

PAPER

## DaNet: dose-aware network embedded with dose-level estimation for low-dose CT imaging

To cite this article: Zhenxing Huang *et al* 2021 *Phys. Med. Biol.* **66** 015005

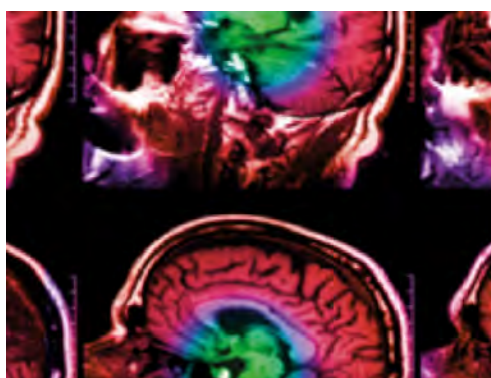
View the [article online](#) for updates and enhancements.

### You may also like

- [Improved discrimination between benign and malignant LDCT screening-detected lung nodules with dynamic over static  \$^{18}\text{F}\$ -FDG PET as a function of injected dose](#)  
Qing Ye, Jing Wu, Yihuan Lu *et al.*
- [Discriminative feature representation: an effective postprocessing solution to low dose CT imaging](#)  
Yang Chen, Jin Liu, Yining Hu *et al.*
- [Low-dose CT with deep learning regularization via proximal forward-backward splitting](#)  
Qiaoqiao Ding, Gaoyu Chen, Xiaoqun Zhang *et al.*

### Recent citations

- [Contrast-enhanced to noncontrast CT transformation via an adjacency content-transfer-based deep subtraction residual neural network](#)  
Xianfan Gu *et al*
- [Correcting motion artifacts in coronary computed tomography angiography images using a dual-zone cycle generative adversarial network](#)  
Fuquan Deng *et al*



**IPEM | IOP**

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics,  
biomedical engineering and related subjects.

Start exploring the collection—download the  
first chapter of every title for free.



## PAPER

## DaNet: dose-aware network embedded with dose-level estimation for low-dose CT imaging

RECEIVED

10 July 2020

REVISED

20 October 2020

ACCEPTED FOR PUBLICATION

29 October 2020

PUBLISHED

8 January 2021

Zhenxing Huang<sup>1,2,3,4</sup> , Zixiang Chen<sup>4</sup> , Jincai Chen<sup>1,2,3,\*</sup> , Ping Lu<sup>1,2,3</sup>, Guotao Quan<sup>6</sup>, Yanfeng Du<sup>6</sup>, Chenwei Li<sup>6</sup>, Zheng Gu<sup>7</sup>, Yongfeng Yang<sup>4,5</sup>, Xin Liu<sup>4,5</sup>, Hairong Zheng<sup>4,5</sup>, Dong Liang<sup>4,5</sup> , and Zhanli Hu<sup>4,5,\*</sup> <sup>1</sup> Wuhan National Laboratory for Optoelectronics, Huazhong University of Science & Technology, Wuhan 430074, People's Republic of China<sup>2</sup> School of Computer Science & Technology, Huazhong University of Science & Technology, Wuhan 430074, People's Republic of China<sup>3</sup> Key Laboratory of Information Storage System, Engineering Research Center of Data Storage Systems and Technology, Ministry of Education of China, Wuhan 430074, People's Republic of China<sup>4</sup> Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, People's Republic of China<sup>5</sup> Chinese Academy of Sciences Key Laboratory of Health Informatics, Shenzhen 518055, People's Republic of China<sup>6</sup> Shanghai United Imaging Healthcare, Shanghai 201807, People's Republic of China<sup>7</sup> Institute of Biomedical Engineering, Shenzhen Bay Laboratory, Shenzhen 518107, People's Republic of China

\* Author to whom any correspondence should be addressed.

E-mail: [zl.hu@siat.ac.cn](mailto:zl.hu@siat.ac.cn) and [jcchen@hust.edu.cn](mailto:jcchen@hust.edu.cn)**Keywords:** low-dose CT, image restoration, dose-level estimation, channel feature transform**Abstract**

Many deep learning (DL)-based image restoration methods for low-dose CT (LDCT) problems directly employ the end-to-end networks on low-dose training data without considering dose differences. However, the radiation dose difference has a great impact on the ultimate results, and lower doses increase the difficulty of restoration. Moreover, there is increasing demand to design and estimate acceptable scanning doses for patients in clinical practice, necessitating dose-aware networks embedded with adaptive dose estimation. In this paper, we consider these dose differences of input LDCT images and propose an adaptive dose-aware network. First, considering a large dose distribution range for simulation convenience, we coarsely define five dose levels in advance as lowest, lower, mild, higher and highest radiation dose levels. Instead of directly building the end-to-end mapping function between LDCT images and high-dose CT counterparts, the dose level is primarily estimated in the first stage. In the second stage, the adaptively learned low-dose level is used to guide the image restoration process as the pattern of prior information through the channel feature transform. We conduct experiments on a simulated dataset based on original high dose parts of American Association of Physicists in Medicine challenge datasets from the Mayo Clinic. Ablation studies validate the effectiveness of the dose-level estimation, and the experimental results show that our method is superior to several other DL-based methods. Specifically, our method provides obviously better performance in terms of the peak signal-to-noise ratio and visual quality reflected in subjective scores. Due to the dual-stage process, our method may suffer limitations under more parameters and coarse dose-level definitions, and thus, further improvements in clinical practical applications with different CT equipment vendors are planned in future work.

**1. Introduction**

The aim of low-dose computed tomography (LDCT) is to estimate normal high-dose CT (HDCT) images from low-dose counterparts to decrease the radiation dose in clinical applications. The wide application of CT in clinics may introduce the potential risk of cancers caused by x-ray exposure, which has aroused much public concern (Brenner and Hall 2007, Pearce *et al* 2012). Under the ALARA (as low as reasonably

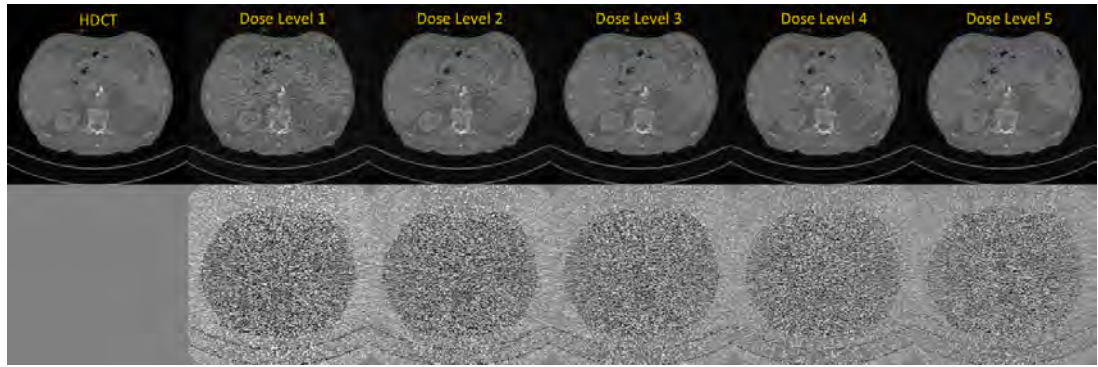
achievable) principle, radiation dose-reduction techniques play an important role in clinical applications (Kalra *et al* 2004, Slovis 2002). Since CT imaging is dominated by photon noise, lower doses of x-ray exposure would result in a degraded image. Thus, the trade-off between image quality and radiation dose needs to be considered in CT planning (Zhang *et al* 2018).

There are two common strategies for lowering the radiation dose: reducing the x-ray exposure photon flux (Chen *et al* 2014, Xu *et al* 2012) or decreasing the projection number of the scanning trajectory (Chen *et al* 2008, Sidky and Pan 2008). In terms of the first strategy, practical operations generally include adjusting the tube voltage or current, which is widely employed in clinical applications but may lead to noisy projections. For the second strategy, sparse-view or limit-angle sampling approaches provide insufficient projection data, which may lead to streaking artifacts (Bian *et al* 2010). In this work, we focus on the first strategy for lowering the radiation dose by considering practical clinical diagnosis. In particular, we apply the degenerative photon flux in data acquisition to obtain different dose-level simulations of low-dose patient data based on original high dose parts of clinical American Association of Physicists in Medicine (AAPM) challenge datasets.

