Deep learning-based projection synthesis for low-dose cone-beam computed tomography imaging in image-guided radiotherapy

Xuzhi Zhao[1], Yi Du[2,3], Haizhen Yue[2], Ruoxi Wang[2], Shun Zhou[2], Hao Wu[2,3], Wei Wang[1], Yahui Peng[1]

[1]School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China; [2]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiation Oncology, Peking University Cancer Hospital & Institute, Beijing, China; [3]Institute of Medical Technology, Peking University Health Science Center, Beijing, China

Contributions: (I) Conception and design: X Zhao, Y Du, Y Peng; (II) Administrative support: H Wu, Y Peng; (III) Provision of study materials or patients: Y Du, H Yue, H Wu; (IV) Collection and assembly of data: X Zhao, Y Du, H Yue, R Wang, S Zhou, W Wang; (V) Data analysis and interpretation: X Zhao, Y Du, W Wang, Y Peng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yi Du, PhD. Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiation Oncology, Peking University Cancer Hospital & Institute, 52 Fucheng Road, Beijing 100142, China; Institute of Medical Technology, Peking University Health Science Center, Beijing, China. Email: yidu_rt@163.com; Yahui Peng, PhD. School of Electronic and Information Engineering, Beijing Jiaotong University, 3 Shangyuan Village, Beijing 100044, China. Email: 13717509106@139.com.

Background: The imaging dose of cone-beam computed tomography (CBCT) in image-guided radiotherapy (IGRT) poses adverse effects on patient health. To improve the quality of sparse-view low-dose CBCT images, a projection synthesis convolutional neural network (SynCNN) model is proposed.

Methods: Included in this retrospective, single-center study were 223 patients diagnosed with brain tumours from Beijing Cancer Hospital. The proposed SynCNN model estimated two pairs of orthogonally direction-separable spatial kernels to synthesize the missing projection in between the input neighboring sparse-view projections via local convolution operations. The SynCNN model was trained on 150 real patients to learn patterns for inter-view projection synthesis. CBCT data from 30 real patients were used to validate the SynCNN, while data from a phantom and 43 real patients were used to test the SynCNN externally. Sparse-view projection datasets with 1/2, 1/4, and 1/8 of the original sampling rate were simulated, and the corresponding full-view projection datasets were restored using the SynCNN model. The tomographic images were then reconstructed with the Feldkamp-Davis-Kress algorithm. The root-mean-square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM) metrics were measured in both the projection and image domains. Five experts were invited to grade the image quality blindly for 40 randomly selected evaluation groups with a four-level rubric, where a score greater than or equal to 2 was considered acceptable image quality. The running time of the SynCNN model was recorded. The SynCNN model was directly compared with the three other methods on 1/4 sparse-view reconstructions.

Results: The phantom and patient studies showed that the missing projections were accurately synthesized. In the image domain, for the phantom study, compared with images reconstructed from sparse-view projections, images with SynCNN synthesis exhibited significantly improved qualities with decreased values in RMSE and increased values in PSNR and SSIM. For the patient study, between the results with and without the SynCNN synthesis, the averaged RMSE decreased by 3.4×10^{-4}, 10.3×10^{-4}, and 21.7×10^{-4}, the averaged PSNR increased by 3.4, 6.6, and 9.4 dB, and the averaged SSIM increased by 5.2×10^{-2}, 18.9×10^{-2} and 33.9×10^{-2}, for the 1/2, 1/4, and 1/8 sparse-view reconstructions, respectively. In expert subjective evaluation, both the median scores and acceptance rates of the images with SynCNN synthesis were higher than those reconstructed from sparse-view projections. It took the model less than 0.01 s to synthesize an inter-view projection. Compared with the three other methods, the SynCNN model obtained the best scores

in terms of the three metrics in both domains.

Conclusions: The proposed SynCNN model effectively improves the quality of sparse-view CBCT images at a low time cost. With the SynCNN model, the CBCT imaging dose in IGRT could be reduced potentially.

Introduction

On-board cone-beam computed tomography (CBCT) has been widely used as the gold standard for online positioning verification in image-guided radiotherapy (IGRT) (1). And recurrent CBCT scans are routinely scheduled over the treatment course (2). For radiotherapy patients with planning target volume (PTV) margins in centimeters, CBCT is usually performed weekly for the first few fractions. However, for patients with head and neck tumors or high-dose fractions, due to the long list of organs-at-risk (OAR) and tight PTV margins, positioning tolerance is very strict. As a result, CBCT is typically scheduled for every fraction. Although the dose of a single CBCT scan is generally low (3), considering that the imaging field is typically much larger than target volumes and in-field OARs are unshielded, adverse radiation dose from recurrent CBCT scans over the whole course can be easily accumulated to a level that may cause radiobiological effects (4,5). Therefore, the non-negligible risk of the accumulative CBCT imaging dose on patients' health in long term has raised growing concerns from researchers (6,7) and professional societies (8).

Fortunately, increasing efforts have been directed to low-dose CBCT. Dose reduction strategies generally fall into several approaches (9), including tube current reduction, optimal selection of tube voltage, and sparse-view sampling. Among these approaches, sparse-view sampling as a straightforward strategy is highly efficient in reducing radiation exposure. However, insufficient projection data induce severe streak artifacts and noise in reconstructed images (10). To address this issue, many iterative reconstruction (IR) algorithms have been developed (11-17) to compensate the image quality deterioration. For instance, Varian commercializes an IR module, iCBCT®, which delivers comparable pelvis images with as much as

33% reduction in projection views and 50% reduction in CT dose index (18). In spite of this, successful deployments of IR on commercial on-board CBCT are limited for several reasons, such as high computation complexity, long time consumption, latent new artifacts patterns, and unnatural "plastic" image textures (19).

Deep learning (DL) achieved breakthroughs in computer vision as well as medical imaging, providing powerful tools for sparse-view CT imaging (20). DL-based methods might be grouped into two types: image-domain refinement (21-30) and projection-domain augmentation (31-37). Image-domain refinement methods are applied to enhance the coarse CT images reconstructed from insufficient projections, which is an intuitive way to improve the quality of sparse-view CT images. Representative convolutional neural network (CNN) models include FBPConvNet (21), which combines the filtered-back-projection (FBP) algorithm with a multiresolution U-Net model, DD-Net (22), which takes advantage of the DenseNet model and deconvolution operation, framing U-Net (23), which is a variant of the original U-Net model, GoogleNet (24), which is characterized by multiscale inception modules, R2-Net (25), which contains several recurrent and recursive stages, SR-CNN (26), which contains symmetric network layers with residual connection, and FRCNN (27), which consists of a residual CNN model with a fractional total-variance loss. The objective of these models is to minimize the pixel-wise L_2 distance between the refined sparse-view and reference full-view CT images. While image noise is efficiently reduced, over-smoothed patterns are yielded (28). To address the over-smoothing issue, Yang *et al.* (28) developed a generative adversarial network (GAN) model with VGG-based perceptual loss. Meanwhile, Li *et al.* (29) developed a 3-dimensional (2D) self-attention CNN model with autoencoder perceptual loss, and Huang *et al.* (30) developed an attributed-augmented Wasserstein GAN

model which took the anatomical prior information into account.

Projection-domain augmentation restores full-view projections from sparse-view projections prior to reconstruction, which outperforms image-domain refinement in suppressing streak artifacts (38). Lee *et al.* (31), Dong *et al.* (32) and Yuan *et al.* (33) developed U-Net models, Liang *et al.* (34) developed a DLI-Net model, and Yin *et al.* (35) developed a SD-Net model for sinogram interpolation where equivalent image qualities were presented. Dong *et al.* (36) improved the quality of sparse-view micro-CT images by restoring full-view sinograms with linear interpolation techniques and a U-Net model. Note that the above-mentioned studies were all in fan-beam CT and could not be applied to CBCT directly. Hu *et al.* (37) proposed a hybrid strategy combining projection augmentation together with an image-domain refinement method. In the projection augmentation stage of their study, inter-view projections were linearly interpolated at first and then refined by a CNN model.

In this study, we propose a novel DL-based projection synthesis model, SynCNN, to restore full-view CBCT projections from sparse-view projections, thereby improving the image quality to an extent that is comparable to the full-view reconstruction but with reduced dose. The proposed SynCNN model is inspired in part by a video frame interpolation technique (39-41). The sequence of CBCT projections is acquired consecutively over a circular scan trajectory, similar to a sequence of video frames in terms of dynamic features between the neighboring projections/frames. The inter-view projections are directly synthesized by the proposed SynCNN model without a pre-step of linear interpolation as adopted in studies (31,36,37). Moreover, the proposed model is fully trained using real patients' data, and its performance is systematically evaluated using an image quality test phantom and real patients' data in both the projection and image domains.

The main contributions of this work are as follows:
(I) The proposed SynCNN model improves the quality of the sparse-view low-dose CBCT images for IGRT using an innovative DL framework to synthesize missing projections.
(II) The originality of the proposed SynCNN model is that the missing CBCT projections are synthesized with local convolution operations between the input neighboring sparse-view projections and orthogonally direction-separable spatial kernels.

The highlights of this work are as follows:
(I) The proposed SynCNN model is fully trained with clinical CBCT scans from real patients, and its performance is evaluated in both the projection and image domains.
(II) Blind and randomized expert scoring is utilized to assess the image quality of authentic full-view reconstructions, sparse-view reconstructions, and composite full-view reconstructions.
(III) The results indicate that the proposed SynCNN model improves the quality of sparse-view CBCT images comparable to that of regular CBCT images.
(IV) The proposed SynCNN model synthesizes the missing projections at a low time cost, which is crucial for enabling online imaging in IGRT.

We present this article in accordance with the TRIPOD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-23-759/rc).

Methods

Sparse-view CBCT imaging chain using DL model

Figure 1 shows the imaging chain for sparse-view CBCT using the proposed SynCNN model. First, the sparse-view projections were acquired by CBCT using the sparse-view scanning protocol. Then, the missing projections were synthesized with the SynCNN model to restore full-view projections. Finally, tomographic CBCT images were reconstructed from the restored full-view projections.

DL model architecture

Figure 2 shows the overall architecture of the proposed SynCNN model. Inspired in part by previous studies (39,42), a U-shaped structure was used as the backbone. The SynCNN model consisted of both an encoder component and a decoder component. Taking neighboring sparse-view projections, P_1 and P_3, as the input, the SynCNN model was designed to synthesize the inter-view projection $\widetilde{P_2}$ similar to the actual middle view projection P_2.

