Computer Methods and Programs in Biomedicine 256 (2024) 108376

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine

An efficient dual-domain deep learning network for sparse-view CT reconstruction

Chang Sun [a,b], Yazdan Salimi [b], Neroladaki Angeliki [c], Sana Boudabbous [c], Habib Zaidi [b,d,e,f,*]

[a] Beijing University of Posts and Telecommunications, School of Information and Communication Engineering, 100876 Beijing, China
[b] Geneva University Hospital, Division of Nuclear Medicine and Molecular Imaging, CH-1211 Geneva, Switzerland
[c] Geneva University Hospital, Division of Radiology, CH-1211, Geneva, Switzerland
[d] Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands
[e] Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark
[f] University Research and Innovation Center, Óbuda University, Budapest, Hungary

ARTICLE INFO

Key words:
CT, reconstruction
Sparse-view
Deep learning
Thoraco-abdominal scans

ABSTRACT

Background and Objective: We develop an efficient deep-learning based dual-domain reconstruction method for sparse-view CT reconstruction with small training parameters and comparable running time. We aim to investigate the model's capability and its clinical value by performing objective and subjective quality assessments using clinical CT projection data acquired on commercial scanners.
Methods: We designed two lightweight networks, namely Sino-Net and Img-Net, to restore the projection and image signal from the DD-Net reconstructed images in the projection and image domains, respectively. The proposed network has small training parameters and comparable running time among dual-domain based reconstruction networks and is easy to train (end-to-end). We prospectively collected clinical thoraco-abdominal CT projection data acquired on a Siemens Biograph 128 Edge CT scanner to train and validate the proposed network. Further, we quantitatively evaluated the CT Hounsfield unit (HU) values on 21 organs and anatomic structures, such as the liver, aorta, and ribcage. We also analyzed the noise properties and compared the signal-to-noise ratio (SNR) and the contrast-to-noise ratio (CNR) of the reconstructed images. Besides, two radiologists conducted the subjective qualitative evaluation including the confidence and conspicuity of anatomic structures, and the overall image quality using a 1–5 likert scoring system.
Results: Objective and subjective evaluation showed that the proposed algorithm achieves competitive results in eliminating noise and artifacts, restoring fine structure details, and recovering edges and contours of anatomic structures using 384 views (1/6 sparse rate). The proposed method exhibited good computational cost performance on clinical projection data.
Conclusion: This work presents an efficient dual-domain learning network for sparse-view CT reconstruction on raw projection data from a commercial scanner. The study also provides insights for designing an organ-based image quality assessment pipeline for sparse-view reconstruction tasks, potentially benefiting organ-specific dose reduction by sparse-view imaging.

1. Introduction

Computed tomography (CT) is an indispensable imaging modality widely used in clinical practice, providing highly detailed axial (cross-sectional) images of internal anatomical structures [1]. CT can be used to image almost any part of the body, making it valuable for a wide range of medical applications. Compared with other imaging modalities, CT scans are fast and can be advantageous in emergency situations where rapid diagnosis and decision-making are crucial. According to Harvard Health Publishing, over 80 million CT scans are performed in the US each year compared to about 3 million in 1980 [2]. However, the possible harmful effects of exposure to ionizing radiation from CT examinations have become a subject of public concern. According to the report issued in 2009 by the National Council on Radiation Protection (NCRP), the largest portion of the increase in medical exposures came from CT scans, accounting for almost half of imaging medical exposures. A study by Gonzalez et al. [3] estimated that CT scans performed in the

* Corresponding author at: Geneva University Hospital, Division of Nuclear Medicine and Molecular Imaging, CH-1211 Geneva, Switzerland.
E-mail address: habib.zaidi@hug.ch (H. Zaidi).

https://doi.org/10.1016/j.cmpb.2024.108376
Received 23 May 2024; Received in revised form 2 August 2024; Accepted 15 August 2024
Available online 16 August 2024

Abbreviations

SNR	signal-to-noise ratio
CNR	contrast-to-noise ratio
CT	computed tomography
FBP	filtered back projection
SIRT	simultaneous iterative reconstruction technique
DL	deep learning
CBCT	Cone-beam computed tomography
ADMIRE	advanced modeled iterative reconstruction
RLFB	residual local feature blocks
ODL	operator discretization library
SART	simultaneous algebraic reconstruction technique
CGLS	conjugate gradient least squares
MAE	mean absolute error
PSNR	peak signal-to-noise ratio
SSIM	structure similarity Index
MedianAE	median absolute error
NPS	noise power spectrum
nNPS	normalized noise power spectrum
ROI	region of interest
ICC	intra-class correlation coefficient
SDE	standard deviation of errors
SD	standard deviation

US in 2007 might be related to approximately 29,000 cases of future cancer. Currently, the importance of radiation protection in medicine and the "as low as reasonably achievable" (ALARA) principle [4] have become an internationally increasing awareness. A large number of risk-reduction efforts have been made in CT radiation protection [5], including justification and optimization of CT procedures [6-8] and the development of low-dose imaging techniques while maintaining sufficient image quality for diagnostic purposes [9]. In this study, we focus on sparse-view CT imaging, which is considered a promising low-dose CT strategy that reduces the number of measurements by expanding the projection-view interval during acquisition.

Reconstructing CT images from sparse-view projections (sinograms) are challenging and ill-posed inverse problems. Conventional reconstruction algorithms include analytical methods, such as filtered back-projection (FBP) and iterative methods, such as Simultaneous Iterative Reconstruction Technique (SIRT) [10]. In recent years, various deep learning (DL)-based methods have been proposed. Most of these methods use large dataset and supervised learning strategy to map the input data to high-quality sinograms or images enabling to achieve high performance in terms of reducing artifacts and restoring structural details compared to iterative methods. Projection-domain [11,12] and image-domain [13,14] DL methods focus on generating the missing projections and post-processing the images reconstructed from conventional methods, respectively. In particular, Zhang et al. [15] designed an effective DenseNet and deconvolution-based network (DD-Net) for image-domain reconstruction with the advantages of high reconstruction performance, small model size, and short running time. However, projection-domain methods lack post-processing of the reconstructed images, while image-domain methods do not fully utilize the raw measurement data from the sensor.

To make full use of the information in both domains, dual-domain DL methods were developed to enhance the sinograms and reconstructed images at the same time, e.g., DRONE [16], CDDCN [17], CLRecon [18], HDNet [19] and DualCNN [20]. However, the main drawbacks of these dual-domain joint learning networks are the large number of trainable parameters, high computational time and the independent optimization of sub-networks. For example, CDDCN has 42.1 M number of trainable parameters and the computational time of DRONE is about 138.48 s with image matrix of 512×512 and 624 detector elements on an NVIDIA TITAN XP GPU. Another popular dual-domain reconstruction scheme consists in directly learning the mapping between the sinogram and image domain [21,22]. On the other hand, Wu et al. [23] proposed a deep embedding attention-refinement (DEAR) network for ultra-sparse-view CT reconstruction. They constructed a new objective function for DEAR with constraints in the data, image, and sparsified transform domains. They also incorporated deep embedding, deep attention, and deep refinement modules based on DNNs for iterative optimization. Pan et al. [24] provided a novel iterative residual optimization network (IRON) to suppress the unobservable artifacts in the reconstructed CT images. They proposed a hybrid structure of the convolutional neural network (CNN) and Transformer called CAT and integrated it as a regularization term to IRON, which alternatively reduces the residual measurements, noise, and artifacts in the image domain and the projection domain at each iteration stage based on the block-coordinate descent. IRON shows competitive reconstruction results and good stabilities when coping with the tiny perturbation in input data. As each iteration applies the forward and back projection to the outcomes of the previous iteration, these iterative optimization methods require large hardware resources. They are time-consuming, especially for real CT projection data with large sizes. In recent years, with the great success of generative models in image-processing tasks, researchers have applied diffusion models to CT reconstruction tasks and achieved outstanding reconstruction performance in denoising and recovering structural details [25-27]. For example, Wu et al. [28] proposed a multi-channel optimization generative model (MOGM) for ultra-sparse-view CT reconstruction, integrating a multi-channel fusion strategy. This multi-channel-based data consistency policy relied on the original data and a multi-channel iteration optimization framework into the score-based generative model (SGM). The proposed method tackled the problem of instability of single-channel-based SGM and enhanced the reconstructed image quality with foundational theoretical supports. Wu et al. [29] also developed an unsupervised CT reconstruction network based on SGM to address the challenge of acquiring high-quality training data. They presented a novel model incorporating the strategy of the wavelet sub-network and the SGM sub-net to reduce the noise perturbations and improve the reconstruction performance effectively. Wang et al. [30] proposed a rapid-sampling strategy for SGM, namely the time-reversion fast-sampling (TIFA) score-based model for limited-angle CT reconstruction. TIFA can achieve nearly ×100 acceleration compared to traditional SGMs while maintaining comparable performance. Li et al. [31] also proposed a dual-domain collaborative diffusion Sampling (DCDS) model that includes sinogram and image domain diffusion processes for multi-source stationary CT reconstruction. They designed a collaborative merging mechanism and an iteration optimization approach to enhance data consistency. However, due to the nature of SGM and the multi-channel design, these methods require more memory and time in the training or testing phase, which leaves room for further enhancing the computational efficiency.

Designing lightweight network models has become a hot research topic enabling to reduce the number of learnable parameters of a DL network and the requirement of large computational memory in the training and testing phases [32]. Lightweight network techniques can be broadly categorized into two types: designing network structures and model compression. Currently, designing specific convolution operations (e.g., group convolution) to reduce the computational complexity of the model has been widely explored in low-level visual tasks, such as image denoising [33], low-light image enhancement [34], and image super-resolution [35]. In medical image enhancement tasks, Li et al. [36] proposed a lightweight VolumeNet based on group convolution and feature aggregation for super-resolution of medical volumetric data.

Cheng et al. [37] proposed a lightweight Unified Fusion Network (UNIFusion) based on the Ghost module and auto-encoder, which can be used for PET and MR image fusion. Ma et al. [38] proposed a lightweight block reconstruction network (LBRN) for CT reconstruction, which contains 708.9 M learnable parameters and can directly learn the mapping between the sinusoidal and image domains. Inspired by these works, we aimed to design lightweight CT reconstruction networks for both projection and image domains to reduce the model size and decrease the computational resource requirements when processing real projection data, thereby offering significant benefits in terms of efficiency and resource utilization.

The lack of clinical data is one of the problems in deep learning-based medical imaging applications. Especially, for the task of sparse-view CT reconstruction, raw projection data of clinical CT images are more difficult to obtain. Therefore, the use of synthetic incomplete-sinograms and sparse-view CT images to train DL networks has been widely explored. The most popular pipeline for incomplete-sinograms synthesis is to calculate the parallel/fan-beam projection data (sinogram) from the image, which includes the following steps: (i) collecting CT images from clinical, experimental, or computerized phantoms; (ii) setting scanning geometry parameters, e.g., the type of projection beams, the number of detector channels and the number of full projections per rotation; (iii) generating full-view sinograms using Radon transform or open-source toolboxes such as, ASTRA [39], TIGRE [40], or ODL (Operator Discretization Library) [41]; (iv) performing noise degradation ("Gaussian + Poisson" noise is widely used) according to the protocol and Beer–Lambert law; (v) down-sampling the projected data to simulate sparse-view sinograms; (vi) reconstructing the corresponding sparse-view CT images from the sparse-view sinograms using FBP or other iterative methods; and (vii) validating the reconstructed images qualitatively (by radiologists for clinical outcomes) and quantitatively (using image-derived metrics).

The advantage of the above-mentioned approach is that it can simulate most scanning geometries and acquisition protocols without the need for real raw projection data. However, the drawback of using synthetic raw projections is that the simulation of CT scanning geometries inherently bears limitations and lack of realism even when using the most sophisticated approaches. This prevents us from understanding the accuracy and the computational cost of these methods on real CT imaging data from commercial medical CT scanners. Moreover, the simulated data may not match the actual data in statistical distribution, which may degrade the algorithm's performance in clinical scenarios. More recently, efforts have been made to use clinical raw CT projection datasets for training and testing DL-based methods in CT reconstruction tasks. Moen et al. [42-44] proposed a low-dose CT image and projection dataset consisting of 299 clinical CT projection data from two different CT manufacturers to facilitate the development and validation of CT reconstruction on real CT data. However, only a few methods were evaluated. Besides, the experimental validation was not sufficiently focused in some previous studies. For example, the number of test subjects was small, the quantitative evaluation metrics were not compared in the region of interest, and diagnostic quality assessment by radiologists lacking.

In this study, we focus on investigating the effectiveness of sparse-view CT reconstruction methods on real clinical CT projection data. We propose a lightweight DL network that performs end-to-end joint learning in both projection and image domains. Different from the mentioned dual-domain methods, the proposed model has a small model size and competitive running time for raw CT data with a large projection size. We trained and evaluated the model on real CT projection data and compared its performance with iterative reconstruction methods and state-of-the-art DL-based methods. The contributions of this study are as follows:

1. We propose an end-to-end dual-domain deep learning network for sparse view CT reconstruction. The proposed network is able to achieve good performance with a small model size (1.08 M training parameters, 4 MB storage size) and competitive runtime (0.3 s per slice on GPU) with joint hybrid domain learning for large projection sizes of real CT projection data.
2. Instead of simulating scan geometry and system noise to generate CT projections of CT images, we developed and validated the reconstruction algorithms using the clinical raw projection data directly from a commercial multi-slice spiral CT scanner.
3. We used a reconstruction platform (ReconCT, Siemens Healthineers, Erlangen, Germany) and the commercial reconstruction method advanced modeled iterative reconstruction (ADMIRE) to generate reference CT images directly from the raw projection data. The quality of the reference image is ensured.
4. In performance evaluation, we not only compared the numerical metrics of the proposed method and state-of-the-art methods, but also conducted organ-based assessment, subjective quality evaluation by radiologists, and computational cost evaluation.

2. Materials and methods

2.1. Patients and datasets

The workflow adopted in this study is given in Fig. 1. Sixty-four patients who underwent thoraco-abdominal oncologic follow-up PET/CT scans were retrospectively included. Acquisitions were performed on a Biograph 128 Edge CT scanner with a helical scanning geometry with a curved panel detector having 736 detector elements along the detector arc, a source-to-object distance of 595 mm, a source-to-detector distance of 1085.6 mm, 2304 projected views uniformly distributed around 360° (one rotation), a gantry rotation time of 500 msec and a collimation of 64 × 0.6 mm. Table 1 reports the scan parameters and patient's characteristics.

2.1.1. Generation of reference CT images

The CT raw projection data were exported from the PET/CT console and loaded into an offline reconstruction platform ReconCT (version 16.0.0.5593; Siemens Healthineers). ReconCT allows reading the CT raw data and performing the same commercially used reconstruction algorithms, such as ADMIRE, which is a widely used statistical iterative reconstruction method developed by Siemens Healthcare, that offers high performance in terms of noise and dose reduction. Reference CT images were reconstructed using the following parameters: 500 mm field-of-view (FOV), 512 × 512 matrix, 0.6 mm slice thickness, 0.6 mm increment, 0.98 mm spacing, ADMIRE (strength 3) and soft-tissue kernel (Br36f). We used 45 patient datasets for training (34,135 images), 6 for validation (440 images) and the remaining 13 for testing (10,713 images).

2.1.2. Generation of multi-slice fan-beam sinograms

After data acquisition, we read the CT raw projection data using a customized MATLAB (MathWorks, MATLAB version R2021b) script and save the raw CT projections and meta-information into a file in HDF5 format. The meta-information is necessary for the generation of the fan-beam sinograms and the subsequent image reconstruction. It includes the physical parameters of the scanner gantry and the CT detector provided by the Siemens Healthineers, the acquisition geometry and scanning parameters selected by the imaging equipment operator, as well as the measurement angle of the projection and its relative position in the longitudinal direction during the acquisition.