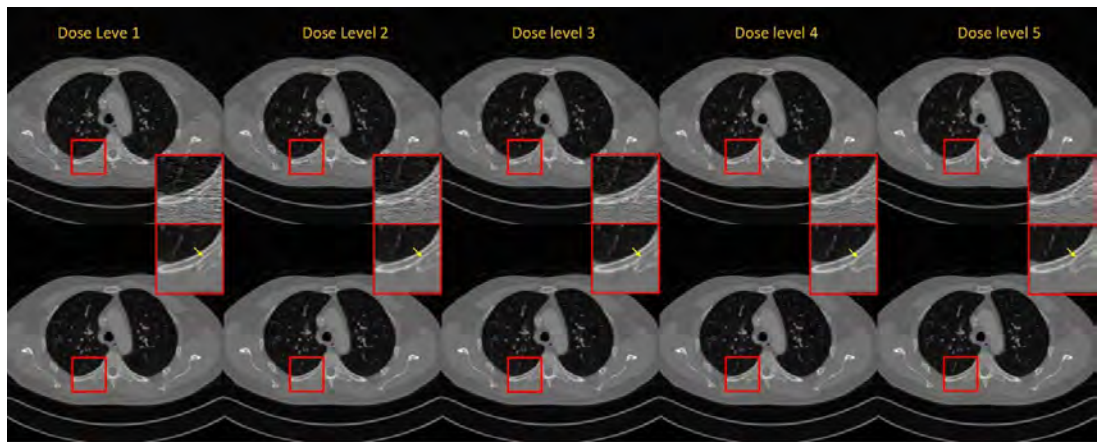
Multiple estimated high-dose CT images are possible for the same input low-dose CT one, producing a typical ill-posed problem. Many methods of low-dose CT imaging have been proposed to solve this ill-posed problem. These methods can generally be grouped into three categories: sinogram restoration methods, iterative reconstruction methods and post-processing methods. Sinogram restoration methods are applied in the sinogram domain to suppress the noise before being input into reconstruction, usually via filtered backprojection (FBP). Typical proposed methods include penalized weighted least-squares (Wang *et al* 2006, Hu *et al* 2016) and bilateral filtering (Manduca *et al* 2009). However, sinogram methods suffer from spatial resolution loss and may introduce extra artifacts in estimated CT images because of the sensitivity to errors caused by sinogram filtering. Iterative reconstruction-based methods tend to optimize a prior regularized objective function. The key insights of these methods are to build image priors. Compressive sensing (Donoho 2006)-based constraints have gradually become a powerful tool for improving CT image reconstruction quality with sparse transforms, such as the total variation (Sidky and Pan 2008, Zhang *et al* 2014, 2016, 2013). Other approaches have tried to explore richer feature knowledge, such as non-local means (NLM) (Chen *et al* 2009, Chen *et al* 2008), dictionary learning (Xu *et al* 2012, Chen *et al* 2014, Hu *et al* 2016), low rank (Cai *et al* 2014, Gao *et al* 2011) and adaptive Markov random fields (Zhang *et al* 2014). Nevertheless, most iterative reconstruction methods are hindered by the intensive computation burden of the projection and back-projection during each iteration. Moreover, the ultimate image quality is highly dependent on the regularized objective function and parameter setting. As the third category, post-processing methods directly apply DICOM-formatted LDCT images without relying on projection data, thus enabling integration in current CT systems. In the past few decades, dictionary learning-based methods have been introduced (Chen *et al* 2013). Li *et al* (2014) introduced the NLM technique to suppress noise for low-dose CT. Block-matching 3D (BM3D), an efficient denoising method in the natural image domain, has also been introduced to estimate high-resolution NDCT images (Feruglio *et al* 2010, Sheng *et al* 2014). Although several post-processing methods provide excellent performance improvements in the image domain, they do not perform well when mottle noise or streak artifacts arise from the back-projection process because such artifacts do not satisfy a special distribution or have strong magnitudes.

More recently, with the rapid development of deep learning (DL) techniques, several methods based on convolutional neural networks (CNNs) have enabled numerous successful applications in medical imaging. Some network architectures have also been proposed for low-dose CT, leading to promising estimation results. Chen *et al* (2017) employed multiple convolutional layers for LDCT images and later proposed residual encoder-decoder convolutional neural networks (REDCNNs) inspired by the autoencoder (AE) (Chen *et al* 2017). Moreover, several U-shaped network architectures (Li and Mueller 2017, Liu and Zhang 2018, Fan *et al* 2019) have been introduced for this task. Kang *et al* (2017) introduced the wavelet transform into CNNs. In addition, several works (Du *et al* 2019, Hu *et al* 2019, Wolterink *et al* 2017, Huang *et al* 2020) based on generative adversarial networks (GANs) have been proposed to produce more structural details. Although these DL-based methods provide excellent performance, many of them recover high-resolution images under a certain radiation dose, which limits their applicability under unknown radiation doses.

In this paper, we rethink the low-dose problem considering the influence of radiation dose differences for image restoration quality and propose an adaptive dose-aware network embedded with dose-level estimation for low-dose CT imaging. The motivation of this work is based on two phenomena: (1) the radiation dose difference greatly impacts the final CT images; (2) a lower dose heightens the restoration challenge. As shown in figure 1, LDCT images from different dose levels are quite different, and the sample with a higher dose level appears closer to the ground truth. As shown in figure 2, a lower radiation dose level generally leads to worse image quality, increasing the complexity of the image restoration process. Table 1 presents five preassigned dose levels and the associated restoration challenge.



**Figure 1.** The first row shows a high-dose CT image and its responding low-dose images under five dose levels. The second row shows the absolute difference results for the different low-dose CT images referenced to the high-dose image. The CT display is the same for the difference outcomes  $[-30, 30]$ .



**Figure 2.** Estimated results on five dose levels. The red boxes denote the regions of interest (ROIs), and the yellow arrows denote the differences. The higher dose level permits the visualization of additional structural details.

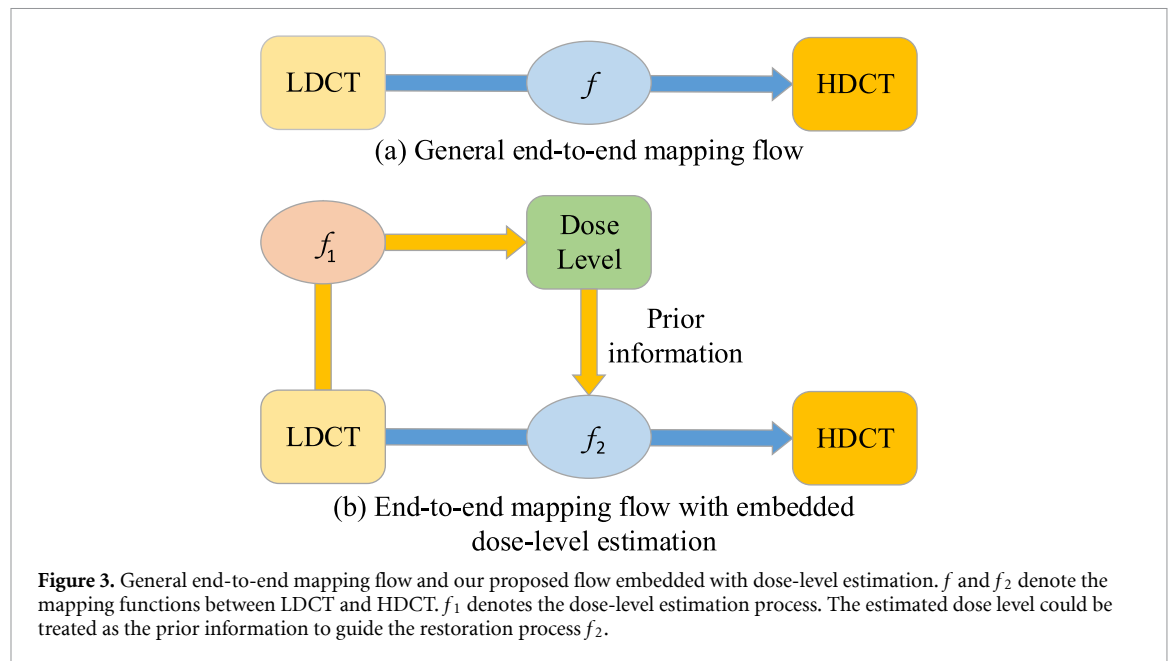
**Table 1.** Illustration of five preassigned dose levels and the associated restoration challenges.

Dose level	Radiation dose	Noise intensity	Restoration challenge
1	Lowest	Strongest	Most hard
2	Lower	Stronger	More hard
3	Mild	Mild	Mild
4	Higher	Slighter	Easier
5	Highest	Slightest	Easiest

Another important consideration is that the CT scan dose may differ among individuals in clinical applications. For instance, the adopted CT scan doses for children and the elderly are generally lower than those for youth. Specially, the routine CT scan radiation dose for adults on head could be 75 mGy, while the dose is set as 34 mGy for children (less than 5 years old). In addition, certain scanning setting specifications are not satisfactory for all practical scenarios, leading to a large radiation dose range in real clinical applications. Due to this dose uncertainty, we divide common doses into multiple levels. In particular, we introduce five coarse preassigned dose levels to consider the influence of dose differences on low-dose CT imaging.

Most DL-based image restoration methods for low-dose CT imaging directly apply the LDCT images from DICOM-formatted files without considering differences in dose level. In this study, we explore the influence of dose-level differences to improve image quality performance and propose an adaptive dose-aware network (DaNet). As shown in figure 3, the significant difference between our end-to-end network compared with the general network flow is that the dose-level estimation task is embedded in our end-to-end network. We summarize the workflow of our method into two stages: the dose-level estimation and the restoration process. First, the dose levels are primarily estimated from unknown input LDCT images. Then, helpful feature information is extracted from the original input samples and estimated dose levels.