The encoder component consisted of five levels of down-sampling steps, which reduced the $H \times W \times 2$ input to a $(H/32) \times (W/32) \times 512$ representation, and the decoder component consisted of five corresponding levels of up-sampling steps to up-sample the encoded representation into four $H \times W \times 51$ spatial kernels. Each
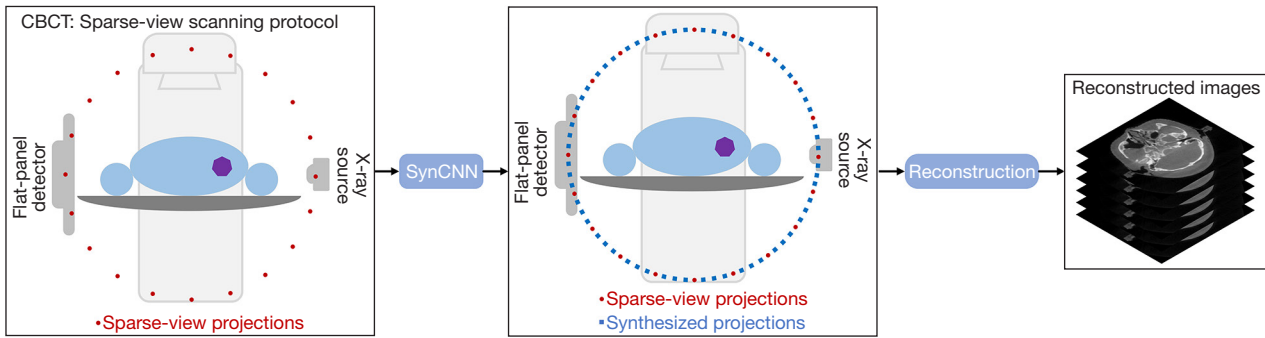
Figure 1 Schematic of deep learning-based projection synthesis for sparse-view CBCT imaging. CBCT, cone-beam computed tomography; SynCNN, synthesis convolutional neural network.
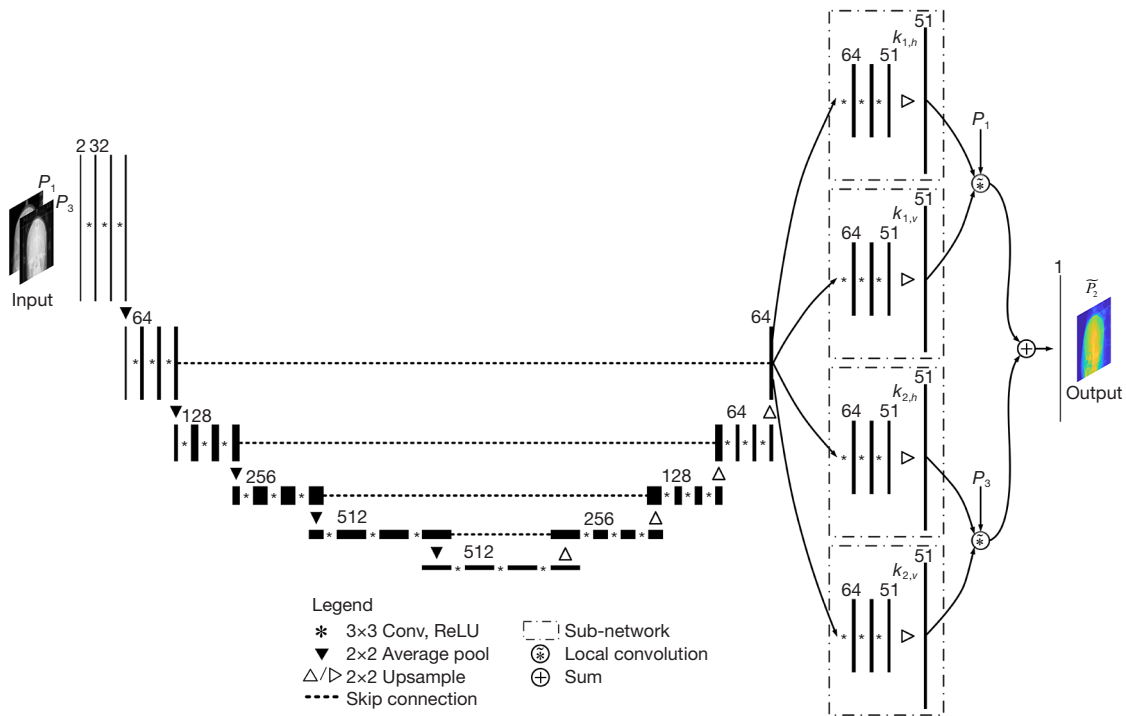

Figure 2 Overall architecture of the proposed SynCNN model for CBCT projection synthesis. CBCT, cone-beam computed tomography; SynCNN, synthesis convolutional neural network.

down or up-sampling step consisted of three consecutive 3×3 convolutions (zero-padded), followed by a rectified linear units (ReLU) activation function and either an average-pooling with a kernel size of 2×2 in the encoder component or up-sampling operations via 2×2 bilinear interpolation in the decoder component. The skip connections (43) were used to let the decoding steps incorporate features from the encoding component of the network, which helped preserve the details of the input projections.

Assuming that the relation between the neighboring input projections and the synthesized projection could be described with local convolution operations (44,45), as

$$\widetilde{P_2} = K_1 \,\tilde{*}\, P_1 + K_2 \,\tilde{*}\, P_3 \qquad [1]$$

where P_1 and P_3 are the neighboring input projections, K_1 and K_2 are associated 2-dimensional (2D) spatial kernels, $\tilde{*}$ is the local convolution operator, and $\widetilde{P_2}$ is the synthesized
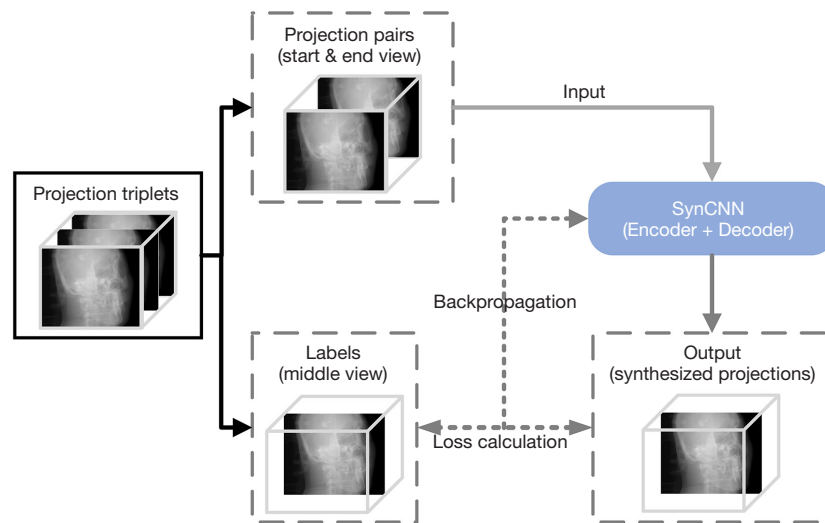
Figure 3 Overall training workflow of the proposed SynCNN model. SynCNN, synthesis convolutional neural network.

projection.

Patients were scanned by CBCT with a circular trajectory, and the resulting consecutive projections were correlated in orthogonal directions (46). These directions were parallel and perpendicular to the trajectory plane, denoted as horizontal and vertical directions, respectively. To take advantage of the circular scan trajectory in CBCT acquisition, the 2D spatial kernels were decomposed into two orthogonal directional (vertical and horizontal) kernels, respectively (39). Therefore, the relation between the neighboring input projections and the synthesized one was approximated with two pairs of direction-separable spatial kernels, as

$$\widetilde{P_2} = K_{1,v} \tilde{*} P_1 + K_{1,h} \tilde{*} P_1 + K_{2,v} \tilde{*} P_3 + K_{2,h} \tilde{*} P_3 \qquad [2]$$

where K_v and K_h are the spatial kernels in the vertical and horizontal directions.

The information flow in the last decoder step in *Figure 2* was directed into four sub-networks, estimating one of the direction-separable spatial kernels, respectively. Finally, the output synthesized projection was calculated as in Eq. [2]. Note that the estimated spatial kernels were applied to the input projections using local convolution operations, which were implemented as dynamic convolution layers (47) in the SynCNN model. This implementation allowed for end-to-end training of the SynCNN model.

Loss function of DL model

To train the SynCNN model, we adopted the mean square error (MSE) loss. The MSE loss is formulated as

$$\mathcal{L}_{MSE} = \frac{1}{MN} \sum_{m=1}^{M} \left\| \mathcal{F}\left(P_{m,1} \circ P_{m,3}, \Theta\right) - P_{m,2} \right\|_2^2 \qquad [3]$$

where $\circ$ represents concatenation of the two input projections, M is the mini-batch size, and N is the number of pixels in a projection.

Data preparation

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional review board (IRB) at Beijing Cancer Hospital and individual consent for this retrospective analysis was waived. The patient inclusion criteria of this study were: (I) age >18 years, and (II) immobilized with double-shell positioning system (MacroMedics, Moordrecht, The Netherlands). The exclusion criteria were: (I) tumour sites on parietal lobe only, and (II) with metal implants. A total of 223 patients diagnosed with brain tumours (glioma and metastases) at the hospital from March 2021 to April 2021 were finally enrolled. The real CBCT scan data were anonymously collected from all 223 patients using an Edge linac (Varian Medical Systems, Inc., Palo Alto, CA, USA). All these scans were performed on the default half-scan head protocol, the details of which are listed in *Table 1*. Note that the imaging centers were all around the nasal midline, and the issue of shoulder-head transition was not applicable. In addition to patients, a CatPhan-504 phantom was also scanned

Table 1 Acquisition parameters of the CBCT head-scan protocol

Acquisition parameters	Values
Scan range (°)	200
Interval (°)	0.4
Tube voltage (kVp)	120
Tube current (mA)	15
Pulse length (ms)	20
Detector array dimension, (R) × (C)	768×1,024
Detector pixel size (mm^2)	0.388×0.388
Source-to-isocenter distance (mm)	1,000
Source-to-detector distance (mm)	1,500

CBCT, cone-beam computed tomography; R, row; C, column.

on the identical linac using the same protocol. The raw scan data were preprocessed in a typical workflow using a validated open-source toolkit, TIGRE-VarianCBCT (48). The preprocessing procedure included dead-pixel removal, beam-hardening correction (49), scatter reduction (50), logarithmic operation, and negative pixel value cutoff.

