framework heavily depends on the pre-trained CNN model. However, the pre-trained PRIOR-Net cannot address various degradations from different datasets because the corresponding high-quality reference images are unavailable in reality to optimize the PRIOR-Net. For example, PRIOR brings fewer improvements in the clinical dataset than the simulated dataset. Therefore, how to improve the robustness of the PRIOR framework is an opening problem. Second, the PRIOR framework still lacks rigorous mathematical theory, and all the explanations are based on the intuitive and rough modification of existing algorithms. So, we cannot guarantee the worst performance of the PRIOR framework because the output of PRIOR-Net is uncontrollable when dealing with real scenarios. Third, in this study, we ignore the motion compensation-based approaches, which are effective for 4D CBCT. Therefore, we could attempt to incorporate these methods into the PRIOR framework in the future.

REFERENCES

[1] D. A. Jaffray, J. H. Siewerdsen, J. W. Wong *et al.*, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 53, no. 5, pp. 1337-1349, 2002.

[2] C. C. Shieh, Y. Gonzalez, B. Li *et al.*, "SPARE: Sparse-view reconstruction challenge for 4D cone-beam CT from a 1-min scan," *Medical Physics*, vol. 46, no. 9, pp. 3799-3811, 2019.

[3] T. Li, L. Xing, P. Munro *et al.*, "Four-dimensional cone-beam computed tomography using an on-board imager," *Medical Physics*, vol. 33, no. 10, pp. 3825-3833, 2006.

[4] S. Zhi, M. Kachelrieß, F. Pan *et al.*, "CycN-Net: A Convolutional Neural Network Specialized for 4D CBCT Images Refinement," *IEEE Transactions on Medical Imaging*, vol. 40, no. 11, pp. 3054-3064, 2021.

[5] L. Dietrich, S. Jetter, T. Tücking *et al.*, "Linac-integrated 4D cone beam CT: first experimental results," *Physics in Medicine Biology*, vol. 51, no. 11, pp. 2939, 2006.

[6] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *Journal of the Optical Society of America A*, vol. 1, no. 6, pp. 612-619, 1984.

[7] J. C. Park, J. S. Kim, S. H. Park *et al.*, "Motion-map constrained image reconstruction (MCIR): Application to four-dimensional cone-beam computed tomography," *Medical Physics*, vol. 40, no. 12, pp. 121710, 2013.

[8] J. Wang, and X. Gu, "Simultaneous motion estimation and image reconstruction (SMEIR) for 4D cone-beam CT," *Medical Physics*, vol. 40, no. 10, pp. 101912, 2013.

[9] M. Brehm, P. Paysan, M. Oelhafen *et al.*, "Artifact-resistant motion estimation with a patient-specific artifact model for motion-compensated cone-beam CT," *Medical Physics*, vol. 40, no. 10, pp. 101913, 2013.

[10] T. Li, A. Koong, and L. Xing, "Enhanced 4D cone-beam CT with inter-phase motion model," *Medical Physics*, vol. 34, no. 9, pp. 3688-3695, 2007.

[11] H. Yan, X. Zhen, M. Folkerts *et al.*, "A hybrid reconstruction algorithm for fast and accurate 4D cone-beam CT imaging(a))," *Medical Physics*, vol. 41, no. 7, Jul, 2014.

[12] J. Liu, X. Zhang, X. Zhang *et al.*, "5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography," *Inverse Problems*, vol. 31, no. 11, pp. 115007, 2015.

[13] J. Liu, Y. Kang, D. Hu *et al.*, "4D-CBCT Reconstruction via Motion Compensation Learning Induced Sparse Tensor Constraint," in 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2019, pp. 1-5.

[14] S. Niu, Y. Gao, Z. Bian *et al.*, "Sparse-view x-ray CT reconstruction via total generalized variation regularization," *Physics in Medicine*, vol. 59, no. 12, pp. 2997, 2014.

[15] Z. Jiang, Y. Chen, Y. Zhang *et al.*, "Augmentation of CBCT Reconstructed From Under-Sampled Projections Using Deep Learning," *IEEE Transactions on Medical Imaging*, vol. 38, no. 11, pp. 2705-2715, 2019.

[16] S. Zhi, M. Kachelrieß, and X. Mou, "High-quality initial image-guided 4D CBCT reconstruction," *Medical Physics*, vol. 47, no. 5, pp. 2099-2115, 2020.

[17] Z. Tian, X. Jia, B. Dong *et al.*, "Low-dose 4DCT reconstruction via temporal nonlocal means," *Medical Physics*, vol. 38, no. 3, pp. 1359-1365, 2011.