Finally, the estimated dose levels are used to guide the image reconstruction process as the pattern of prior information on the case, and we accomplish the fusion of these features through the channel feature transform (CFT).

The rest of this paper is organized as follows. We describe our method in section 2. We first provide an overview of our method, including network architecture details such as the dose-level estimation, CFT and cascade fusion module. Then, the network parameter selection is elaborated. In the next section, the datasets and network training implementation are described, along with the evaluation methods. In section 4, the experimental results and quantitative evaluation are presented. In section 5, ablation studies are conducted. In section 6, limitations and future work are discussed. Finally, we draw conclusions in section 7.

## 2. Methods

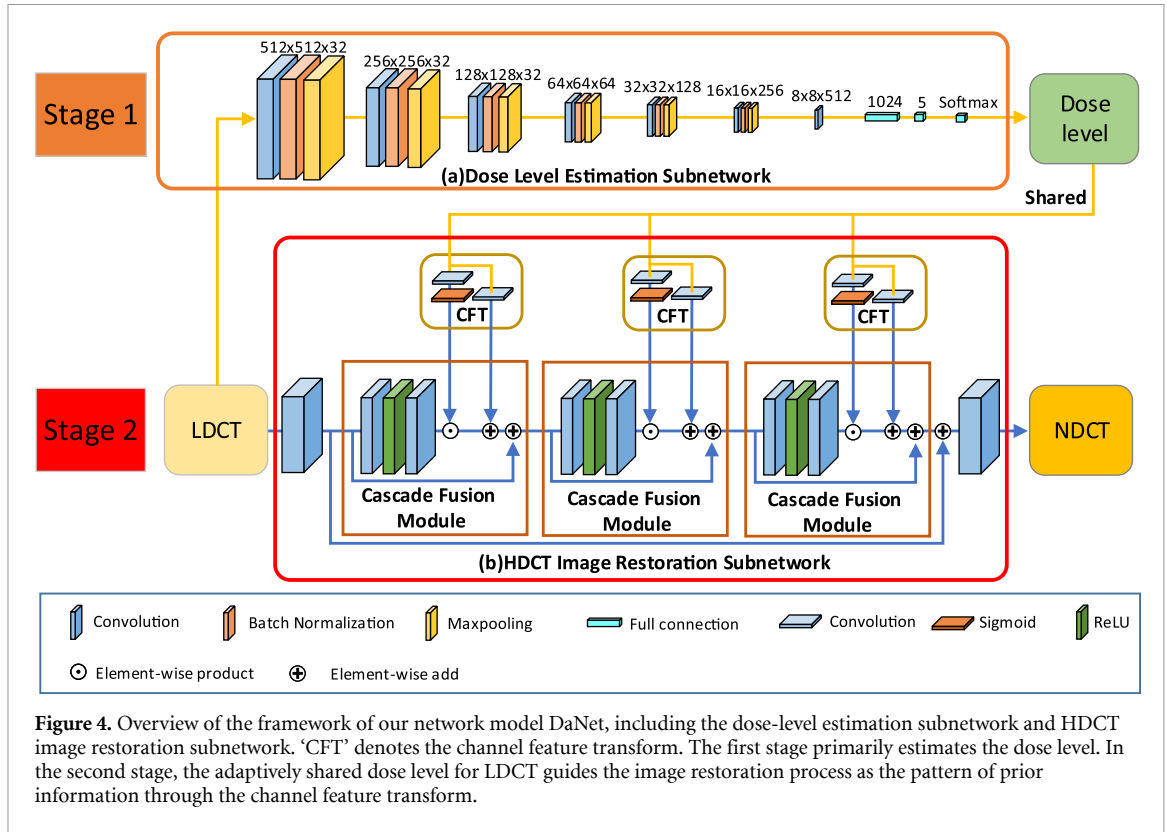
In this section, details of our method are given to solve the low-dose CT problem. First, we present an overview of the framework, including dose-level estimation, CFT and the cascade fusion module. Finally, the network parameter selection is elaborated.

### 2.1. Method overview

As shown in figure 4, the framework consists of three parts: dose-level estimation, CFT and the cascade fusion module. The first part corresponds to the dose-level estimation subnetwork, and the latter two parts form the HDCT image restoration subnetwork. The aim of the first part is to estimate the closest given dose level from LDCT images obtained under an unknown radiation dose. In real clinical applications, the radiologist may apply different radiation doses for specific patients. In this study, we define 5 coarse dose levels of radiation, including lowest, lower, mild, higher and highest, as shown in table 1. The second part is the CFT based on the estimated dose level. Finally, contextual feature information is extracted and fused with the transform information from the dose-level prior information in each cascade module in the third part.

### 2.2. Dose-level estimation

Although some DL-based approaches (Ma *et al* 2019, Dong and Xing 2020) have been proposed for dose calculations, they depend on certain treatment planning systems, such as Varian Medical Systems. Our dose-level estimation is directly applied to CT images, and the framework flow is completed in an end-to-end network. The network architecture of the dose-level estimation subnetwork is shown in figure 4 and contains both convolution and full connection layers. This part estimates the closest dose level for the input LDCT images, as shown in table 1. We employ several convolutions with a kernel size of  $3 \times 3$  on multi-scale feature maps as well as other operations, such as rectified linear unit (ReLU) and batch normalization (BN). Max-pooling operations are utilized for downsampling. In addition, full connection layers follow the convolution outputs and are used to estimate the ultimate dose-level softmax activation function. Before training the whole network, the dose-level estimation subnetwork requires pre-training with the cross-entropy loss function.



We argue that the classification of the categorical prior  $\psi$  based on the dose level improves the capability. The categorical prior could be probability map  $P$  for each dose level. The outcome of this dose-level estimation process is represented by the following:

$$\psi = P = (P_1, P_2, \dots, P_K) \quad (1)$$

where  $K$  denotes the number of given dose levels, which we fix as 5 in our study.

### 2.3. Channel feature transform (CFT)

Based on the categorical prior  $\psi$ , the CFT module estimates a mapping function  $g$ . Inspired by spatial feature transform (SFT) in Wang *et al* (2018), the output of this process is a modular parameter pair  $(\omega, \beta)$ , which adaptively applies affine transformation on the channel dimensionality of intermediate feature maps. The modular parameter pair is formulated as follows:

$$(\omega, \beta) = g(\psi). \quad (2)$$

For  $(\omega, \beta)$  obtained from the same estimated dose level, the transformation is performed with scaling and shifting operations. In addition, a convolution layer and softmax activation function are used to produce the scaling operation  $\omega$ , which ranges from 0 to 1. For calculation convenience, the channel dimensions of  $\omega$  and  $\beta$  should be consistent with the input feature map. This transformation can be formulated as follows:

$$CFT(F|\omega, \beta) = \omega \odot F + \beta \quad (3)$$

where  $F$  denotes the input feature map and  $\odot$  denotes the element-wise product. Since the spatial dimensions are preserved, the CFT module performs on the channel-wise transformation.

### 2.4. Cascade fusion module

In this module, we primarily extract basic feature information  $F_b$  from the previous module with two convolution layers. Then, we apply the basic feature to join the calculation with the CFT. Finally, the skip connection is utilized to reduce the training difficulty. In total, we cascade  $T$  cascade fusion modules for the whole framework. For the  $i$ th cascade fusion module  $M^i$ , we use the basic feature information  $F_b^i$ . Consequently, the process of this module can be formulated as follows:

$$M^i = CFT(F_b^i|\omega_i, \beta_i) + M^{i-1} \quad (4)$$

where  $(\omega_i, \beta_i)$  is the modular parameter pair for  $i$ th cascade fusion module.

**Table 2.** Network parameter details for the dose-level estimation in the first stage.

Components	Kernel size	Image size
Convolution layer 1	$3 \times 3$	$512 \times 512 \times 32$
Batch norm layer 1	—	$512 \times 512 \times 32$
Max pooling layer 1	—	$256 \times 256 \times 32$
Convolution layer 2	$3 \times 3$	$256 \times 256 \times 32$
Batch norm layer 2	—	$256 \times 256 \times 32$
Max pooling layer 2	—	$128 \times 128 \times 32$
Convolution layer 3	$3 \times 3$	$128 \times 128 \times 32$
Batch norm layer 3	—	$128 \times 128 \times 32$
Max pooling layer 3	—	$64 \times 64 \times 32$
Convolution layer 4	$3 \times 3$	$64 \times 64 \times 64$
Batch norm layer 4	—	$64 \times 64 \times 64$
Max pooling layer 4	—	$32 \times 32 \times 64$
Convolution layer 5	$3 \times 3$	$32 \times 32 \times 128$
Batch norm layer 5	—	$32 \times 32 \times 128$
Max pooling layer 5	—	$16 \times 16 \times 128$
Convolution layer 6	$3 \times 3$	$16 \times 16 \times 256$
Batch norm layer 6	—	$16 \times 16 \times 256$
Max pooling layer 6	—	$8 \times 8 \times 256$
Convolution layer 7	$3 \times 3$	$8 \times 8 \times 512$
Fully connection layer 1		1024
Fully connection layer 2		5

**Table 3.** Network parameter details for the channel feature transform.

Components	Kernel size	Input size	Output size
Convolution layer 1	$1 \times 1$	$1 \times 5$	$1 \times 1 \times 1 \times 64$
Convolution layer 2	$1 \times 1$	$1 \times 5$	$1 \times 1 \times 1 \times 64$

**Table 4.** Network parameter details for the cascade fusion module;  $h, w$  denotes the image size of the input LDCT image.