We applied a sinogram rebinning method [45] to first convert the original raw helical CT projection from a curved panel detector geometry to a flat panel detector geometry, and then rebinned the helical projection into multi-slice circular fan-beam projections using a Python (version 3.9) script developed by Wagner et al. [43,46]. As a result, each raw helical CT projection data was converted into a stack of flat-panel detector-based fan-beam sinograms. The resolution of the fan-beam

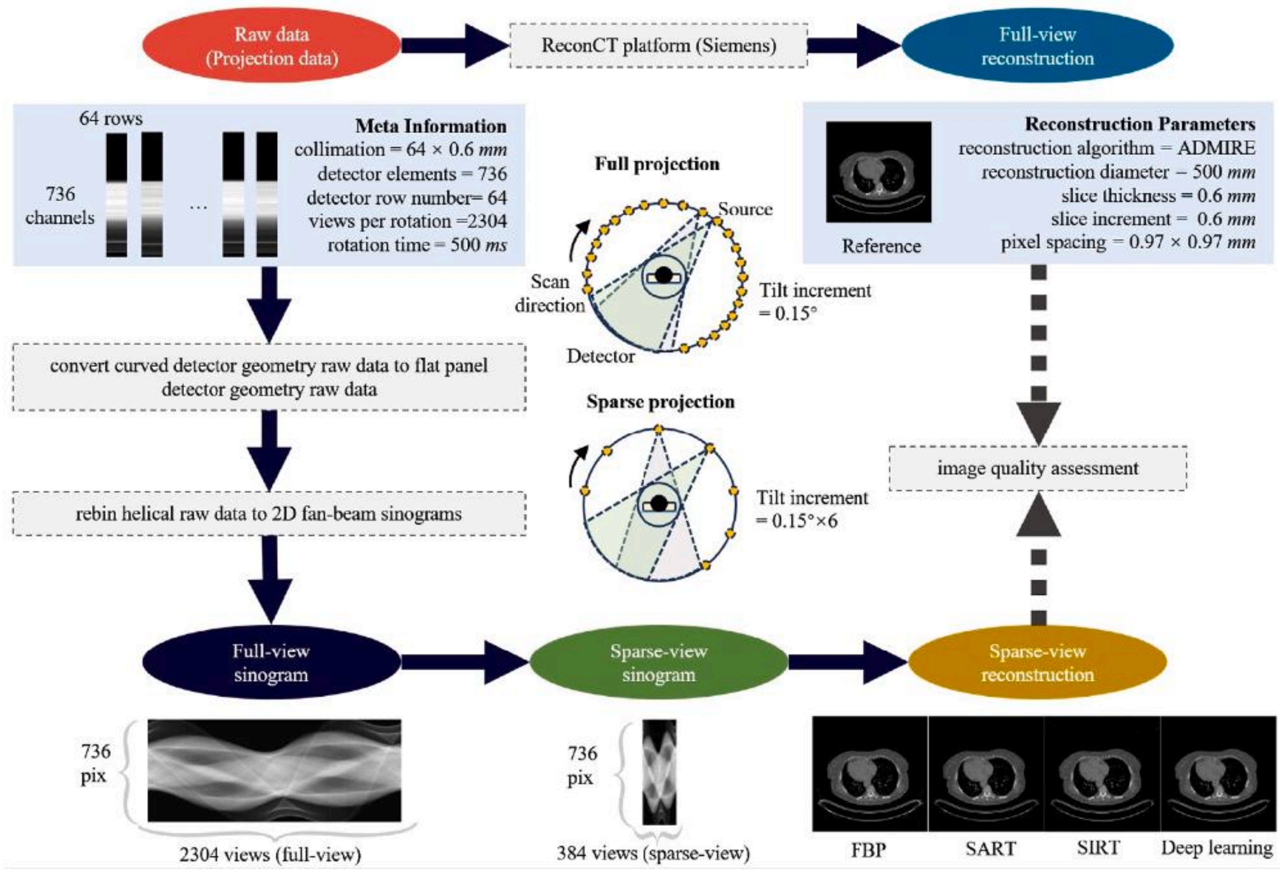

Fig. 1. Workflow adopted in this study, including CT raw data acquisition, full-view CT reconstruction using ReconCT software, full-view sinograms rebinning, sparse-view sinograms generation, sparse-view CT reconstruction and evaluation of reconstruction methods.

Table 1
Summary of scan parameters and patient's characteristics. (Values are expressed as mean ± standard deviation or number (percentage)).

	Training	Validation	Testing
Number of cases	45	6	13
Age (years)	56± 14	62± 11	66± 10
Gender			
Female	24 (53 %)	3 (50 %)	4 (31 %)
Male	21 (47 %)	3 (50 %)	9 (69 %)
Pitch			
0.6	40 (89 %)	4 (67 %)	13 (100 %)
0.8	2 (4 %)	1 (17 %)	0 (0 %)
0.95	3 (7 %)	1 (17 %)	0 (0 %)
kVp			
80	1 (2 %)	0 (0 %)	0 (0 %)
100	38 (84 %)	5 (83 %)	12 (92 %)
120	6 (13 %)	1(17 %)	1(8 %)
Tube current (mAs)	270.7± 129.6	373.5± 90.1	309.0± 99.6

sinogram was 736×2304 pixels. After generating the fan-beam sinograms, we sparsely down-sampled the sinograms to 384 views to simulate the sparse view scanning geometry with a sampling rate of 1/6.

2.2. Proposed network

2.2.1. Conceptual design of the network

The purpose of this study was to develop an efficient deep learning method for reconstructing sparse-view CT images. We began with an image-domain post-processing method called DD-Net [15], which is a smaller model compared to other classic image-domain-based methods, such as Improved GoogLeNet [47], Tight frame U-Net [48], and RED-CNN [49]. DD-Net achieves similar reconstruction performance by using a combination of U-shape architecture, dense connections, and effective deconvolution operations. However, it does not preserve the original measured projection data, leaving room for further improvement in image quality. Therefore, we need a Sino-Net to ensure data fidelity in the sinogram domain. Additionally, we require an Img-Net to maintain consistency with reference images and correct errors in the image domain. Sino-Net and Img-Net are designed to enhance the reconstruction model within a dual-domain reconstruction framework. We aim to find a network structure that balances reconstruction performance and computational overhead. Many lightweight networks developed for image restoration tasks may meet this criterion, as mentioned in the Introduction section [32]. In this study, we only used residual local feature blocks (RLFBs) [50] in Sino-Net and Img-Net. RLFB is the core module of the Residual Local Feature Network (RLFN), which won first place in the main track of the NTIRE 2022 efficient super-resolution challenge [51]. We chose RLFB because it uses three convolutional layers for residual local feature learning to simplify feature aggregation and applies an enhanced spatial attention (ESA) [52] layer with a large reception field to focus on important spatial content and reconstruct rich edge and texture information. The network includes widely used operations in lightweight network architecture design strategies, such as strided convolutions and max-pooling with large window sizes."

2.2.2. Network structure

We developed a dual-domain DL-based network for sparse-view CT reconstruction. As shown in Fig. 2, our method includes three steps: Firstly, we suppress the streaking artifacts in the FBP-reconstructed image using a classical image-domain reconstruction network DD-Net

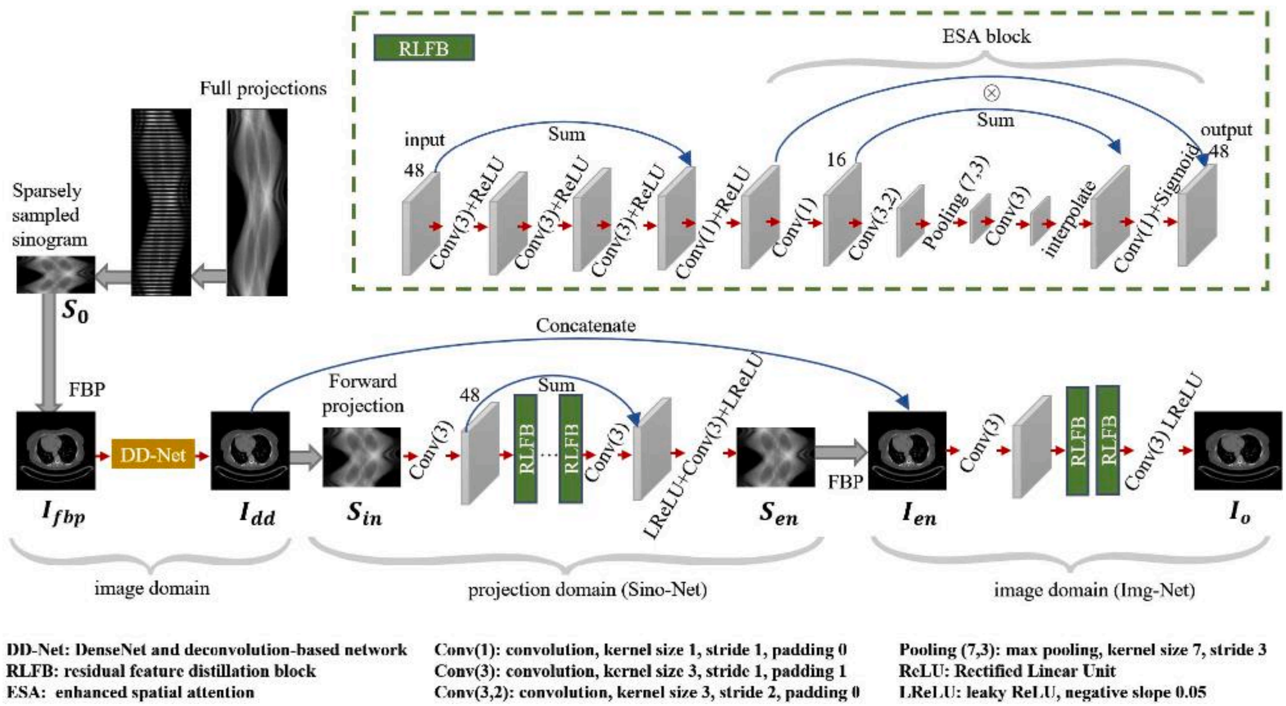

Fig. 2. Proposed dual-domain DL-based sparse-view CT reconstruction network.

[15]. Secondly, we forward project the enhanced images to dense-view sinograms (512 views) and then use an efficient CNN (Sino-Net) to enhance the sinograms. Finally, we reconstruct the enhanced sinograms using FBP, concatenate them with the output of DD-Net, and finally use a lightweight CNN (Img-Net) to further improve image quality.

In the proposed method, as the FBP image I_{fbp} reconstructed from the sparsely sampled sinogram is highly degraded, it is first fed to DD-Net to recover an initial refined image I_{dd}. As mentioned earlier, DD-Net is a well-known image-domain reconstruction network consisting of a U-Net based architecture with four times pooling operations to extract deep features and combining the dense connections between different layers and deconvolution operations to improve the reconstruction quality. Effectiveness of DD-Net is the motivation behind its use. Compared with other image-domain reconstruction networks, it presents advantages in terms of computational speed compared to Framing U-Net [48], which performs wavelet transforms and is more lightweight with only 0.56 M training parameters compared with Improved GoogLeNet [47], which has 1.25 M training parameters.

Despite the performance improvement achieved by DD-Net, the intrinsic information of the raw projection is not directly used. Therefore, we perform sinogram recovery after the initial image-domain refinement to ensure the data consistency of the projection data. Specifically, we forward project the initial refined image I_{dd} to a dense-view sinogram S_{in} with 512 views using the same fan-beam geometry. Then, we propose using Sino-Net lightweight model for sinogram enhancement. We project the image to dense-view instead of full-views (2304 views) due to the balance between the performance improvement and the computational burden of the dual-domain network. We also designed the Sino-Net architecture based on this principle.

Inspired by recent DL-based image super-resolution approaches that achieve high running speed while maintaining image quality [50], we designed the Sino-Net based on residual local feature blocks (RLFBs), which is the core module of the Residual Local Feature Network (RLFN) [50] that won the first place in the main track of NTIRE 2022 efficient super-resolution challenge [53]. As shown in Fig. 2, the dense-view sinogram S_{in} first uses a 3×3 convolutional layer to extract the spatial feature f_s, followed by n cascaded RLFBs. The RLFB first uses three stacked 3×3 convolutional layer+ReLU layers to perform residual learning; at this time, the refined feature is obtained. Using the residual feature learning has advantages in suppressing the training instability and improving the trainability. Then, a 1×1 convolutional layer+ReLU layer is employed to further combine the feature information across channels. Subsequently, a powerful enhanced spatial attention (ESA) block is used at the end, with a large receptive field to get more representative features for signal recovery. The ESA first reduces channel dimensions using a 1×1 convolutional layer, and then enlarges the receptive field by combining a strided 3×3 convolutional layer (stride 2) and a max pooling layer (kernel size 7, stride 3). After further extracting the information from the feature using a 3×3 convolutional layer, an interpolate operation is applied to recover the spatial dimensions. A skip connection is applied to forward the feature obtained before the spatial dimension reduction. Finally, the attention mask is generated using a 1×1 convolutional layer+Sigmoid layer.

After n stacked RLFBs, a 3×3 convolutional layer is applied and the residual feature f_r is obtained. Then, the spatial feature f_s and the residual feature f_r are fused and reconstructed to an enhanced dense-view sinogram S_{en} using a cascaded leaky ReLU+3×3 convolutional layer+leaky ReLU layer. Finally, the enhanced dense-view sinogram S_{en} is inputted into an FBP layer to reconstruct an enhanced image I_{en}.

After acquiring the initial refined image I_{dd} recovered by DD-Net, and the enhanced image I_{en} recovered by Sino-Net, we concatenate them together and use a small model, img-Net, to reconstruct the structure details in the CT image. Similar to Sino-Net, we also apply RLFB to build img-Net, but with fewer layers to reduce the computation cost. The architecture of img-Net is shown in Fig. 2.

2.2.3. Training

The proposed network was trained in an end-to-end manner using an overall mixed loss function as follows:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{M}\sum_{i=1}^{N}|x(i,j) - \widehat{x}(i,j)|}{MN} \quad (1)$$

where I_{gt} denotes the reference CT images, S_{gt} denotes the forward projection of I_{gt} with 512 dense views. I_{en}, S_{en}, and I_o denote the output of

DD-Net, Sino-Net and Img-Net, respectively. M_o is the binary mask for image regions belonging to the patient body. We empirically set α = β = 1 to achieve optimal performance. We used PyTorch [54] for network training and ODL [41] for the implementation of FBP and forward projection. The training parameters were as follows: optimizer=Adam [55], initial learning rate=2e-4, batch size=1, and training epoch=10. The network was trained on a PC equipped with a NVIDIA 2080Ti GPU.

Prior to the training, CT values were clipped to [−1000, 2000] HUs and then normalized to [0, 1]. Sinogram values were clipped to [0, 9] and then normalized to [0, 1]. The number of RLFB in Sino-Net and Img-Net was set to 6 and 2, respectively. The channel numbers of the proposed network, empirically chosen according to [50], are shown in Fig. 2.

2.3. Evaluation strategy

We compared the performance of the proposed method with FBP, three iterative reconstruction methods, including simultaneous algebraic reconstruction technique (SART) [56], SIRT [10], and the conjugate gradient least squares (CGLS) [57], and three DL-based methods, including the image-domain postprocessing method DD-Net [15], which has been a common benchmark model for DL-based methods, a projection-to-image DL-based method LRR-CED [22], and a dual-domain learning approach, namely CLRecon [18]. The motivation behind the choice of LRR-CED and CLRecon was that the two methods have been tested on simulated or real CT data with projection geometry similar to the dataset used in this study. For example, the same fan-beam geometry with large projection size (2304×736) or (1152×736) was used. Besides, these two methods perform an end-to-end network training, which is easy to implement and is advantageous in terms of computational cost.

The ODL library was used to implement FBP and iterative reconstruction methods. The number of iterations for SART, SIRT and CGLS were 300, 200 and 18, respectively. FBPConvNet [13] was used as the base network for CLRecon, as used in the original paper. For LRR-CED, we used the U-Net-based structure LRR-CED(U) as it performed better than DenseNets-based network LRR-CED(D) in terms of PSNR and SSIM in the experiment of real CT data in the original study. The training parameters of DD-Net, LRR-CED(U), and CLRecon were similar to implementations reported in the original papers. All comparisons were implemented on the same PC as the proposed network.

