One half-scan dual-energy CT imaging using the Dual-domain Dual-way Estimated Network (DoDa-Net) model

Yizhong Wang, Ailong Cai, Ningning Liang, Xiaohuan Yu, Xinyi Zhong, Lei Li, Bin Yan

Henan Key Laboratory of Imaging and Intelligent Processing, PLA Strategic Support Force Information Engineering University, Zhengzhou, China

Contributions: (I) Conception and design: Y Wang, L Li; (II) Administrative support: B Yan; (III) Provision of study materials or patients: N Liang; (IV) Collection and assembly of data: X Yu, X Zhong; (V) Data analysis and interpretation: Y Wang, A Cai; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Lei Li. Henan Key Laboratory of Imaging and Intelligent Processing, PLA Strategic Support Force Information Engineering University, Zhengzhou, China. Email: leehotline@163.com.

Background: Compared with single-energy computed tomography (CT), dual-energy CT (DECT) can distinguish materials better. However, most DECT reconstruction theories require two full-scan projection datasets of different energies, and this requirement is hard to meet, especially for cases where a physical blockage disables a full circular rotation. Thus, it is critical to relax the requirements of data acquisition to promote the application of DECT.

Methods: A flexible one half-scan DECT scheme is proposed, which acquires two projection datasets on two-quarter arcs (one for each energy). The limited-angle problem of the one half-scan DECT scheme can be solved by a reconstruction method. Thus, a dual-domain dual-way estimation network called DoDa-Net is proposed by utilizing the ability of deep learning in non-linear mapping. Specifically, the dual-way mapping Generative Adversarial Network (DM-GAN) was first designed to mine the relationship between two different energy projection data. Two half-scan projection datasets were obtained, the data of which was twice that of the original projection dataset. Furthermore, the data transformation from the projection domain to the image domain was realized by the total variation (TV)-based method. In addition, the image processing network (Im-Net) was employed to optimize the image domain data.

Results: The proposed method was applied to a digital phantom and real anthropomorphic head phantom data to verify its effectiveness. The reconstruction results of the real data are encouraging and prove the proposed method's ability to suppress noise while preserving image details. Also, the experiments conducted on simulated data show that the proposed method obtains the closest results to the ground truth among the comparison methods. For low- and high-energy reconstruction, the peak signal-to-noise ratio (PSNR) of the proposed method is as high as 40.3899 and 40.5573 dB, while the PSNR of other methods is lower than 36.5200 dB. Compared with FBP, TV, and other GAN-based methods, the proposed method reduces root mean square error (RMSE) by, respectively, 0.0124, 0.0037, and 0.0016 for low-energy reconstruction, and 0.0102, 0.0028, and 0.0015 for high-energy reconstruction.

Conclusions: The developed DoDa-Net model for the proposed one half-scan DECT scheme consists of two stages. In stage one, DM-GAN is used to realize the dual map of projection data. In stage two, the TV-based method is employed to transform the data from the projection domain to the image domain. Furthermore, the reconstructed image is processed by the Im-Net. According to the experimental results of qualitative and quantitative evaluation, the proposed method has advantages in detail preservation, indicating the potential of the proposed method in one half-scan DECT reconstruction.

Keywords: Dual-energy CT (DECT); one half-scan; limited-angle problem; deep learning

Introduction

Dual-energy computed tomography (DECT) has been widely applied in different fields, including industrial non-destructive testing (1), clinical medical diagnosis (2-4), and safety inspections (5,6). Compared with traditional single-energy CT, DECT can explore the interdependence of X-ray attenuation and photon energy and thus has a greater ability to identify substances (7).

Three DECT systems for commercial use have been developed (8): sequential acquisition, rapid voltage switching, and dual-source CT. The sequential acquisition system requires the least hardware effort and involves the sequential acquisition of two datasets at different tube voltages, but the disadvantages are long acquisition time and high radiation dose. The rapid voltage switching system uses a control device to alternate the tube voltage between a high value and a low value and collects two transmission datasets for each projection. However, the important disadvantage of this method is the limited photon output at low voltage, which leads to high noise. Another system is dual-source CT with two tubes operating at different voltages, which requires nearly twice the hardware investment. In this paper, a flexible DECT imaging scheme was designed to reduce the total radiation dose's hardware cost. As shown in *Figure 1*, the X-ray tube is adjusted to change the tube voltage at a 90° angle and obtain the projection dataset of two quarter scans at a high- and low-energy spectrum. This DECT is referred to as one half-scan DECT, and it can be performed on traditional CT scanning equipment, requiring only one source/detector pair. Also, an effective method to relax the data acquisition requirements of DECT by using the correlation between two different X-ray energies is proposed in this paper.

The DECT system is generally considered an imaging system that utilizes two effective X-ray spectrums to obtain projection data and performs material decomposition. The existing DECT reconstruction methods can be divided into three categories: one-step reconstruction (9,10), two-step reconstruction based on projection domain (11,12), and image domain (13,14). For the conventional methods of these approaches, high-quality DECT reconstruction images all require a full-angle projection dataset (15,16).

The missing scanning angles in the two different energies may disturb the material decomposition, resulting in the distortion of the decomposition results. Thus, developing methods to obtain high-quality DECT images without a full-angle projection dataset has become a research hotspot.

Recently, various methods have been proposed to improve DECT images in the case of each energy without full-angle scanning, including software-based (17,18) and hardware-based methods. As for the software-based method, Shen *et al.* proposed a segmental scanning imaging method to realize multi-energy CT imaging on traditional CT systems (19). As for the hardware-based method, Petrongolo *et al.* used a grid plate to modulate and filter the X-ray energy spectrum and proposed a primary modulation dual-energy sparse-view projection data reconstruction method (20). The above methods have produced beneficial effects for DECT, but the total scan angle of the multi-energy imaging still needs to perform a full-angle scan of the circumference, which is invalid in the imaging scheme without full-angle scanning. To solve this problem, Zhang and Xing proposed a joint clustering prior sparsity regularization (CPSR) model, which exploits the coherence of all the data at different energies to address the limited-angle reconstruction problem (21). Zhang *et al.* subsequently proposed a reconstruction method based on the image domain with a half-scan plus a second limited-angle scan (22). Although these methods help to improve the quality of DECT images without full-angle scanning, the reconstruction results may lose some details.

Recently, deep learning (DL) technology has been extensively used in CT imaging due to its non-linear feature extraction and modeling capabilities (23-27). In the field of material decomposition, researchers have tried to apply DL to DECT to solve the problem of noise amplification in material decomposition (28,29). Clark *et al.* used the existing U-Net (30) architecture to obtain good material decomposition performance (31). In 2019, Zhang *et al.* used DL to solve the ill-posed solution of the material decomposition model and developed a model-based butterfly network to decompose DECT in the image domain (32). DL has mainly been used for normal-dose DECT in these methods, and low-dose or incomplete-angle DECT has not been investigated. In 2020, we conducted
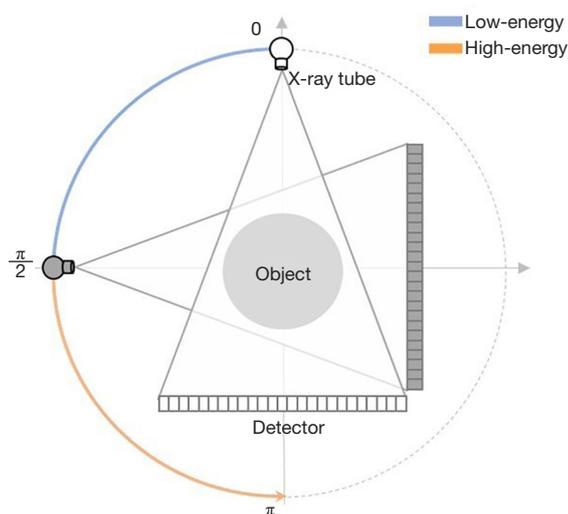
Figure 1 One half-scan DECT scheme. Low energy scan range (0°, 90°]; high energy scan range (90°, 180°]. DECT, dual-energy CT.

a preliminary study of the one half-scan DECT scheme and obtained good DECT image quality by using the Generative Adversarial Network (GAN) method to process projection domain data (33). However, the method ignores image domain information and the correlation between low- and high-energy projection domain information. Recently, Zhang *et al.* used the redundancy of projection and image domain (dual-domain) data to propose a comprehensive domain network for the problem of DECT sparse-view reconstruction, which improved the image quality of DECT imaging (34). This means that the dual-domain DL method has the potential to solve the problem of incomplete-angle DECT. Inspired by these ideas, the proposed method in this paper aims to solve the limitations of our previous work (the details of the proposed method are described in section Methods).

The main challenges of the one half-scan DECT reconstruction are as follows. First, an unreliable solution can be obtained by using standard CT reconstruction methods on the projection data collected by the one half-scan DECT, and there are usually serious artifacts in the reconstruction image. Second, the small changes of the original projection data can significantly impact the reconstruction image quality. To overcome these challenges, an efficient method is proposed in this study to solve the one half-scan DECT scheme by simultaneously exploring the relationship between different energy data and applying DL to non-linear mapping. Inspired by CycleGAN (35,36),

a novel dual-way mapping GAN model called DM-GAN is proposed. The cycle-consistent structure design of the DM-GAN enables the network to learn the mapping relationship between low- and high-energy projection datasets.

Moreover, the projection domain loss may also play an important role in this task. The projection domain loss requires paired training data, which is different from common unsupervised methods (such as CycleGAN). Then, the trained neural network is exploited to estimate the projection data with two different energies. Based on this, two half-scan projection datasets with two different energies can be obtained, which are twice the size of the original projection datasets. Subsequently, the total variation (TV)-based method (37) is exploited to realize the data translation from the projection domain to the image domain. Finally, an image processing network (Im-Net) is adopted to eliminate the artifacts caused by the small inconsistency in the projection generated by DM-GAN. Following the two stages, the end-to-end network for mapping from the projection domain to the image domain is obtained, which has advantages in exploring the mapping relationship between different energy projection data by DM-GAN and considering the projection error and image error in the process of DM-GAN and Im-Net supervision training. The evaluation results indicate that the proposed method can solve the one half-scan DECT reconstruction problem, effectively suppressing the artifacts of the reconstruction image and ensuring the accuracy of material decomposition. The dual-domain dual-way estimation network is referred to as DoDa-Net in this paper.

We present the following article following the MDAR checklist (available at https://dx.doi.org/10.21037/qims-21-441).

Methods

Reconstruction algorithm and angular sampling strategy

The mathematical model of CT imaging for reconstructing images from projection data can be represented approximately by the following discrete linear equations:

$$\vec{p}_s = A_s \vec{u}_s , \qquad [1]$$

where $p_s \in \mathbb{R}^{N_U N_V}$, $A_s \in \mathbb{R}^{N_U N_V \times N_W N_H}$, and $u_s \in \mathbb{R}^{N_W N_H}$ denote the projection data, the system matrix, and the discrete CT image (linear attenuation map) at energy E_s, respectively. In this paper, $E_s = E_l$ or $E_s = E_h$, and l and h represent low-energy and high-energy, respectively. N_U and N_V represent the number of detector pixels and projection

views, respectively. N_W and N_H represent the width and height of the discrete CT image, respectively.

