Article

Low Dose CT Image Denoising: A Comparative Study of Deep Learning Models and Training Strategies

Heng Zhao [1], Like Qian [1], Yaqi Zhu [1] and Dingcheng Tian [1,2,]*

[1] Research Institute for Medical and Biological Engineering, Ningbo University, Ningbo 315211, China

[2] College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110016, China

* Correspondence: 2310520@stu.neu.edu.cn

Abstract: Low-dose computed tomography (LDCT) denoising is an important topic in CT image research. Compared with normal-dose CT images, LDCT can reduce the radiation dose of X-rays, decreasing the radiation burden on the human body, which is beneficial to human health. However, quantum noise caused by low-dose rays will reduce the quality of CT images, thereby decreasing the accuracy of clinical diagnosis. In recent years, deep learning-based denoising methods have shown promising advantages in this field. Researchers have proposed some optimized models for low-dose CT image denoising. These methods have enhanced the application of low-dose CT image denoising from different aspects. From the perspective of experimental research, this paper investigates and evaluates some top deep learning models proposed in the field of low-dose image denoising in recent years, with the aim of determining the best models and training strategies for this task. We conducted experiments on seven deep learning models (REDCNN, EDCNN, QAE, OCTNet, UNet, WGAN, CTformer) on the AAPM dataset and the Piglet dataset. Our research shows that UNet has the best denoising effect among the models, obtaining PSNR = 33.06 (AAPM dataset) and PSNR = 31.21 (Piglet dataset), and good generalization capacity is also observed. However, UNet has a large number of parameters, and the time it takes to process an image is about 8 ms, while EDCNN takes about 4.8 ms to process an image, and its average PSNR is ranked second after UNet. EDCNN strikes a balance between denoising performance and processing efficiency, making it ideal for low-dose CT image denoising tasks.

Keywords: deep learning; low dose CT; image denoising; convolutional neural network

1. Introduction

Computed tomography imaging system, as a non-invasive imaging device, has been widely applied in medical diagnosis and treatment [1]. However, excessive CT scans may cause some cancers and diseases, towing to the effect of the radiation dose [2,3]. Therefore, in clinical diagnosis, it is advocated to adhere to the ALARA (As Low As Reasonably Achievable) principle [4], that is, to minimize the damage of X-rays to the human body on the premise of ensuring that the quality of CT images meets the diagnostic needs. However, during low-dose imaging, the radiation dose will affect the density distribution of X-ray photons, which will increase quantum noise, causing noise and stripe artifacts in the reconstructed image, and further, will lead to disconnected edges, smooth the target subtle structures and lack of X-ray photons resulting in low-contrast visual effects, impairing the quality of CT images and affecting the accuracy of clinical diagnosis. Since Naidich et al. [5] first proposed the concept of low-dose CT (LDCT) denoising in 1990, the issue of effectively suppressing noise and artifacts in

LDCT images has attracted much attention. Researchers have continuously optimized the design scheme from different perspectives and have achieved some outstanding results [6–12].

Early in the field of LDCT image denoising research, some studies focused on applying denoising techniques directly to the raw sinogram data [13,14]. The sinogram denoising algorithm [15,16] relies on the projection data and uses the characteristic that noise obeys the Poisson distribution in the projection data to eliminate the noise in the projection data [17]. The iterative reconstruction algorithms operate on raw data and reconstructed CT image [18,19]. These methods transfer the raw data between the image domain and the projection domain multiple times and each time update and modify the results to obtain clear CT images. In practice, raw data from commercial scanners are difficult to obtain. Therefore, many studies directly denoise the reconstructed CT images [20,21]. These methods do not require raw data and can be easily integrated into the workflow. These methods are usually based on techniques such as filtering [22,23], wavelet transform [24], dictionary learning [25], etc, to improve image quality and reduce the impact of noise. Sparse representation and non-local means have been applied to remove noise from LDCT images [8,26]. The state-of-the-art Block Matching 3D (BM3D) [27] is also employed in multiple studies to perform this task with successful results [28].

In recent years, deep learning methods have achieved advanced performance in LDCT image denoising [29–32]. Deep residual networks and convolutional Neural networks (CNN) [33] are early applications of LDCT denoising. Chen et al. [34] first proposed an LDCT denoising method based on deep neural network, this method can convert LDCT images into normal-dose CT images. Compared with traditional denoising methods, the model improved the denoising effect and computation time. Zhang et al. [35] combined dense blocks and deconvolution structures to build a lightweight network that can reuse image features to improve image quality. In addition to supervised learning methods, unsupervised learning [36,37] and semi-supervised learning models [38,39] have also achieved significant accomplishments. Generative adversarial network (GAN) are also used to improve the quality of LDCT images [40,41]. Xin et al. [42] added a sharpness detection network to the GAN network to guide the training process. The processed images have minimal resolution loss and achieve advanced performance. Yang et al. [43] proposed a CT image denoising method based on the GAN with Wasserstein distance and perceptual loss, the network reduces noise while maintaining the key information of the image. Autoencoders [44–47] achieve image denoising by learning to encode input data into a low-dimensional representation and decoding it back to the original data space. Self-supervised learning [48] and unsupervised learning use the characteristics of the data itself for training and do not require a large amount of labeled data. The semi-supervised learning method combines labeled data and unlabeled data [49], it can better utilize the information of the data and improve the generalization ability of the model and the effect of image denoising. Recent years, the emergence of Transformer [50] has also achieved remarkable results in the field of medical image processing. Luthra et al. [51] proposed a Transformer model based on edge enhancement to build the encoder and decoder. The network uses the self-attention mechanism to learn the relationship between pixels and other pixels in image blocks containing non-overlapping windows. By integrating the features of all positions to generate image details, and introducing a trainable Sobel operator to enhance the edges of the image, it provides higher performance on the AAPM dataset. In addition, a large number of combinations of deep learning and traditional denoising algorithms have also been proposed and achieved excellent results [52,53].