2.3.1. Objective overall quality evaluation

All CT images were cropped to cover only the patients' body contour to eliminate the effect of the bed and background air. The mean absolute error (MAE), peak signal-to-noise ratio (PSNR), and structure similarity (SSIM) were used. They are defined as follows:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}|x(i,j) - \widehat{x}(i,j)|}{MN} \tag{2}$$

$$\mathrm{PSNR} = 20 \times \log_{10}\left(\frac{MN \parallel x\|_\infty}{\parallel x - \widehat{x}\|}\right) \tag{3}$$

$$\mathrm{SSIM} = \frac{(2\mu_x\mu_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + C_1)(\sigma_x^2\sigma_{\hat{x}}^2 + C_2)} \tag{4}$$

where x and $\widehat{x}$ denote the reconstructed image and reference image, respectively. M and N denote the numbers of pixels for the row and column, respectively. μ_x and $\mu_{\hat{x}}$ are the mean of images x and $\widehat{x}$, σ_x^2 and $\sigma_{\hat{x}}^2$ are the variance of x and $\widehat{x}$, $\sigma_{x\hat{x}}$ is the covariance of x and $\widehat{x}$, $C_1 = 0.01L$ and $C_2 = 0.03L$ are constants with L denotes the dynamic range of the reference image. Before evaluation, the intensity of the reconstructed image and reference image were clipped to [−1000, 2000] HUs, and further normalized to [0, 1] to measure the PSNR and SSIM. The MAE and PSNR metrics were measured in only the region containing the patient's body.

2.3.2. Organ-based assessment

In addition, we performed an organ-based assessment by comparing the reconstruction performance in different anatomical structures. Firstly, an automated segmentation network [58] was applied on five test cases to segment the body into 21 regions including the whole heart, left/right atrium, left/right ventricle, myocardium, lungs, adrenal, spleen, liver, kidneys, pancreas, stomach, esophagus, colon, small bowel, aorta, autochthonous dorsal musculature (autochthon), clavicles, ribcage and vertebrae. Fig. 3a displays the segmentation results of one case. Subsequently, the median of all the absolute differences between the predicted CT values and the reference CT values, denoted as median absolute error (MedianAE), was used to measure the average prediction performance of the models in different anatomical structures because it is insensitive to outliers compared to MAE. The standard deviation of the prediction errors, SDE, was used to evaluate the spread of the prediction errors. The MeidianAE and the SDE were defined as follows:

$$MedianAE = median\ (|y_1 - \widehat{y}_1|,\ \cdots, |y_n - \widehat{y}_n|) \tag{5}$$

$$SDE = standard\ deviation\ (y_1 - \widehat{y}_1,\ \cdots, y_n - \widehat{y}_n) \tag{6}$$

where y_i and $\widehat{y}_i$ denote thei^{th} reconstructed CT value and the reference CT value of the evaluated organ area,respectively. n denotes the number of voxels of the organ being evaluated.

2.3.3. Analysis of lung lesions and liver lesions

The signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) were used as metrics for evaluating the reconstruction quality in the lung and liver area. They were compared between images reconstructed using different methods. One radiologist with over 10 years experience segmented the lung and liver lesions manually on the test dataset. In the liver, the HU value of the lesion was obtained using a region of interest (ROI) as large as possible (range 13–97 mm^2), drawn on the slice containing the lesion. To measure the mean HU value of the liver, two circular ROIs (each 107 mm^2) were placed within 30 mm of the lesion, avoiding the vessel and intestinal gas. Similarly for the lung, a ROI as large as possible (range, 30–106 mm^2) was drawn on the image containing the lesion. Besides, three circular ROIs (each 69 mm^2) were placed in the lung to measure the HU value of the lung. For the comparison of images reconstructed by different methods, the ROIs were first defined on the reference images and then pasted to other images. The radiologist evaluated the 13 cases from the test datasets. After excluding cases with no liver lesions or small lesions (<10 mm^2) in the lung, 5 and 4 cases were included for calculating the CNR and SNR of the liver and lung, respectively.

The SNR of the liver and the CNR of the lesion to the liver was calculated using the following formulas:

$$\mathrm{SNR}_{liver} = ROI_{liver}/\sigma_{liver} \tag{7}$$

$$\mathrm{CNR}_{lesion(liver)} = \left|ROI_{lesion(liver)} - ROI_{liver}\right| \Big/ \sqrt{\left(\sigma^2_{lesion(liver)} + \sigma^2_{liver}\right)\Big/2} \tag{8}$$

where ROI_{liver} and σ_{liver} denote the mean and the standard deviation of CT HU value within the liver ROIs, respectively. $ROI_{lesion(liver)}$ and $\sigma_{lesion(liver)}$ denote the mean and the standard deviation of CT HU value within the liver lesion ROIs, respectively. In the same way, the CNR of the lesion in the lung was calculated using the following formula:

$$\mathrm{CNR}_{lesion(lung)} = \left|ROI_{lesion(lung)} - ROI_{lung}\right| \Big/ \sqrt{\left(\sigma^2_{lesion(lung)} + \sigma^2_{lung}\right)\Big/2} \tag{9}$$

where ROI_{lung} and σ_{lung} denote the mean and the standard deviation of CT HU value in the lung ROIs, respectively. $ROI_{lesion(lung)}$ and $\sigma_{lesion(lung)}$

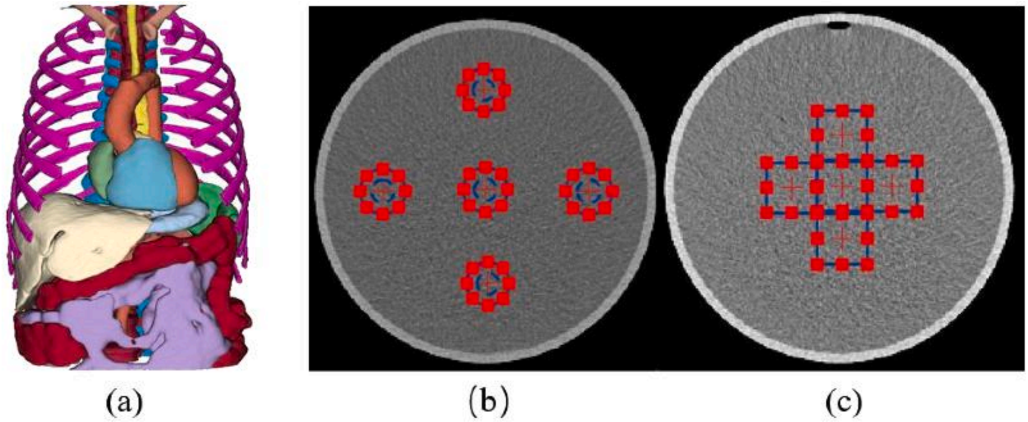

Fig. 3. Example of organ segmentation results and ROI placement for analyzing noise properties. (a) organ segmentation result. (b) ROI placement for measuring the noise magnitude. (c) ROI placement for calculating noise power spectrum.

denote the mean and the standard deviation of CT HU value in the lung lesion ROIs, respectively.

2.3.4. Analysis of noise properties

We conducted a phantom experiment to investigate the noise properties according to the techniques reported in AAPM Task Group Report 233 [59]. A water phantom was scanned using the same scanner and protocol as the dataset (tube current of 75 mAs, tube voltage of 120 kVp, pitch of 0.8, exposure time of 0.5 s, slice thickness of 0.6 mm, pixel spacing of 0.98 mm, reconstruction field-of-view of 500 mm). One-dimensional Noise Power Spectrum (NPS) and normalized NPS (nNPS, divided by total noise power) were measured in the uniform section of the phantom to comprehensively analyze the image noise magnitude and texture.

The ROIs for the measurement of noise magnitude were placed in five areas (center, 3, 6, 9, and 12 o'clock), shown in Fig. 3b. The diameter of the ROI was approximately 22 pixels (21.5 mm). The noise magnitude was measured as the average standard deviation of pixel values across the ROIs for five images. The ROIs for calculating NPS were at the center of the images shown in Fig. 3c using 50 slices. The squared ROIs were approximately 28 × 28 pixels (~27.3 × 27.3 mm) in size. Mean frequencies of NPS f_{av} were measured to describe the noise texture, as fine texture usually indicates NPS has higher mean and peak frequencies.

2.3.5. Subjective quality assessment

A total of 13 sparse-view test images reconstructed with various algorithms including our proposed methodology were presented to two radiologists blinded to patient IDs or reconstruction algorithms. An in-house image display and reviewing program was designed and the radiologists scored images in terms of image quality, diagnostic confidence, and conspicuity in scale of 1 to 5 (1: unacceptable, 2: poor, 3, average, 4: high, and 5: excellent). The scores were compared between the reviewers and reconstruction algorithms.

The interobserver variability of the two radiologists was assessed using the Spearman correlation coefficient, ρ, and the intra-class correlation coefficient (ICC), with 95 % confidence interval and the *p*-value. ρ is widely used in the image quality assessment (IQA) field [60,61] to describe the monotonic relationship between the ratings of two observers, and can be interpreted as describing anything between no association ($\rho = 0$) to a perfect monotonic relationship ($\rho = -1$ or $\rho = +1$). ICC is also a typically used metric in medical imaging research to assess the consistency of the rating scores of physicians [62,63]. An ICC < 0.40 indicates poor agreement, 0.40–0.59 fair agreement, 0.60–0.74 good agreement, and 0.75–1.0 excellent agreement, following the guidelines [64].

The number of test images was limited and the distribution of scores was not normally distributed. Hence, we used non-parametric tests. To compare the subjective scores of the proposed method with other reconstruction methods, a one-sided paired Wilcoxon signed rank test was performed to analyze the statistical significance of the rating results. *p*-value <0.05 was considered statistically significant in all analyses.

3. Results

3.1. Objective overall quality evaluation

Table 2 reports the means and the standard deviations of MAE, PSNR and SSIM results of different methods. The proposed network achieved an MAE of 29.80±3.86 HUs (i.e. 68 %, 57 %, 42 % and 43 % reduction compared to FBP, SART, SIRT and CGLS, respectively), a PSNR of 32.80 ±1.41 dB (i.e. 35 %, 23 %, 17 % and 16 % improvement compared to FBP, SART, SIRT and CGLS, respectively), and SSIM of 0.968±0.008 (i.e. 14 %, 9 %, 5 % and 5 % improvement compared to FBP, SART, SIRT and CGLS, respectively). Dual-domain-based DL methods LRR-CED(U) and CLRecon show superior performance to DD-Net image-domain-based DL method. The proposed method achieved competitive results compared to LRR-CED(U) and CLRecon.

3.2. Organ-based assessment

Table 3 summarizes the MedianAE results achieved by the various reconstruction methods in different body regions. Among iterative reconstruction methods, SIRT achieved smaller MedianAE than SART and CGLS in almost all assessed body parts. The DL-based methods achieved a smaller MedianAE value than FBP and iterative reconstruction methods. The proposed method achieved smaller MedianAE compared to DD-Net (*p*-value<1e-6), LRRCED (*p*-value<1e-6), and CLRecon (*p*-value<1e-4). Indeed, we observed that all methods have relatively larger MedianAE values in the clavicle, ribcage and vertebrae

Table 2
Comparation of quantitative results achieved by the different methods on the testing dataset with 384 views.

Method	MAE (HU)	PSNR (dB)	SSIM
FBP	94.39 ± 10.73	24.25 ± 1.47	0.846 ± 0.045
SART	69.09 ± 9.85	26.61 ± 1.26	0.883 ± 0.046
SIRT	51.17 ± 6.36	28.13 ± 1.14	0.925 ± 0.017
CGLS	51.92 ± 6.02	28.17 ± 1.06	0.923 ± 0.019
DD-Net	37.74 ± 3.49	31.14 ± 1.17	0.950 ± 0.012
LRR-CED(U)	34.33 ± 4.42	31.46 ± 1.41	0.958 ± 0.011
CLRecon	32.18 ± 4.69	32.23 ± 1.40	0.965 ± 0.008
Ours	29.80 ± 3.86	32.80 ± 1.41	0.968 ± 0.008

Table 3
Organ-based evaluation results (MedianAE, HU) of different methods. Evaluation was performed on five testing cases with 384 views.

Body Part	FBP	SART	SIRT	CGLS	DD-Net	LRR-CED(U)	CLRecon	Ours
WholeHeart	73.4	54.0	29.2	32.2	24.8	21.0	19.6	19.6
Atrium(R)	73.2	54.6	31.2	33.8	24.6	22.4	20.6	19.8
Atrium(L)	79.0	56.6	29.2	34.2	24.4	19.4	18.4	18.4
Ventricle(L)	73.0	52.2	27.0	29.2	22.0	17.2	16.4	16.2
Ventricle(R)	72.2	53.2	28.2	30.4	22.8	20.8	19.6	19.2
Myocardium	74.0	54.8	30.4	33.4	27.4	20.0	19.8	19.6
Lungs	68.2	55.2	38.2	39.6	28.0	29.0	35.2	25.2
Adrenal Glands	82.4	59.4	31.6	33.6	28.2	35.2	29.8	28.8
Spleen	76.6	54.2	29.4	32.0	24.8	20.2	17.4	17.2
Liver	78.4	56.6	28.4	30.8	22.0	19.6	16.4	16.4
Kidneys	78.8	57.0	32.2	34.0	28.0	27.4	24.0	23.2
Pancreas	77.8	55.4	29.0	31.2	25.2	23.6	22.0	21.4
Stomach	78.0	55.8	30.8	32.0	26.2	22.4	21.6	21.0
Esophagus	73.4	54.8	35.2	37.0	28.0	29.4	25.4	24.4
Colon	79.4	60.0	44.2	45.6	36.4	39.6	37.2	34.4
Small Bowel	77.8	57.6	36.4	38.2	30.2	28.4	26.8	25.4
Aorta	74.8	54.2	32.2	35.4	25.4	23.0	20.8	20.0
Autochthon	73.6	52.4	28.2	29.0	26.2	20.2	18.8	18.4
Clavicles	100.7	90.2	112.6	104.8	76.4	71.6	64.2	56.8
RibCage	125.2	111.2	121.6	124.2	78.0	76.0	64.2	56.6
Vertebrae	92.8	73.6	60.0	60.6	53.4	51.6	46.2	43.0

bones of the skeleton system. Organs with relatively smooth (low frequency) regions have relatively smaller MedianAE, e.g., liver and spleen, compared to organs with many regions with a lot of CT intensity variation (high frequency), e.g., lungs, colon and small bowel. We can also observe similar results on the SDE shown in Table 4. To compare the performance of the proposed model in different body regions more intuitively, Fig. 4 illustrates the linear regression plots (reference CT value vs. predicted CT value) and the residual plots of the proposed method in different anatomical regions.

3.3. Qualitative visual assessment

Fig. 5 shows the reconstructed lung images of different methods under 384 views and the corresponding ROIs. Below each reconstruction result, the difference between the reference and reconstructed images obtained by the different methods is displayed (the lighter color represents the smaller error). It can be observed that FBP and iterative reconstruction results suffer from noise and streaking artifacts. DL-based methods can remove the artifacts, preserve organ structure, and produce more smoothed images. Zoomed-in images depict that the proposed method can reconstruct fine structures e.g., muscle structures and the shape of the esophagus, and restore more high-frequency feature details, e.g., bronchial tubes and the edge of the spine. It can be observed in the difference images that the FBP results have the most obvious errors among all methods. Iterative methods improve the image quality with less noise, but the errors of structure edges are still significant. DL-based methods produce slight errors in most soft-tissue regions, e.g., fat and muscles. The proposed method has the least errors in the structure edges among DL-based methods.