Usually, the projection data provided by limited-angle scans is not enough to be reconstructed by Eq. [1]. An effective method is to combine the compressed sensing theory with a sparse prior. With a certain degree of sparsity, it is found that Eq. [1] can be accurately solved by the TV-based method (37):

$$\min_{\vec{u}_s} \sum_i \left\| D_i \vec{u}_s \right\|_1,$$
$$s.t. \ A_s \vec{u}_s - \vec{p}_s = 0, \ \vec{u}_s \geq 0 \qquad [2]$$

where $\|\cdot\|_1$ represents the L1-norm, and D_i represents the discrete direction gradient operator of direction $_i$. However, when the projection data is seriously insufficient, it is difficult to obtain an accurate solution through the TV-based method.

For DECT, there is a correlation between low- and high-energy data of the same object. Thus, researchers have provided different ideas when facing seriously insufficient projection data in DECT. In 2016, Zhang *et al.* used DL to establish the correspondence between the attenuation coefficient at the same position of low- and high-energy images for limited-angle scan DECT imaging (38). In 2020, we used the GAN method to process projection domain data and realized limited-angle scan DECT imaging (33). These methods only focus on the information of one domain (image domain or projection domain) and ignore the information of the dual-domain. Based on this idea, we propose the DoDa-Net method to achieve one half-scan DECT imaging. Let θ_s be the viewing angle range of the tube potential E_s. Generally, the sufficient condition for an excellent reconstruction is that the angle of each energy coverage θ_s is at least 180° + fan angle, and the requirements of the proposed method can be relaxed to $\sum_s \theta_s = 180°$.

Dual-domain dual-way estimation network (DoDa-Net) reconstruction

The symbols θ_1 and θ_2 are used to represent the blue and orange scan ranges in *Figure 1*, respectively. As for the designed one half-scan DECT reconstruction, we have:

$$\vec{p}_l^{\theta_1} = A_l^{\theta_1} \vec{u}_l^{\theta_1},$$
$$\vec{p}_h^{\theta_2} = A_h^{\theta_2} \vec{u}_h^{\theta_2} \qquad [3]$$

To solve Eq. [3], a novel DoDa-Net reconstruction method is proposed in this paper, and the overall framework of the DoDa-Net method is shown in *Figure 2*. The overall

process is divided into two stages, and the details of each stage are as follows.

Stage 1: dual-way mapping GAN (DM-GAN)

In stage one, the entire DM-GAN framework consists of two generators G and F, which respectively map low-energy data to high-energy data and high-energy data to low-energy data in the projection domain, i.e., realizing two mapping relationships: $G:E_l \rightarrow E_h, F:E_h \rightarrow E_l$. The input of the network is the dual-quarter projection datasets $\vec{p}_l^{\theta_1}$ and $\vec{p}_h^{\theta_2}$ obtained by the one half-scan DECT scheme. The architecture of DM-GAN is derived from CycleGAN (DM-GAN as a whole is different due to the projection domain loss), and the direction of network data flow is dual-way. Specifically, $\vec{p}_l^{\theta_1}$ is input to the trained generator G to obtain $\hat{\vec{p}}_h^{\theta_1}$, and then the trained generator F converts $\hat{\vec{p}}_h^{\theta_1}$ to $\hat{\hat{\vec{p}}}_l^{\theta_1}$. Meanwhile, $\vec{p}_h^{\theta_2}$ is input to the trained generator F to obtain $\hat{\vec{p}}_l^{\theta_2}$, and then the trained generator G converts $\hat{\hat{\vec{p}}}_l^{\theta_2}$ to $\vec{p}_h^{\theta_2}$. In this way, the dual-energy projection datasets of $\vec{p}_l^{\text{whole}} = \begin{pmatrix} \vec{p}_l^{\theta_1} \\ \hat{\vec{p}}_l^{\theta_2} \end{pmatrix}$ and $\vec{p}_h^{\text{whole}} = \begin{pmatrix} \hat{\vec{p}}_h^{\theta_1} \\ \vec{p}_h^{\theta_2} \end{pmatrix}$ are obtained through generators G and F, which doubles the amount of data compared with the original data.

Furthermore, the purpose of discriminators D_l and D_h is to identify samples from the real projection dataset instead of the generated projection dataset. The generator and the discriminator are trained alternately to improve the quality of the network output, which is the fundamental of the "game" in the GAN training process. In this study, the specific architecture of the generator and discriminator in the network model is described as follows.

Generator

As shown in *Figure 3*, the overall structure of the generator is composed of an encoder, decoder, and residual module. In the generator, the input projection image is a 512×512 gray image. The generator works as follows. First, the encoder extracts features from the input projection image. Then, the residual module diversifies the feature extraction to avoid the problem of network degradation. Subsequently, it converts the feature vector extracted from the encoder to the target domain. Overall, the residual module plays an important role in feature extraction diversification and the conversion domain. Finally, the decoder obtains the projection image by transposed convolution. The skip-connections connect the corresponding encoder and decoder layers to help the decoder better recover the details of the projection data. Specifically, the encoder consists of three layers.
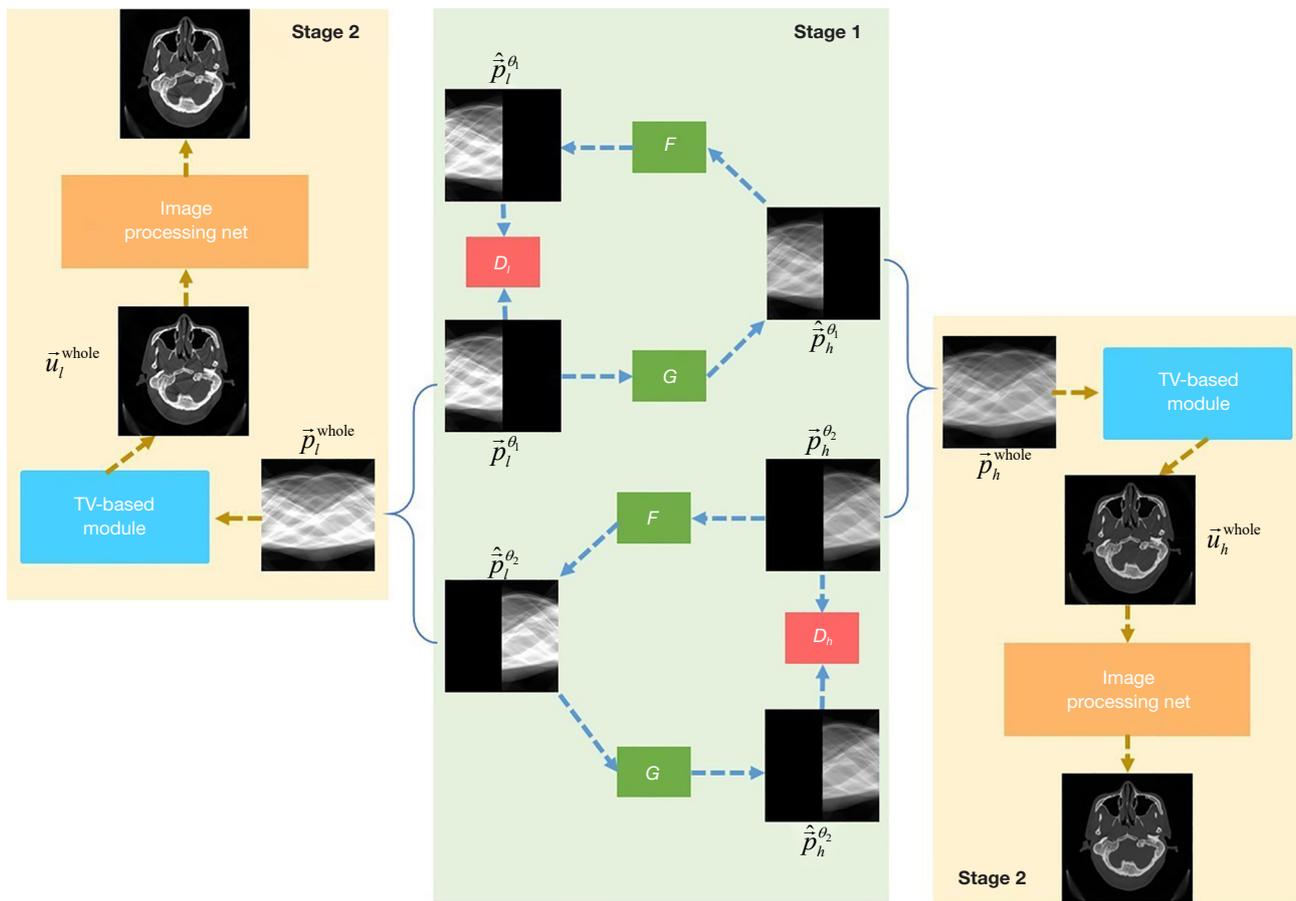
Figure 2 The overall framework of the proposed network model. The process of the model includes two stages, where stage one realizes dual mapping of projection data, and stage two realizes high-quality restoration of the reconstructed image. F and G are generators; D_l and D_h are detectors.

In the first layer, the padding operation is performed. Then, the convolution (Conv) operation is performed with a step size of 2 units and a kernel size of 4×4. In the second and third layers, the Conv operation is performed first with a step size of 2 units and a Conv kernel size of 4×4, followed by the execution of the instance norm (IN) and leaky rectified linear units (Leaky ReLU). The numbers of filters in these three layers are 64, 128, and 256, respectively. The residual module is a residual network (ResNet) (39,40), which contains 6 ResNet blocks (41). Each ResNet block consists of two convolutional layers, and each layer contains 256 filters. The decoder consists of three layers. The first two are deconvolution (DeConv) layers, which perform the DeConv operation with a step size of 2 units and a Conv kernel size of 4×4, followed by the IN and the rectified linear unit (ReLU). In the last layer, the padding operation is performed. Then, the

DeConv operation is performed with a step size of 2 units and a kernel size of 4×4, followed by the Tanh function. The numbers of filters in these three layers are 128, 64, and 1, respectively.