[18] L. Ritschl, S. Sawall, M. Knaup *et al.*, "Iterative 4D cardiac micro-CT image reconstruction using an adaptive spatio-temporal sparsity prior," *Physics in Medicine Biology*, vol. 57, no. 6, pp. 1517, 2012.

[19] 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," *Medical Physics*, vol. 35, no. 2, pp. 660-663, 2008.

[20] S. Zhi, M. Kachelrieß, and X. Mou, "A novel 4D-CBCT reconstruction Framework via a Joint of Spatiotemporal Dictionary Learning and Prior Constraint," in 6th International Conference on Image Formation in X-Ray Computed Tomography, Regensburg, Germany, 2020, pp. 4.

[21] H. Chen, Y. Zhang, M. K. Kalra *et al.*, "Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524-2535, 2017.

[22] K. Hammernik, T. Klatzer, E. Kobler *et al.*, "Learning a variational network for reconstruction of accelerated MRI data," *Magnetic Resonance in Medicine*, vol. 79, no. 6, pp. 3055-3071, Jun, 2018.

[23] Z. Han, H. Shangguan, X. Zhang *et al.*, "A Dual-Encoder-Single-Decoder Based Low-dose CT Denoising Network," *IEEE Journal of Biomedical and Health Informatics*, pp. 1-1, 2022.

[24] Y. Zhou, K. Yu, M. Wang *et al.*, "Speckle Noise Reduction for OCT Images Based on Image Style Transfer and Conditional GAN," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 139-150, 2022.

[25] S. Majee, T. Balke, C. A. J. Kemp *et al.*, "Multi-Slice Fusion for Sparse-View and Limited-Angle 4D CT Reconstruction," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 448-462, 2021.

[26] Y. Han, and J. C. Ye, "Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1418-1429, 2018.

[27] J. Schlemper, J. Caballero, J. V. Hajnal *et al.*, "A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 491-503, 2018.

[28] B. Zhu, J. Z. Liu, S. F. Cauley *et al.*, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487-492, 2018.

[29] Z. Huang, X. Liu, R. Wang *et al.*, "Learning a Deep CNN Denoising Approach Using Anatomical Prior Information Implemented With Attention Mechanism for Low-Dose CT Imaging on Clinical Patient Data From Multiple Anatomical Sites," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3416-3427, 2021.

[30] M. Wu, W. Chen, Q. Chen *et al.*, "Noise Reduction for SD-OCT Using a Structure-Preserving Domain Transfer Approach," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3460-3472, 2021.

[31] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention, 2015, pp. 234-241.
[32] J. Beaudry, P. L. Esquinas, and C.-C. Shieh, "Learning from our neighbours: a novel approach on sinogram completion using bin-sharing and deep learning to reconstruct high quality 4DCBCT," in Medical Imaging 2019: Physics of Medical Imaging, 2019, pp. 1094847.
[33] X. Huang, Y. Zhang, L. Chen *et al.*, "U-net-based deformation vector field estimation for motion-compensated 4D-CBCT reconstruction," *Medical physics*, vol. 47, no. 7, pp. 3000-3012, 2020.
[34] H. Gupta, K. H. Jin, H. Q. Nguyen *et al.*, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1440-1453, 2018.
[35] K. Zhang, Y. Li, W. Zuo *et al.*, "Plug-and-Play Image Restoration with Deep Denoiser Prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-1, 2021.
[36] W. Wu, D. Hu, W. Cong *et al.*, "Stabilizing deep tomographic reconstruction: Part A. Hybrid framework and experimental results," *Patterns*, vol. 3, no. 5, pp. 100474, 2022.
[37] W. Wu, D. Hu, W. Cong *et al.*, "Stabilizing deep tomographic reconstruction: Part B. Convergence analysis and adversarial attacks," *Patterns*, vol. 3, no. 5, pp. 100475, 2022/05/13/, 2022.
[38] Z. Zhang, X. Liang, X. Dong *et al.*, "A Sparse-View CT Reconstruction Method Based on Combination of DenseNet and Deconvolution," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1407-1417, 2018.
[39] K. He, X. Zhang, S. Ren *et al.*, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
[40] G. Chen, Y. Zhao, Q. Huang *et al.*, "4D-AirNet: a temporally-resolved CBCT slice reconstruction method synergizing analytical and iterative method with deep learning," *Physics in Medicine Biology*, vol. 65, no. 17, pp. 175020, 2020.
[41] D. P. Kingma, and J. Ba, "Adam: A Method for Stochastic Optimization," in ICLR, 2015.
[42] S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in International conference on machine learning, 2015, pp. 448-456.