Fig. 6 displays the abdomen images reconstructed by different methods and ROIs defined on the liver, spleen, stomach and aorta. We observed that in the zoomed-in liver, spleen and stomach regions, FBP and iterative methods results have noise signals and streak artifacts. DD-Net results are clearer but still have little noise. LRCEED(U) and CLRecon suppress the noise but produce overly-smoothed images. The proposed method improved image quality by providing more details in the organ areas and fewer errors in the structure edges. Similar results can be observed in the aorta region. LRCEED(U) and CLRecon reconstructed vessels are blurrier than those reconstructed using the proposed method.

Fig. 7 displays the abdomen images reconstructed by different methods and ROIs of the kidney, colon and small bowel. In the zoomed-in kidney area, the FBP and iterative methods results have severe

Table 4
Organ-based evaluation results (SDE, HU) of different methods. Evaluation was performed on five testing cases with 384 views.

Body Part	FBP	SART	SIRT	CGLS	DD-Net	LRR-CED(U)	CLRecon	Ours
WholeHeart	109.5	80.2	45.4	48.5	35.4	35.5	34.4	33.6
Atrium(R)	111.9	86.5	58.6	59.9	46.7	47.9	45.3	42.9
Atrium(L)	117.1	83.5	42.3	46.1	32.8	32.6	31.7	31.5
Ventricle(L)	106.8	75.9	35.6	40.9	24.7	23.9	23.7	24.1
Ventricle(R)	106.9	77.3	40.1	43.6	31.8	31.4	30.7	30.5
Myocardium	111.6	85.5	52.2	55.9	43.8	41.9	39.9	41.5
Lungs	112.7	99.1	90.8	89.3	71.3	72.4	66.5	61.6
Adrenal Glands	121.4	87.8	46.5	49.3	42.4	40.7	39.5	37.1
Spleen	113.6	79.6	45.1	48.2	33.4	31.7	30.3	30.0
Liver	118.5	82.5	41.1	43.8	28.7	27.2	26.3	26.1
Kidneys	130.7	84.2	48.9	51.2	41.3	41.2	39.6	36.7
Pancreas	114.7	81.0	42.1	44.5	34.8	34.7	34.4	33.0
Stomach	117.0	84.5	51.4	53.5	41.7	42.5	40.7	38.8
Esophagus	109.6	82.1	69.5	66.0	53.8	56.4	47.2	40.7
Colon	146.6	109.2	99.7	97.6	87.2	92.1	85.7	78.2
Small Bowel	155.8	99.8	83.8	82.8	70.3	71.8	66.5	63.0
Aorta	117.9	85.8	67.4	66.9	47.4	48.5	44.2	40.0
Autochthon	122.8	76.7	39.6	42.2	30.8	30.7	29.8	28.7
Clavicles	172.2	161.5	187.7	178.5	160.0	157.0	142.1	129.0
RibCage	195.8	172.7	171.6	172.4	132.7	131.9	117.5	109.4
Vertebrae	154.1	123.9	119.7	116.4	106.0	111.5	101.8	92.8

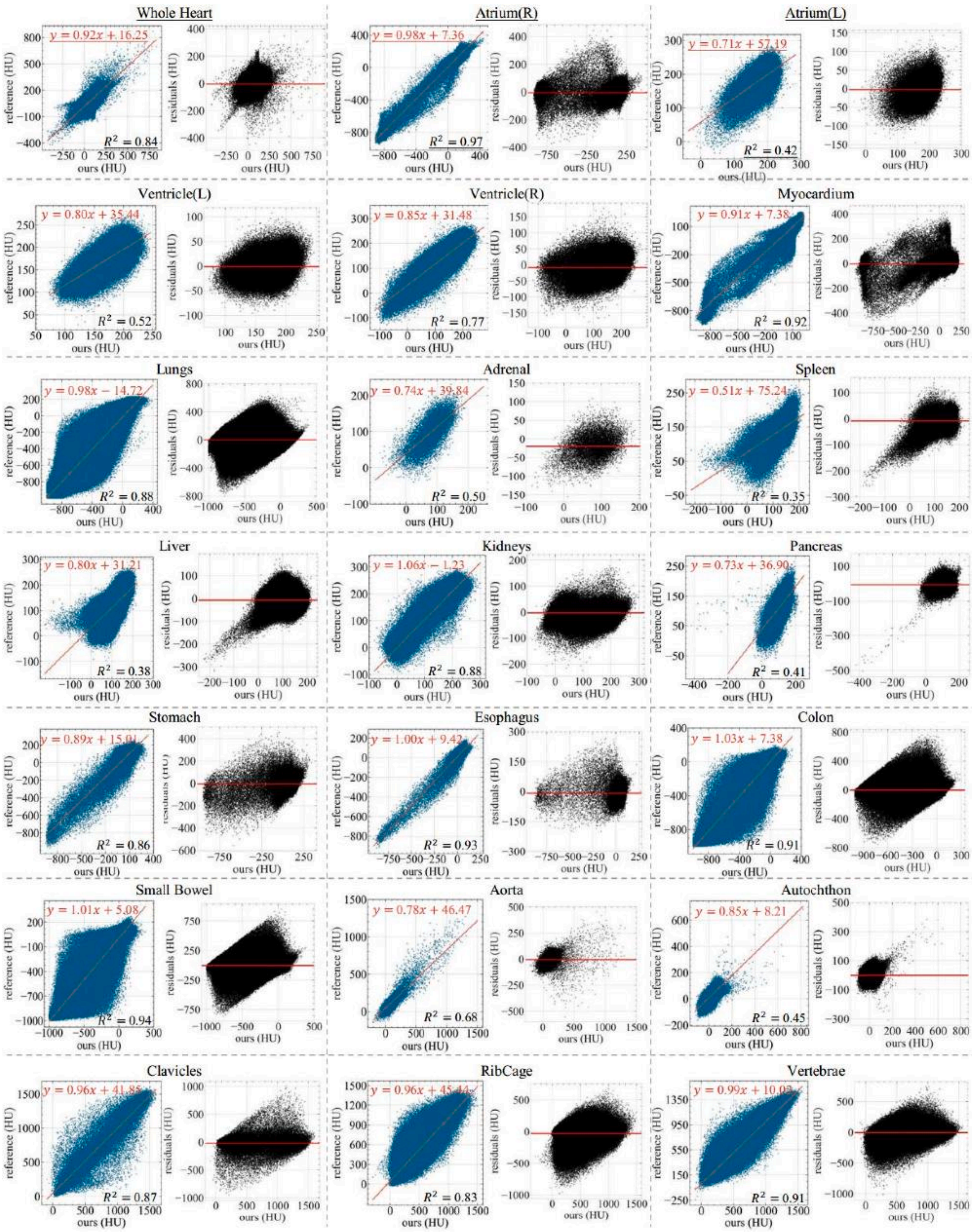

Fig. 4. Linear regression plots (actual reference CT value VS. predict CT value by the proposed method) and residual plots of different anatomical structures of one test case.

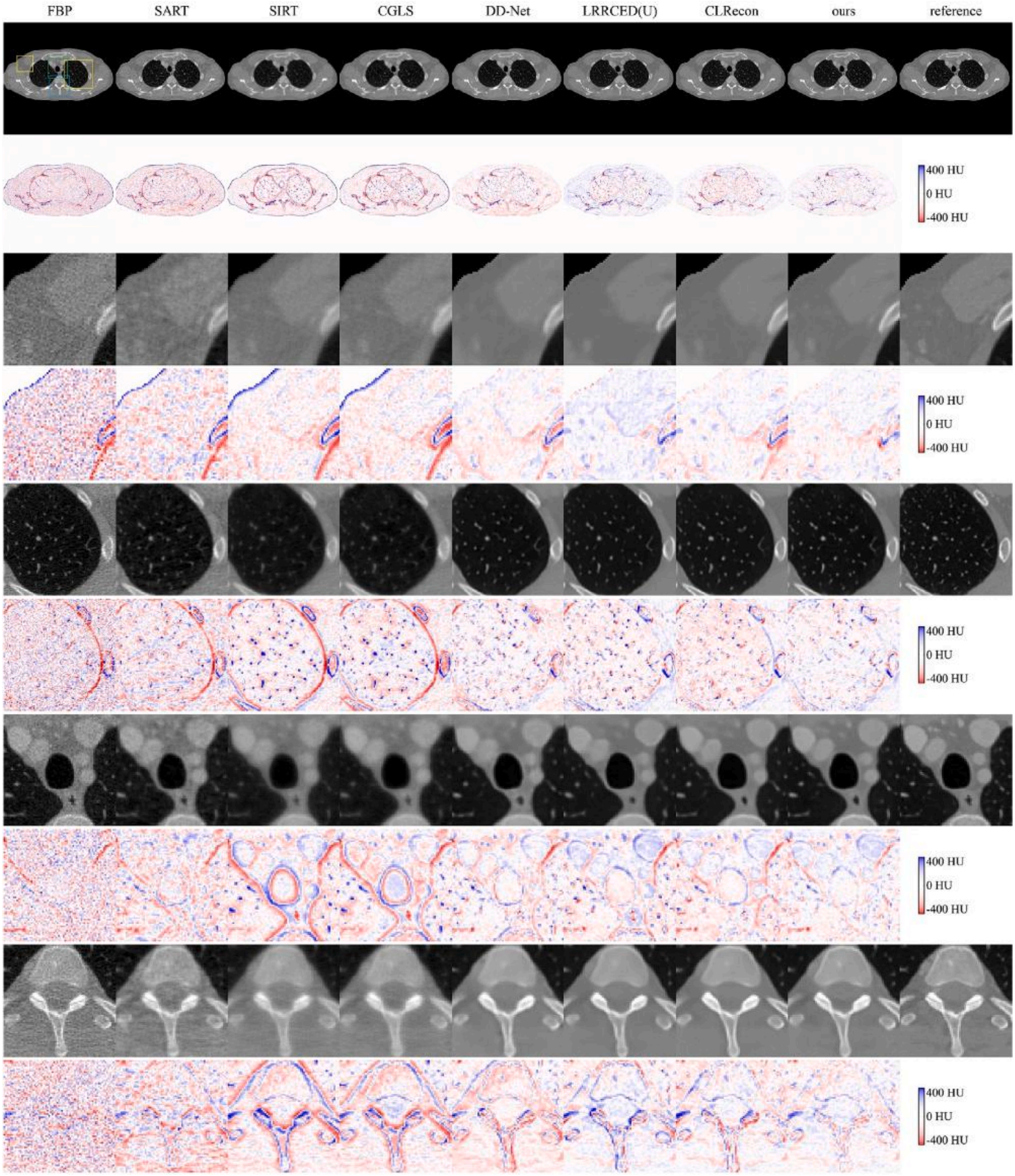

Fig. 5. Visual comparison of sparse-view reconstructions of different methods (display window [−1000, 1000] HUs). The images in the third, fifth, seventh and ninth rows represent zoom-in regions (yellow, blue and green boxes). Below each reconstruction result, the difference images obtained by subtracting the reference image from the reconstructed image (display window [−400 400] HUs).

degradation. DL-based methods significantly reduce the artifacts, recover organ shape and enhance the contrast. The proposed method gets closer to the reference image regarding the contours and internal details of the kidney. In the zoomed-in colon and small bowel region, the FBP reconstructed results contain severe streak artifacts and loss of information regarding organ structures with a very low signal-to-noise ratio. Iterative methods suppress the artifacts slightly whereas DL-based methods further eliminate streaking artifacts and reproduce well organ structures. However, it can be observed that although the proposed method can recover most organ shapes, reconstructing rich features and details in the colon and small bowel area is still challenging.

Fig. 8 compares the coronal planes of one testing sample reconstructed using different methods with lung window [−1000 200] HUs and abdomen window [−160 240] HUs, respectively. We observe that

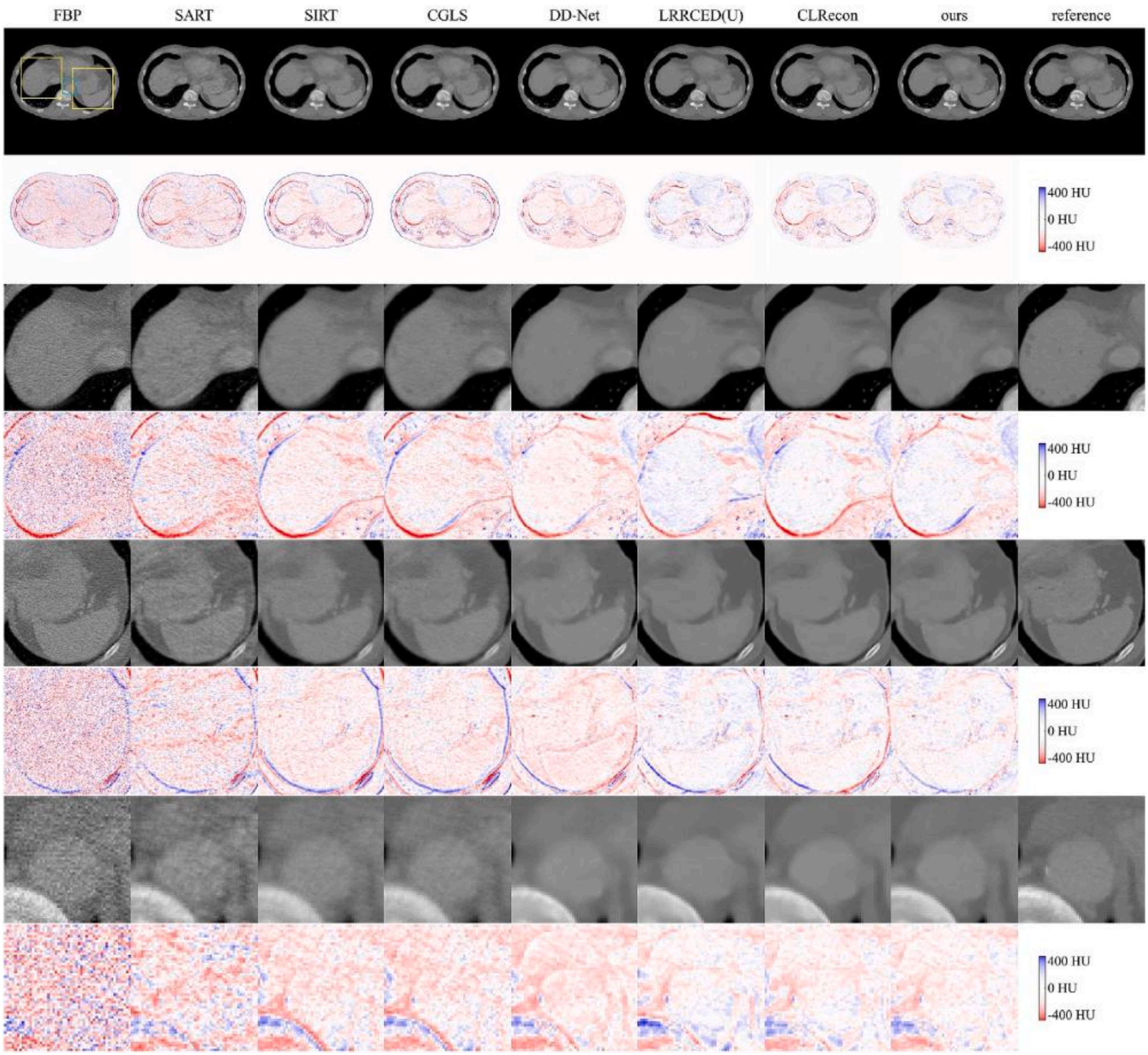

Fig. 6. Visual comparison of sparse-view reconstructions of different methods (display window [−800, 1000] HUs). The images in the third, fifth, seventh and ninth rows represent zoom-in regions (yellow, blue and green boxes). Below each reconstruction result, the difference images obtained by subtracting the reference image from the reconstructed image (display window [−400 400] HUs).

the proposed method achieves better results in the recovery of structure details of organs compared to other methods, e.g., the artery in the lung region and the veins in the liver. The difference images show that the proposed method provides the lowest reconstruction error compared to other methods, especially in the edges of the rib cage and contours of organs.