Components	Kernel size	Input size	output size
Convolution layer 1	$3 \times 3$	$h \times w \times 64$	$h \times w \times 64$
Convolution layer 2	$3 \times 3$	$h \times w \times 64$	$h \times w \times 64$

### 2.5. Network parameter selection

The proposed dose-aware network estimates high-resolution CT images from low-dose counterparts in an end-to-end network. It consist of two stages: one that conducts dose-level estimation and one that recovers high-dose CT images. In addition, the dose-level prior information produced by the first stage is fused in each cascade fusion module in the second stage. For the first stage, the output size is described for each scale feature map in table 2. Seven convolution layers with a kernel size of  $3 \times 3$  are applied. In addition, the node numbers of the fully connected layers are 1024 and 5. For the second stage, the parameter details of the CFT and cascade fusion module are given in tables 3 and 4. In the cascade fusion module, we apply two convolution layers with a filter size of  $3 \times 3$  to obtain the basic feature maps. Moreover, the number of cascade fusion modules  $T$  is fixed as 15, as illustrated in section 5.

## 3. Experiments

In this section, we evaluate the performance of our method. First, the simulation datasets based on original high dose parts of AAPM challenge datasets from the Mayo Clinic are elaborated. Next, the network training implementation details are presented. Finally, we describe evaluation metrics and several related methods.

### 3.1. The simulation datasets

We obtain the simulation datasets from AAPM challenge datasets where the original high dose CT images are employed to simulate multiple low dose CT counterparts under different dose levels. The scanning geometry parameters for routine high-dose protocols of simulated datasets are shown in table 5. For low-dose CT images with multiple dose levels, the high-dose CT noise-free projection sinogram data are obtained under a

**Table 5.** Scanning geometry parameters on routine high-dose protocols for the simulation datasets.

Parameters	The simulation datasets
Source-to-detector distance	570 mm
Source-to-object distance	106.23 cm
Detector element size	$1.2856 \times 1.0947 \text{ mm}^2$
Voltage	120 kVp
Current	200–500 mA
Thickness	1.0 mm
Matrix size	$512 \times 512$

fan-beam geometry through the Michigan Image Reconstruction Toolbox (MIRT) (Shi *et al* 2020)<sup>8</sup>. The parameter setting for MIRT is in accordance with the scanning geometry shown in table 5. The compound Poisson model could accurately carve the statistical properties of CT transmission data due to the poly-energetic nature x-ray (Whiting 2002, Liu *et al* 2019, Gupta *et al* 2018, Elbakri *et al* 2002). Several works have demonstrated the fact that the statistical model of CT transmission data with energy integration detection could be represented by an independent Poisson and Gaussian distribution (Elbakri *et al* 2002, Lasio *et al* 2007, Xu and Tsui 2009). For this low dose simulation, the noise projection sinogram data are injected with Poisson and Gaussian noise. Similar to the common simulation process in several works (Liu *et al* 2019, Gupta *et al* 2018, Huang *et al* 2020), the incident fluxes are fixed to obtain different dose levels CT images. To make it easier to understand for five simulated dose levels, assuming that the radiation dose of routine abdominal CT examination for adults is 25 mGy, the radiation dose under five dose levels could be 1.3, 1.9, 2.5, 3.1 and 3.8 mGy in a descending order, respectively. We acquire simulated LDCT images from clinical patients to avoid scanning the patient multiple times under different dose levels. Since the low dose noise projection sinogram data is simulated, the noise LDCT images are reconstructed through FBP algorithm under each dose level. In total, 5800 high-dose CT images are selected from our simulated dataset to build the norma dose CT images for the simulated datasets. From the selected images, 4000 and 800 images are utilized for training and validation of the proposed networks under each low-dose level, respectively. Then, 1000 images are set as the test dataset for each low-dose level; these images are invisible during training process.

### 3.2. Network training implementation

Our proposed method consists of two stages corresponding to two subnetworks: the dose-level estimation subnetwork and the restoration subnetwork. For the two subnetworks, all dose-level data joins the training process instead of inputting each dose-level data to gain a subnetwork, which makes it possible for the network model training only once for a task. The training details are as follows.

#### 3.2.1. Training on the dose-level estimation subnetwork

Before training the restoration subnetwork in the second stage, the dose-level estimation subnetwork in the first stage needs to be trained. For the dose-level estimation subnetwork, the cross entropy is employed as the loss function, and the learning rate is set as 0.0001. Then, the ADAM optimizer is employed. This subnetwork is trained with 1000 epochs for the simulation datasets.

#### 3.2.2. Training on the restoration subnetwork

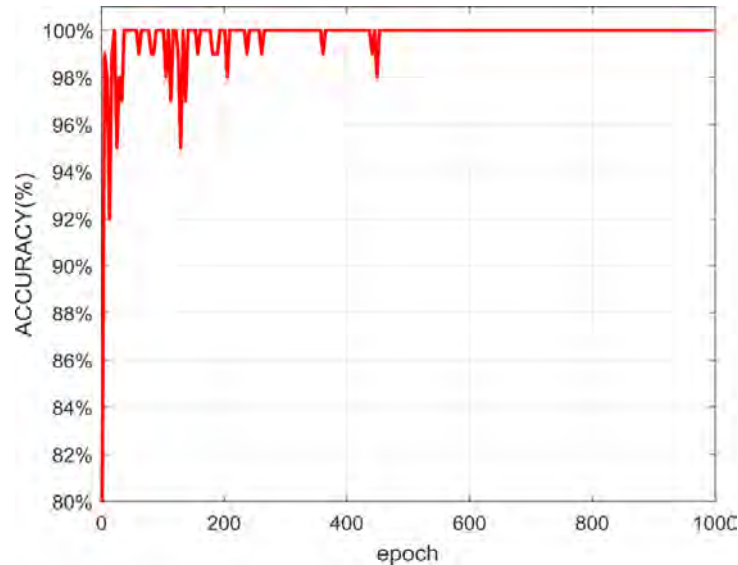
The pre-trained frozen dose-level estimation subnetwork, which offers dose-level prior information, is added to the training process of the restoration subnetwork. Several data augmentation strategies are introduced, for example, random flips in both the horizontal and vertical directions as well as rotation by 90°, 180° and 270°. To reduce the training time, we apply LDCT image patches with an image size of  $64 \times 64$ . To minimize the mean square error (MSE) loss function, the ADAM optimizer is also applied, and the initial learning rate is fixed as 0.0001. After every 200 epochs, the learning rate is fixed as half. We train this subnetwork network with 1000 epochs, and the whole network is implemented in Pytorch on a TITAN 2080Ti GPU.

### 3.3. Evaluation metrics

Two common quantitative metrics, the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM), are used to calculate quantitative outcomes. Besides, an objective evaluation with expert human readers such as radiologists is introduced. For details, the measure metric of subjective quantitative scores for CT images is employed in diagnostic imaging research. To evaluate performance, several other

<sup>8</sup> The code is available at <https://web.eecs.umich.edu/~fessler/code/>.





**Figure 5.** Accuracy of dose-level prediction on simulation validation datasets during the network training for dose-level estimation.

DL-based methods, such as CNN and REDCNN, are included in the comparison. To ensure fair comparisons, the same training and test datasets are adopted for these methods. For the CNN, we apply 30 convolution layers with  $3 \times 3$  filter size. In terms of REDCNN, the convolution and deconvolution layers with  $3 \times 3$  filter size are equipped with stride 1 and zero-padding strategy following (Chen *et al* 2017). The channel number of convolution and deconvolution is set as 64. In addition, the parameters of these methods are kept constant as much as possible. Furthermore, these network models are trained with the same epochs with MSE loss function.

## 4. Results

In this section, we give both several generated visual results and quantitative values. First, the evaluation of the dose-level estimation is described. Then, more visual results and quantitative results are given.