Discriminator

The generated projection image and the corresponding real projection image are respectively input into the discriminator, and the output of both is the 64×64 patch. The sigmoid function is used to calculate all the elements in the two patches, and then the average value is taken as the basis to judge the authenticity of the generated projection image. Specifically, the structure of the discriminator is a fully convolutional network with five layers. The numbers of filters in the convolutional layer are 64, 128, 256, 512, and 1, respectively. These filters are used to extract features of different levels. In the first layer, the Conv is performed with a kernel size of 4×4 and a stride of 2 units, followed
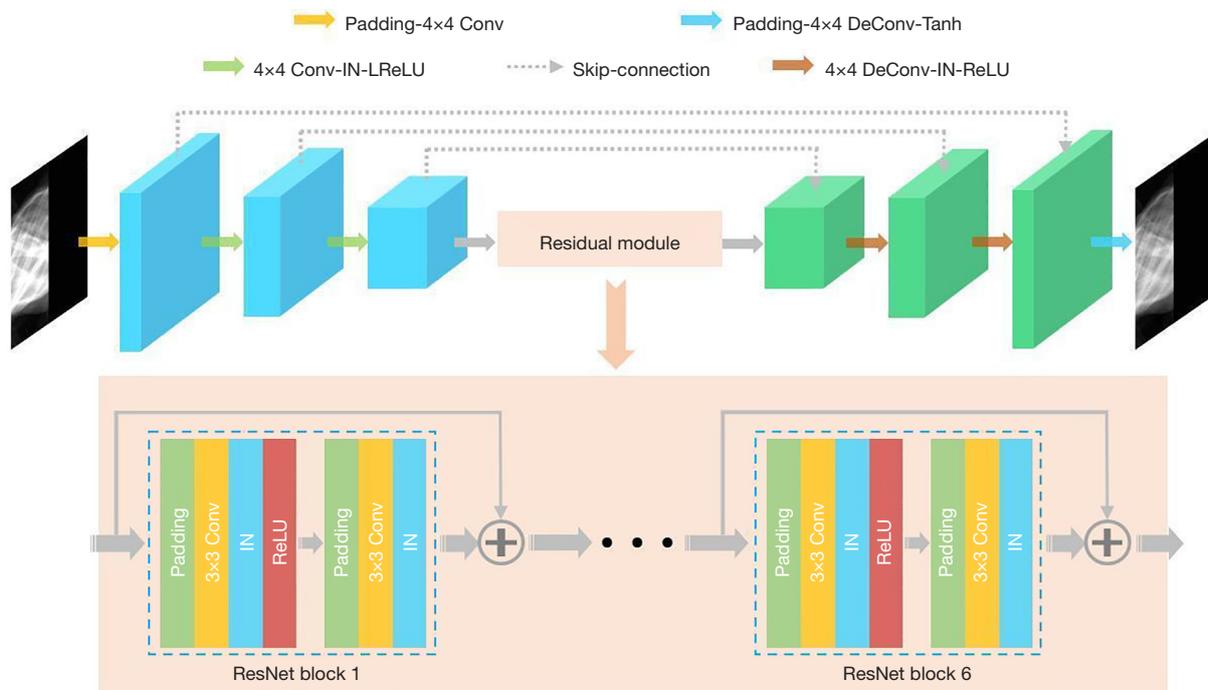
Figure 3 The architecture of DM-GAN generators. The generators consist of three parts: encoder, decoder, and residual module. The residual module is a residual network composed of six residual blocks. Conv, the convolution operation; DeConv, the deconvolution operation; IN, the instance norm; ReLU, rectified linear unit; Leaky ReLU, leaky rectified linear unit.

by the execution of a Leaky ReLU. In the second and third layers, the Conv is performed with a kernel size of 4×4 and a stride of 2 units, followed by the execution of an IN-Leaky ReLU. In the fourth layer, the Conv is performed with a kernel size of 4×4 and a stride of 1 unit, followed by the execution of an IN-Leaky ReLU. In the last layer, the Conv is performed with a kernel size of 4×4 and a stride of 1 unit, followed by the execution of a sigmoid function.

Stage 2: image processing network (Im-Net)

In stage two, the projection datasets $\vec{p}_l^{\text{whole}}$ and $\vec{p}_h^{\text{whole}}$ obtained in stage one are first used to obtain two reconstruction images $\vec{u}_l^{\text{whole}}$ and $\vec{u}_h^{\text{whole}}$ by using the TV-based reconstruction algorithm. It is well known that minor inconsistency in projection data can cause serious distortion of the reconstructed image. Therefore, the reconstructed image is further processed to achieve a high-quality reconstruction. The input of the network is the reconstruction image $\vec{u}_l^{\text{whole}}$ or $\vec{u}_h^{\text{whole}}$. The target of the network is the corresponding image reconstructed from the full-angle projection data under different energies. Since the data distribution of images under different energies is inconsistent, two network models are used to train the

reconstructed images at different energies independently. As shown in *Figure 4*, the modified U-Net is used as the Im-Net, which includes an encoder that down-samples the image to extract representative features and a decoder that up-samples the features to restore the image. The architecture of the Im-Net is described in detail as follows.

The input of the Im-Net is a 512×512 gray image. Its structure is mainly composed of twelve modules, where the encoder and decoder have six modules. In the encoder, each module performs two Conv-ReLU operations on the input data, and the numbers of filters in the Conv layers of each module are 32, 64, 128, 256, 512, and 1,024, respectively. Meanwhile, the size of the Conv kernel is 3×3 and the step size is 1 unit. Moreover, max-pooling with a step size of 2 units is used for down-sampling to extract abstract features between every two modules. As for the decoder, it has a similar structure to the encoder. In the decoder, the numbers of filters in the Conv layers of each module are 512, 256, 128, 64, 32, and 1, respectively. In the first five modules of the decoder, up-sampling is first performed on the feature maps to reduce the number of feature maps by half. Then, the skip-connection between the feature mapping of the coding layer and the decoding layer is
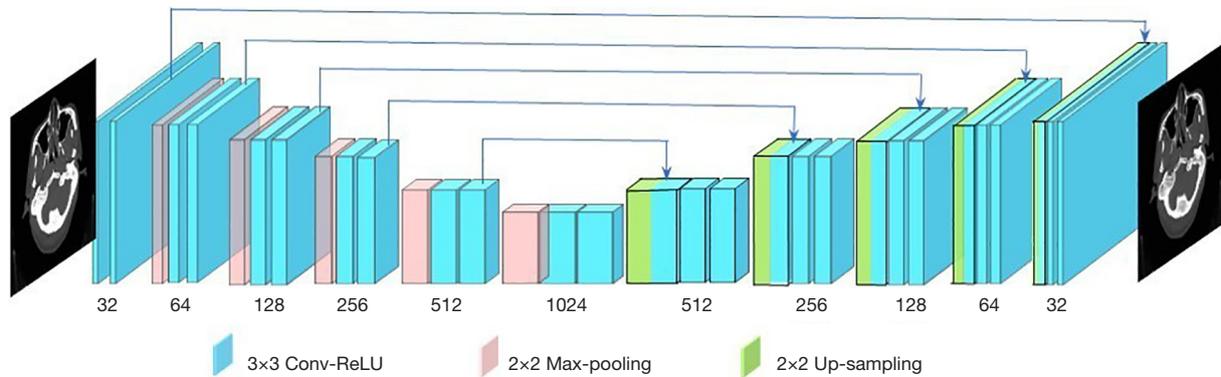
Figure 4 The architecture of the Im-Net. The blue, pink, and green rectangles represent 3×3 Conv-ReLU, 2×2 Max-pooling, and 2×2 Up-sampling, respectively. Conv, the convolution operation; ReLU, rectified linear unit.

introduced to obtain the high-resolution details of the corresponding module. Moreover, two Conv-ReLUs with a Conv kernel size of 3×3 and one unit stride of 1 unit are repeatedly applied to the connected data. The last module of the encoder uses a 1×1 Conv to map the 32-channel feature vector to the high- or low-energy image.

Loss function

In the first stage of the DoDa-Net model, it is desired to learn the mapping function between the projection images at two different energies E_l and E_h of a given training sample. For the convenience of explanation, the E_l and E_h domains are respectively replaced with X and Y domains. Also, the discriminators D_l and D_h correspond to D_X and D_Y, respectively. To promote high-quality dual-energy image generation, the network structure proposed in this paper combines the following loss functions.

Adversarial loss

The generators G and F mainly realize the conversion between domain Y and domain $X(G:X{\rightarrow}Y,F:Y{\rightarrow}X)$. The adversarial loss is applied to these two mapping functions (42) so that the generated projection data obeys the empirical distribution in the X or Y domain. Thus, the adversarial loss can be defined as:

$$\mathcal{L}_{GAN}\left(G,D_Y\right)=\mathbb{E}_{y\sim P_{data}(y)}\left[\log D_Y\left(y\right)\right]+\mathbb{E}_{x\sim P_{data}(x)}\left[\log\left(1-D_Y\left(G(x)\right)\right)\right]$$

$$\mathcal{L}_{GAN}\left(F,D_X\right)=\mathbb{E}_{x\sim P_{data}(x)}\left[\log D_X\left(x\right)\right]+\mathbb{E}_{y\sim P_{data}(y)}\left[\log\left(1-D_X\left(F(y)\right)\right)\right]$$

[4]

where $x{\in}X$ and $y{\in}Y$ represent the training samples in the X

and Y domain, respectively, and P_{data} is the projection data distribution.

Cycle consistency loss

To prevent the degradation of adversarial learning (43), the cycle consistency loss is adopted:

$$\mathcal{L}_{cyc}\left(G,F\right)=\mathbb{E}_{x\sim P_{data}(x)}\left[\left\|F\left(G(x)\right)-x\right\|_1\right]+\mathbb{E}_{y\sim P_{data}(y)}\left[\left\|G\left(F(y)\right)-y\right\|_1\right]$$

[5]

Projection domain loss

The accuracy of the generated projection image is crucial, which makes direct constraints on the generator necessary. In the process of projection image mapping, the projection domain loss is introduced to ensure the fidelity of input projection data and the 180° reference projection data. Previous research shows that mean absolute error (MAE) loss is commonly used in CycleGAN, which is conducive to pixel-level image approximation (35,44). Thus, the projection domain loss function is constructed based on the MAE loss, and it is expressed as:

$$\mathcal{L}_{prj}\left(G,F\right)=\mathbb{E}_{x\sim P_{data}(x)}\left[\left\|G(x)-y\right\|_1\right]+\mathbb{E}_{y\sim P_{data}(y)}\left[\left\|F(y)-x\right\|_1\right]$$

[6]

Therefore, to train the proposed model, the total loss function is:

$$\mathcal{L}_{total}\left(G,F,D_X,D_Y\right)=\mathcal{L}_{GAN}\left(G,D_Y\right)+\mathcal{L}_{GAN}\left(F,D_X\right)+\lambda_1\mathcal{L}_{cyc}\left(G,F\right)+\lambda_2\mathcal{L}_{prj}\left(G,F\right)$$

[7]

where λ_1 and λ_2 adjust the relative importance of the three target losses. Based on this, the optimization problem can

be described as:

$$G*, F* = \arg \min_{G,F} \max_{D_X, D_Y} \mathcal{L}_{total}(G, F, D_X, D_Y). \qquad [8]$$

One half-scan DECT using the DoDa-Net reconstruction method

In summary, the proposed reconstruction method for the one half-scan DECT scheme includes two stages. In stage one, DM-GAN is exploited to realize dual-way mapping of the low- and high-energy projection data from different scanning angles. In this way, two half-scan projection datasets with two different energies are obtained. In stage two, the TV-based method is adopted to convert the data from the projection domain to the image domain. Then, the reconstruction image is processed by the Im-Net to obtain a relatively high-quality DECT image.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of Henan Provincial People's Hospital, and written informed consent was obtained.