Fig. 9 compares the sagittal planes of one testing sample reconstructed using different methods. The difference images show that the proposed method outperforms other methods in reducing the noise and restoring the edges of the spine. The visual performance of the proposed method is further demonstrated in the 1D intensity profiles shown in Fig. 10, where the results achieved by the proposed method are closer to the reference compared to other methods.

3.4. Lung and liver lesions analysis

In the quantitative analysis results shown in Table 5, the standard deviations (SDs) of CT HUs measured in the liver on DL-reconstructed images are lower than those of FBP, iterative methods and reference images. In addition, DL-reconstructed images show higher SNR_{liver} than images reconstructed from iterative methods. These results indicate the high ability of DL-based methods to suppress the noise caused by sparse-view projections. In addition, the $CNRs_{lesion(liver)}$ of DL reconstructed images are higher than those of iterative methods, and the $CNR_{lesion(liver)}$ of the images reconstructed by the proposed method (3.85±0.88) is close to $CNR_{lesion(liver)}$ of reference images (4.04±1.50). Besides, the $CNR_{lesion(lung)}$ of DL reconstructed images are also higher than those of iterative methods. LRR-CED(U) achieved the highest $CNR_{lesion(lung)}$ among DL-based reconstruction methods (24.58± 6.39), but the value is still lower than the reference images (26.27±2.76).

3.5. Noise properties analysis

Fig. 11a shows the noise magnitude of different methods. The average noise magnitude is generally lower for DL-based methods than FBP and iterative methods. It can also be observed in the noise texture images shown in Fig. 11b that artifacts and noise are still retained in the FBP and iterative reconstructed images, while images reconstructed

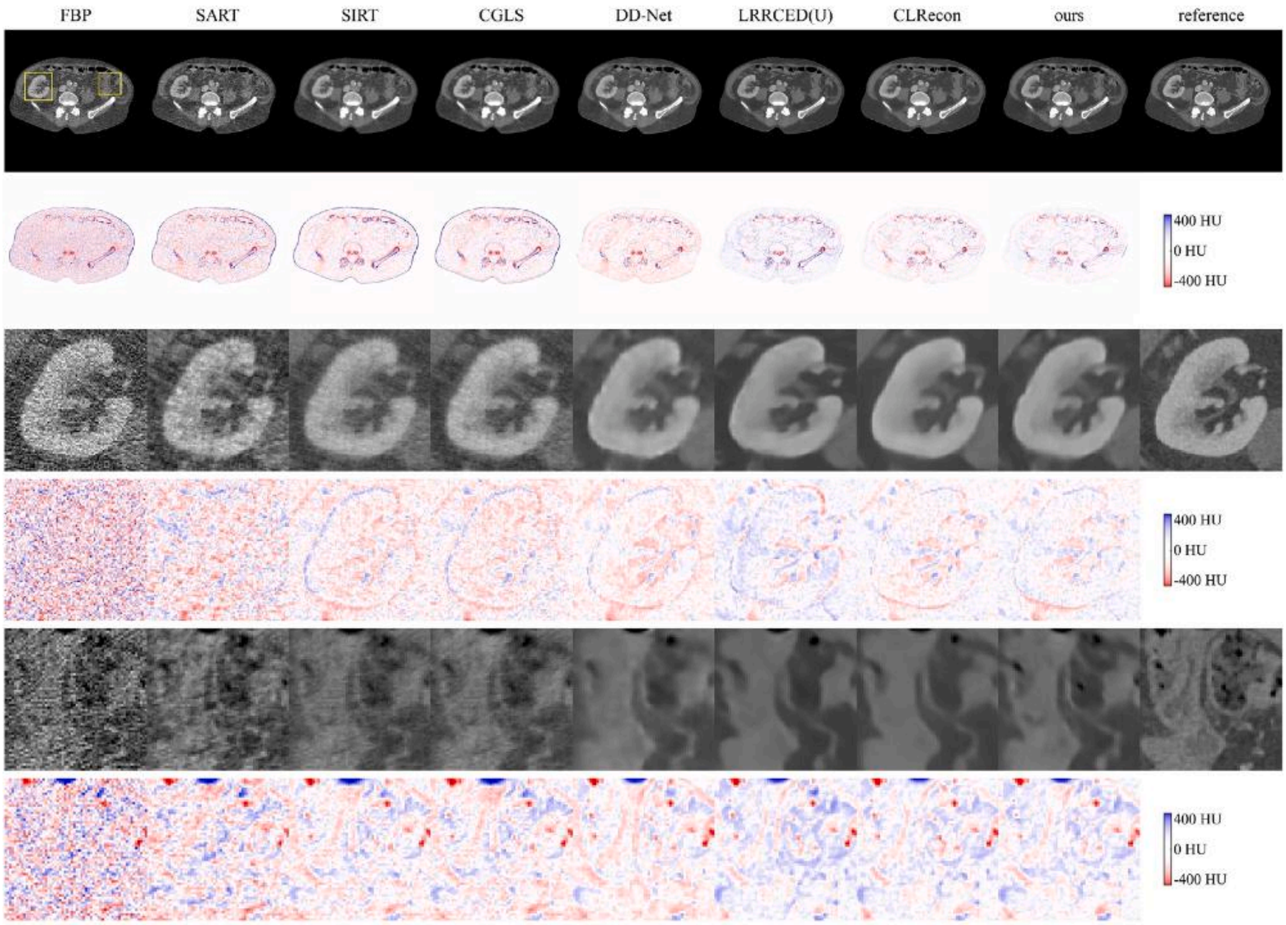

Fig. 7. Visual comparison of sparse-view reconstructions of different methods (display window [−300, 500] HUs). The images in the third, fifth, seventh and ninth rows represent zoom-in regions (yellow, blue and green boxes). Below each reconstruction result, the difference images obtained by subtracting the reference image from the reconstructed image (display window [−400 400] HUs).

using DL-based methods are smoother. Fig. 11c and d show the NPS and nNPS curves, respectively. Overall, the mean frequencies of NPS f_{av} of the DL methods shift to the low-frequency domain. This also indicates that the DL methods suppress the noise and generate smoother images as the low-frequency noise occupies a larger fraction than the high-frequency noise. The reference images have higher f_{av} compared to images of DL-based methods. As mentioned earlier, fine texture usually indicates that NPS has higher f_{av}. The proposed method has 11 %, 25 % and 25 % higher f_{av} than DD-Net, LRR-CED(U) and CLRecon, respectively.

3.6. Subjective quality assessment

The radiologists achieved fair and good agreements in the rating of the confidence, conspicuity, and overall image quality with $0.70<\rho<0.73$, $0.56<\text{ICC}<0.72$ and p-value<0.001 (Table 6.).

Table 7 shows the subjective scores of different methods (represented as mean ± SD), whereas Fig. 12 shows the distribution of the scores. Compared to FBP, SART has lower mean scores for diagnostic features; on the other hand, SIRT and DD-Net improved the confidence and conspicuity of the anatomic structures as well as the overall image quality. However, the mean scores of the overall image quality of FBP, SIRT, SART, and DD-Net are <3, which fails to fully meet the diagnostic requirement. The mean scores of CGLS, LRR-CED(U), CLRecon, and the proposed method in all aspects are greater than 3, which shows better visual performance. The proposed method achieved the highest mean scores in all aspects of the subjective evaluation. Significant improvement of subjective image quality scores is found in the proposed method compared to FBP, SART, SIRT and DD-Net (all p-value<0.001) (Table 8) The subjective scores of the proposed method are not significantly greater than CGLS (p-value>0.05). For the overall image quality scores, the proposed method is not significantly better than LRR-CED(U) and CLRecon (p-value>0.05).

3.7. Computational cost and memory requirements

We assessed the convergence of SIRT, SART and CGLS iterative reconstruction methods. We plotted the running time vs. iteration number and the MAE vs. iteration number curves for iterative reconstruction methods. The running time and the MAE metric were calculated by averaging the reconstruction time and the MAE value on 300 test images. As mentioned earlier, all methods were tested on the same PC with NVIDIA 2080Ti GPU. Fig. 13a shows the running time vs. iteration number curves of iterative methods. The reconstruction time of SIRT, SART and CGLS increases linearly with the number of iterations. CGLS runs slightly longer than SIRT per iteration, while SART has the shortest computation time per iteration. Fig. 13b shows the MAE vs. iteration number curves. The MAE of SIRT and SART converge after 200 and 300 iterations, respectively. The MAE value of CGLS increases after 18 iterations. CGLS and SIRT achieve smaller MAE than SART after convergence.

Table 9 summarizes the computational complexity of all reconstruction methods. Fig. 13c shows the scatter plot of MAE vs. running time of the different methods. It can be observed that CGLS is the most efficient among iterative reconstruction algorithms, achieving a similar MAE to SIRT but roughly nine times faster. Among the DL-based

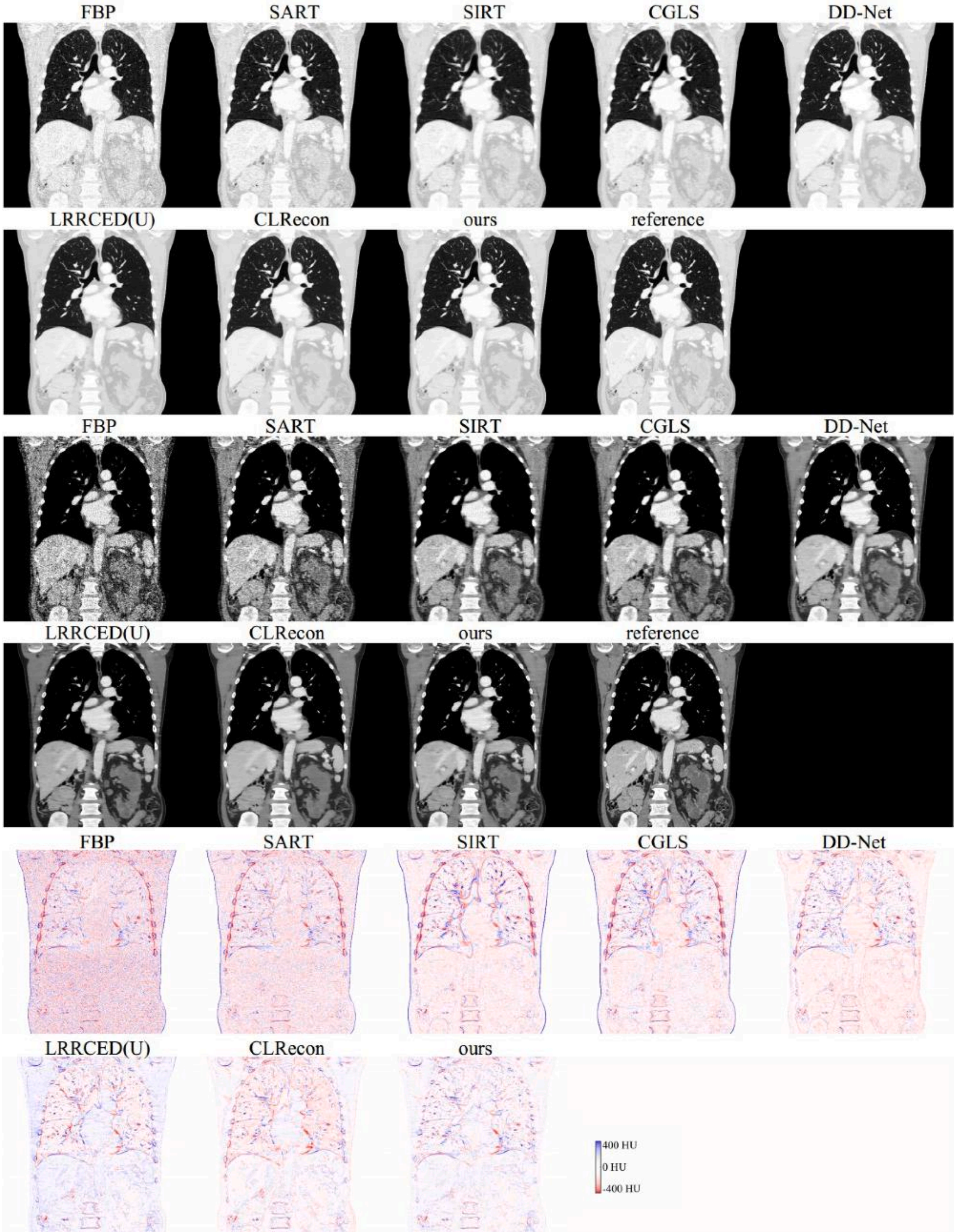

Fig. 8. Visual comparison of sparse-view reconstructions of different methods (coronal plane). The images in the first and second row are displayed with lung widow [−1000 200] HUs. The images in the third and fourth row are displayed with abdomen window [−160 240] HUs. Below the reconstruction results, the difference images obtained by subtracting the reference image from the reconstructed image (display window [−400 400] HUs).

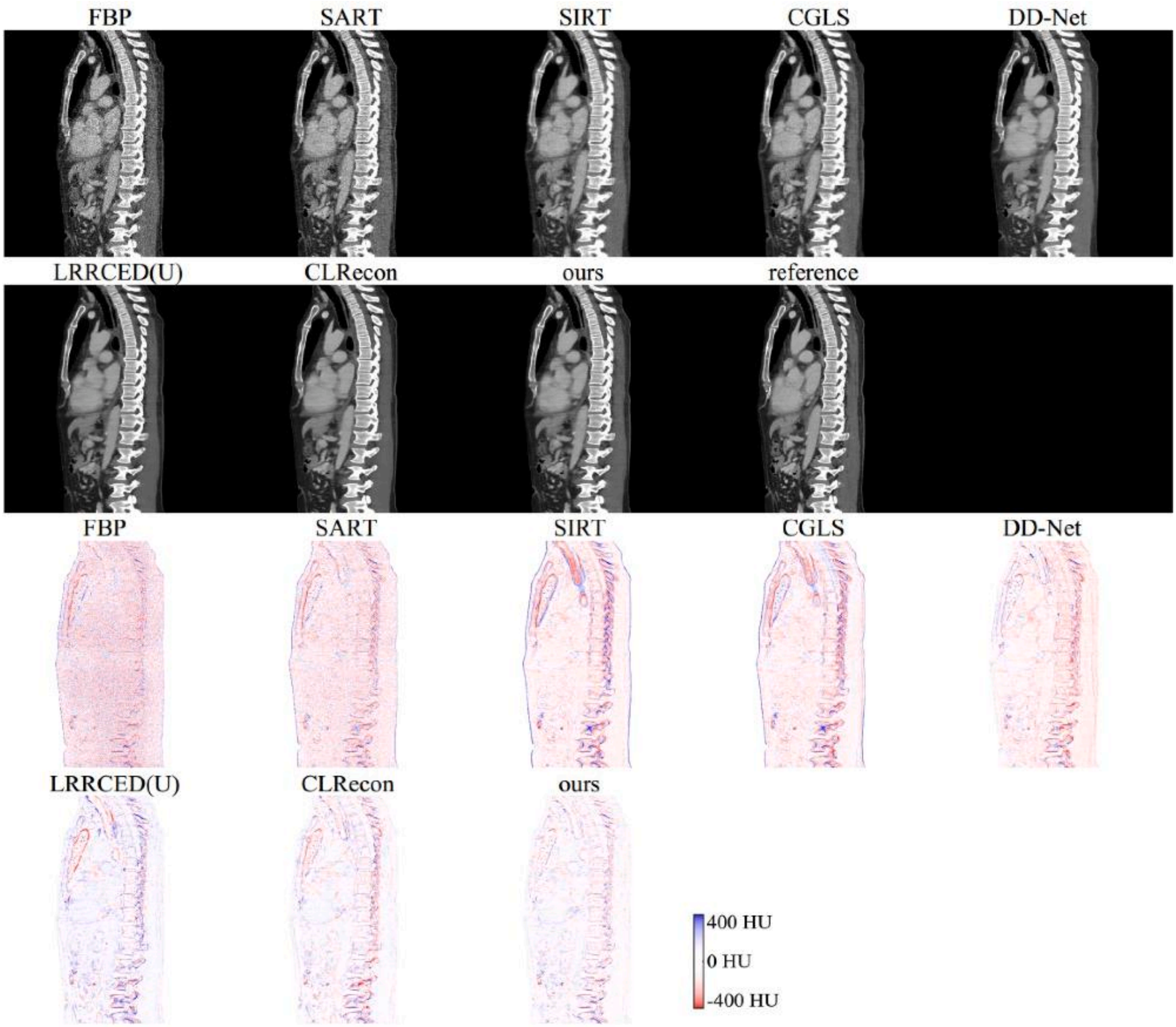

Fig. 9. Visual comparison of sparse-view reconstructions obtained using the different methods (sagittal plane, display window [−160 400] HUs). Below the reconstruction results, the difference images obtained by subtracting the reference image from the reconstructed image (display window [−400 400] HUs).