### 4.1. Evaluation of dose-level classification

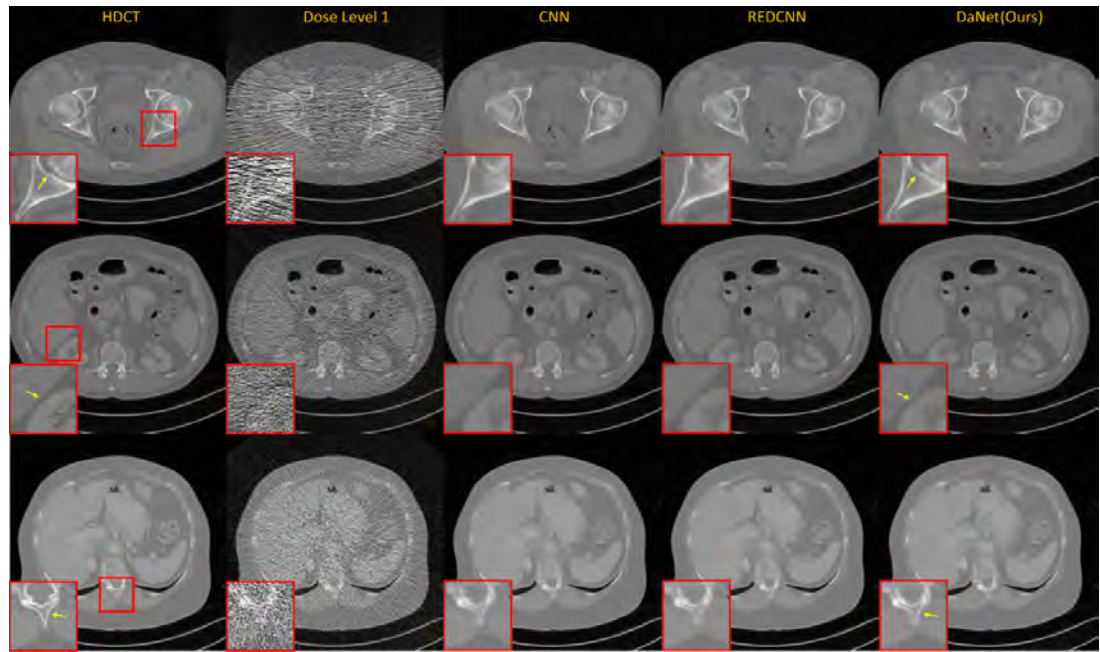
We train our dose-level estimation subnetwork in the first stage with 1000 epochs. The purpose of this dose-level estimation subnetwork is to predict the dose level for each input low-dose CT image, which could be treated as prior information to improve the performance of the restoration subnetwork. Figure 5 shows the accuracy of dose-level prediction on the validation datasets. We find that the performance of dose-level prediction reaches a steady good state after 600 epochs. The accuracy of dose-level classification for test datasets is 99.96%.

### 4.2. Evaluation of restoration results

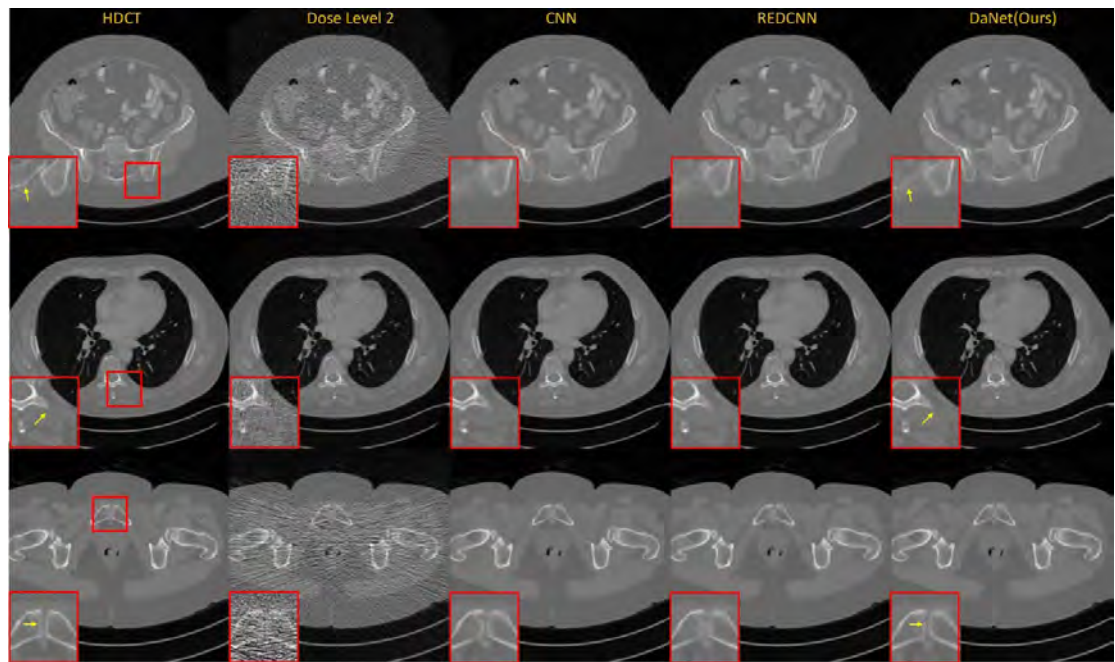
Visual results and quantitative results for the different methods on the five dose levels are given in this section. Figure 6 presents the results for dose level 1 for the different methods and regions of interest (ROIs). In the first row of figure 6, an obvious body of ischium is seen for DaNet, while blurring is apparent in the other two results. In the second row, the boundary of right kidney appears to be clearer for DaNet. In the last row, the boundary of spinal cord is compared, and our method obtains a better visual effect.

As shown in figure 7, among the results in the first row, the sacral vertebrae is most clearly produced by DaNet. For the results in the second row, the corner of the spinal cord is obviously different from reality for REDCNN. In the third row, our DaNet method estimated the pubic symphysis, while blurring is observed in the other two methods.

Figure 8 presents the results of the different methods for dose levels 3, 4 and 5. In the first row, our method clearly generates the pubic symphysis, which is not clear in CNN and REDCNN. In the second row, we can distinguish between two granular tissues in the referenced high-dose CT image and the result generated by DaNet. In the last row for dose level 5, DaNet appears to clearly estimate the sacroiliac joint, while blurring is observed for the other methods.



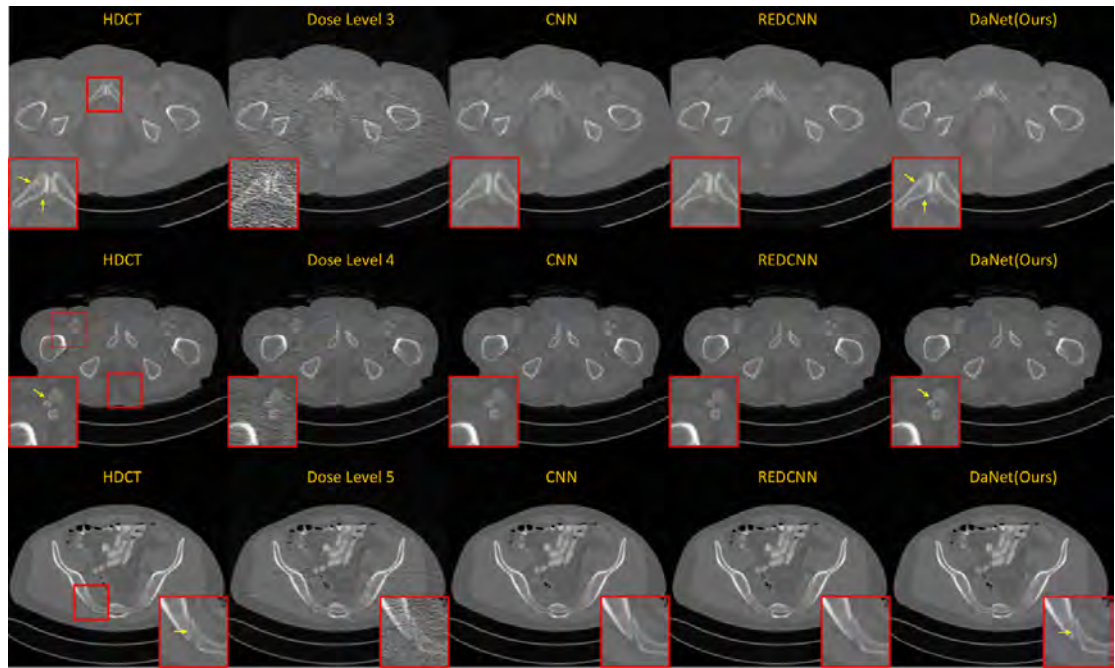
**Figure 6.** Estimated results on simulation datasets for the three methods on dose level 1. The red boxes denote the regions of interest (ROIs), and the yellow arrows denote the differences.



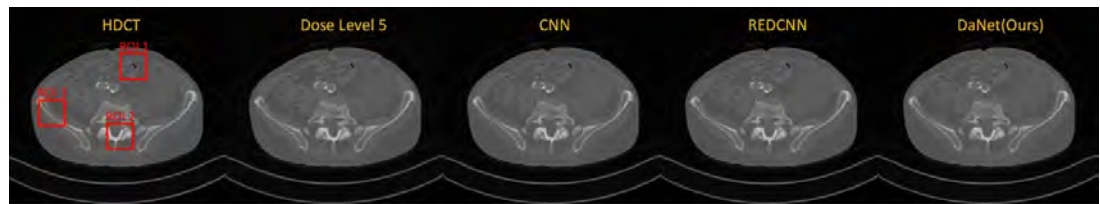
**Figure 7.** Estimated results on simulation datasets for the three methods on dose level 2. The red boxes denote the regions of interest (ROIs), and the yellow arrows denote the differences.