Evaluation

Experimental data and training details

Digital cranial cavity experiment data

Obtained from the clinical dataset, the dataset used in the simulation experiment contained 1,603 cranial cavity CT images from 6 patients. The size of the CT image was 512×512. The training data for training the network consisted of 1,391 CT images obtained from 5 patients. A total of 212 CT images were obtained from another patient for preparing the test dataset. The specific steps for preparing the training dataset are described below.

To train the network, the projection dataset at two different energies was needed. Thus, the patient's basic material images, including bone and tissue images were obtained with the assistance of the local hospital radiologist. Radiologists manually corrected each pair of bone and tissue images using the bone removal function of Syngo.Via equipped in Siemens SOMATOM Definition Flash CT scanner. Taking the decomposed bone and tissue images as input, dual-energy projections were generated by the discrete projection model (45). The mass attenuation coefficient of the basic material was obtained from the database of the National Institute of Standards

and Technology (https://physics.nist.gov/PhysRefData/XrayMassCoef/tab4.html). Specifically, SpekCalc software was adopted to generate 80 and 140 kVp polychromatic spectra with an energy sampling interval of 1 keV.

Furthermore, Sidden's ray tracing algorithm was applied to simulate the geometry of the fan beam (46). The distances from the source to the object and the source to the detector were set to 1,000 and 1,500 mm, respectively. The dual-energy projections were uniformly sampled in 180 views rotated by 180°. The projected samples in each view were collected by a linear detector, which consisted of 512 bins and each pixel was 0.408 mm. Thus, the size of the projection data for simulating the projection process was 512×180. Then, the low- and high-energy 180° projection datasets were processed to obtain a low-energy projection dataset (0°–90°) and a high-energy projection dataset (90°–180°) as the DM-GAN input, and a high-energy projection dataset (0°–90°) and a low-energy projection dataset (90°–180°) as the DM-GAN target. The input projection dataset was labeled and adjusted to 512×512 for feature extraction during DM-GAN training.

Additionally, simulated Poisson noise was added to the projection to illustrate the practicability of the proposed method:

$$\vec{\vec{p}}_i = \frac{\lambda^k}{k!} e^{-\lambda}, \lambda = I\left(\exp\left(-\vec{p}_0\right)\right) \qquad [9]$$

where $\vec{\vec{p}}_i$ represents the number of added noise photons collected by the detector unit i, and $\vec{p}_0$ is the logarithmic projection data. k is the index of the projection detector. The initial intensity I of the incident photon is set to 1×10^5.

Furthermore, the TV-based reconstruction algorithm was used to reconstruct the 180° projection dataset output by the DM-GAN, and then the low- and high-energy reconstruction image datasets obtained were used as Im-Net input. The target of the Im-Net was the corresponding image reconstructed from the full-angle projection dataset under different energies. The reconstruction image dataset was labeled and adjusted to 512×512 for feature extraction during Im-Net training.

Anthropomorphic head phantom experiment data

In the actual projection collection process, the data collected by the CT scanner is often affected by photon scattering and system noise. Although the data generation process of CT scanning can be simulated, it is challenging to completely reduce the difference of statistical distribution between real data and simulated data (47). Thus, obtaining

Figure 5 Real anthropomorphic head data experimental phantom: Chengdu Dosimetric Phantom.

the real data from a practical CT scanner is essential to adjust the network parameters for the trained network. In this work, a laboratory physical CT scanner was used to scan the anthropomorphic head phantom (as shown in *Figure 5*) (48) at tube voltages of 80 and 140 kVp. The physical CT scanner was mainly composed of an X-ray source (Hawkeye130, Thales) and a flat panel detector (Varian 4030E). The distance from the source to the object and the detector was 623.09 and 865.05 mm. The detector was composed of 512 bins, and the size of each bin was 0.83 mm. A total of 180 projections were collected within 180° at a sampling interval of 1°. The center slice of each projection was extracted for two-dimensional DECT reconstruction, and the size of the reconstruction image was 512×512. The anthropomorphic head phantom data source was obtained from a real CT scanner, which is different from the simulation data. However, for the process of generating the training dataset, the anthropomorphic head phantom experiment was the same as the simulation experiment. Similarly, the anthropomorphic head phantom training dataset generation was the same as that of the simulation experiment. As for the anthropomorphic head phantom data, 747 and 76 slices of physical phantoms were randomly selected as the training dataset and test dataset for real data experiments, respectively.

Training details

DM-GAN and Im-Net training respectively used the Pytorch toolbox (ver. 0.4.1) and TensorFlow (ver. 1.9.0)

toolbox running on an AMAX workstation equipped with two Intel(R) Xeon(R) E5-2640 v4 @ 2.4 GHz CPUs. Four NVIDIA GeForce GTX 1080 Ti GPUs were used for training and testing. The parameters were updated by the Adam adaptive moment optimizer (49). The exponential decay rate β_1 and β_2 were respectively set to 0.8 and 0.999, and the learning rate was set to 0.0002. The total training times of DM-GAN and Im-Net were 40 and 15 hours, respectively.

Performance evaluation

In this study, the performance of the proposed network in the projection domain and image domain was qualitatively evaluated. In the projection domain, the GAN-based (50) and CycleGAN-based (30) methods were adopted for performance comparison. The network architecture of the GAN-based method consisted of a generator and a discriminator, and the structure of the generator was like that of U-Net. In the process of network training, we used the 90° projection data as the network input, and the corresponding 180° projection data as the network target to generate fake 180° projection data. In addition, DM-GAN and CycleGAN had the same architecture in this work, which was composed of two generators and one discriminator. The difference between the two was that we had improved the generator of DM-GAN. The DM-GAN consists of three parts: encoder, decoder, and residual module, while the structure of the CycleGAN is similar to

Figure 6 The average RMSE of the projection data for different values of λ_1 and λ_2. RMSE, root mean square error.

Results

Study of parameters

Parameter selection is based on the simulated data with Poisson noise. With the total loss function, different RMSE losses in the training process of DM-GAN are determined jointly by the weight parameters λ_1 and λ_2, where λ_1 and λ_2 control the weight of cycle consistency loss and the weight of loss between the generated projection dataset and the 180° reference projection dataset, respectively. An improper parameter selection will affect the performance of the network model and lead to the degradation of the network output. Thus, the combination of different weight parameter values was tested in this study. As shown in *Figure 6*, the effects of λ_1 and λ_2 were quantitatively determined by calculating the average RMSE of the projection data, and the RMSE of different parameter combinations are represented by different square colors. It can be observed from the square colors in *Figure 6* that when $\lambda_1=0.001$ and $\lambda_2=10$, the RMSE of the generated projection dataset is the smallest.

Simulated data results

Four representative projection datasets were used to investigate the performance of the proposed DM-GAN method, and the projection restoration results obtained by different methods on these datasets are illustrated in *Figure 7*. The first two rows and the last two rows represent the projection datasets of different slices. The first and third rows are low-energy projections, and the second and fourth rows are high-energy projections. As for the projection data in the first and second rows of *Figure 7*, the reconstructed DECT images obtained by different methods for the data are shown in *Figure 8*. *Figure 9* shows the profiles of the pixel value line of the gray dotted line in *Figure 8*. *Figure 10* shows the decomposed basis material corresponding to the reconstructed DECT images.

Figure 7 shows the results of projection dataset recovery with different methods, which further evaluate DM-GAN performance. The results indicate that the DM-GAN captures more projection information. Compared with the 180° reference projection dataset, we can intuitively find that the projection dataset produced by DM-GAN has high accuracy. *Figure 8* shows the reconstructed results obtained by different methods. From left to right are the ground truth and reconstructed results based on FBP, TV, GAN + TV, CycleGAN + TV, DM-GAN + TV, and DoDa-Net.

that of U-Net. Then, the projection domain results of DM-GAN and Cycle-GAN could reflect the performance of the improved generator.

Meanwhile, in the image domain, the filtered back projection (FBP) (51) and TV-based methods were used to reconstruct images from a limited-angle projection dataset. Furthermore, the TV-based method was used to reconstruct the images obtained by GAN, CycleGAN, and the proposed method on the processed projection dataset. Then, the reconstruction images were compared. To achieve high-quality DECT imaging, the Im-Net was adopted to perform artifact suppression processing on the reconstruction image obtained by the DM-GAN method. Finally, the material decomposition method was further exploited to obtain the decomposition results of the reconstructed DECT images (52). We set the regularization parameter λ to be the same value at the same experimental background. In the simulation and real experiments, λ was 1.4×10^{-5} and 2.1×10^{-5}, respectively. In addition, the optimization of the decomposition method required approximately 200 iterations under a strict relative error tolerance of 1.0×10^{-12}.

To quantitatively evaluate the performance of the proposed method, the reconstruction images were evaluated by three image quality indicators: peak signal-to-noise ratio (PSNR), structural similarity (SSIM) (53), and root mean square error (RMSE).

Figure 7 The projection restoration results of simulated data obtained by different methods. (A) The reference; (B) limited-angle projection data; (C) estimated projection data by the GAN; (D) estimated projection data by the CycleGAN; (E) estimated projection data by the DM-GAN; (F) error maps of (A,C); (G) error maps of (A,D); (H) error maps of (A,E). "1" and "2" represent different slices. "l" and "h" represent low and high energy, respectively. The display window of (a-e) is [0, 3]. The display window of (f-h) is [–1, 1]. GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network.

Furthermore, regions of interest (ROIs) of the reconstructed images were selected for analysis, as shown in the third and sixth rows in *Figure 8*. It can be observed that the quality of the reconstructed DECT images from the FBP and TV-based methods is relatively poor, and the limited-angle artifacts seriously affect the details of the reconstructed image. The quality of the DECT images reconstructed by the network-based method was improved. This is expected because the FBP and TV-based methods only use 90° projection data, while the network-based method uses the generated 180° projection data. Compared with other network-based methods, the DM-GAN + TV has a better ability to recover image detail. To further optimize the reconstructed image, the Im-Net was employed to process DM-GAN + TV as shown in *Figure 8*. Reconstruction results similar to the ground truth can be obtained based on the proposed method.

Furthermore, it can be seen from the ROIs that the results based on FBP and TV suffer from a large number of limited-angle artifacts. The GAN + TV and CycleGAN + TV methods can eliminate the artifacts to a certain extent and restore the general structure of the image. However, the details of the image shown by the arrow are still blurred. In the reconstruction of CycleGAN + TV, DM-GAN + TV, and DoDa-Net, the sharp edge structure shown by the red arrow is blurred. However, compared with CycleGAN + TV, the DM-GAN + TV and DoDa-Net methods can better restore the image details in the ROI, especially the proposed DoDa-Net method. This indicates that the proposed method has advantages in maintaining microstructure and reducing limited-angle artifacts. The partial line profiles are shown in *Figure 9*, which are drafted from the 80th pixel to the 310th pixel along the gray dotted line in *Figure 8*. Among all the comparisons, the line profile of the proposed method is closest to the ground truth. Also, the area indicated by the arrow in *Figure 9* indicates the accuracy of the proposed method for detailed repair.