The full-view projections were acquired at an interval of 0.4 degrees, yielding 501 views. Sparse-view projections were down-sampled from the full-view projections. For each full-view projection dataset, denoted as S_1, three sparse-view projection datasets were decimated and denoted as $S_{1/2}$, for 251 projections sampled at 0.8 degrees, $S_{1/4}$, for 126 projections at 1.6 degrees, and $S_{1/8}$, for 63 projections at 3.2 degrees.

For the model training, a total of 150 patients were randomly selected from the whole patient cohort (150/223, 67%) using the Fisher-Yates shuffle (51). For each patient, every three adjacent projections in either S_1, $S_{1/2}$ or $S_{1/4}$ were independently grouped into a triplet, resulting in 167, 83 and 42 triplets, respectively. Therefore, a total of 25,050, 12,490, and 6,300 triplets were generated from all S_1, $S_{1/2}$ and $S_{1/4}$ in the training set. An additional 30 patients (30/223, 13%) were set aside as the validation set, which was used for hyperparameter optimization and to prevent overfitting of the model. The CatPhan phantom and the rest 43 patients were used to test the SynCNN model.

Model training and implementation details

Figure 3 shows the workflow of the model training. Triplets were fed to train the proposed model. For each triplet, the start and end view projections served as the input, and the middle view projection served as the corresponding ground truth label. The model encoded the input projections to gain high-level feature representation and then decoded back to synthesize the inter-view projections. The model weights were updated iteratively till the synthesized projections were close enough to the ground truth labels.

Three different models, referred to as SynCNN$_{1/2}$, SynCNN$_{1/4}$, and SynCNN$_{1/8}$, were trained using all triplets from S_1, $S_{1/2}$ and $S_{1/4}$ in the training set, respectively. These models were independently implemented using Python V3.8.13, Pytorch V1.10.1, and Cuda V11.1.74. The Adam optimizer (52) was utilized with an initial learning rate 10^{-5}, β_1 =0.9, and β_2 =0.999. The mini-batch size was set to 6 triplets. For the training process, a learning rate scheduler was employed. If the average loss on the validation set did not decrease over 10 consecutive epochs, the learning rate was reduced to one-tenth of its previous value. Furthermore, an early stopping mechanism was incorporated to prevent potential overfitting. If the average loss on the validation set did not decrease over 20 consecutive epochs, the training was terminated. The SynCNN1/2, SynCNN1/4, and SynCNN1/8 models were all trained with a maximum of 200 epochs. All experiments were carried out on a single NVIDIA RTX3090 GPU with 24 GB memory.

Performance evaluation

The performance of the proposed SynCNN model was evaluated at three levels, i.e., objective assessment with a phantom data, objective assessment with real patients' data, and subjective assessment given by experts. Besides, the running time of the SynCNN models was recorded as an indicator of computation complexity.

Evaluation metrics

For objective evaluation, quantitative metrics including root-mean-square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM) (53) were used.

The RMSE and PSNR are defined below:

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\hat{I}_n - I_n\right)^2} \quad [4]$$

$$PSNR = 10\log_{10}\left(\frac{\max^2(I)}{RMSE^2}\right) \quad [5]$$

where N is the number of pixels, $\hat{I}_n$ and I_n are the values of pixel n in the evaluated image and reference image, and
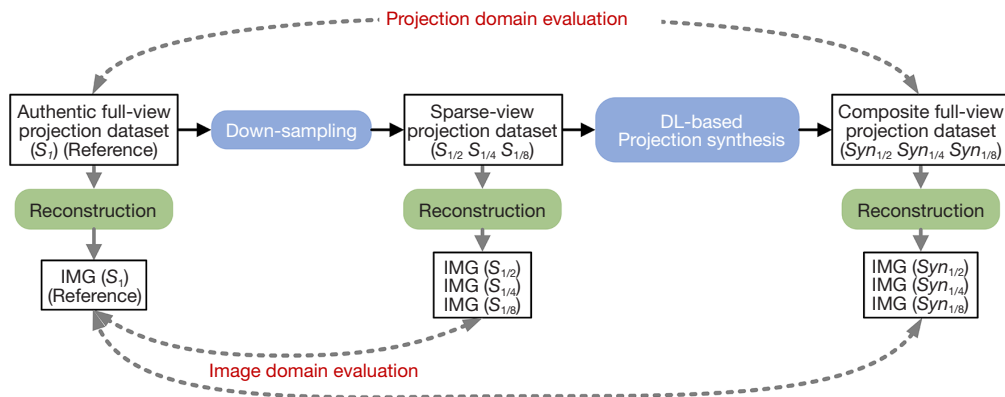
Figure 4 Overall performance evaluation workflow of the proposed SynCNN model. DL, deep learning; SynCNN, synthesis convolutional neural network.
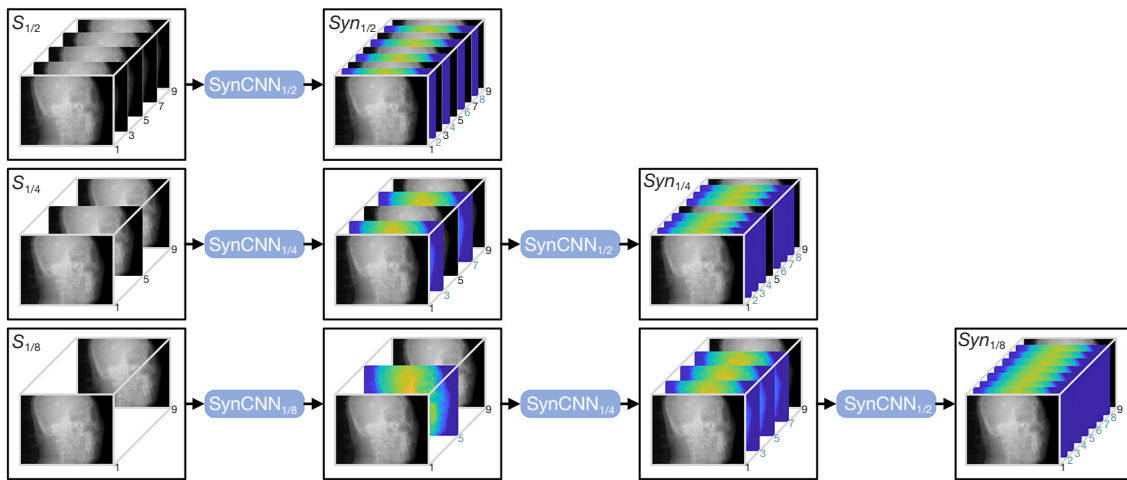

Figure 5 Schematic of generating composite full-view projection datasets by feeding sparse-view projection datasets to the trained SynCNN models. SynCNN, synthesis convolutional neural network.

max(I) is the maximum pixel value of the reference image.

The SSIM is defined as

$$SSIM = \frac{\left(2\mu_i\mu_I + C_1\right)\left(2\sigma_{iI} + C_2\right)}{\left(\mu_i^2 + \mu_I^2 + C_1\right)\left(\sigma_i^2 + \sigma_I^2 + C_2\right)} \quad [6]$$

where μ_i and μ_I are the mean values of local windows for the evaluated image and reference image, respectively, σ_i and σ_I are the respective standard deviations, σ_{iI} is the covariance, and C_1 and C_2 are constants of $(0.01 \times L)^2$ and $(0.03 \times L)^2$, respectively, where L equals to the dynamic range of pixel-values in the reference image (53). The SSIM computation employs an 11×11 window size.

Phantom and patient studies

The CatPhan phantom and 43 patients were used to evaluate the performance of the proposed model in both the projection and image domains. *Figure 4* shows the corresponding workflow. In the projection domain, the composite full-view projection dataset was generated by feeding the sparse-view projection dataset to the trained models. The corresponding projection synthesis process is shown in *Figure 5*. Specifically, the $S_{1/2}$ was fed into the trained SynCNN$_{1/2}$ model to obtain the composite full-view projection dataset denoted as $Syn_{1/2}$. The $S_{1/4}$ was fed into a cascade of trained SynCNN$_{1/4}$ and SynCNN$_{1/2}$ models to

Table 2 Descriptions of comparison methods

Methods	Descriptions
Linear interpolation	Linear interpolation from $S_{1/4}$ to full-view projections, followed by FDK reconstruction
FBPConvNet (21)	FDK reconstruction with $S_{1/4}$, followed by 2D U-Net refinement
IDU-Net (37)	FDK reconstruction with $S_{1/4}$, followed by 3D U-Net refinement
SynCNN (proposed)	SynCNN synthesis from $S_{1/4}$ to full-view projections, followed by FDK reconstruction

FDK, Feldkamp-Davis-Kress algorithm. SynCNN, synthesis convolutional neural network.

obtain the composite full-view projection dataset denoted as $Syn_{1/4}$. The $S_{1/8}$ was fed into a cascade of trained $SynCNN_{1/8}$, $SynCNN_{1/4}$ and $SynCNN_{1/2}$ models to obtain the composite full-view projection dataset denoted as $Syn_{1/8}$. The projections in the authentic full-view projection dataset, S_1, were used as the reference benchmark, and the synthesized projections in the composite full-view projection datasets, $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$, were compared with the projections in S_1, both qualitatively and quantitatively. The RMSE, PSNR and SSIM values were calculated.

In the image domain, seven 3D tomographic CBCT images were reconstructed from projection datasets S_1, $S_{1/2}$, $S_{1/4}$, $S_{1/8}$, $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$. These 3D images were denoted as $IMG(S_1)$, $IMG(S_{1/2})$, $IMG(S_{1/4})$, $IMG(S_{1/8})$, $IMG(Syn_{1/2})$, $IMG(Syn_{1/4})$, and $IMG(Syn_{1/8})$, respectively. The Feldkamp-Davis-Kress algorithm was used for reconstruction (48,54), and the resulting images had a matrix size of 512×512×93 and a voxel size of 0.511×0.511×1.990 mm^3. The axial images in $IMG(S_1)$ were used as the reference benchmark, and the axial images in $IMG(S_{1/2})$, $IMG(S_{1/4})$, $IMG(S_{1/8})$, $IMG(Syn_{1/2})$, $IMG(Syn_{1/4})$, and $IMG(Syn_{1/8})$ were qualitatively and quantitatively compared with the images in $IMG(S_1)$. The RMSE, PSNR and SSIM values were calculated.

Blind and randomized expert scoring

Five radiation oncology experts were invited herein. Three experts were radiation oncologists (physicians) with more than 11 years of experience, and the other two were senior medical physicist with more than 8 years of experience. The qualities of reconstructed CBCT images were subjectively assessed by the five experts. The expert scoring was organized in a blind and randomized fashion as follows:

❖ Step 1: twenty patients were randomly selected from the 43 test patients, i.e., (20/43).