As shown in figure 9, three ROIs for the different methods are selected for dose level 5. We give the statistical indicators of PSNR and SSIM in figure 10. Based on our observations, our method achieves better statistical results. Moreover, the PSNR for ROI3 is obviously better with our method than with the other two methods.

We also calculate the PSNR and SSIM for several test datasets on whole images under five dose levels, as shown in table 6. Our method shows superior performance compared with the other methods. In particular, our method easily produces better quantitative results at lower doses, such as dose levels 1 and 2. At all five dose levels, our method performs better in terms of both PSNR and SSIM.



**Figure 8.** Estimated results on simulation datasets for the three methods on dose levels 3, 4 and 5. The red boxes denote the regions of interest (ROIs), and the yellow arrows denote the differences.



**Figure 9.** Estimated results on simulation datasets for the three methods on dose level 5. Three regions of interest (ROIs) are marked in red boxes.

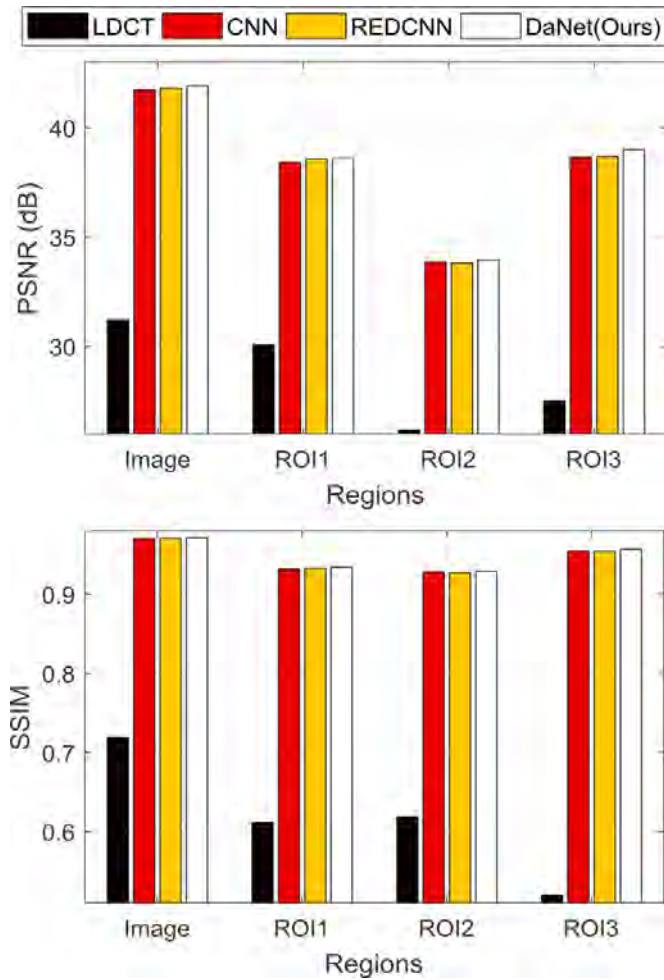
**Table 6.** Quantitative results (mean  $\pm$  std) for PSNR and SSIM on simulation datasets for the test datasets. The best results are marked in bold.

Dose level	Metrics	LDCT	CNN	REDCNN	DaNet(Ours)
1	PSNR	17.58 $\pm$ 2.57	36.63 $\pm$ 2.13 <sup>a</sup>	37.84 $\pm$ 2.13 <sup>a</sup>	<b>38.16 <math>\pm</math> 2.09<sup>a</sup></b>
	SSIM	0.2665 $\pm$ 0.1139	0.9480 $\pm$ 0.0145 <sup>a</sup>	0.9482 $\pm$ 0.0146 <sup>a</sup>	<b>0.9493 <math>\pm</math> 0.0143<sup>a</sup></b>
2	PSNR	20.73 $\pm$ 2.74	39.55 $\pm$ 1.74 <sup>a</sup>	39.71 $\pm$ 1.74 <sup>a</sup>	<b>39.83 <math>\pm</math> 1.74<sup>a</sup></b>
	SSIM	0.3403 $\pm$ 0.1173	0.9550 $\pm$ 0.0123 <sup>a</sup>	0.9555 $\pm$ 0.0123 <sup>a</sup>	<b>0.9559 <math>\pm</math> 0.0123<sup>a</sup></b>
3	PSNR	22.90 $\pm$ 2.61	40.26 $\pm$ 1.55 <sup>a</sup>	40.40 $\pm$ 1.56 <sup>a</sup>	<b>40.48 <math>\pm</math> 1.58<sup>a</sup></b>
	SSIM	0.3954 $\pm$ 0.1168	0.9578 $\pm$ 0.0114 <sup>a</sup>	0.9582 $\pm$ 0.0114 <sup>a</sup>	<b>0.9586 <math>\pm</math> 0.0114<sup>a</sup></b>
4	PSNR	24.28 $\pm$ 2.54	40.61 $\pm$ 1.52 <sup>a</sup>	40.74 $\pm$ 1.53 <sup>a</sup>	<b>40.81 <math>\pm</math> 1.56<sup>a</sup></b>
	SSIM	0.4374 $\pm$ 0.1154	0.9594 $\pm$ 0.0111 <sup>a</sup>	0.9598 $\pm$ 0.0111 <sup>a</sup>	<b>0.9601 <math>\pm</math> 0.0111<sup>a</sup></b>
5	PSNR	25.30 $\pm$ 2.49	40.85 $\pm$ 1.49 <sup>a</sup>	40.98 $\pm$ 1.50 <sup>a</sup>	<b>41.04 <math>\pm</math> 1.52<sup>a</sup></b>
	SSIM	0.4711 $\pm$ 0.1135	0.9604 $\pm$ 0.0108 <sup>a</sup>	0.9608 $\pm$ 0.0108 <sup>a</sup>	<b>0.9611 <math>\pm</math> 0.0108<sup>a</sup></b>

<sup>a</sup> denotes  $P < 0.05$ , corresponding to a significant difference.

#### 4.3. Evaluation of subjective scores

In order to better illustrate the visual effect, we apply the quantitative results using statistical subjective scores. Two radiologists with expert experiences are invited to give subjective scores ranging from 1 to 10 for different CT images. Corresponding 10 images are selected from each image category to join this evaluation under each dose level on two datasets. As is shown in table 7, the subjective scores of our method on the simulation datasets gain better performance than other methods. Based on observation, the subjective score would improve with the increasing dose level for each method as well as low-dose CT images.



**Figure 10.** Statistical indicators on simulation datasets of PSNR and SSIM on simulation dataset for the three ROIs marked in figure 9.

**Table 7.** Subjective scores (mean  $\pm$  std) for test data on simulation datasets under 5 low dose levels. The best results are marked in bold.

Dose level	LDCT	HDCT	CNN	REDCNN	DaNet(Ours)
1	1.90 $\pm$ 0.31 <sup>a</sup>	8.20 $\pm$ 0.89	5.55 $\pm$ 0.60 <sup>a</sup>	5.95 $\pm$ 0.76 <sup>a</sup>	<b>6.75 <math>\pm</math> 0.64<sup>a</sup></b>
2	2.35 $\pm$ 0.67 <sup>a</sup>	8.20 $\pm$ 0.89	5.85 $\pm$ 0.81 <sup>a</sup>	6.00 $\pm$ 0.65 <sup>a</sup>	<b>7.05 <math>\pm</math> 0.76<sup>a</sup></b>
3	2.85 $\pm$ 0.59 <sup>a</sup>	8.20 $\pm$ 0.89	5.90 $\pm$ 0.64 <sup>a</sup>	6.15 $\pm$ 0.49 <sup>a</sup>	<b>7.30 <math>\pm</math> 0.47<sup>a</sup></b>
4	2.90 $\pm$ 0.79 <sup>a</sup>	8.20 $\pm$ 0.89	6.00 $\pm$ 0.46 <sup>a</sup>	6.20 $\pm$ 0.62 <sup>a</sup>	<b>7.50 <math>\pm</math> 0.51<sup>a</sup></b>
5	3.15 $\pm$ 0.59 <sup>a</sup>	8.20 $\pm$ 0.89	6.25 $\pm$ 0.55 <sup>a</sup>	6.50 $\pm$ 0.51 <sup>a</sup>	<b>7.80 <math>\pm</math> 0.59<sup>a</sup></b>

<sup>a</sup> denotes  $P < 0.05$ , corresponding to a significant difference.