Since material decomposition is very sensitive to image artifacts and noise, a small error can yield a large deviation in the final decomposition result. To verify the performance
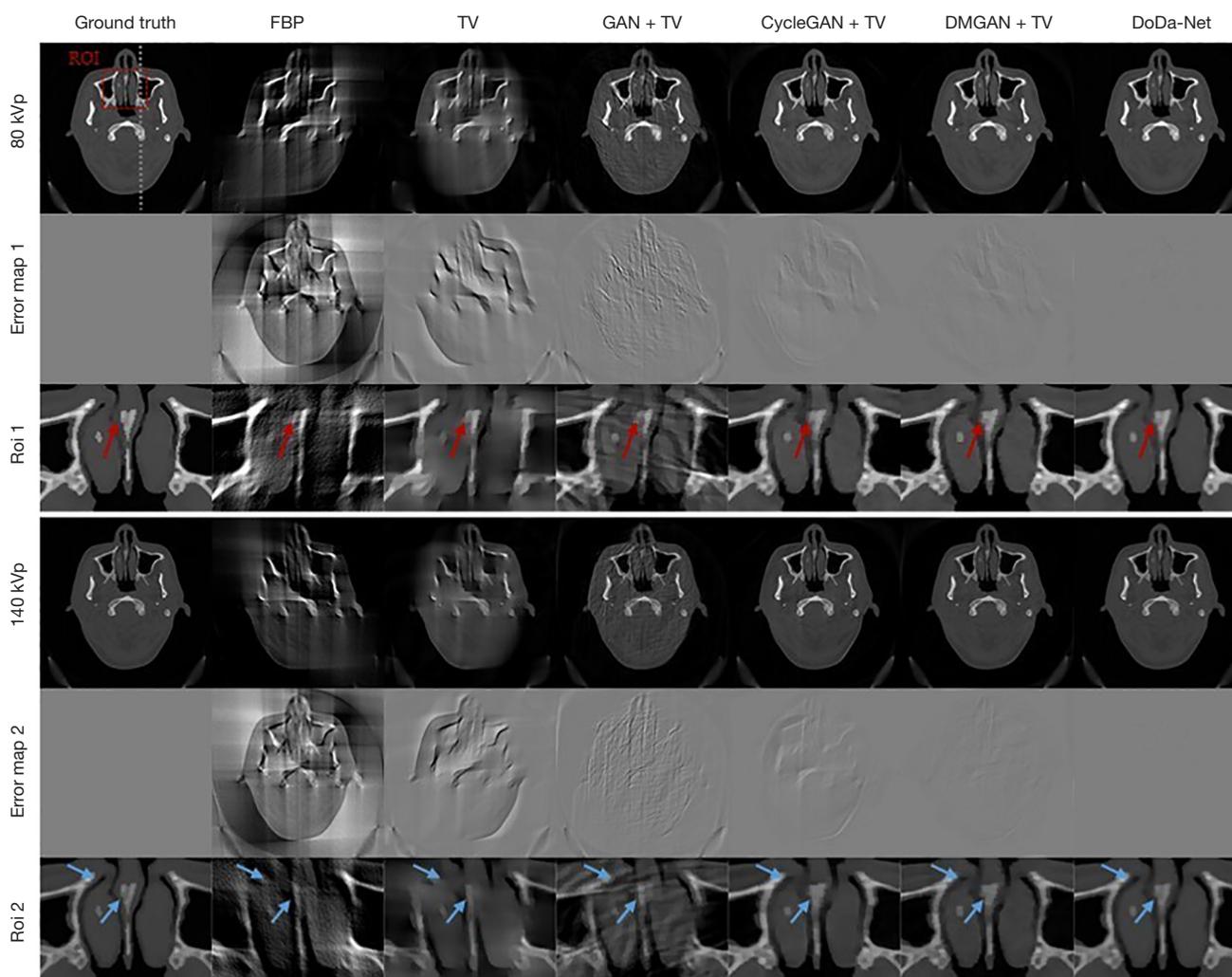
Figure 8 Reconstruction results of the first and second projection data obtained by different methods. The first and fourth rows represent low- and high-energy reconstructed images. The second and fifth rows represent the error maps between the ground truth and the reconstructed image. The third and sixth rows are ROIs of the reconstructed image. The columns from left to right represent the ground truth and the results of FBP, TV, GAN + TV, CycleGAN + TV, DM-GAN + TV, and DoDa-Net. The display windows of reconstructed images and ROIs are [0, 0.07]. The display window of the error maps is [–0.03, 0.03]. ROIs, regions of interest; FBP, filtered back projection reconstruction algorithm; TV, total variation reconstruction algorithm; GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.

of the Im-Net, the reconstructed images of DM-GAN + TV and DoDa-Net were further decomposed to generate the decomposition results of tissue and bone materials, and the results are shown in *Figure 10*. It can be seen that the results of DM-GAN + TV deviate from the ground truth. Specifically, the results of tissue decomposition still retain bone components. In contrast, the basis material decomposed from the reconstructed image by DoDa-Net is

closer to the ground truth, showing higher image quality.

Table 1 shows the quantitative comparison results of reconstructed images. The SSIM of the proposed method reaches 0.9013 and 0.9035 at 80 and 140 kVp, respectively, while the SSIM of other methods is lower than 0.8900. Meanwhile, the RMSEs of the proposed method are 0.0007 and 0.0005 at 80 and 140 kVp, respectively. Compared with the FBP-based method, the proposed method reduces
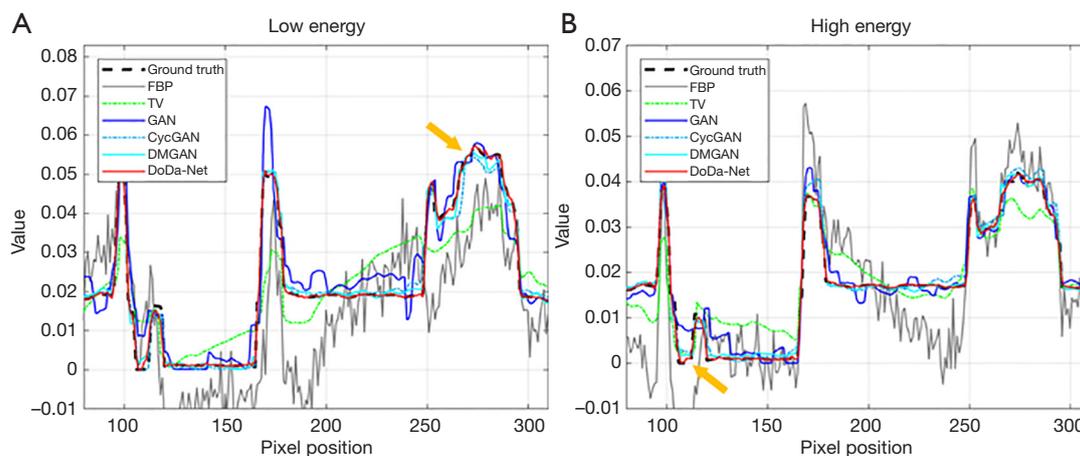
Figure 9 The line profiles of different reconstruction methods. (A) Low energy; (B) high energy. FBP, filtered back projection reconstruction algorithm; TV, total variation reconstruction algorithm; GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.
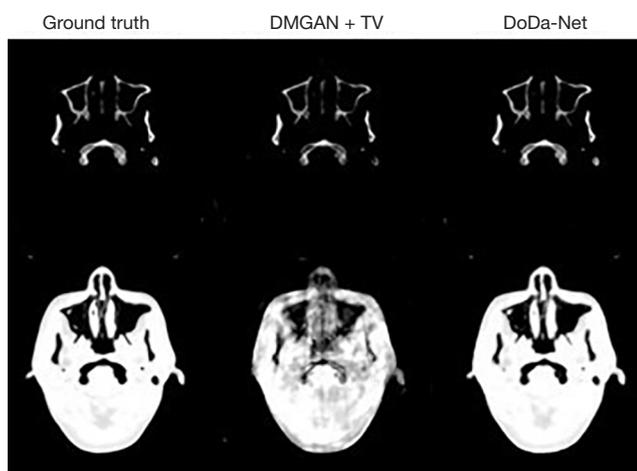
Figure 10 Decomposition results based on the reconstruction images. The first and second rows represent the decomposed bone and tissue material images, respectively. The display window of decomposed materials is [0.1, 1]. TV, total variation reconstruction algorithm; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.

the RMSE by two orders of magnitude. Furthermore, the PSNR of the proposed method is better than that of the other methods, and this result is consistent with the visual analysis of the reconstructed images. In summary, the results of qualitative and quantitative experiments demonstrate the effectiveness of the proposed DoDa-Net method applied to the proposed one half-scan DECT scheme, compared to

other reconstruction methods applied to this scheme.

Real data results

To verify the reliability of the proposed method in practical applications, the method was applied to the real data experiment of the anthropomorphic head phantom. Four test slices were selected to show the projection recovery results exhibited in *Figure 11*. *Figure 12* shows the different methods to reconstruct the DECT images of the first two rows in *Figure 11*. Moreover, we took the full-scan reconstruction result as ground truth. We also selected ROIs to evaluate the performance of different methods to reconstruct image details.

The projection datasets estimated by different methods and their error maps are shown in *Figure 11*, wherefrom left to right are the reference results, limited-scan projection datasets, and the recovery by GAN, CycleGAN, and DM-GAN. It can be seen that the results of GAN and CycleGAN have larger errors, and the DM-GAN method obtains the smallest error. Meanwhile, the error map is further verified, and the projection data obtained by the DM-GAN method is most similar to the reference. The reconstruction results of different methods on the real data are illustrated in *Figure 12*. The first and fourth rows present the reconstructed images, and the second and fifth rows present the error maps between the reconstructed image and the ground truth. It can be observed that in the error maps, the reconstruction results of DM-GAN + TV

666

Wang et al. DECT reconstruction using deep learning model

Table 1 Quantitative results of reconstructed images obtained by different methods on the simulated data (80 testing images)

Results	Metrics	FBP	TV	GAN + TV	CycleGAN + TV	DMGAN + TV	DoDa-Net
80 kVp	Avg. PSNR	19.1082	24.5278	30.8014	33.3288	36.3151	40.3899
	Avg. SSIM	0.4508	0.7771	0.7965	0.8248	0.8814	0.9013
	Avg. RMSE	0.0131	0.0044	0.0023	0.0017	0.0012	0.0007
140 kVp	Avg. PSNR	19.8774	25.9555	29.4841	33.7371	36.5195	40.5573
	Avg. SSIM	0.4914	0.7838	0.7916	0.8388	0.8768	0.9035
	Avg. RMSE	0.0107	0.0033	0.0020	0.0012	0.0010	0.0005