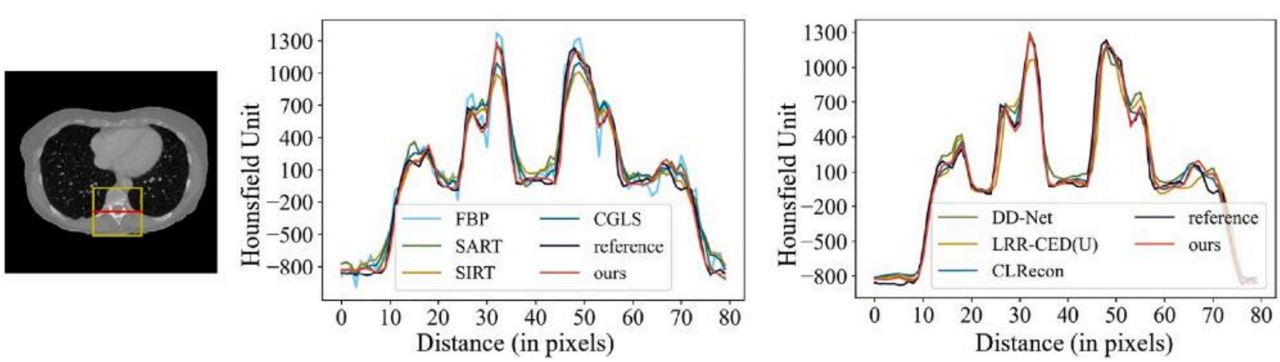

Fig. 10. Comparison of horizontal intensity plot profiles for the region marked in red on the CT image (left).

methods, DD-Net and LRR-CED(U) are the fastest, with similar reconstruction times to CGLS. The proposed method has a comparable reconstruction time to CLRecon. Fig. 13d shows the scatter plot of MAE vs. parameter numbers of DL-based methods. We observe that the proposed method achieved the lowest MAE, and the model size of the proposed method is roughly 11 times and 96 times smaller than LRR-CED(U) and CLRecon, respectively.

Table 5
The SD of CT HUs within the liver ROIs, SNR of liver and the CNR of the lesion to the liver, and the CNR of the lesion to the lung.

Method	SD_{liver}	SNR_{liver}	CNR_{lesion} (liver)	CNR_{lesion} (lung)
FBP	111.37± 19.24	1.36± 0.33	0.77± 0.29	8.65± 1.21
SART	67.90± 17.79	2.32± 0.84	1.22± 0.37	10.53± 1.62
SIRT	33.74± 8.43	4.62± 1.43	2.06± 0.76	15.20± 2.73
CGLS	34.16± 8.19	4.38± 1.14	2.00± 0.68	15.16± 2.70
DD-Net	13.98± 4.35	11.67± 4.66	3.70± 1.56	20.01± 4.91
LRR-CED (U)	9.56± 2.10	12.91± 4.56	3.49± 0.30	24.58± 6.39
CLRecon	8.24± 2.98	19.96± 13.29	3.40± 1.17	21.39± 4.88
Ours	11.28± 4.78	13.31± 6.01	3.85± 0.88	21.50± 3.58
Reference	21.85± 4.97	6.32± 1.80	4.04± 1.50	26.27± 2.76

3.8. Model and performance trade-offs

In the proposed network, a key module is the efficient RLFBs. Thus, we investigated how the number of RLFBs used in Sino-Net and Img-Net would influence the reconstruction performance. Fig. 14a plots the average MAE of the validation dataset when the trained model uses 5, 6 and 7 RLFBs in Sino-Net. We observed that increasing the number of RLFBs from 5 to 6 slightly decreases the average MAE value, while further increasing the number of RLFBs from 6 to 7 shows little influence on the MAE metric. Considering the computational efficiency, the number of RLFBs in Sino-Net is set to 6. Fig. 14b plots the average MAE of the validation dataset when the trained model uses 1, 2 and 3 RLFBs in Img-Net. Similarly, considering the trade-off between reconstruction performance and computational complexity, the number of RLFBs in Img-Net is set to 2.

3.9. Ablation study

The proposed network used in this study is based on an image-domain reconstruction network (DD-Net), and combines it with a sinogram-domain refinement network (Sino-Net) and a lightweight image-domain network (Img-Net). To validate the effectiveness of the design of dual-domain learning and the role of these three modules, we trained several model variants and compared them against the proposed method. The quantitative results (MAE, PSNR and SSIM) of the testing dataset, reconstruction time per image and the number of parameters of different model variants are summarized in Table 10. We use the Sino-Net as the baseline, the combination of DD-Net and Sino-Net can improve the reconstruction performance, but the number of training parameters is doubled. By comparing the results of 'DD-Net+SinoNet' and DD-Net+SinoNet+ImgNet', we observe that using the lightweight Img-Net can improve the reconstruction performance slightly without compromising the reconstruction time.

Table 6
Inter-observer variability in the subjective quality assessment of CT images.

	ρ	ICC	Confidence interval		p-value
Confidence	0.72	0.56	0.08	0.78	<0.001
Conspicuity	0.71	0.60	0.19	0.79	
Image quality	0.70	0.71	0.59	0.79	

Table 7
Subjective evaluation of different reconstruction algorithms represented as mean ± SD.

	FBP	SART	SIRT	CGLS	DD-Net
Confidence	2.15 ± 1.03	1.85± 0.72	3.00± 0.78	3.77± 0.58	3.04± 0.59
Conspicuity	2.04± 0.94	1.73± 0.59	2.92± 0.73	3.73± 0.65	2.96± 0.65
Image quality	1.77± 0.75	1.46± 0.57	2.65± 0.73	3.15± 0.72	2.73± 0.59
	LRR-CED (U)	CLRecon	Ours	Reference	
Confidence	3.42± 0.69	3.62± 0.49	3.81± 0.62	5.00± 0.00	
Conspicuity	3.42± 0.69	3.50± 0.50	3.81± 0.68	5.00± 0.00	
Image quality	3.19± 0.68	3.38± 0.68	3.38± 0.56	4.92± 0.27	

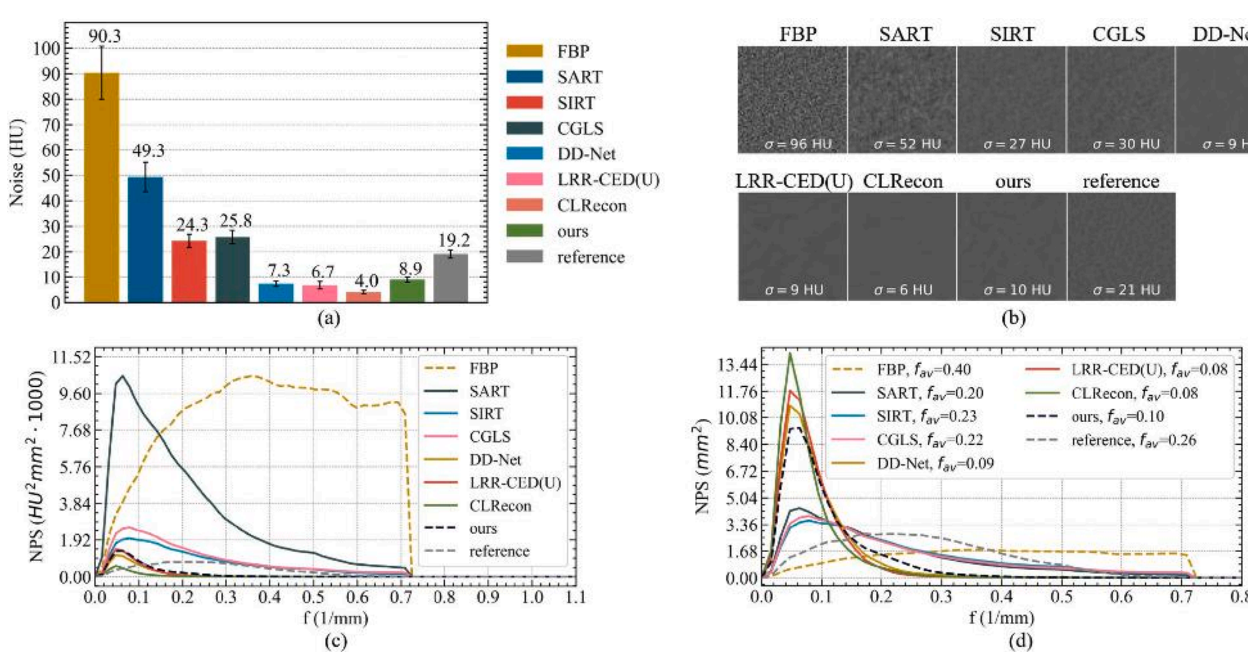

Fig. 11. Noise properties analysis. (a) Comparison of noise magnitude. (b) Comparison of noise texture. (c) Comparison of one-dimensional NPS. (d) Comparison of normalized NPS (nNPS, divided by total noise power).

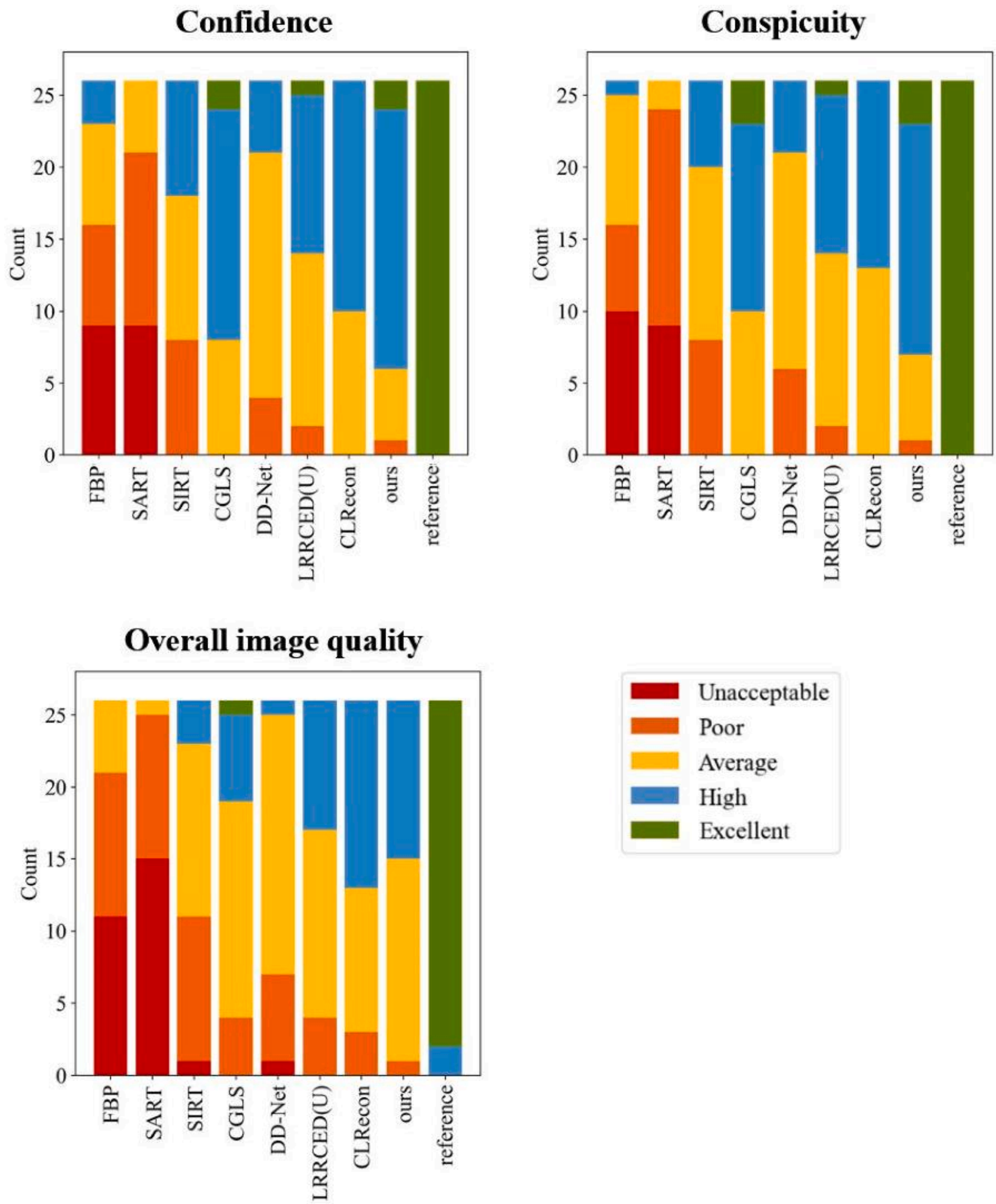

Fig. 12. Subjective evaluation results achieved by the different reconstruction algorithms.

Table 8
P-value results of statistical significance testing in subjective evaluation. # means no significant statistical differences.

	Confidence	Conspicuity	Image quality
Ours vs. FBP	<0.001	<0.001	<0.001
Ours vs. SART	<0.001	<0.001	<0.001
Ours vs. SIRT	<0.001	<0.001	<0.001
Ours vs. CGLS	0.391#	0.319#	0.083#
Ours vs. DD-Net	<0.001	<0.001	<0.001
Ours vs. LRR-CED(U)	0.006	0.006	0.113#
Ours vs. CLRecon	0.083#	0.026	0.500#

4. Discussion

In this study, we proposed a dual-domain deep learning-based reconstructed method for sparse-view CT reconstruction. The method split the CT reconstruction problem into three stages: Firstly, DD-Net [15] was used to recover initial refined images from highly degraded FBP images reconstructed from sparse-view sinograms. Secondly, we forward projected the refined images to dense-view sinograms and proposed a lightweight network, Sino-Net, to enhance sinogram signals. Finally, we reconstructed the enhanced sinograms to enhanced images, concatenated them with the initial refined images, and then used a small Img-Net to further preserve the structural details of image content.