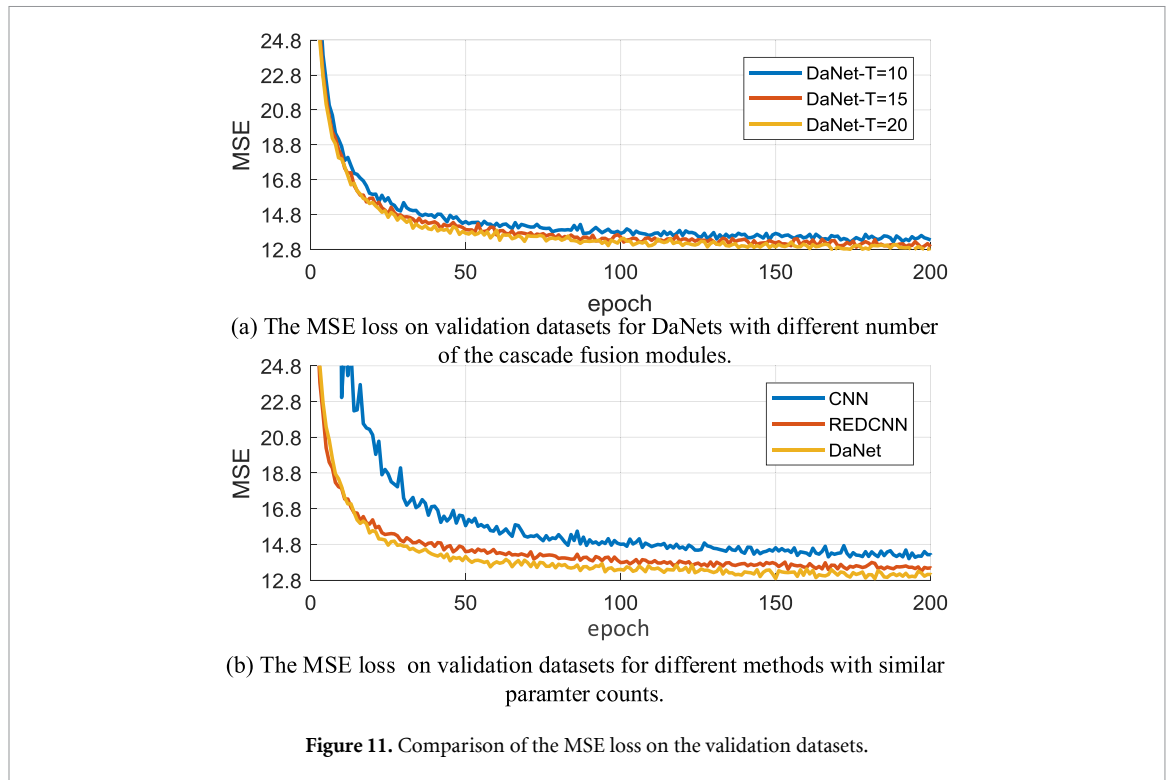
## 5. Ablation studies

In this section, we conduct several ablation studies on simulation datasets to select better parameters for network design and validate the effectiveness of our method. In these studies, we train our DaNet as well as CNN and REDCNN with 200 epochs, and the MSE loss is recorded.

### 5.1. Number selection of the cascade fusion module

Figure 11(a) presents the MSE loss on validation datasets for DaNets with 10, 15 and 20 cascade fusion modules. The downward trend of the loss function is quite similar when the number of cascade fusion modules  $T$  is fixed at 15 and 20. To use fewer network parameter counts, we set the number of cascade fusion modules to 15. To ensure a fair comparison, CNN and REDCNN are configured with 30 middle convolution layers in addition to the similar first and last convolution layers.





**Table 8.** Quantitative results (mean  $\pm$  std) on PSNR and SSIM for baseline and our proposed DaNet. The baseline model is a simple version of DaNet without dose-level estimation. The best results are marked in red.

Dose level	Metrics	LDCT	Baseline	DaNet(Ours)
1	PSNR	$17.16 \pm 2.19$	$36.70 \pm 1.98^a$	<b><math>37.26 \pm 1.96^a</math></b>
	SSIM	$0.2384 \pm 0.0900$	$0.9406 \pm 0.0151^a$	<b><math>0.9416 \pm 0.0150^a</math></b>
2	PSNR	$20.15 \pm 2.17$	$38.82 \pm 1.51^a$	<b><math>38.94 \pm 1.51^a</math></b>
	SSIM	$0.3078 \pm 0.0972$	$0.9488 \pm 0.0134^a$	<b><math>0.9492 \pm 0.0134^a</math></b>
3	PSNR	$22.13 \pm 2.05$	$39.53 \pm 1.36^a$	<b><math>39.59 \pm 1.35^a</math></b>
	SSIM	$0.3596 \pm 0.1000$	$0.9521 \pm 0.0128^a$	<b><math>0.9523 \pm 0.0127^a</math></b>
4	PSNR	$23.44 \pm 1.98$	$39.91 \pm 1.33^a$	<b><math>39.94 \pm 1.33^a</math></b>
	SSIM	$0.4000 \pm 0.1009$	$0.9539 \pm 0.0124^a$	<b><math>0.9541 \pm 0.0124^a</math></b>
5	PSNR	$24.40 \pm 1.96$	$40.16 \pm 1.31^a$	<b><math>40.17 \pm 1.31^a</math></b>
	SSIM	$0.4330 \pm 0.1010$	$0.9551 \pm 0.0121^a$	<b><math>0.9553 \pm 0.0121^a</math></b>

<sup>a</sup> denotes  $P < 0.05$ , corresponding to a significant difference.

**Table 9.** Network parameter counts and average running times for each test CT image.

Methods	CNN	REDCNN	Baseline	DaNet(Ours)
Parameters	1.1 M	1.1 M	1.1 M	1.6 + 1.1 M
Running times	185.70 ms	195.81 ms	187.96 ms	221.97 ms

## 5.2. The effectiveness of whole DaNet framework

Figure 11(b) presents the MSE loss on the validation datasets for the different methods. We find that our DaNet achieves lower MSE during the training process.

## 5.3. The effectiveness of dose-level estimation

The derived network model Baseline is designed based on our proposed DaNet. The Baseline model is a simple version of DaNet that omits the dose-level estimation. Quantitative results are calculated on PSNR and SSIM under five dose levels for the test CT images. As shown in table 8, our DaNet method performs better under the five dose levels than the Baseline version, thus validating the effectiveness of the performance improvement in dose-level estimation.



## 6. Limitations and future work

Although our method provides better performance, there are two main limitations: the coarse classification of five dose levels and the greater number of parameter counts. Here we illustrate these drawbacks and outline future work.

### 6.1. The coarse classification of five dose levels

In this paper, we validate the effectiveness of the dose-aware network by using five coarse dose levels in the experiments. However, given the complex scanning parameters used in clinical practice, five dose levels are likely insufficient. To include more dose levels, much more data for each dose level must be collected. Although the low-dose CT simulation is widely used, the specific dose value can be directly obtained by changing the scanning protocols through the real CT system rather than simulation on incident flux, which would better meet clinical needs. We will continue exploring this model with support from clinical data in future work. Data from different CT manufacturers will also be considered.

### 6.2. More parameter counts and running time

Our DaNet consists of two subnetworks: the dose-level estimation subnetwork and the image restoration subnetwork. The parameter counts and running times for different methods are presented in table 9. Due to the limitation of the framework flow for the dual subnetworks, our model has the most parameter counts and longest running time among the compared methods. However, if we only consider the restoration process excluding the extra dose-level estimation in our method, the parameter count is quite similar to those of the other methods. In the future, we intend to reduce the parameter counts of the whole network framework and combine these two subnetworks into one united network via dual tasks.

## 7. Conclusion

We have proposed a novel dose-aware network that focuses on the influence of different radiation dose levels on restoration performance in low-dose CT imaging. Many recently proposed low-dose CT DL-based image restoration methods for estimating high-resolution CT images depend on a certain radiation dose or ignore the influence of different doses. These methods may suffer from performance loss when applied on images of unknown radiation doses. Our method considers the dose-level prior information, and the proposed network model is divided into two subnetworks: the dose-level estimation subnetwork and the HDCT image restoration subnetwork with feature transform on the previous subnetwork. One of the most important targets of the first subnetwork is to estimate the unknown dose level. Then, the dose-level prior information is fused into the feature maps through the scaling and shifting operations. Thus, the dose-level information plays a regulatory role in the image feature maps to guide the restoration process. Experiments demonstrate that the extra dose-level estimation provides improved performance. Our method is superior to several DL-based methods. Due to the limitation of the framework flow with dual stages, future work is needed to optimize the parameter counts and define the strict classification of multiple dose levels based on clinical experiments by radiologists. Better performance under more complex scanning conditions is also needed.

## Acknowledgments

The authors would like to thank the editor and anonymous reviewers for their constructive comments and suggestions. This work was supported by the National Natural Science Foundation of China (Grant Nos. 61672246, 61272068, 32022042, 81871441), the Fundamental Research Funds for the Central Universities, HUST (Grant No. 2016YXMS018), the Natural Science Foundation of Guangdong Province in China (Grant No. 2020A1515010733), the Guangdong International Science and Technology Cooperation Project of China (Grant No. 2018A050506064), the Shenzhen International Cooperation Research Project of China (Grant No. GJHZ20180928115824168), the Guangdong Special Support Program of China (Grant No. 2017TQ04R395) and the Chinese Academy of Sciences Key Laboratory of Health Informatics in China (Grant No. 2011DP173015).