Quantitative testing of simulated data. PSNR, peak signal-to-noise ratio; SSIM, structural similarity; RMSE, root mean square error. FBP, filtered back projection reconstruction algorithm; TV, total variation reconstruction algorithm; GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.
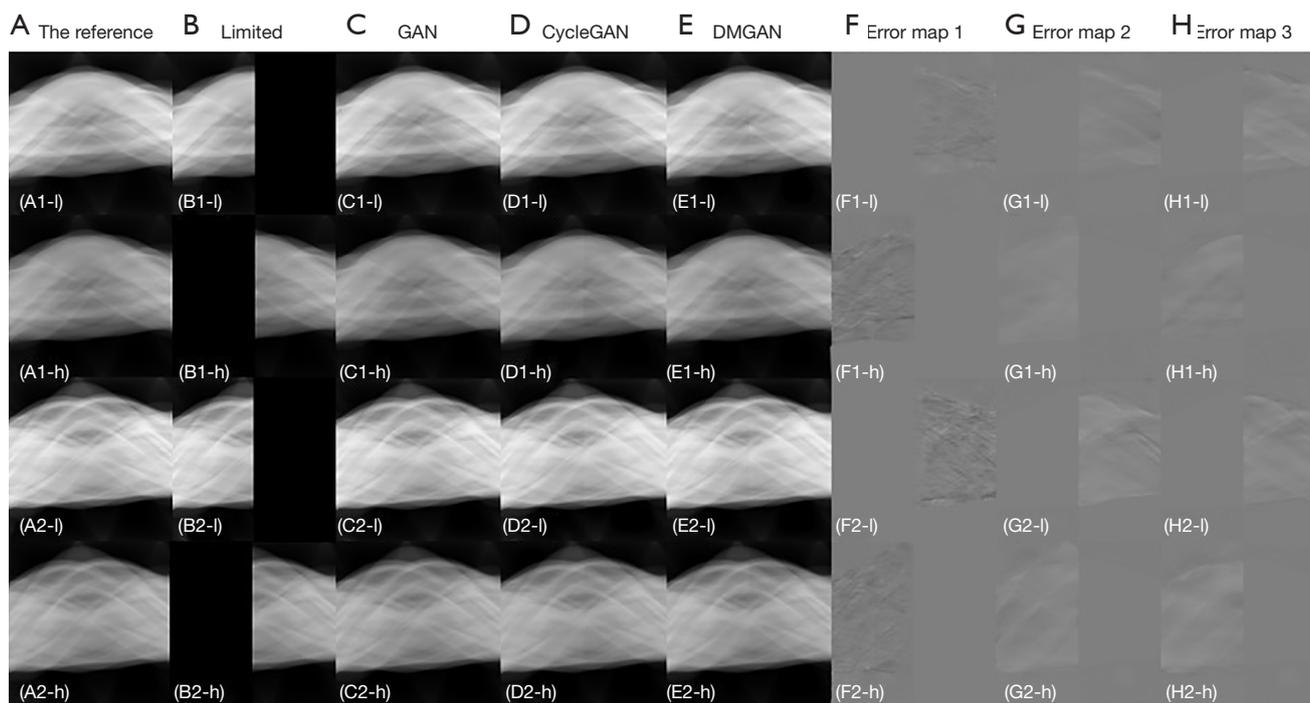
Figure 11 The projection restoration results of real data obtained by different methods. (A) The reference; (B) limited-angle projection data; (C) estimated projection data by the GAN; (D) estimated projection data by the CycleGAN; (E) estimated projection data by the DM-GAN; (F) error maps of (A,C); (G) error maps of (A,D); (H) error maps of (A,E). "1" and "2" represent different slices. "l" and "h" represent low and high energy, respectively. The display window of (A-E) is [0, 1]. The display window of (f-h) is [–0.2, 0.2]. GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network.

and DoDa-Net are closest to the ground truth, especially the latter. To better analyze the restoration of image details, the enlarged ROIs are shown in the third and sixth rows of *Figure 12*. The FBP and TV-based methods fail to reconstruct the detailed structure of the image from the limited-angle projection data. The GAN + TV and CycleGAN + TV methods can eliminate the limited-angle artifacts to a certain extent, but they still fail to handle the fine areas. Compared with the previous methods, DM-GAN + TV and DoDa-Net can suppress the limited-angle artifacts more effectively and obtain better reconstruction results. In particular, the proposed DoDa-Net method can accurately
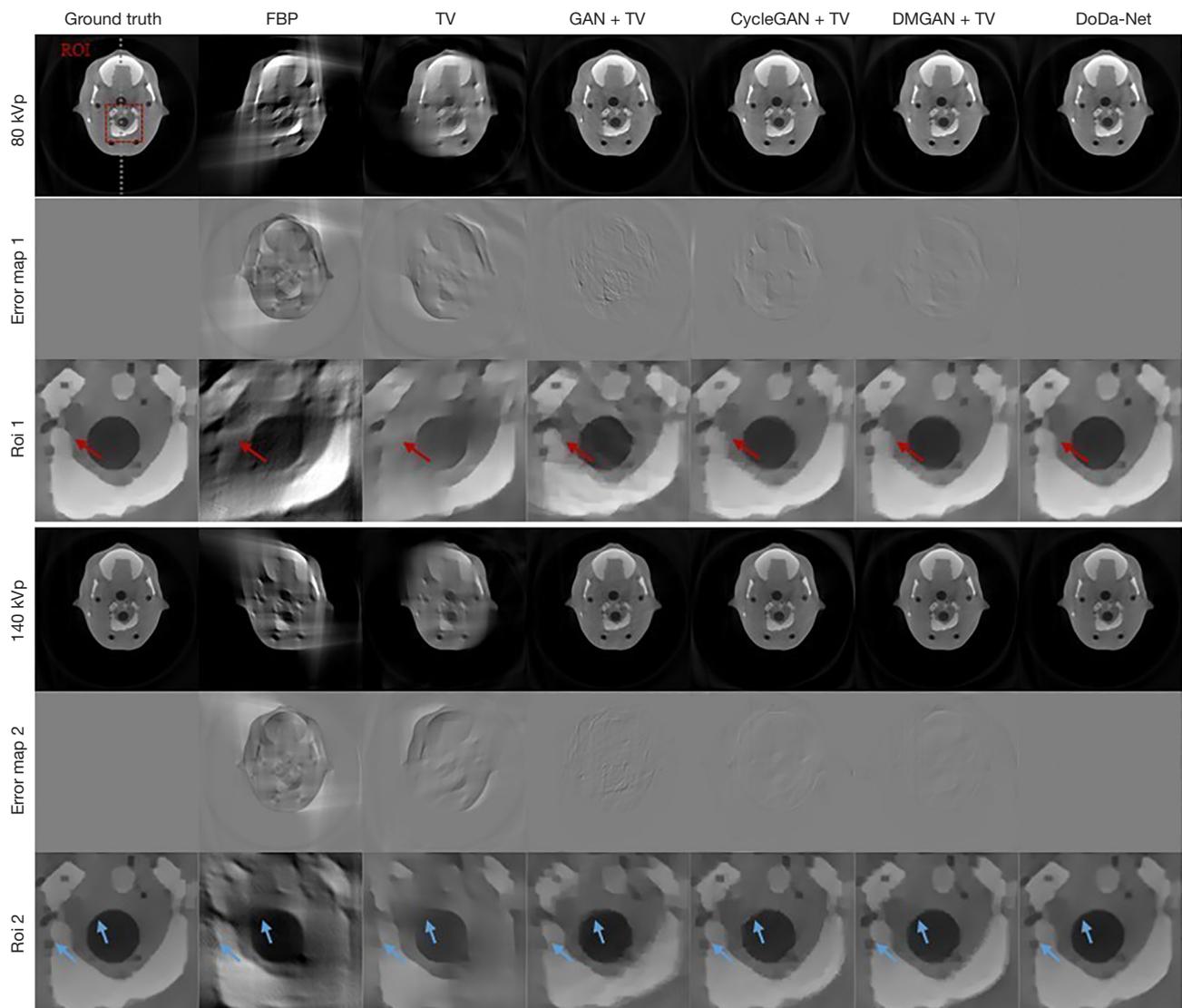
Figure 12 Reconstruction results of the first and second projection data with different methods. The first and fourth rows represent low- and high-energy reconstructed images. The second and fifth rows represent the error maps between the ground truth and the reconstructed image. The third and sixth rows are ROIs of the reconstructed image. Left to right columns represent the ground truth and the results of FBP, TV, GAN + TV, CycleGAN + TV, DM-GAN + TV, and DoDa-Net. The display windows of reconstructed images and ROIs are [0, 0.07]. The display window of error maps is [–0.03, 0.03]. ROIs, regions of interest. FBP, filtered back projection reconstruction algorithm; TV, total variation reconstruction algorithm; GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.

restore the edge and detail area, as shown by the arrow in *Figure 12*. The partial line profiles of the gray dotted line from the 60th pixel to the 335th pixel in *Figure 12* are shown in *Figure 13*. The profile provided by the proposed method is the closest to the ground truth. Also, in the region with a complex structure (area indicated by

arrow), the proposed method obtains more accurate pixel values than other methods.

To verify the effectiveness of the Im-Net, the decomposition results of the reconstructed images in *Figure 12* are illustrated in *Figure 14*. The decomposition result obtained by DM-GAN + TV fails to preserve the internal
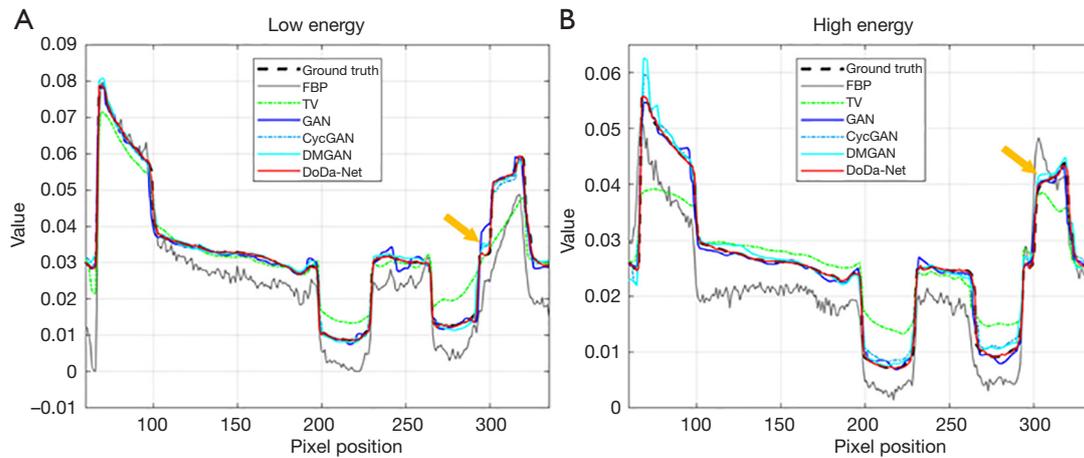
Figure 13 The line profiles of different reconstruction methods. (A) Low energy; (B) high energy. FBP, filtered back projection reconstruction algorithm; TV, total variation reconstruction algorithm; GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.
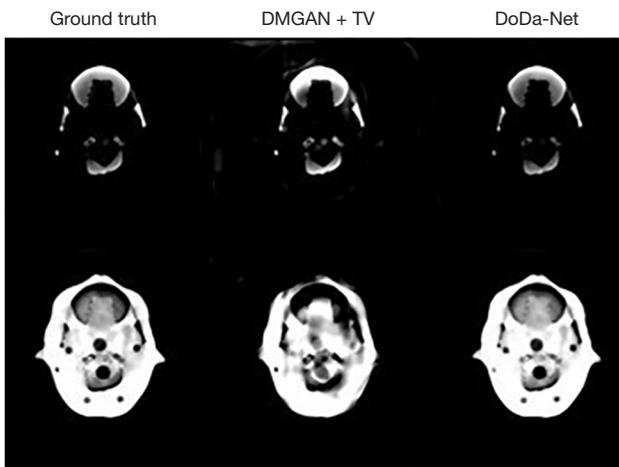
Figure 14 Decomposition results of the reconstruction images. The first and second rows represent the decomposed bone and tissue material images, respectively. The display window of decomposed materials is [0.1, 1]. TV, total variation reconstruction algorithm; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.

images are listed in *Table 2*. Among the methods, DoDa-Net performs best in terms of SSIM, indicating that the results obtained by the proposed method are most similar to the ground truth in structure. Also, the PSNR of the proposed method is higher than 41 dB, while those of the other methods are lower than 36 dB. Furthermore, the RMSE indicates that the proposed method has superior noise suppression ability. The real data experimental results further verify the effectiveness of the proposed method for the one half-scan DECT scheme, which is consistent with the evaluation results of the simulation data.