❖ Step 2: for each patient, two axial locations were randomly selected for 93 cross-section images, i.e., (2/93).

❖ Step 3: for each selected location, seven axial images from $IMG(S_1)$, $IMG(S_{1/2})$, $IMG(S_{1/4})$, $IMG(S_{1/8})$, $IMG(Syn_{1/2})$, $IMG(Syn_{1/4})$, and $IMG(Syn_{1/8})$ were combined to form one evaluation group, i.e., a total of 40 evaluation groups were generated.

❖ Step 4: within each evaluation group, the seven images were shuffled and labelled from (a) to (g).

❖ Step 5: the 40 evaluation groups were packed in a random order and sent to one expert for scoring.

❖ Step 6: repeat Step 4 to 5 till the five experts all received their own evaluation dataset.

The image quality was scored using a four-level rubric: excellent (score 3), good (score 2), suboptimal (score 1), and poor (score 0). Images quality with a score ≥2 were regarded as acceptable. The median scores and acceptance rates of the seven image categories were compared.

Comparison study

The proposed SynCNN model was directly compared with three other methods for sparse-view CBCT reconstruction using our patient CBCT scan data: the projection-domain linear interpolation method, and the image-domain refinement methods utilizing the FBPConvNet (21) and IDU-Net (37) models. The focus of these comparisons was on the 1/4 sparse-view reconstructions, given the encouraging results from SynCNN, which delivered image quality comparable to that of full-view reconstructions (as detailed in the *Expert scoring* section). *Table 2* provides the descriptions of different methods. The source codes for both the FBPConvNet and IDU-Net models were not accessible. They were re-implemented and fine-tuned according to the descriptions provided in their original papers (21,37). To ensure consistency, the training, validation, and test sets used for the FBPConvNet and IDU-Net models were
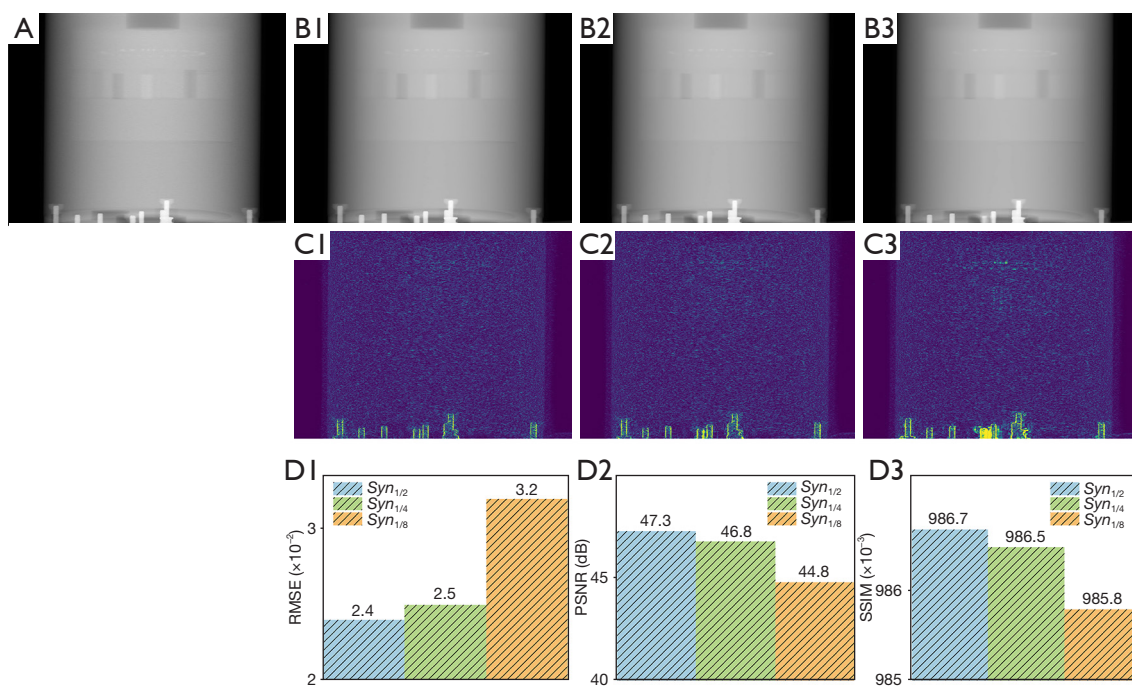
Figure 6 Representative projections on CatPhan phantom: (A) reference projection in authentic full-view projection dataset S_1; (B1-B3) synthesized projections in composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$, respectively; (C1-C3) absolute difference maps between (B1-B3) and (A); (D1-D3) RMSE, PSNR and SSIM values of synthesized projections in $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$ with projection in S_1 as the reference. The display gray scale of (A,B1-B3) is [0, 5.554] and the display grayscale of (C1-C3) is [0, 0.162]. RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.

identical to those used for the proposed SynCNN model.

Results

Phantom study

Projection domain

Figure 6 shows the representative synthesized projections of the CatPhan phantom and the corresponding absolute differences from the reference projection. The RMSE, PSNR and SSIM values of the synthesized projections in composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$ were also given [*Figure 6* (D1-D3)]. The synthesized projections in $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$ [*Figure 6* (B1-B3)] were close to the reference projection in S_1 (*Figure 6A*). As depicted in *Figure 6* (C1-C3), the disparity between the synthesized projections and the reference projection mainly resided around the sharp edges. When the down-sampling rate decreased, the disparity became greater, along with the

deterioration of the RMSE, PSNR and SSIM values.

Image domain

Figure 7 shows the representative axial images of the CatPhan phantom, reconstructed from the projection datasets S_1, $S_{1/2}$, $S_{1/4}$, $S_{1/8}$, $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$, respectively. The RMSE, PSNR and SSIM values of the images in IMG($S_{1/2}$), IMG($S_{1/4}$), IMG($S_{1/8}$), IMG($Syn_{1/2}$), IMG($Syn_{1/4}$), and IMG($Syn_{1/8}$) were also given. As depicted in *Figure 7* (B1-B3,C1-C3), when the number of projections was reduced, the radial streak artifacts and noise induced by under-sampling became greater, and the RMSE, PSNR and SSIM values deteriorated accordingly. However, with the aid of the proposed SynCNN model, the overall image qualities of the composite full-view reconstructions were significantly improved upon the original sparse-view reconstructions, accompanied by the improvement of the RMSE, PSNR and SSIM values.
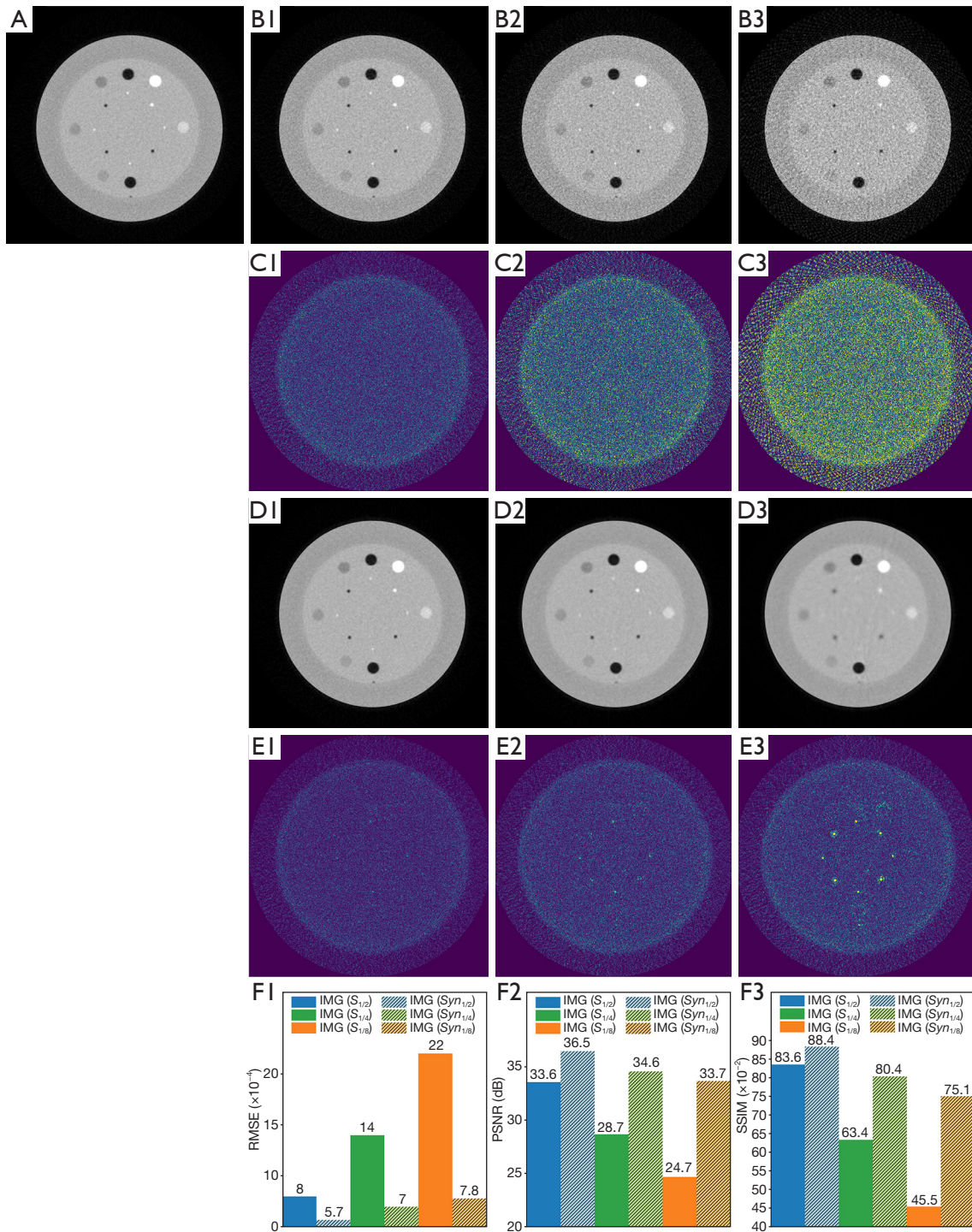
240

Zhao et al. DL-based projection synthesis for LD CBCT imaging in IGRT

Figure 7 Representative axial images on CatPhan phantom: (A) reference image reconstructed from authentic full-view projection dataset S_1; (B1-B3) images reconstructed from sparse-view projection datasets $S_{1/2}$, $S_{1/4}$, and $S_{1/8}$; (C1-C3) absolute difference maps between (B1-B3) and (A); (D1-D3) images reconstructed from composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$; (E1-E3) absolute difference maps between (D1-D3) and (A); (F1-F3) RMSE, PSNR and SSIM values of images in IMG($S_{1/2}$), IMG($S_{1/4}$), IMG($S_{1/8}$), IMG($Syn_{1/2}$), IMG($Syn_{1/4}$), and IMG($Syn_{1/8}$) with image in IMG(S_1) as the reference. The display gray scale of (A,B1-B3,D1-D3) is [0, 0.030] and the display grayscale of (C1-C3,E1-E3) is [0, 0.004]. RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.