In the past few years, various dual-domain methods and DL-based

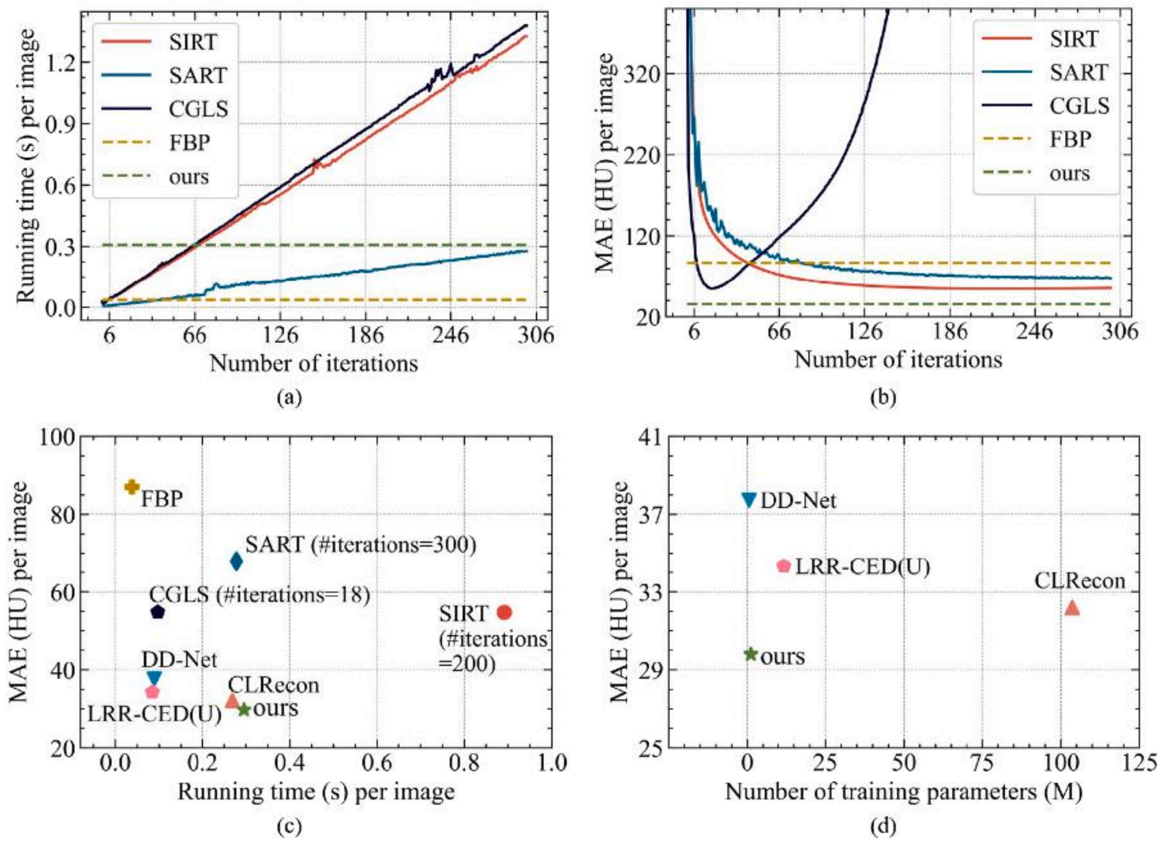

Fig. 13. Comparison of computational cost of different methods. (a) Running time vs. iteration number curves of iterative methods. (b) MAE vs. iteration number curves of iterative methods. (c) Scatter plot of MAE vs. running time of different methods. (d) Scatter plot of MAE vs. storage size of DL-based methods.

Table 9
Quantitative comparison of computational efficiency of different methods.

Method	Run time (s)/slice	Params (M)
FBP	0.0378	–
SART	0.2779	–
SIRT	0.8915	–
CGLS	0.0965	–
DD-Net	0.0887	0.56
LRR-CED(U)	0.0842	11.78
CLRecon	0.2684	103.65
Ours	0.2951	1.08

iterative reconstruction methods have been developed for sparse-view CT reconstruction to make full use of the information in both domains, e.g. DRONE [16], CDCNN [17], CLRecon [18], HDNet [19], DualCNN [20], SWISTA-Nets [65], DEAR [23], IRON [24] and MOGM [28]. However, the main shortcomings of these joint learning networks are the large number of trainable parameters and the requirement of large hardware resources. They are also time-consuming which may restrict their applications on real CT projection data with large projection sizes.

In this work, we made several efforts to reduce the computational complexity of dual-domain learning. Firstly, in image-domain processing, we choose to use DD-Net, which is sufficiently lightweight with only 0.56 M training parameters and has a fast-running speed. Secondly, in

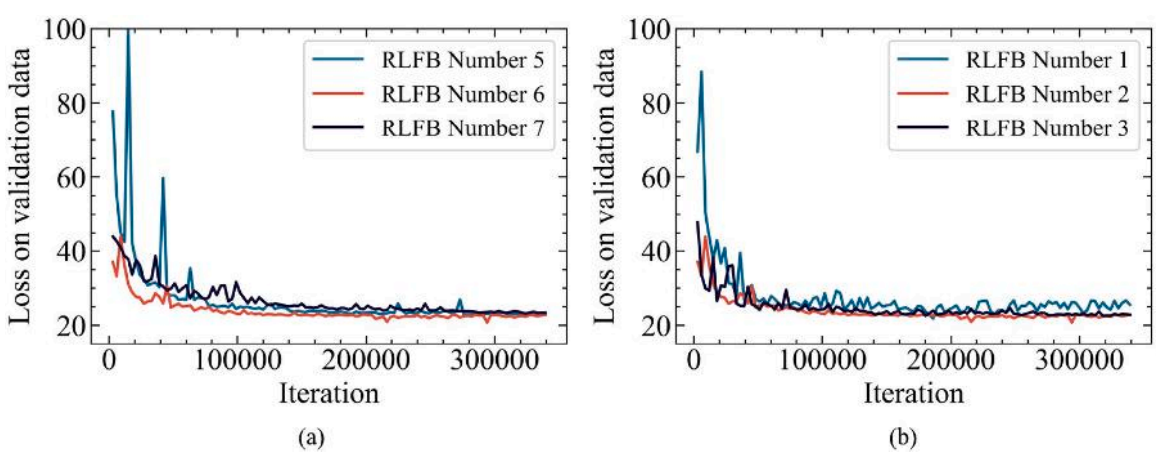

Fig. 14. MAE plots on the validation dataset for the proposed method with different network parameters. (a) RLFB number in Sino-Net. (b) RLFB number in Img-Net.

Table 10
Quantitative performance and computational cost of different ablation networks.

DD-Net	Sino-Net	Img-Net	Params (M)	Run time (s)/slice	MAE	PSNR	SSIM
	√		0.45	0.2758	34.79 ± 4.97	31.80 ± 1.39	0.960 ± 0.013
√	√		1.01	0.2865	30.77 ± 3.95	32.50 ± 1.36	0.967 ± 0.009
√	√	√	1.08	0.2951	29.80 ± 3.86	32.80 ± 1.41	0.968 ± 0.008

the sinogram-domain recovery stage, we projected the refined images to dense-view sinograms (512 views) instead of full-views (2304 views), which benefits the required GPU memory and decreases the running time. Thirdly, we designed the network structure following the lightweight principle. We used RLFBs, which have been previously demonstrated in the efficient image super-resolution task by its great superiority in inference speed and model capability. Fourthly, we performed global optimization of the networks in the three stages with an end-to-end training fashion using differential forward projection and FBP layers, which highly reduced the training complexity. As shown in Table 9, the proposed method has 1.08 M network parameters and requires 0.3 s to generate a single image with 512 × 512 pixels from a sparse sinogram with 736 detector elements at 384 views on a GPU.

Recently, DL-based models that contain iterative optimization, such as SWISTA-Nets [65], DEAR [23], and IRON [24] have demonstrated outstanding performance, particularly in low-sparse-view CTs (with fewer than 120 views). Due to variations in scanning geometry, implementation details, and the calculation formulas used to assess image-derived metrics in different research studies, it's not fair to compare the absolute quantitative results such as PSNR. However, the common characteristic of these methods is their higher computational costs for reconstructing 512×512 images due to the iterative optimization they involve. For instance, SWISTA-Nets was trained using an RTX 2080Ti GPU to reconstruct CT images from 60 views and from 120 views. The computation time for SWISTA-Nets was approximately 2.7 and 5.4 s per slice for 60 and 120 views, respectively. DEAR was implemented on an NVIDIA TITAN XP GPU for reconstructing CT images from projections sized 60×1120, taking 167.90 s to reconstruct one case. IRON, accelerated by an NVIDIA TITAN RTX GPU, reconstructed sparse-view projections sized 114×1120 views with an inference time of around 256 s for 391 test images. In comparison to these methods, the proposed method significantly reduces the inference time. As a matter of fact, a recently proposed dual-domain reconstruction method, CLRecon [18] (FBPConvNet [13]), which performs end-to-end closed-loop learning, can also achieve comparable reconstruction time. However, our model reduces the model size significantly. The evaluation results show that CLRecon tends to reconstruct over-smoothing edges compared to the proposed method (Figs. 5-7). The median absolute error obtained using our method is smaller in different organ regions, and the noise texture of the images is closer to the reference image. Another efficient DL-based reconstruction method, LRR-CED(U) [22], which directly reconstructs the image from the sparse-view sinogram, has a competitive inference speed. Compared to LRR-CED(U), the proposed method further reduces the number of training parameters and achieved better reconstruction performance by reducing noise and artifacts and reconstructing anatomic structure details.

The key to network design is how to integrate a lightweight strategy within the reconstruction framework. In this study, we utilized lightweight models suitable for dual-domain-based end-to-end reconstruction networks. This is a customized architecture designed for a practical CT reconstruction task on real projection data. We provide new ideas on how to effectively apply the dual-domain reconstruction algorithm to actual projection data and the ablation studies demonstrated the effectiveness of the design of dual-domain learning and the role of these three models. In fact, we believe that other advanced lightweight strategies, such as neural architecture search (NAS) [66], network pruning [67], low-rank decomposition [68], low-bit quantization [69], and knowledge distillation [70], are worth exploring for application in sparse-view CT reconstruction tasks. Investigating the combination of the lightweight strategy with different categories of reconstruction frameworks (e.g. direct reconstruction and deep-unrolling methods) to enhance the efficiency of the reconstruction methods using real projection data would be an interesting research direction.

One of the motivations behind this study is to demonstrate the feasibility of the proposed method on real clinical projection data from commercial CT scanners. As mentioned in the flowchart illustrated in Fig. 1, we collected clinical raw projection data of thoraco-abdominal diagnostic CT scans acquired on a SOMATOM Edge CT scanner with 128 rows (Siemens Healthineers, Erlangen, Germany), analyzed the helical imaging geometry, and refined the projection data to obtain sinograms. Moreover, we utilized a propriety platform (ReconCT) to produce reference images using ADMIRE reconstruction algorithm.

The commonly used quantitative evaluation method for CT reconstruction performance compares the average of MAE, PSNR or other quantitative metrics on reconstructed 2D slices, which are globally averaged results that neglect the image content and details. To tackle this problem, we conducted more comprehensive and detailed evaluation experiments. Firstly, we performed an organ-specific assessment on the reconstructed thoraco-abdominal CT scans. Organ-based assessment has been widely used in evaluating the quality of CT images. For example, the European Guidelines on Quality Criteria for CT classify anatomical image criteria into six categories: cranium, face and neck, spine, chest, abdomen and pelvis, bones and joints [71]. Additionally, measuring the CT value of specific organs, such as the aorta, liver, pancreas, and spleen, has become a widely used objective method for quantitatively comparing the quality of CT images acquired using different protocols and analyzing the impact of dose-related parameters and reconstruction protocols on image quality [72,73]. Moreover, an automated technique has been proposed for measuring Hounsfield units in specific organs (lung tissue, liver, aorta, and spine) in clinical chest CT images to monitor the quality of clinical images [74]. In this study, an intuitive sense is that the difficulty of reconstructing the organs is different due to the varying sizes, shapes, and textures of different organs. Organ-based assessment allows for a quantitative evaluation of the effectiveness of different reconstruction methods on various organs. From a clinical perspective, this assessment provides crucial information, indicating which organs present a higher degree of uncertainty in reconstructive outcomes. This knowledge can guide clinicians to exercise greater caution when making diagnoses in these areas, thereby enhancing the accuracy of their clinical decisions. Specifically, we segmented the reconstructed thoraco-abdominal CT images into 21 anatomical structures: whole heart, left/right atrium, left/right ventricle, myocardium, lungs, adrenal, spleen, liver, kidneys, pancreas, stomach, esophagus, colon, small bowel, aorta, autochthonous dorsal musculature (autochthon), clavicles, ribcage and vertebrae, and performed quantitative analysis of CT HU values in the different organs regions. Secondly, we not only compared the transverse section of the reconstructed CT images but also displayed the coronal and sagittal planes to fully compare the reconstruction results of different methods. Thirdly, we measured the SNR in the liver, the CNR of the lesion to the liver and the CNR of the lesion to the lung to quantitively evaluate overall image quality. Finally, we conducted a phantom study to analyze the noise magnitude and texture of the reconstructed images.

The comprehensive evaluation shows that the proposed method can provide high-quality CT images with clear structure edges and fine anatomic details under 384 sparse views (1/6 sampling rate). The proposed method achieves high SNR in the liver, indicating a high ability to suppress noise and artifacts. The images reconstructed by the proposed method achieve a high CNR (lesion to the liver) close to the CNR of reference images. By observing the difference images, the proposed method achieves the lowest total error compared to other methods and shows better detail preservation abilities, such as reconstructing the artery in the lung region and the veins in the liver. However, reconstructing the anatomic structures in abdominal CT images is still challenging for DL-based methods. Compared to reference images, DL-reconstructed methods still tend to over-smooth the images, as shown in Fig. 7. More efforts should be made to effectively restore the details in the organ area with rich structural information and high-frequency features, e.g., in the colon and small bowel region.

In this study, two radiologists conducted subjective scoring of CT scans reconstructed using different methods using a 1–5 scoring system. The subjective evaluation included the confidence and conspicuity of anatomic structures, and overall image quality. The results show that the proposed method achieved the highest scores in confidence (3.81$\pm$0.62) and conspicuity (3.81$\pm$0.68) of anatomic structures. The overall image quality scores were similar among images reconstructed using CGLS (iterative reconstruction), LRR-CED(U), CLRecon, and the proposed method (3.15$\pm$0.72, 3.19$\pm$0.68, 3.38$\pm$0.68 and 3.38$\pm$0.56 for CGLS, LRR-CED(U), CLRecon, and the proposed method, respectively). In the future, more subjective image quality indicators can be included to fully evaluate the ability of reconstruction methods to meet diagnostic requirements.

The current work inherently bears a number of limitations. Similar to previous studies based on supervised learning, additional models need to be trained for different scanning protocols, which increases the required training time and computational resources. One solution would be to apply transfer learning to reduce the re-training time. Another possible solution is to involve scanning protocols as input of the reconstruction network as prior knowledge. Another limitation is that we have studied only clinical studies obtained on a single scanner in one center. In our future studies, we will investigate the proposed method on more datasets acquired using different protocols on different scanners of multiple institutions. More sparse sampling strategies will be also investigated. Finally, the proposed method works well with high-contrast organs but has limitations when reconstructing small organs or adjacent organs with similar textures. To address this, one possible strategy could be to incorporate prior knowledge of organs and explore organ-specific reconstruction methods. This could help improving the subjective quality of the reconstructed images, making it closer to the reference image for accurate clinical diagnosis. Here, we introduce some ideas to achieve this objective. For instance, due to the diverse shapes and textures of organs, we can utilize different spatial feature descriptors like spatially variant gradients [75], learnable texture priors [76] and learnable filters/transformers [77] as image priors to effectively capture texture features for CT reconstruction tasks. An organ-segmentation mask/context map [78,79] representing the pixel-level categories in an image based on their organ/content differences may also be useful for applying the best reconstruction/enhancement action to different organs/pixels. Additionally, the degradation mask [80], representing the difference between the input noisy image and the reference image, is helpful for guiding the network to make additional efforts in regions with significant degradations.

5. Conclusion

The proposed dual-domain deep learning-based framework provides a new paradigm for efficient sparse-view CT reconstruction. Compared with other dual-domain methods that also perform end-to-end training, the proposed network reduces the training parameters with competitive running time, which decreases the requirement for training resources. The organ-based evaluation of the reconstruction performance on real clinical projection data demonstrates that the proposed method is competitive in reducing artifacts and preserving anatomic details.

Data and code availability

Trained models and code will be made available on GitHub.

CRediT authorship contribution statement

Chang Sun: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Yazdan Salimi:** Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Neroladaki Angeliki:** Writing – review & editing, Validation, Investigation, Data curation, Conceptualization. **Sana Boudabbous:** Writing – review & editing, Validation, Methodology, Data curation. **Habib Zaidi:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This work was supported by the Euratom Research and Training Programme 2019−2020 under grant agreement No. 945196. Chang Sun acknowledges the support from the BUPT Excellent Ph.D. Students Foundation (CX2022203).