## ORCID iDs

Zhenxing Huang  <https://orcid.org/0000-0001-5012-1992>

Zixiang Chen  <https://orcid.org/0000-0001-6015-5010>

Jincai Chen  <https://orcid.org/0000-0001-7368-1677>

Dong Liang  <https://orcid.org/0000-0003-1358-9777>

Zhanli Hu  <https://orcid.org/0000-0003-0618-6240>

## References

- Bian J, Siewerdsen J H, Han X, Sidky E Y, Prince J L, Pelizzari C A and Pan X 2010 Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT *Phys. Med. Biol.* **55** 6575–99
- Brenner D J and Hall E J 2007 Computed tomography—an increasing source of radiation exposure *N. Engl. J. Med.* **357** 2277–84
- Cai J, Jia X, Gao H, Jiang S B, Shen Z and Zhao H 2014 Cine cone beam CT reconstruction using low-rank matrix factorization: algorithm and a proof-of-principle study *IEEE Trans. Med. Imaging* **33** 1581–91
- Chen G-H, Tang J and Leng S 2008 Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic ct images from highly undersampled projection data sets *Med. Phys.* **35** 660–3
- Chen H, Zhang Y, Kalra M K, Lin F, Chen Y, Liao P, Zhou J and Wang G 2017 Low-dose CT with a residual encoder-decoder convolutional neural network *IEEE Trans. Med. Imaging* **36** 2524–35
- Chen H, Zhang Y, Zhang W, Liao P, Li K, Zhou J and Wang G 2017 Low-dose CT via convolutional neural network *Biomed. Opt. Express* **8** 679–94
- Chen Y, Gao D, Nie C, Luo L, Chen W, Yin X and Lin Y 2009 Bayesian statistical reconstruction for low-dose x-ray computed tomography using an adaptive-weighting nonlocal prior *Comput. Med. Imaging Graph.* **33** 495–500
- Chen Y, Ma J, Feng Q, Luo L, Shi P and Chen W 2008 Nonlocal prior bayesian tomographic reconstruction *J. Math. Imaging Vis.* **30** 133–46
- Chen Y, Shi L, Feng Q, Yang J, Shu H, Luo L, Coatrieux J and Chen W 2014 Artifact suppressed dictionary learning for low-dose ct image processing *IEEE Trans. Med. Imaging* **33** 2271–92
- Chen Y, Yin X, Shi L, Shu H, Luo L, Coatrieux J-L and Toumoulin C 2013 Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing *Phys. Med. Biol.* **58** 5803–20
- Dong P and Xing L 2020 Deep dosenet: a deep neural network for accurate dosimetric transformation between different spatial resolutions and/or different dose calculation algorithms for precision radiation therapy *Phys. Med. Biol.* **65** 035010
- Donoho D L et al 2006 Compressed sensing *IEEE Trans. Inf. Theory* **52** 1289–306
- Du W, Chen H, Liao P, Yang H, Wang G and Zhang Y 2019 Visual attention network for low-dose CT *IEEE Signal Process. Lett.* **26** 1152–6
- Elbakri I A and Fessler J A 2002 Statistical image reconstruction for polyenergetic x-ray computed tomography *IEEE Trans. Med. Imaging* **21** 89–99
- Fan F, Shan H, Kalra M K, Singh R, Qian G, Getzin M, Teng Y, Hahn J and Wang G 2019 Quadratic autoencoder (q-AE) for low-dose CT denoising *IEEE Trans. Med. Imaging* **39** 2035–50
- Feruglio P F, Vinegoni C, Gros J, Sbarbati A and Weissleder R 2010 Block matching 3D random noise filtering for absorption optical projection tomography *Phys. Med. Biol.* **55** 5401–15
- Gao H, Yu H, Osher S and Wang G 2011 Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM) *Inverse Problems* **27** 115012
- Gupta H, Jin K H, Nguyen H Q, McCann M T and Unser M 2018 Cnn-based projected gradient descent for consistent CT image reconstruction *IEEE Trans. Med. Imaging* **37** 1440–53
- Hu Z, Jiang C, Sun F, Zhang Q, Ge Y, Yang Y, Liu X, Zheng H and Liang D 2019 Artifact correction in low-dose dental CT imaging using Wasserstein generative adversarial networks *Med. Phys.* **46** 1686–96
- Hu Z, Liu Q, Zhang N, Zhang Y, Peng X, Wu P Z, Zheng H and Liang D 2016 Image reconstruction from few-view CT data by gradient-domain dictionary learning *J. X-Ray Sci. Technol.* **24** 627–38
- Hu Z, Zhang Y, Liu J, Ma J, Zheng H and Liang D 2016 A feature refinement approach for statistical interior CT reconstruction *Phys. Med. Biol.* **61** 5311
- Huang Z et al 2020 Cagan: A cycle-consistent generative adversarial network with attention for low-dose CT imaging *IEEE Trans. Computat. Imaging* **6** 1203–18
- Kalra M K, Maher M M, Toth T L, Hamberg L M, Blake M A, Shepard J-A and Saini S 2004 Strategies for CT radiation dose optimization *Radiology* **230** 619–28
- Kang E, Min J and Ye J C 2017 A deep convolutional neural network using directional wavelets for low-dose x-ray CT reconstruction *Med. Phys.* **44** e360–e375
- Lasio G M, Whiting B R and Williamson J F 2007 Statistical reconstruction for x-ray computed tomography using energy-integrating detectors *Phys. Med. Biol.* **52** 2247
- Li H and Mueller K 2017 Low-dose ct streak artifacts removal using deep residual neural network in *Proc. Fully 3D Conf.* vol 2017 pp 191–4
- Li Z, Yu L, Trzasko J D, Lake D S, Blezek D J, Fletcher J G, McCollough C H and Manduca A 2014 Adaptive nonlocal means filtering based on local noise level for CT denoising *Med. Phys.* **41** 011908
- Liu J et al 2019 Deep iterative reconstruction estimation (DIRE): approximate iterative reconstruction estimation for low dose CT imaging *Phys. Med. Biol.* **64** 135007
- Liu Y and Zhang Y 2018 Low-dose CT restoration via stacked sparse denoising autoencoders *Neurocomputing* **284** 80–9
- Ma M, Buyyounouski M K, Vasudevan V, Xing L and Yang Y 2019 Dose distribution prediction in isodose feature-preserving voxelization domain using deep convolutional neural network *Med. Phys.* **46** 2978–87
- Manduca A, Yu L, Trzasko J D, Khaylova N, Kofler J M, McCollough C M and Fletcher J G 2009 Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT *Med. Phys.* **36** 4911–19
- Pearce M S et al 2012 Radiation exposure from ct scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study *The Lancet* **380** 499–505
- Sheng K, Gou S, Wu J and Qi S X 2014 Denoised and texture enhanced MVCT to improve soft tissue conspicuity *Med. Phys.* **41** 101916
- Shi L, Liu B, Yu H, Wei C, Wei L, Zeng L and Wang G 2020 Review of CT image reconstruction open source toolkits *J. X-Ray Sci. Technol.* **28** 619–39
- Sidky E Y and Pan X 2008 Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization *Phys. Med. Biol.* **53** 4777–807
- Slovits T L 2002 The alara concept in pediatric CT: myth or reality? *Radiology* **223** 5–6

- Wang J, Li T, Lu H and Liang Z 2006 Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography *IEEE Trans. Med. Imaging* **25** 1272–83
- Wang X, Yu K, Dong C and Change Loy C 2018 Recovering realistic texture in image super-resolution by deep spatial feature transform in *The Conf. on Computer Vision and Pattern Recognition (CVPR)*
- Whiting B R 2002 Signal statistics in x-ray computed tomography *Int. Society for Optics and Photonics* vol 4682 pp 53–60
- Wolterink J M, Leiner T, Viergever M A and Isgum I 2017 Generative adversarial networks for noise reduction in low-dose CT *IEEE Trans. Med. Imaging* **36** 2536–45
- Xu J and Tsui B M 2009 Electronic noise modeling in statistical iterative reconstruction *IEEE Trans. Image Process.* **18** 1228–38
- Xu Q, Yu H, Mou X, Zhang L, Hsieh J and Wang G 2012 Low-dose x-ray CT reconstruction via dictionary learning *IEEE Trans. Med. Imaging* **31** 1682–97
- Zhang H, Han H, Wang J, Ma J, Liu Y, Moore W and Liang Z 2014 Deriving adaptive mrf coefficients from previous normal-dose CT scan for low-dose image reconstruction via penalized weighted least-squares minimization *Med. Phys.* **41** 041916
- Zhang Y, Wang Y, Zhang W, Lin F, Pu Y and Zhou J 2016 Statistical iterative reconstruction using adaptive fractional order regularization *Biomed. Opt. Express* **7** 1015–29
- Zhang Y, Zhang W, Lei Y and Zhou J 2014 Few-view image reconstruction with fractional-order total variation *J. Opt. Soc. Am. A* **31** 981–95
- Zhang Y, Zhang W-H, Chen H, Yang M-L, Li T-Y and Zhou J-L 2013 Few-view image reconstruction combining total variation and a high-order norm *Int. J. Imaging Syst. Technol.* **23** 249–55
- Zhang Z, Liang X, Dong X, Xie Y and Cao G 2018 A sparse-view ct reconstruction method based on combination of densenet and deconvolution *IEEE Trans. Med. Imaging* **37** 1407–17