Table 3 Quantitative evaluation of synthesized projections in composite full-view projection datasets for all patient studies in the test set

Projection dataset	RMSE (×10^{-2})	PSNR (dB)	SSIM (×10^{-2})
$Syn_{1/2}$*	2.6±1.2	46.2±1.6	98.3±0.6
$Syn_{1/4}$	2.8±1.2	45.7±1.6	98.1±0.6
$Syn_{1/8}$	3.3±1.2	44.2±1.7	97.8±0.7

All values are reported as mean ± standard deviation. Best results are indicated with an asterisk (*). RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.
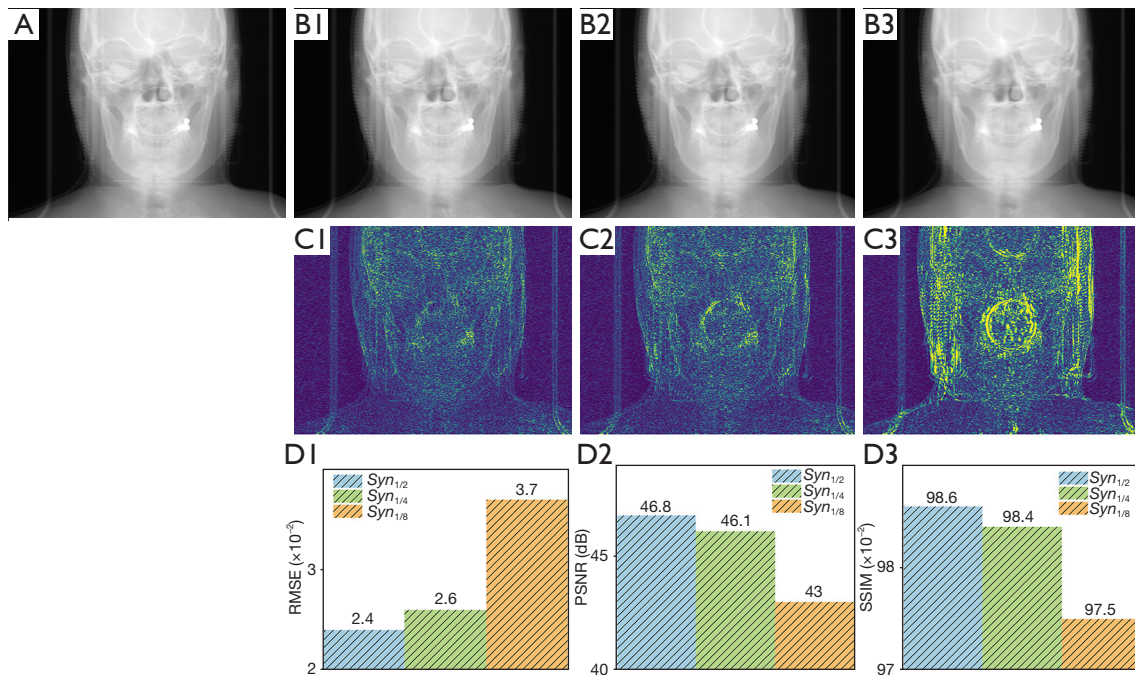

Figure 8 Representative patient projections: (A) reference projection in authentic full-view projection dataset S_1; (B1-B3) synthesized projections in composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$, respectively; (C1-C3) absolute difference maps between (B1-B3) and (A); (D1-D3) RMSE, PSNR and SSIM values of synthesized projections in $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$ with projection in S_1 as the reference. The display gray scale of (A,B1-B3) is [0, 4.718] and the display grayscale of (C1-C3) is [0, 0.086]. RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.

Patient study

Projection domain

The composite full-view projection datasets $Syn_{1/2}$ achieved the best RMSE, PSNR and SSIM values, shown in *Table 3*, indicating that the synthesized projections in $Syn_{1/2}$ were closest to the reference projections. Besides, as the down-sampling rate decreased, the accuracies of the synthesized projections in $Syn_{1/4}$ and $Syn_{1/8}$ also decreased in a sequential manner.

Figure 8 shows the representative projections of a patient. The absolute difference maps [*Figure 8* (C1-C3)] exhibited

that the major discrepancies between the synthesized projections and the reference projection were around the structure edges, similar to the patterns observed in the phantom study.

Image domain

A few trends can be observed from the numerical results shown in *Table 4*. First of all, the image qualities of the sparse-view reconstructions deteriorated as the down-sampling rate decreased. Second, the image qualities of sparse-view reconstructions improved with DL-based

Table 4 Quantitative evaluation of spare-view reconstructions with or without DL-based projection synthesis for all patient studies in the test set

Reconstructed image	RMSE ($\times10^{-4}$)	PSNR (dB)	SSIM ($\times10^{-2}$)
IMG($S_{1/2}$)	10.5±2.9	36.1±2.5	88.9±2.8
IMG($S_{1/4}$)	19.6±5.0	30.6±2.5	71.2±5.4
IMG($S_{1/8}$)	33.1±7.6	26.0±2.4	52.9±6.0
IMG($Syn_{1/2}$)*	7.1±2.2	39.5±2.7	94.1±1.9
IMG($Syn_{1/4}$)	9.3±2.6	37.2±2.6	90.1±3.2
IMG($Syn_{1/8}$)	11.4±3.0	35.4±2.6	86.8±4.2

All values are reported as mean ± standard deviation. Best results are indicated with an asterisk (*). DL, deep learning; RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.

projection synthesis, and the improvement increased as the down-sampling rate decreased. The difference between IMG($Syn_{1/2}$) and IMG($S_{1/2}$), IMG($Syn_{1/4}$) and IMG($S_{1/4}$), and IMG($Syn_{1/8}$) and IMG($S_{1/8}$) were 3.4×10^{-4}, 10.3×10^{-4}, and 21.7×10^{-4} for the averaged RMSE, 3.4, 6.6, and 9.4 dB for the averaged PSNR, and 5.2×10^{-2}, 18.9×10^{-2} and 33.9×10^{-2} for the averaged SSIM, respectively.

Figures 9-11 show the representative axial, coronal, and sagittal images of three different patients. Sparse-view reconstructions with or without DL-based projection synthesis showed substantially different properties of streak artifacts and noise, as depicted in *Figure 9* (B1-B3,D1-D3), *Figure 10* (B1-B3,D1-D3), and *Figure 11* (B1-B3,D1-D3). In addition, the integrity of the patient anatomical structures in sparse-view reconstructions was effectively preserved by using the proposed model. In all three patients, despite slight blurring of bone or other tissues, images reconstructed from $Syn_{1/4}$ delivered reasonable integrity of anatomical details and exhibited good contrast against streak artifacts and noise.

Expert scoring

Figure 12 shows the results of blind and randomized expert scoring for the seven categories of reconstructed images. The median scores were 3, 2, 3, 1, 3, 0, and 2, respectively, for IMG(S_1), IMG($S_{1/2}$), IMG($Syn_{1/2}$), IMG($S_{1/4}$), IMG($Syn_{1/4}$), IMG($S_{1/8}$), and IMG($Syn_{1/8}$). The utilization of DL-based projection synthesis in sparse-view reconstructions achieved higher median scores compared to those without this technique. Moreover, the proposed model increased the rates of image quality acceptance from 69.5% to 99%, 17% to 95.5%, 1% to 61.5%, respectively, for images reconstructed from half, quarter, and one-eighth sparse-view projections. The acceptance rate for IMG(S_1)

was 95.5%, which was even lower than that of IMG($S_{1/2}$).

Time cost

Table 5 shows the average running time of the SynCNN models. The projection synthesis from sparse-view projection datasets $S_{1/2}$, $S_{1/4}$, and $S_{1/8}$ to composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$ took 1.8, 2.6 and 2.9 s, respectively. Moreover, the projection synthesis per frame required less than 0.01 s.

Comparison study

Tables 6,7 show the quantitative comparison results of different methods in the projection and image domains, respectively. The SynCNN model obtained the best scores in terms of the three metrics in both domains.

Discussion

CBCT is an important imaging modality in IGRT (1). The accumulative CBCT imaging dose poses a risk to patient health (2-5). Sparse-view sampling is a reduced-dose strategy that results in compromised image quality (10). In this study, we propose a DL-based model, SynCNN, that can synthesize missing CBCT projections to improve the image quality. The SynCNN architecture takes advantage of the circular scan trajectory in CBCT acquisition, where projection synthesis is formulated as local convolution operations between the input neighboring sparse-view projections and orthogonally direction-separable spatial kernels.