Ablation study

In order to study the influence of the projection domain processing network and the image domain processing network in the proposed DoDa-Net model, we gradually modified the baseline model and compared their differences. Thirty samples were randomly selected from the test dataset of simulation data to assess the performance of different models. We verified the reconstruction results from both PSNR and RMSE to quantitatively evaluate the effects of the main modifications in the proposed DoDa-Net model. The overall quantitative comparison is shown in *Table 3*. The 1st to 4th represent different models, and their configurations are shown in the second to third rows. The 2nd and 3rd show the performance of image domain and projection domain processing networks separately. The improved effect is more obvious in the model of the

structure and edge information and erroneously decomposes the bone and tissue material images. However, the DoDa-Net method obtains better results and maintains a clearer image edge and internal structure. The region of different substrates can be determined through the decomposition results of DoDa-Net.

The quantitative comparison results of the reconstructed

Table 2 Quantitative results of reconstructed images obtained by different methods on the real data (25 testing images)

Results	Metrics	FBP	TV	GAN + TV	CycleGAN + TV	DMGAN + TV	DoDa-Net
80 kVp	Avg. PSNR	20.9740	24.8941	32.8195	34.3788	34.6581	41.1025
	Avg. SSIM	0.4792	0.7112	0.8079	0.8564	0.8704	0.8983
	Avg. RMSE	0.0094	0.0044	0.0022	0.0015	0.0015	0.0008
140 kVp	Avg. PSNR	22.1091	24.9711	33.1150	34.3013	35.8911	41.2132
	Avg. SSIM	0.5102	0.7288	0.8102	0.8479	0.8742	0.8991
	Avg. RMSE	0.0076	0.0038	0.0020	0.0017	0.0010	0.0005

Quantitative testing of real data. PSNR, peak signal-to-noise ratio; SSIM, structural similarity; RMSE, root mean square error; FBP, filtered back projection reconstruction algorithm; TV, total variation reconstruction algorithm; GAN, generative adversarial network; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.

Table 3 Effects of each major modification in the proposed DoDa-Net model

Results	1st	2nd	3rd	4th
Projection domain processing?	×	×	√	√
Image domain processing?	×	√	×	√
Avg. PSNR (80 kVp)	24.5131	25.2111	36.3011	40.3871
Avg. RMSE	0.0047	0.0036	0.0014	0.0008
Avg. PSNR (140 kVp)	25.9566	26.3302	36.5177	40.5561
Avg. RMSE	0.0035	0.0030	0.0010	0.0005

Ablation test. PSNR, peak signal-to-noise ratio; RMSE, root mean square error.

projection domain processing network. From the 2nd to 4th in *Table 3*, we can observe that the projection domain processing network and the image domain processing network help suppress noise and maintain structure.

Discussion and conclusion

The DECT reconstruction theory usually requires two full-scan projection datasets with different X-ray energies. However, this requirement cannot be met, especially when a full circular rotation is disabled by a physical blockage. Thus, the study of the limited-angle scan DECT scheme is conducive to broadening its application scenarios. To this end, this study designed a novel flexible DECT imaging scheme (i.e., one half-scan DECT scheme) to reduce the radiation dose and simplify the imaging scheme. Currently, the relationship between different energy data is not well studied for the limited-angle scan DECT problem. Also, the utilization of dual-domain information is not sufficient. In the process of studying this problem, it was found that the application of a joint dual-domain method to DECT image

reconstruction is a feasible way to improve the effective information of DECT images. Thus, the DoDa-Net method was proposed to solve the problem. First, inspired by the CycleGAN for image conversion, the DM-GAN is designed to explore the relationship between different energy projection data to obtain two half-scan projection datasets. The TV-based method was exploited to realize the conversion of data from the projection domain to the image domain. However, the half-scan projection data obtained through DM-GAN is limited because the angle coverage requirement of 180° + fan angle is a theoretical derivation and is theoretically guaranteed. Therefore, the Im-Net was further employed to eliminate image artifacts caused by minor inconsistencies in the DM-GAN restoration projection.

The experimental results show that the proposed method is better than other comparison methods. Moreover, the experimental results also show the main contributions of different technologies in DoDa-Net. Among them, the main contribution of DM-GAN lies in the restoration of projection data. The comparison between the results of

DM-GAN and CycleGAN in the projection domain results mainly shows the ability of the improved CycleGAN to recover projection data. The main contribution of the Im-Net lies in image restoration. The comparison of DM-GAN + TV and DoDa-Net in the image reconstruction results and the material decomposition results shows the ability of the Im-Net to perform image processing. The experimental results show that the proposed method is better than other comparison methods.

As for the proposed DoDa-Net model, proper hyperparameters can improve the output quality of the network to a certain extent. Especially in the DM-GAN, the different weights of the total loss function significantly impact network training. In Section "Study of parameters", a large number of experiments were conducted on combinations of different weights. By adjusting the weights to make the different losses relatively balanced, the generated projection dataset can have higher accuracy. Also, the two-stage network structure contained in DoDa-Net was studied. In stage one, the generator of the DM-GAN consists of three parts: encoder, residual module, and decoder. The Conv layer between the encoder and decoder in U-Net is replaced by a residual module, which utilizes the ResNet to realize the accurate mapping between the features of the source domain and the target domain.

Meanwhile, the two generators in the DM-GAN realize dual-way mapping (G realizes the mapping from low- to high-energy, F is the opposite), which can continuously optimize the network parameters and improve the quality of network training. In stage two, the main purpose of the Im-Net is to eliminate the artifacts of the estimated image, and the U-Net structure is used to suppress the artifacts of the input image. Since U-Net has skip-connections, it can extract the multi-scale features of the input image and ensure high-quality image restoration. The experimental results indicate that the DM-GAN can restore the projection dataset with high accuracy, and the Im-Net can suppress the image artifacts to obtain high-quality reconstructed images. In addition, the proposed method was tested with fewer-angle data, using low- and high-energy projection data of less than 90° to obtain the corresponding projection data of the sum of the two by DM-GAN. The projection data obtained in this way has serious artifacts in the image reconstructed by the TV-based method. Although the work of this manuscript also uses the Im-Net to eliminate the difference between the reconstructed image of 180° projection data and the ground truth, the effect of using Im-Net to process the reconstructed image of less than 180° projection data is not good. *Figure 15* shows the results of 120°, 140°, 160°, and 180° high-energy projection data reconstruction using the DM-GAN + TV and DoDa-Net methods. We need to retrain the Im-Net model to obtain better results. It may also be necessary to adjust the Im-Net structure.

In the process of network training, it is observed that the test results are not satisfactory when simulated datasets are used for training and testing with real data in the experiment. This result may be caused by the rule difference between the simulated and real projection data distribution. Thus, it is necessary to use the projection dataset collected by the real CT system to conduct additional training for improving network performance. Besides, there are differences in the data distribution among different real CT imaging systems. Therefore, multiple models can be trained for different real CT imaging systems for practical applications.

Furthermore, different objects can be scanned to expand the dataset and make it suitable for a specific real CT system. With the increase of projection datasets, the robustness of the trained network model can be enhanced. The proposed method requires training with paired data, i.e., well-matching projections measured for two energies. For moving objects such as humans, paired ground truth data acquisition is difficult, so one may need to scan static phantoms. It is unclear whether the learned method can generalize between different scanned objects (acquired on the same scanner), such as from phantom scans to human scans.

Although the proposed method has achieved convincing reconstruction results, some limitations still exist. First, the proposed DoDa-Net method needs a long training time due to the concatenation of two networks in this model. Also, for DM-GAN + TV reconstructed images of different projection data, the Im-Net model needs to be retrained. In the future, we hope to study a more efficient architecture to solve this problem. The Im-Net model can be adaptively trained to obtain good results for different limited-angle projection data. Second, there is no data transfer between the two networks in DoDa-Net. To solve this problem, future work will explore the data transposition module to realize a bidirectional transmission of information between the projection domain and the image domain. The improved network model can further enhance the reconstruction details. Third, the work of this paper adopts a supervised training strategy. As an outlook, adapting the network to other data with unsupervised training strategies can be considered.
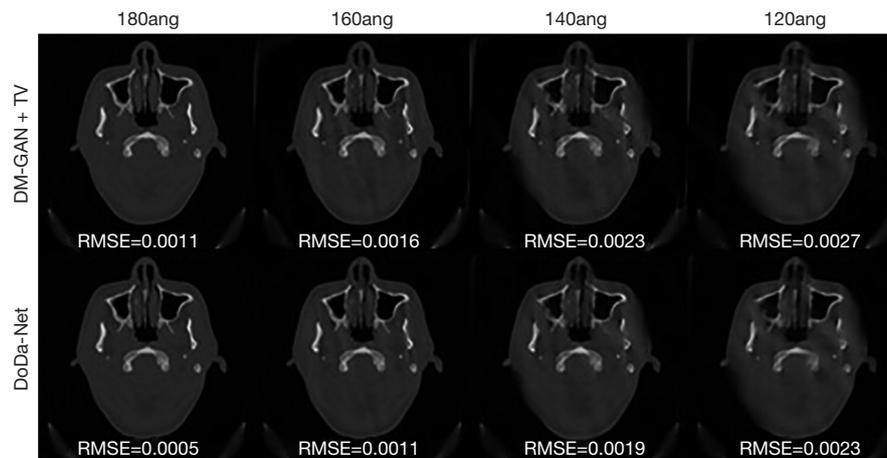
Figure 15 The reconstruction results with different angle data. The RMSEs of the reconstruction results are denoted at the bottom of the reconstructed images. The display windows of reconstructed images are [0, 0.07]. RMSE, root mean square error; TV, total variation reconstruction algorithm; DMGAN, dual-way mapping generative adversarial network; DoDa-Net, dual-domain dual-way estimated network.