References

[1] C.H. McCollough, P.S. Rajiah, Milestones in CT: past, present, and future, Radiology 309 (2023) e230803.
[2] Radiation Risk from Medical Imaging (2021).
[3] A.B. De González, M. Mahesh, K.-P. Kim, M. Bhargavan, R. Lewis, F. Mettler, C. Land, Projected cancer risks from computed tomographic scans performed in the United States in 2007, Arch. Intern. Med. 169 (2009) 2071–2077.
[4] T.L. Slovis, The ALARA concept in pediatric CT: myth or reality? Radiology 223 (2002) 5–6.
[5] J. Malone, X-rays for medical imaging: radiation protection, governance and ethics over 125 years, Physica Medica 79 (2020) 47–64.
[6] A. Ferrero, N. Takahashi, T.J. Vrtiska, A.E. Krambeck, J.C. Lieske, C. H. McCollough, Understanding, justifying, and optimizing radiation exposure for CT imaging in nephrourology, Nat. Rev. Urol. 16 (2019) 231–244.
[7] Y. Salimi, I. Shiri, A. Akavanallaf, Z. Mansouri, H. Arabi, H. Zaidi, Fully automated accurate patient positioning in computed tomography using anterior-posterior localizer images and a deep neural network: a dual-center study, Eur. Radiol. 33 (2023) 3243–3252.
[8] Y. Salimi, I. Shiri, A. Akhavanallaf, Z. Mansouri, A. Saberi Manesh, A. Sanaat, M. Pakbin, D. Askari, S. Sandoughdaran, E. Sharifipour, H. Arabi, H. Zaidi, Deep learning-based fully automated Z-axis coverage range definition from scout scans to eliminate overscanning in chest CT imaging, Insights Imaging 12 (2021) 162.
[9] K.S.H. Kulathilake, N.A. Abdullah, A.Q.M. Sabri, K.W. Lai, A review on deep learning approaches for low-dose computed tomography restoration, Complex Intell. Syst. 9 (2023) 2713–2745.
[10] P. Gilbert, Iterative methods for the three-dimensional reconstruction of an object from projections, J. Theor. Biol. 36 (1972) 105–117.
[11] T. Würfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, A. K. Maier, Deep learning computed tomography: learning projection-domain weights from image domain in limited angle problems, IEEE Trans. Med. Imaging 37 (2018) 1454–1463.
[12] Y. Li, K. Li, C. Zhang, J. Montoya, G.-H. Chen, Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions, IEEE Trans. Med. Imaging 38 (2019) 2469–2481.
[13] K.H. Jin, M.T. McCann, E. Froustey, M. Unser, Deep convolutional neural network for inverse problems in imaging, IEEE Trans. Image Process. 26 (2017) 4509–4522.
[14] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, G. Wang, Low-dose CT via convolutional neural network, Biomed. Opt. Express 8 (2017) 679–694.
[15] Z. Zhang, X. Liang, X. Dong, Y. Xie, G. Cao, A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution, IEEE Trans. Med. Imaging 37 (2018) 1407–1417.
[16] W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabhuti, G. Wang, DRONE: dual-domain residual-based optimization network for sparse-view CT reconstruction, IEEE Trans. Med. Imaging 40 (2021) 3002–3014.
[17] Q. Li, R. Li, T. Wang, Y. Cheng, Y. Qiang, W. Wu, J. Zhao, D. Zhang, A cascade-based dual-domain data correction network for sparse view CT image reconstruction, Comput. Biol. Med. 165 (2023) 107345.

[18] Y. Guo, Y. Wang, M. Zhu, D. Zeng, Z. Bian, X. Tao, J. Ma, Dual domain closed-loop learning for sparse-view CT reconstruction, in: 7th International Conference on Image Formation in X-Ray Computed Tomography, SPIE, 2022, pp. 130–136.
[19] D. Hu, J. Liu, T. Lv, Q. Zhao, Y. Zhang, G. Quan, J. Feng, Y. Chen, L. Luo, Hybrid-domain neural network processing for sparse-view CT reconstruction, IEEE Trans. Radiat. Plasma Med. Sci. 5 (2020) 88–98.
[20] L. Chao, Z. Wang, H. Zhang, W. Xu, P. Zhang, Q. Li, Sparse-view cone beam CT reconstruction using dual CNNs in projection domain and image domain, Neurocomputing 493 (2022) 536–547.
[21] J. He, Y. Wang, J. Ma, Radon inversion via deep learning, IEEE Trans. Med. Imaging 39 (2020) 2076–2087.
[22] V. Kandarpa, A. Perelli, A. Bousse, D. Visvikis, LRR-CED: low-resolution reconstruction-aware convolutional encoder–decoder network for direct sparse-view CT image reconstruction, Phys. Med. Biol. 67 (2022) 155007.
[23] W. Wu, X. Guo, Y. Chen, S. Wang, J. Chen, Deep embedding-attention-refinement for sparse-view CT reconstruction, IEEE Trans. Instrum. Meas. 72 (2022) 1–11.
[24] J. Pan, H. Yu, Z. Gao, S. Wang, H. Zhang, W. Wu, Iterative residual optimization network for limited-angle tomographic reconstruction, IEEE Trans. Image Process. 33 (2024) 910–925.
[25] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, Adv. Neural Inf. Process. Syst. 33 (2020) 6840–6851.
[26] Z. Li, Y. Wang, J. Zhang, W. Wu, H. Yu, Two-and-a-half order score-based model for solving 3D ill-posed inverse problems, Comput. Biol. Med. 168 (2024) 107819.
[27] Y. Song, J. Sohl-Dickstein, D.P. Kingma, A. Kumar, S. Ermon, B. Poole, Score-based generative modeling through stochastic differential equations, arXiv preprint arXiv:2011.13456, (2020).
[28] W. Wu, J. Pan, Y. Wang, S. Wang, J. Zhang, Multi-channel optimization generative model for stable ultra-sparse-view CT reconstruction, IEEE Trans. Med. Imaging (2024) 1–11.
[29] W. Wu, Y. Wang, Q. Liu, G. Wang, J. Zhang, Wavelet-improved score-based generative model for medical imaging, IEEE Trans. Med. Imaging 43 (3) (2023) 966–979.
[30] Y. Wang, Z. Li, W. Wu, Time-reversion fast-sampling score-based model for limited-angle CT reconstruction, IEEE Trans. Med. Imaging (2024) 1–11.
[31] Z. Li, D. Chang, Z. Zhang, F. Luo, Q. Liu, J. Zhang, G. Yang, W. Wu, Dual-domain collaborative diffusion sampling for multi-source stationary computed tomography reconstruction, IEEE Trans. Med. Imaging (2024).
[32] F. Chen, S. Li, J. Han, F. Ren, Z. Yang, Review of lightweight deep convolutional neural networks, Arch. Comput. Methods Eng. 31 (2024) 1915–1937.
[33] J. Wang, Y. Lu, G. Lu, Lightweight image denoising network with four-channel interaction transform, Image Vis. Comput. 137 (2023) 104766.
[34] X. Liu, Z. Wu, A. Li, F.-A. Vasluianu, Y. Zhang, S. Gu, L. Zhang, C. Zhu, R. Timofte, Z. Jin, NTIRE 2024 challenge on low light image enhancement: methods and results, arXiv preprint arXiv:2404.14248, (2024).
[35] X. Luo, Y. Qu, Y. Xie, Y. Zhang, C. Li, Y. Fu, Lattice network for lightweight image restoration, IEEE Trans. Pattern Anal. Mach. Intell. 45 (2022) 4826–4842.
[36] Y. Li, Y. Iwamoto, L. Lin, R. Xu, R. Tong, Y.-W. Chen, VolumeNet: a lightweight parallel network for super-resolution of MR and CT volumetric data, IEEE Trans. Image Process. 30 (2021) 4840–4854.
[37] C. Cheng, X.-J. Wu, T. Xu, G. Chen, Unifusion: a lightweight unified image fusion network, IEEE Trans. Instrum. Meas. 70 (2021) 1–14.
[38] G. Ma, X. Zhao, Y. Zhu, H. Zhang, Projection-to-image transform frame: a lightweight block reconstruction network for computed tomography, Phys. Med. Biol. 67 (2022) 035010.
[39] W. Van Aarle, W.J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K.J. Batenburg, J. Sijbers, Fast and flexible X-ray tomography using the ASTRA toolbox, Opt. Express 24 (2016) 25129–25147.
[40] A. Biguri, M. Dosanjh, S. Hancock, M. Soleimani, TIGRE: a MATLAB-GPU toolbox for CBCT image reconstruction, Biomed. Phys. Eng. Express 2 (2016) 055010.
[41] J.a.K.H.a.Ö. Adler, Ozan, Operator discretization library (ODL). Software available from, 2017.
[42] T.R. Moen, B. Chen, D.R. Holmes III, X. Duan, Z. Yu, L. Yu, S. Leng, J.G. Fletcher, C. H. McCollough, Low-dose CT image and projection dataset, Med. Phys. 48 (2021) 902–911.
[43] C. McCollough, B. Chen, D. Holmes, X. Duan, Z. Yu, L. Xu, S. Leng, J. Fletcher, Low Dose CT Image and Projection Data (LDCT-and-Projection-data) (Version 6) [Data set], in: The Cancer Imaging Archive, 2020. https://doi.org/10.7937/9NPB-2637.
[44] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository, J. Digit. Imaging 26 (2013) 1045–1057.
[45] F. Noo, M. Defrise, R. Clackdoyle, Single-slice rebinning method for helical cone-beam CT, Phys. Med. Biol. 44 (1999) 561.
[46] F. Wagner, M. Thies, L. Pfaff, O. Aust, S. Pechmann, D. Weidner, N. Maul, M. Rohleder, M. Gu, J. Utz, On the benefit of dual-domain denoising in a self-supervised low-dose CT setting, in: 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), IEEE, 2023, pp. 1–5.
[47] S. Xie, X. Zheng, Y. Chen, L. Xie, J. Liu, Y. Zhang, J. Yan, H. Zhu, Y. Hu, Artifact removal using improved GoogLeNet for sparse-view CT reconstruction, Sci. Rep. 8 (2018) 6700.
[48] Y. Han, J.C. Ye, Framing U-Net via deep convolutional framelets: application to sparse-view CT, IEEE Trans. Med. Imaging 37 (2018) 1418–1429.
[49] H. Chen, Y. Zhang, M.K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, G. Wang, Low-dose CT with a residual encoder-decoder convolutional neural network, IEEE Trans. Med. Imaging 36 (2017) 2524–2535.
[50] F. Kong, M. Li, S. Liu, D. Liu, J. He, Y. Bai, F. Chen, L. Fu, Residual local feature network for efficient super-resolution, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 766–776.
[51] F.S. Khan, S. Khan, Ntire 2022 challenge on efficient super-resolution: methods and results, in: Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022, pp. 1061–1101.
[52] J. Liu, W. Zhang, Y. Tang, J. Tang, G. Wu, Residual feature aggregation network for image super-resolution, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2359–2368.
[53] F. S. Khan, S.Khan, Ntire 2022 challenge on efficient super-resolution: Methods and results[C], Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2022: 1061-1101.
[54] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, A. Lerer, Automatic differentiation in PyTorch. In NIPS Workshop, 2017. URL https://openreview.net/pdf?id=BJJsrmfCZ.
[55] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980, (2014).
[56] A.H. Andersen, A.C. Kak, Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm, Ultrason. Imaging 6 (1984) 81–94.
[57] Å. Björck, T. Elfving, Z. Strakos, Stability of conjugate gradient and Lanczos methods for linear least squares problems, SIAM J. Matrix Anal. Appl. 19 (1998) 720–736.
[58] Y. Salimi, I. Shiri, Z. Mansouri, H. Zaidi, Deep learning-assisted multiple organ segmentation from whole-body CT images, medRxiv (2023), 2023.2010.2020.23297331.
[59] E. Samei, D. Bakalyar, K.L. Boedeker, S. Brady, J. Fan, S. Leng, K.J. Myers, L. M. Popescu, J.C. Ramirez Giraldo, F. Ranallo, Performance evaluation of computed tomography systems: summary of AAPM Task Group 233, Med. Phys. 46 (2019) e735–e756.
[60] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (2004) 600–612.
[61] H.R. Sheikh, A.C. Bovik, Image information and visual quality, IEEE Trans. Image Process. 15 (2006) 430–444.
[62] K. Ohashi, Y. Nagatani, M. Yoshigoe, K. Iwai, K. Tsuchiya, A. Hino, Y. Kida, A. Yamazaki, T. Ishida, Applicability evaluation of full-reference image quality assessment methods for computed tomography images, J. Digit. Imaging 36 (2023) 2623–2634.
[63] C. Qi, S. Wang, H. Yu, Y. Zhang, P. Hu, H. Tan, Y. Shi, H. Shi, An artificial intelligence-driven image quality assessment system for whole-body [18F] FDG PET/CT, Eur. J. Nucl. Med. Mol. Imaging 50 (2023) 1318–1328.
[64] D.V. Cicchetti, Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology, Psychol. Assess 6 (1994) 284.
[65] B. Lu, L. Fu, Y. Pan, Y. Dong, SWISTA-Nets: subband-adaptive wavelet iterative shrinkage thresholding networks for image reconstruction, Comput. Med. Imaging Graph. 113 (2024) 102345.
[66] X. Chu, B. Zhang, R. Xu, Moga: searching beyond mobilenetv3, in: ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2020, pp. 4042–4046.
[67] Y. He, P. Liu, Z. Wang, Z. Hu, Y. Yang, Filter pruning via geometric median for deep convolutional neural networks acceleration, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4340–4349.
[68] X. Zhang, J. Zou, K. He, J. Sun, Accelerating very deep convolutional networks for classification and detection, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2015) 1943–1955.
[69] M. Nagel, M.v. Baalen, T. Blankevoort, M. Welling, Data-free quantization through weight equalization and bias correction, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 1325–1334.
[70] C. Buciluǎ, R. Caruana, A. Niculescu-Mizil, Model compression, in: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, 2006, pp. 535–541.
[71] European Guidelines on Quality Criteria for Computed Tomography, Report EUR 16262 (Brussels: EU), 2000.
[72] L. Cao, X. Liu, T. Qu, Y. Cheng, J. Li, Y. Li, L. Chen, X. Niu, Q. Tian, J. Guo, Improving spatial resolution and diagnostic confidence with thinner slice and deep learning image reconstruction in contrast-enhanced abdominal CT, Eur. Radiol. 33 (2023) 1603–1611.
[73] J. Leipsic, T.M. Labounty, B. Heilbron, J.K. Min, G.J. Mancini, F.Y. Lin, C. Taylor, A. Dunning, J.P. Earls, Adaptive statistical iterative reconstruction: assessment of image noise and image quality in coronary CT angiography, Am. J. Roentgenol. 195 (2010) 649–654.
[74] E. Abadi, J. Sanders, E. Samei, Patient-specific quantification of image quality: an automated technique for measuring the distribution of organ Hounsfield units in clinical chest CT images, Med. Phys. 44 (2017) 4736–4746.
[75] T.S. Cho, N. Joshi, C.L. Zitnick, S.B. Kang, R. Szeliski, W.T. Freeman, A content-aware image prior, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010, pp. 169–176.
[76] L. Zhu, D. Ji, S. Zhu, W. Gan, W. Wu, J. Yan, Learning statistical texture for semantic segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 12537–12546.
[77] D. Marcos, M. Volpi, D. Tuia, Learning rotation invariant convolutional filters for texture classification, in: 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE, 2016, pp. 2012–2017.

[78] C. Liu, H. Yang, J. Fu, X. Qian, 4D LUT: learnable context-aware 4d lookup table for image enhancement, IEEE Trans. Image Process. 32 (2023) 4742–4756.
[79] L. Liao, J. Xiao, Z. Wang, C.-W. Lin, S.i. Satoh, Image inpainting guided by coherence priors of semantics and textures, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 6539–6548.
[80] M. Suin, K. Purohit, A. Rajagopalan, Degradation aware approach to image restoration using knowledge distillation, IEEE J. Sel. Top. Signal Process. 15 (2020) 162–173.