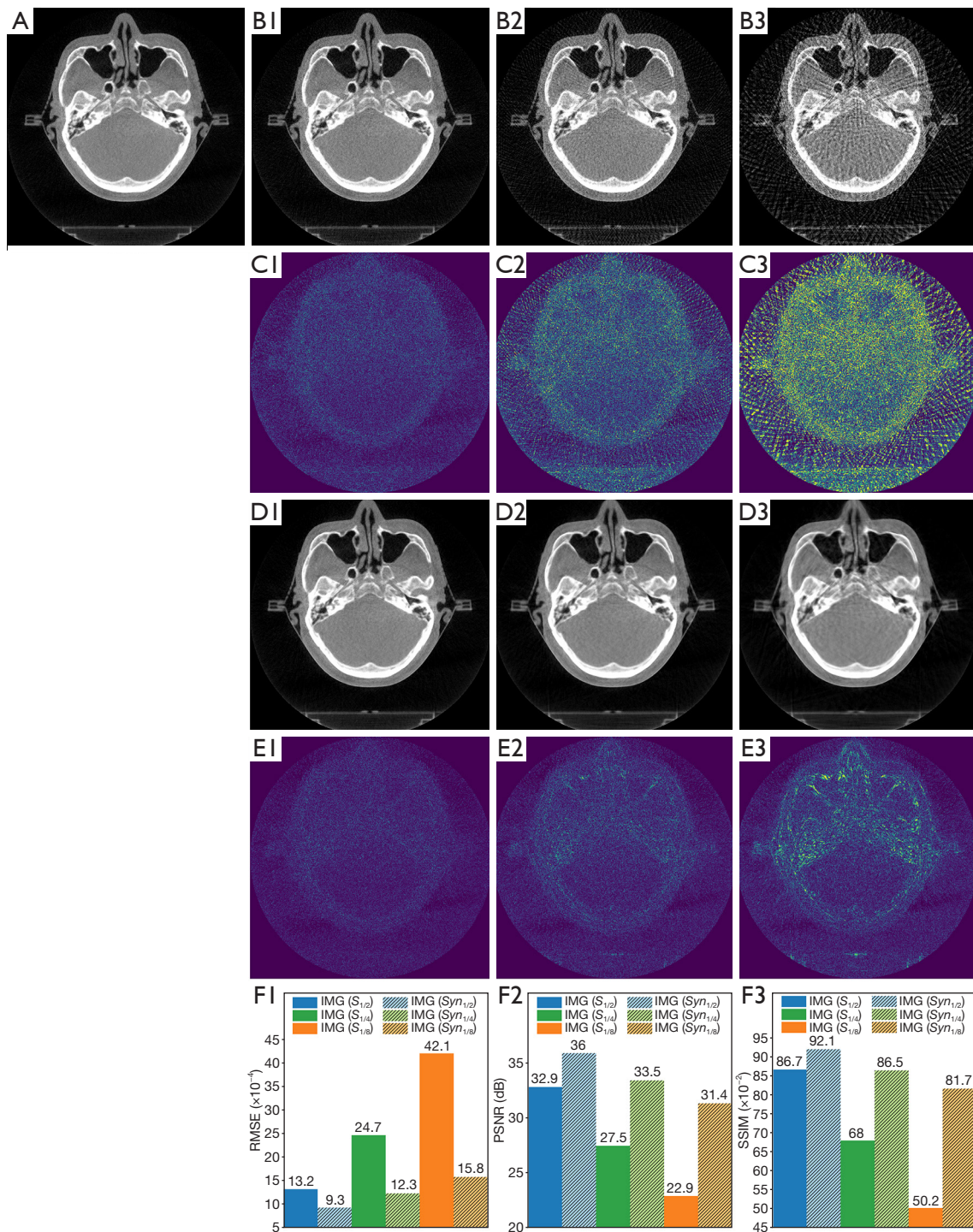The direct comparisons highlighted the superior performance of the proposed SynCNN model over the

Figure 9 Representative patient axial images: (A) reference image reconstructed from authentic full-view projection dataset S_1; (B1-B3) images reconstructed from sparse-view projection datasets $S_{1/2}$, $S_{1/4}$, and $S_{1/8}$; (C1-C3) absolute difference maps between (B1-B3) and (A); (D1-D3) images reconstructed from composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$; (E1-E3) absolute difference maps between (D1-D3) and (A); (F1-F3) RMSE, PSNR and SSIM values of images in IMG($S_{1/2}$), IMG($S_{1/4}$), IMG($S_{1/8}$), IMG($Syn_{1/2}$), IMG($Syn_{1/4}$), and IMG($Syn_{1/8}$) with image in IMG(S_1) as the reference. The display gray scale of (A,B1-B3,D1-D3) is [0, 0.041] and the display grayscale of (C1-C3,E1-E3) is [0, 0.009]. RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.
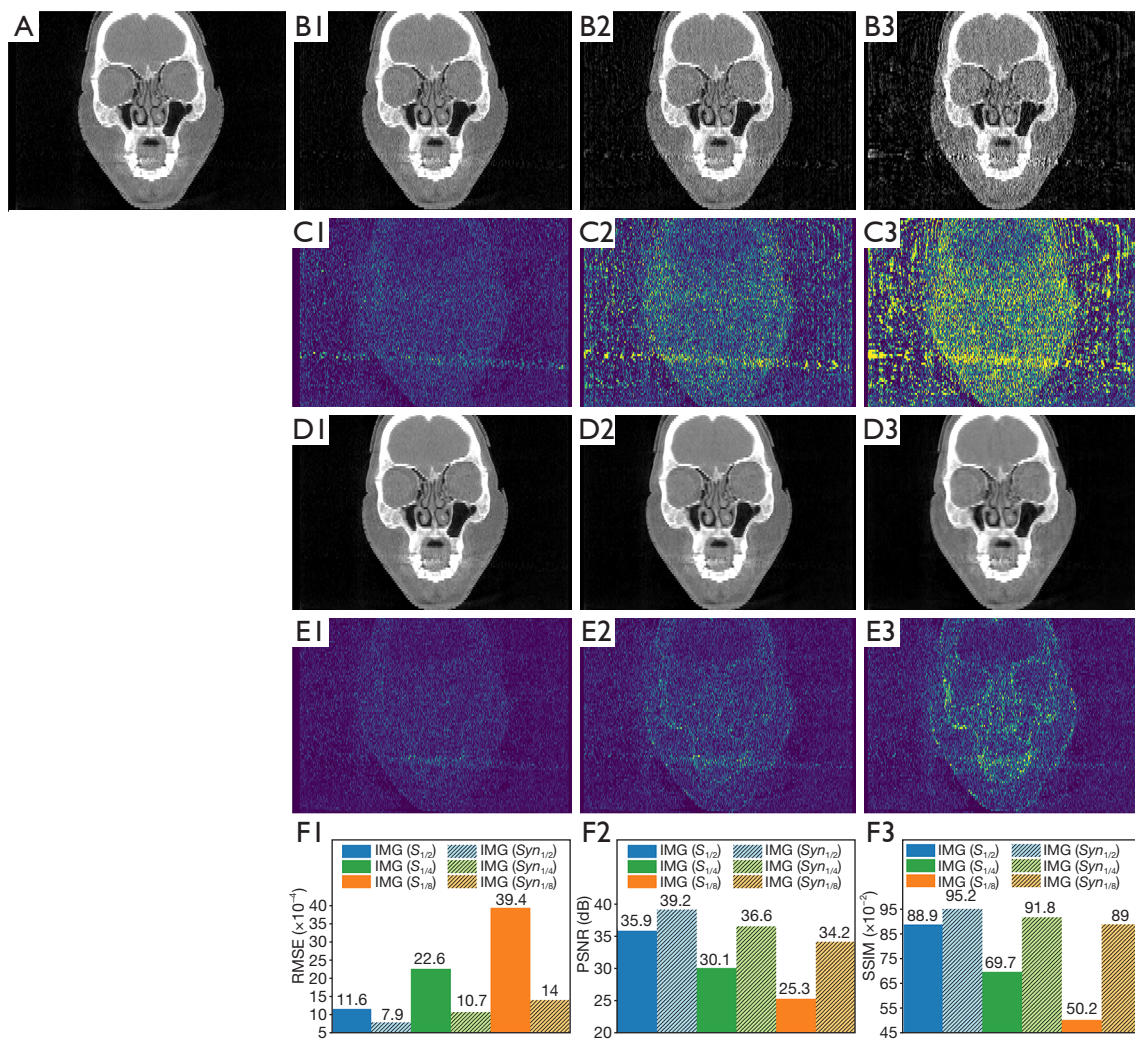
Figure 10 Representative patient coronal images: (A) reference image reconstructed from authentic full-view projection dataset S_1; (B1-B3) images reconstructed from sparse-view projection datasets $S_{1/2}$, $S_{1/4}$, and $S_{1/8}$; (C1-C3) absolute difference maps between (B1-B3) and (A); (D1-D3) images reconstructed from composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$; (E1-E3) absolute difference maps between (D1-D3) and (A); (F1-F3) RMSE, PSNR and SSIM values of images in IMG($S_{1/2}$), IMG($S_{1/4}$), IMG($S_{1/8}$), IMG($Syn_{1/2}$), IMG($Syn_{1/4}$), and IMG($Syn_{1/8}$) with image in IMG(S_1) as the reference. The display gray scale of (A,B1-B3,D1-D3) is [0, 0.042] and the display grayscale of (C1-C3,E1-E3) is [0, 0.007]. RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.

three other methods in sparse-view CBCT reconstruction. As shown in *Table 6*, the SynCNN model achieved a lower RMSE and higher PSNR and SSIM values than the linear interpolation method. This suggested that the proposed model was not only more accurate in pixel-level projection reproduction, but also ensured better perceptual quality of the projections. Furthermore, as detailed in *Table 7*, the SynCNN model demonstrated the power and adaptability of DL-based projection-domain augmentation method

specifically tailored for CBCT modality. It had the capability of capturing valuable information present in the cone-beam X-ray projections and this capability translates to an enhancement in the image quality of reconstructions.