In conclusion, this study designed a one-half-scan DECT scheme and proposed an effective method to obtain high-quality DECT images by utilizing dual-way learning of the projection data and optimizing the image domain data. The proposed method effectively broadens the application of the DECT imaging system and has great potential in reducing the X-ray radiation dose and hardware cost.

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at https://dx.doi.org/10.21037/qims-21-441

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/qims-21-441). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of Henan Provincial People's Hospital, and written informed consent was obtained.

References

1. Yu Z, Leng S, Li Z, McCollough CH. Spectral prior image constrained compressed sensing (spectral PICCS) for photon-counting computed tomography. Phys Med Biol 2016;61:6707-32.
2. Kalender WA, Perman WH, Vetter JR, Klotz E. Evaluation of a prototype dual-energy computed tomographic apparatus. I. Phantom studies. Med Phys 1986;13:334-9.
3. McCollough CH, Leng S, Yu L, Fletcher JG. Dual- and

672

Wang et al. DECT reconstruction using deep learning model

Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications. Radiology 2015;276:637-53.

4. Yang X, Sun W, Huang D, Li H, Zhao Y, Li P, Liu Y. Quantitative spectral CT evaluation of kidney tumors with the stretched-exponential nonlinear regression analysis model. Quant Imaging Med Surg 2021;11:676-84.

5. Rebuffel V, Dinten JM. Dual-energy X-ray imaging: benefits and limits. Insight: Non-Destructive Testing & Condition Monitoring 2007;49:589-94.

6. Ying Z, Naidu R, Crawford CR. Dual energy computed tomography for explosive detection. J Xray Sci Technol 2006;14:235-56.

7. Pelc N, McCollough C, Yu L, Schmidt T. We-e-18c-01: Multi-energy ct: Current status and recent innovations. Med Phys 2014;41:512-3.

8. Johnson TR. Dual-energy CT: general principles. AJR Am J Roentgenol 2012;199:S3-8.

9. Long Y, Fessler JA. Multi-material decomposition using statistical image reconstruction for spectral CT. IEEE Trans Med Imaging 2014;33:1614-26.

10. De Man B, Nuyts J, Dupont P, Marchal G, Suetens P. An iterative maximum-likelihood polychromatic algorithm for CT. IEEE Trans Med Imaging 2001;20:999-1008.

11. Alvarez RE, Macovski A. Energy-selective reconstructions in X-ray computerized tomography. Phys Med Biol 1976;21:733-44.

12. Zhao X, Chen P, Wei J, Qu Z. Spectral CT imaging method based on blind separation of polychromatic projections with Poisson prior. Opt Express 2020;28:12780-94.

13. Xue Y, Jiang YK, Yang CL, Lyu QH, Wang J, Luo C, Zhang LH, Desrosiers C, Feng K, Sun SN, Hu XH, Sheng K, Niu TY. Accurate multi-material decomposition in dual-energy CT: A phantom study. IEEE Trans Comput Imaging 2019;5:515-29.

14. Jiang Y, Xue Y, Lyu Q, Xu L, Luo C, Yang P, Yang C, Wang J, Hu X, Zhang X, Sheng K, Niu T. Noise Suppression in Image-Domain Multi-Material Decomposition for Dual-Energy CT. IEEE Trans Biomed Eng 2020;67:523-35.

15. Zhang W, Liang N, Wang Z, Cai A, Wang L, Tang C, Zheng Z, Li L, Yan B, Hu G. Multi-energy CT reconstruction using tensor nonlocal similarity and spatial sparsity regularization. Quant Imaging Med Surg 2020;10:1940-60.

16. Dong X, Niu T, Zhu L. Combined iterative reconstruction and image-domain decomposition for dual energy CT using total-variation regularization. Med Phys

2014;41:051909.

17. Wang T, Zhu L. Dual energy CT with one full scan and a second sparse-view scan using structure preserving iterative reconstruction (SPIR). Phys Med Biol 2016;61:6684-706.

18. Chen B, Zhang Z, Sidky EY, Xia D, Pan X. Image reconstruction and scan configurations enabled by optimization-based algorithms in multispectral CT. Phys Med Biol 2017;62:8763-93.

19. Shen L, Xing Y. Multienergy CT acquisition and reconstruction with a stepped tube potential scan. Med Phys 2015;42:282-96.

20. Petrongolo M, Zhu L. Single-Scan Dual-Energy CT Using Primary Modulation. IEEE Trans Med Imaging 2018;37:1799-808.

21. Zhang H, Xing YX. Limited-angle multi-energy CT using joint clustering prior and sparsity regularization. Spie Medical Imaging 2016.

22. Zhang W, Wang L, Li L, Niu T, Li Z, Liang N, Xue Y, Yan B, Hu G. Reconstruction method for DECT with one half-scan plus a second limited-angle scan using prior knowledge of complementary support set (Pri-CSS). Phys Med Biol 2020;65:025005.

23. Deng K, Sun C, Liu Y, Yang HW. Real-Time Limited-View CT Inpainting and Reconstruction with Dual Domain Based on Spatial Information. 2021. Available online: https://arxiv.org/abs/2101.07594

24. Li Z, Cai A, Wang L, Zhang W, Tang C, Li L, Liang N, Yan B. Promising Generative Adversarial Network Based Sinogram Inpainting Method for Ultra-Limited-Angle Computed Tomography Imaging. Sensors (Basel) 2019;19:3941.

25. Anirudh R, Kim H, Thiagarajan JJ, Mohan KA, Champley KM, Bremer T. Lose the views: Limited angle CT reconstruction via implicit sinogram completion. Available online: https://ieeexplore.ieee.org/document/8578762

26. Wozniak M, Sika J, Wieczorek M. Deep neural network correlation learning mechanism for CT brain tumor detection. Available online: https://link.springer.com/article/10.1007/s00521-021-05841-x

27. Wozniak M, Wieczorek M, Silka J, Polap D. Body pose prediction based on motion sensor data and Recurrent Neural Network. Available online: https://ieeexplore.ieee.org/document/9165920

28. Poirot MG, Bergmans RHJ, Thomson BR, Jolink FC, Moum SJ, Gonzalez RG, Lev MH, Tan CO, Gupta R. Physics-informed Deep Learning for Dual-Energy Computed Tomography Image Processing. Sci Rep

2019;9:17709.

29. Xu Y, Yan B, Zhang J, Chen J, Zeng L, Wang L. Image Decomposition Algorithm for Dual-Energy Computed Tomography via Fully Convolutional Network. Comput Math Methods Med 2018;2018:2527516.

30. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Available online: https://link.springer.com/chapt er/10.1007/978-3-319-24574-4_28

31. Clark DP, Holbrook M, Badea CT. Multi-energy CT decomposition using convolutional neural networks. Available online: https://www.spiedigitallibrary.org/ conference-proceedings-of-spie/10573/105731O/Multi-energy-CT-decomposition-using-convolutional-neural-networks/10.1117/12.2293728.short

32. Zhang W, Zhang H, Wang L, Wang X, Hu X, Cai A, Li L, Niu T, Yan B. Image domain dual material decomposition for dual-energy CT using butterfly network. Med Phys 2019;46:2037-51.

33. Wang Y, Zhang W, Cai A, Wang L, Tang C, Feng Z, Li L, Liang N, Yan B. An effective sinogram inpainting for complementary limited-angle dual-energy computed tomography imaging using generative adversarial networks. J Xray Sci Technol 2021;29:37-61.

34. Zhang YK, Lv TL, Ge RJ, Zhao QL, Hu DL, Zhang L, Liu J, Zhang Y, Liu QG, Zhao W, Chen Y. CD-Net: Comprehensive Domain Network with Spectral Complementary for DECT Sparse-View Reconstruction. IEEE Trans Comput Imaging 2021;7:436-47.

35. Zhu J, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. Available online: https://ieeexplore.ieee.org/document/8237506

36. Yuan Y, Liu S, Zhang J, Zhang Y, Dong C, Lin L. Unsupervised Image Super-Resolution Using Cycle-in-Cycle Generative Adversarial Networks. Available online: https://openaccess.thecvf.com/content_cvpr_2018_workshops/papers/w13/Yuan_Unsupervised_Image_Super-Resolution_CVPR_2018_paper.pdf

37. Yu H, Wang G. Compressed sensing based interior tomography. Phys Med Biol 2009;54:2791-805.

38. Zhang HY, Xing YX. Reconstruction of limited-angle dual-energy CT using mutual learning and cross-estimation (MLCE). Available online: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9783/978344/Reconstruction-of-limited-angle-dual-energy-CT-using-mutual-learning/10.1117/12.2211224.short

39. Ibrahim DM, Elshennawy NM, Sarhan AM. Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. Comput Biol Med 2021;132:104348.

40. Kang E, Min J, Ye JC. Wavelet Domain Residual Network (WavResNet) for Low-Dose X-ray CT Reconstruction. arXiv preprint arXiv:1703.01383, 2017.

41. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Available online: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf

42. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Ben-gio Y. Generative adversarial nets. In NIPS 2014. Available online: https://papers.nips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf

43. Goodfellow I. NIPS 2016 Tutorial: Generative Adversarial Networks. arXiv preprint arXiv:1701.00160, 2016.

44. Qiao T, Zhang W, Zhang M, Ma Z, Xu D. Ancient Painting to Natural Image: A New Solution for Painting Processing. arXiv 2019; DOI: 10.1109/WACV.2019.00061.

45. Chen B, Zhang Z, Xia D, Sidky EY, Pan X. Algorithm-enabled partial-angular-scan configurations for dual-energy CT. Med Phys 2018;45:1857-70.

46. Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. Med Phys 1985;12:252-5.

47. Li Y, Li K, Zhang C, Montoya J, Chen GH. Learning to Reconstruct Computed Tomography Images Directly From Sinogram Data Under A Variety of Data Acquisition Conditions. IEEE Trans Med Imaging 2019;38:2469-81.

48. ICRU. Phantoms and Computational Models in Therapy, Diagnosis and Protcction. Report No 48; ICRU: Bethesda 1992.

49. Kingma D, Ba J. Adam: A method for stochastic optimization. Available online: https://www.researchgate.net/publication/269935079_Adam_A_Method_for_Stochastic_Optimization

50. Zhao J, Chen Q, Zhang L, Jin X. Unsupervised Learnable Sinogram Inpainting Network (SIN) for Limited Angle CT reconstruction. arXiv preprint arXiv:1811.03911, 2018.

51. Turbell H. Cone-Beam Reconstruction Using Filtered Backprojection. Available online: http://people.csail.mit.edu/bkph/courses/papers/Exact_Conebeam/Turbell_Thesis_FBP_2001.pdf

52. Niu T, Dong X, Petrongolo M, Zhu L. Iterative image-domain decomposition for dual-energy CT. Med Phys 2014;41:041901.

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