The evaluation results in the projection domain demonstrated that the missing projections were accurately synthesized using the SynCNN model. As observed from *Figures 4,6* and *Table 2*, the synthesized projections in different composite full-view projection datasets were all
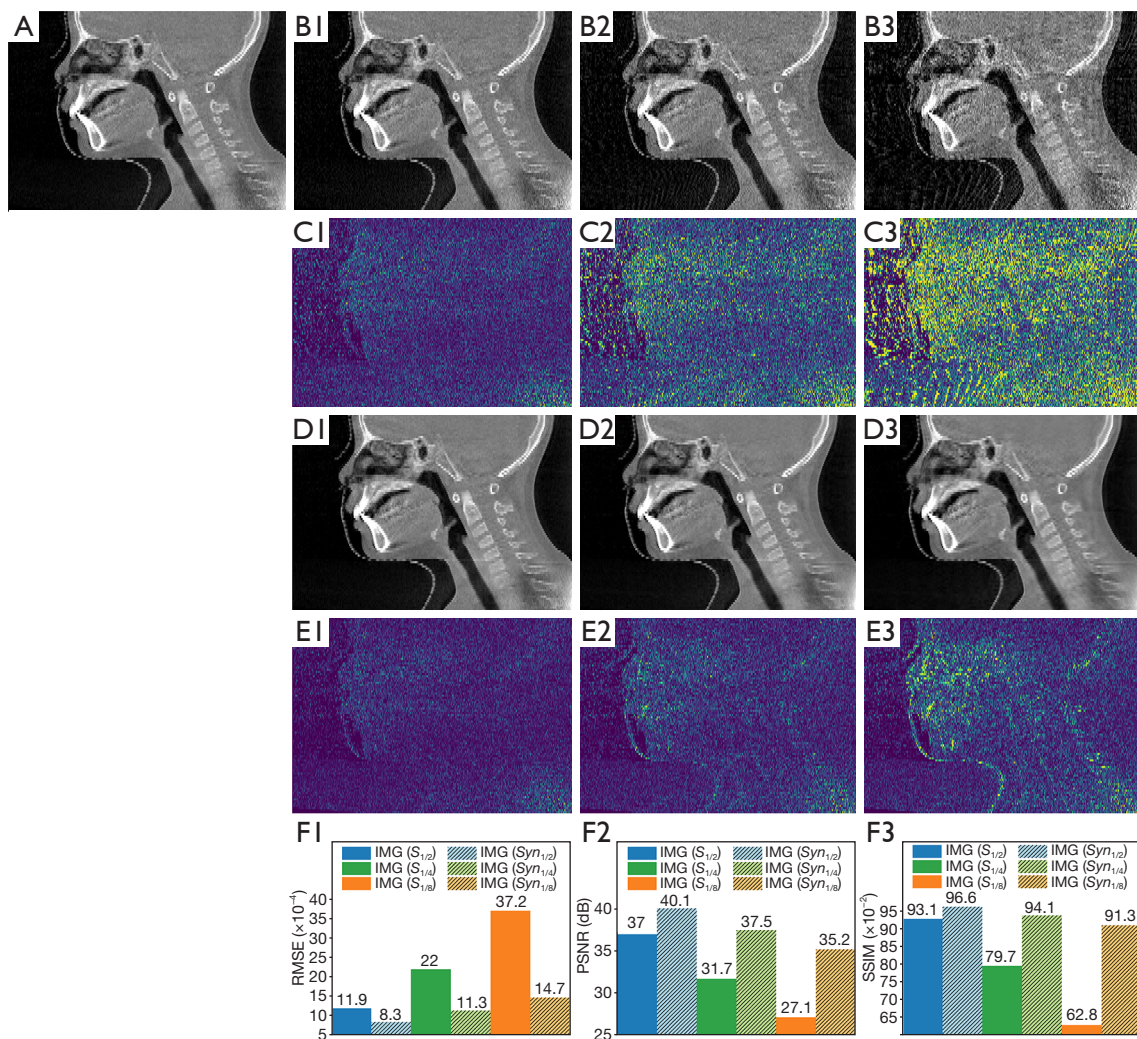
Figure 11 Representative patient sagittal images: (A) reference image reconstructed from authentic full-view projection dataset S_1; (B1-B3) images reconstructed from sparse-view projection datasets $S_{1/2}$, $S_{1/4}$, and $S_{1/8}$; (C1-C3) absolute difference maps between (B1-B3) and (A); (D1-D3) images reconstructed from composite full-view projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$; (E1-E3) absolute difference maps between (D1-D3) and (A); (F1-F3) RMSE, PSNR and SSIM values of images in IMG($S_{1/2}$), IMG($S_{1/4}$), IMG($S_{1/8}$), IMG($Syn_{1/2}$), IMG($Syn_{1/4}$), and IMG($Syn_{1/8}$) with image in IMG(S_1) as the reference. The display gray scale of (A,B1-B3,D1-D3) is [0, 0.042] and the display grayscale of (C1-C3,E1-E3) is [0, 0.006]. RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity.

close to the reference projections, from both qualitative and quantitative standpoints. As the input projections became sparser, the accuracy of the synthesized projections deteriorated due to the cascaded synthesis process.

The evaluation results in the image domain demonstrated that with the SynCNN model, the streak artifacts and noise in the reconstructed images could be mitigated substantially, leading to a significant improvement in overall image quality of sparse-view reconstructions.

This finding was consistent with the literature (31-37). The quality of images reconstructed from $Syn_{1/4}$ was found to be comparable with that of the reference images reconstructed from S_1, indicating a well-balanced trade-off between dose reduction and image quality preservation. Some issues such as edge differences and blurring of fine structures existed in IMG($Syn_{1/2}$), IMG($Syn_{1/4}$) and IMG($Syn_{1/8}$), indicating that the interpolation of missing projections using the SynCNN model could not completely eliminate the aliasing effect
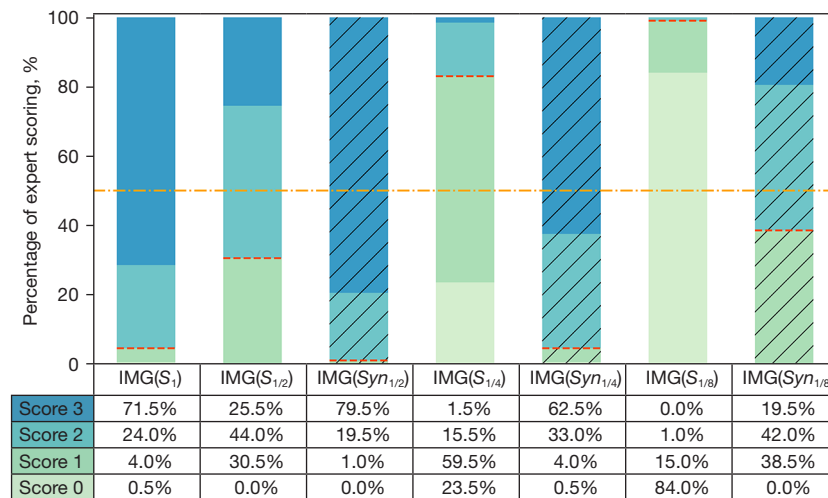
Figure 12 Subjective evaluation of image quality for seven CBCT images that were reconstructed from the full-view projection dataset S_1, sparse-view projection datasets $S_{1/2}$, $S_{1/4}$, and $S_{1/8}$, and composite projection datasets $Syn_{1/2}$, $Syn_{1/4}$, and $Syn_{1/8}$. In each bar plot, the score value corresponding to the area where the orange dash-dotted line is located represents the median score. The red dashed line in each bar plot represents the boundary of acceptable image quality. CBCT, cone-beam computed tomography.

Table 5 Average running time of the SynCNN models for projection synthesis from sparse-view projection dataset to composite full-view projection dataset

Projection synthesis	SynCNN$_{1/8}$	SynCNN$_{1/4}$	SynCNN$_{1/2}$	Subtotal
$S_{1/8} \rightarrow Syn_{1/8}$	0.4	0.8	1.7	2.9
$S_{1/4} \rightarrow Syn_{1/4}$	–	0.9	1.7	2.6
$S_{1/2} \rightarrow Syn_{1/2}$	–	–	1.8	1.8

Note that the running time is measured in seconds. SynCNN, synthesis convolutional neural network.

Table 6 Quantitative evaluation of synthesized projections in composite full-view projection datasets restored from 1/4 sparse-view projection datasets using different methods for all patient studies in the test set

Methods	RMSE (×10^{-2})	PSNR (dB)	SSIM (×10^{-2})
Linear interpolation	4.0±1.6	42.4±2.0	96.7±1.1
SynCNN (proposed)*	2.8±1.2	45.7±1.6	98.1±0.6

All values are reported as mean ± standard deviation. Best results are indicated with an asterisk (*). RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity; SynCNN, synthesis convolutional neural network.

Table 7 Quantitative evaluation of 1/4 spare-view reconstructions using different methods for all patient studies in the test set

Methods	RMSE (×10^{-4})	PSNR (dB)	SSIM (×10^{-2})
Linear interpolation	13.1±3.5	34.1±2.4	84.4±4.2
FBPConvNet	10.0±2.8	36.4±2.5	88.9±3.4
IDU-Net	10.1±2.7	36.4±2.5	88.9±3.4
SynCNN (proposed)*	9.3±2.6	37.2±2.6	90.1±3.2

All values are reported as mean ± standard deviation. Best results are indicated with an asterisk (*). RMSE, root-mean-square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity; SynCNN, synthesis convolutional neural network.

caused by under-sampling.

In IGRT, CBCT is a task-specific modality for online positioning verification, that requires clinicians to make

critical decisions. Therefore, experts should participate in image quality evaluation. In this study, we conducted blind and randomized expert scoring, and the corresponding

results demonstrated the task-specific merit of the SynCNN model. Images reconstructed from $Syn_{1/2}$, $Syn_{1/4}$ and $Syn_{1/8}$ were preferred over those reconstructed from $S_{1/2}$, $S_{1/4}$ and $S_{1/8}$. Both $IMG(Syn_{1/2})$ and $IMG(Syn_{1/4})$ received median scores identical to $IMG(S_1)$, and their acceptance rates were also comparable to $IMG(S_1)$. These findings suggested that the image quality of 1/4 sparse-view reconstructions using the SynCNN model had the ability to perform positioning verification tasks on par with regular dose images. When the down-sampling rate went to as low as 1/8, the results between $IMG(S_{1/8})$ and $IMG(Syn_{1/8})$ indicated that the substantial potential of the proposed model for image quality enhancement in ultra-sparse down-sampling scenarios.

The SynCNN model restored full-view projections from sparse-view projections at a low time cost, which was crucial for enabling online imaging in IGRT. Results showed that the average running time required for restoring full-view projections from half, quarter, and one-eighth sparse-view projections was less than 3 s. With further optimization, the process of projection synthesis could be accelerated.

The SynCNN model operates solely in the projection domain, which facilitates integrating the model into the existent CBCT imaging chain without major workflow revision. When using the SynCNN model as a generator network coupled with a discriminator network, it can be transformed into a GAN. This GAN configuration, when subjected to a generative-adversarial training strategy, holds promise in enhancing the precision of projection synthesis (55). Additionally, the flexible nature of the SynCNN model means it can be coupled with other DL-based image-domain refinement methods, offering another avenue to improve the image quality of sparse-view CBCT reconstructions.

There are three limitations of this work. First, the proposed SynCNN models synthesized the missing projections in a cascaded fashion, which worsens the blurring effect particularly in $IMG(Syn_{1/8})$. Future efforts will be directed to modifying the model architecture to synthesize multi-frame projections in one-step and incorporating attention mechanism modules to improve model performance in edge-enhancement and deblurring. Second, limited by the scope of the IRB approval, the patient scans used for model training, validation, and test were all acquired in half-scan head protocols. A new clinical study should be launched to include a variety of scan protocols and sites to further test the generalization of the proposed model. Third, the projections synthesized using the SynCNN model did not derive from the Radon transform, and this DL-based projection synthesis method might introduce new artifacts. While a rigorous mathematical explanation we believe is very necessary, we left this end loose due to a lack of strong mathematical background.

Conclusions

The proposed SynCNN model for projection synthesis is capable of improving the quality of sparse-view CBCT images at a low time cost. With the SynCNN model, the CBCT imaging dose in IGRT could be reduced potentially.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-23-759/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-23-759/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the institutional review board at Beijing Cancer Hospital. Individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the

formal publication through the relevant DOI and the license).
See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Bell K, Licht N, Rübe C, Dzierma Y. Image guidance and positioning accuracy in clinical practice: influence of positioning errors and imaging dose on the real dose distribution for head and neck cancer treatment. Radiat Oncol 2018;13:190.
2. Nabavizadeh N, Elliott DA, Chen Y, Kusano AS, Mitin T, Thomas CR Jr, Holland JM. Image Guided Radiation Therapy (IGRT) Practice Patterns and IGRT's Impact on Workflow and Treatment Planning: Results From a National Survey of American Society for Radiation Oncology Members. Int J Radiat Oncol Biol Phys 2016;94:850-7.
3. Alaei P, Spezi E. Imaging dose from cone beam computed tomography in radiation therapy. Phys Med 2015;31:647-58.
4. Zhou L, Bai S, Zhang Y, Ming X, Zhang Y, Deng J. Imaging Dose, Cancer Risk and Cost Analysis in Image-guided Radiotherapy of Cancers. Sci Rep 2018;8:10076.
5. Rehani MM, Melick ER, Alvi RM, Doda Khera R, Batool-Anwar S, Neilan TG, Bettmann M. Patients undergoing recurrent CT exams: assessment of patients with non-malignant diseases, reasons for imaging and imaging appropriateness. Eur Radiol 2020;30:1839-46.
6. Brambilla M, Vassileva J, Kuchcinska A, Rehani MM. Multinational data on cumulative radiation exposure of patients from recurrent radiological procedures: call for action. Eur Radiol 2020;30:2493-501.
7. Ibbott GS. Patient doses from image-guided radiation therapy. Phys Med 2020;72:30-1.
8. Ding GX, Alaei P, Curran B, Flynn R, Gossman M, Mackie TR, Miften M, Morin R, Xu XG, Zhu TC. Image guidance doses delivered during radiotherapy: Quantification, management, and reduction: Report of the AAPM Therapy Physics Committee Task Group 180. Med Phys 2018;45:e84-99.
9. Liu Y, Shangguan H, Zhang Q, Zhu H, Shu H, Gui Z. Median prior constrained TV algorithm for sparse view low-dose CT reconstruction. Comput Biol Med 2015;60:117-31.
10. Chen GH, Tang J, Leng S. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. Med Phys 2008;35:660-3.
11. Sidky EY, Kao CM, Pan X. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. J Xray Sci Technol 2006;14:119-39.
12. Sidky EY, Jørgensen JH, Pan X. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. Phys Med Biol 2012;57:3065-91.
13. Sidky EY, Pan X, Reiser IS, Nishikawa RM, Moore RH, Kopans DB. Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms. Med Phys 2009;36:4920-32.
14. Bian J, Siewerdsen JH, Han X, Sidky EY, Prince JL, Pelizzari CA, Pan X. Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. Phys Med Biol 2010;55:6575-99.
15. Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Phys Med Biol 2008;53:4777-807.
16. Han X, Bian J, Ritman EL, Sidky EY, Pan X. Optimization-based reconstruction of sparse images from few-view projections. Phys Med Biol 2012;57:5245-73.
17. Song Y, Zhang W, Zhang H, Wang Q, Xiao Q, Li Z, Wei X, Lai J, Wang X, Li W, Zhong Q, Gong P, Zhong R, Zhao J. Low-dose cone-beam CT (LD-CBCT) reconstruction for image-guided radiation therapy (IGRT) by three-dimensional dual-dictionary learning. Radiat Oncol 2020;15:192.
18. Cai B, Laugeman E, Mazur TR, Park JC, Henke LE, Kim H, Hugo GD, Mutic S, Li H. Characterization of a prototype rapid kilovoltage x-ray image guidance system designed for a ring shape radiation therapy unit. Med Phys 2019;46:1355-70.
19. Geyer LL, Schoepf UJ, Meinel FG, Nance JW Jr, Bastarrika G, Leipsic JA, Paul NS, Rengo M, Laghi A, De Cecco CN. State of the Art: Iterative CT Reconstruction Techniques. Radiology 2015;276:339-57.
20. Wang G, Ye JC, De Man B. Deep learning for tomographic image reconstruction. Nature Machine Intelligence 2020;2:737-48.
21. Kyong Hwan Jin, McCann MT, Froustey E, Unser M. Deep Convolutional Neural Network for Inverse Problems in Imaging. IEEE Trans Image Process 2017;26:4509-22.
22. Zhang Z, Liang X, Dong X, Xie Y, Cao G. A Sparse-View CT Reconstruction Method Based on Combination of DenseNet and Deconvolution. IEEE Trans Med Imaging 2018;37:1407-17.
23. Han Y, Ye JC. Framing U-Net via Deep Convolutional Framelets: Application to Sparse-View CT. IEEE Trans Med Imaging 2018;37:1418-29.

24. Xie S, Zheng X, Chen Y, Xie L, Liu J, Zhang Y, Yan J, Zhu H, Hu Y. Artifact Removal using Improved GoogLeNet for Sparse-view CT Reconstruction. Sci Rep 2018;8:6700.

25. Shen T, Li X, Zhong Z, Wu J, Lin Z. R^2-Net: Recurrent and Recursive Network for Sparse-View CT Artifacts Removal. Medical Image Computing and Computer Assisted Intervention – MICCAI 2019. 22nd International Conference: Shenzhen, China, October 13-17, 2019.

26. Jiang Z, Chen Y, Zhang Y, Ge Y, Yin FF, Ren L. Augmentation of CBCT Reconstructed From Under-Sampled Projections Using Deep Learning. IEEE Trans Med Imaging 2019;38:2705-15.

27. Chen M, Pu YF, Bai YC. Low-dose CT image denoising using residual convolutional network with fractional TV loss. Neurocomputing 2021;452:510-20.

28. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, Kalra MK, Zhang Y, Sun L, Wang G. Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. IEEE Trans Med Imaging 2018;37:1348-57.

29. Li M, Hsu W, Xie X, Cong J, Gao W. SACNN: Self-Attention Convolutional Neural Network for Low-Dose CT Denoising With Self-Supervised Perceptual Loss Network. IEEE Trans Med Imaging 2020;39:2289-301.

30. Huang Z, Liu X, Wang R, Chen J, Lu P, Zhang Q, Jiang C, Yang Y, Liu X, Zheng H, Liang D, Hu Z. Considering anatomical prior information for low-dose CT image enhancement using attribute-augmented Wasserstein generative adversarial networks. Neurocomputing 2021;428:104-15.

31. Lee H, Lee J, Kim H, Cho B, Cho S. Deep-Neural-Network-Based Sinogram Synthesis for Sparse-View CT Image Reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences 2019;3:109-19.

32. Dong J, Fu J, He Z. A deep learning reconstruction framework for X-ray computed tomography with incomplete data. PLoS One 2019;14:e0224426.

33. Yuan H, Jia J, Zhu Z. SIPID: A deep learning framework for sinogram interpolation and image denoising in low-dose CT reconstruction. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA: IEEE; 2018:1521-4.

34. Liang K, Yang H, Kang K, Xing Y. Improve angular resolution for sparse-view CT with residual convolutional neural network. Medical Imaging 2018: Physics of Medical Imaging. SPIE; 2018.

35. Yin X, Zhao Q, Liu J, Yang W, Yang J, Quan G, Chen Y, Shu H, Luo L, Coatrieux JL. Domain Progressive 3D Residual Convolution Network to Improve Low-Dose CT Imaging. IEEE Trans Med Imaging 2019;38:2903-13.

36. Dong X, Vekhande S, Cao G. Sinogram interpolation for sparse-view micro-CT with deep learning neural network. Medical Imaging 2019: Physics of Medical Imaging. SPIE; 2019.

37. Hu D, Liu J, Lv T, Zhao Q, Zhang Y, Quan G, Feng J, Chen Y, Luo L. Hybrid-Domain Neural Network Processing for Sparse-View CT Reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences 2021;5:88-98.

38. Bertram M, Wiegert J, Schafer D, Aach T, Rose G. Directional view interpolation for compensation of sparse angular sampling in cone-beam CT. IEEE Trans Med Imaging 2009;28:1011-22.

39. Niklaus S, Mai L, Liu F. Video Frame Interpolation via Adaptive Separable Convolution. 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE; 2017:261-70.

40. Bao W, Lai WS, Zhang X, Gao Z, Yang MH. MEMC-Net: Motion Estimation and Motion Compensation Driven Neural Network for Video Interpolation and Enhancement. IEEE Trans Pattern Anal Mach Intell 2021;43:933-48.

41. Liu YL, Liao YT, Lin YY, Chuang YY. Deep Video Frame Interpolation Using Cyclic Frame Generation. Proceedings of the AAAI Conference on Artificial Intelligence 2019;33:8794-802.

42. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv 2015. arXiv:1505.04597.

43. Bishop CM. Pattern Recognition and Machine Learning. Springer New York; 2006.

44. Sironi A, Tekin B, Rigamonti R, Lepetit V, Fua P. Learning Separable Filters. IEEE Trans Pattern Anal Mach Intell 2015;37:94-106.

45. Xue T, Wu J, Bouman KL, Freeman WT. Visual dynamics: Probabilistic future frame synthesis via cross convolutional networks. NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems. 2016.

46. Zhang Y, Yin FF, Pan T, Vergalasova I, Ren L. Preliminary clinical evaluation of a 4D-CBCT estimation technique using prior information and limited-angle projections. Radiother Oncol 2015;115:22-9.

47. De Brabandere B, Jia X, Tuytelaars T, Van Gool L. Dynamic filter networks. NIPS'16: Proceedings of the 30th International Conference on Neural Information

Processing Systems. 2016.

48. Du Y, Wang R, Biguri A, Zhao X, Peng Y, Wu H. TIGRE-VarianCBCT for on-board cone-beam computed tomography, an open-source toolkit for imaging, dosimetry and clinical research. Phys Med 2022;102:33-45.

49. Hsieh J. Computed tomography: principles, design, artifacts, and recent advances. Bellingham, WA, USA: SPIE Press; 2009.

50. Sun M, Star-Lack JM. Improved scatter correction using adaptive scatter kernel superposition. Phys Med Biol 2010;55:6695-720.

51. Fisher RA, Yates F. Statistical tables for biological, agricultural and medical research. Hafner Publishing

Company; 1953.

52. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. arXiv 2014. arXiv:14126980.

53. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600-12.

54. Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. J Opt Soc Am A Opt Image Sci Vis 1984;1:612-9.

55. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. Med Image Anal 2019;58:101552.