Although there are so many deep learning methods for LDCT image denoising [54–56], previous studies have major differences in dataset, training strategy, and performance indicators, making it impossible to evaluate the results of the models in a relatively fair manner. This paper addresses the issue through investigation and fair comparison of seven deep learning models. We conduct experiments on all models based on two datasets (AAPM and Piglet, Figure 1 shows some examples from the two datasets.) and implement seven deep learning models (REDCNN [34], EDCNN [57], QAE [58], OCTNet [59], UNet [59], WGAN [43], CTformer [60]). We systematically evaluate them from several aspects to find the best model and training strategy for the LDCT image denoising task. In summary, the work and contributions of this paper are as follows: (1) We evaluate the performance of seven deep learning models under the same metrics (PSNR, SSIM, and RMSE). (2) We conduct cross-experiments on the models based on different datasets to examine the generalization ability of the models. (3) The experiments evaluate the denoising performance of the models on CT images at various dose levels. (4) To measure the efficiency of the models, we calculate and compare the training cost and processing speed of the models.

The rest of this paper is organized as follows. Section II details the dataset and performance metrics. Section III presents the experimental results and data analysis, including the experimental setup and evaluation methods. In Section IV, we analyze and discuss the experimental results. Section V concludes the paper and presents future work.
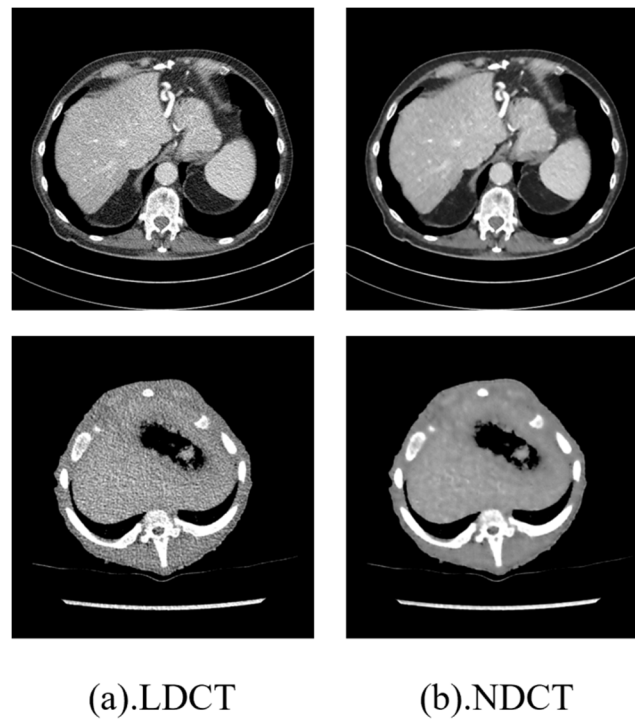
(a).LDCT   (b).NDCT

Figure 1. The dataset we used. The first row is the LDCT image and NDCT image of the AAPM dataset, and the second row is the LDCT image and NDCT image of the Piglet dataset.

2. Method

Standard datasets are crucial for model training. The datasets in this article use the Piglet dataset and the AAPM dataset. The LDCT image of the Piglet dataset is obtained by reducing the tube current, and the LDCT image of the AAPM dataset is obtained by adding Poisson noise to the original image.

2.1. Dataset

2.1.1. Piglet Dataset

The real dataset uses the Piglet dataset [42]. The LDCT image of this dataset is obtained by using a GE scanner (Discovery CT750 HD), setting the source potential and slice thickness to 100 kVp and 0.625 mm, and adjusting the tube current (or voltage), obtained by X-ray scanning with different intensities. Among them, the radiation dose when the tube current is 300 mA is the normal dose, and the tube current is reduced to 50%, 25%, 10% and 5%, respectively to obtain 4 different dose LDCT images. With different X-ray radiation doses, reconstructed CT images are subject to varying degrees of noise and artifacts. In the experiment, 720 CT images are selected from the dataset as the training dataset, and 180 images are used as the test dataset. The images of the Piglet dataset during training are one-dimensional. We extracted 48,000 pairs of image patches from the 720 CT images as input and Label, size is 64×64. Notably, 11,520 pairs of image patches were extracted from another 180 CT images for testing. The Table 1 below provides details of the dataset. The Piglet dataset is available from the original author's GitHub repository: https://github. com/xinario/SAGAN (accessed on 2 August 2024) [42].

Table 1. Dose used for the Piglet dataset. In all 5 series, tube potential was 100 KV with 0.625 mm slice thickness. Tube current decreased to 50, 25, 10 and 5% of full-dose tube current (300 mAs) to obtain images with different doses. CTDI is the CT dose index and DLP is the dose-length product.

Dose Level	FULL	50%	25%	10%	5%
Tube current (mAs)	300	150	75	30	15
$CTDI_{vol}$ (mGy)	30.83	15.41	7.71	3.08	1.54
DLP (mGy-cm)	943.24	471.62	235.81	94.32	47.1
Effective dose (mSv)	14.14	7.07	3.54	1.41	0.71

2.1.2. AAPM Dataset

The simulation dataset is from "the 2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge" [61]. The dataset contains 2378 slices from 10 anonymous patients, with a slice thickness of 1.0 mm, including LDCT images and NDCT images. The dataset is contrast-enhanced abdominal CT patient scans, each acquired during the portal venous phase using a Siemens SOMATOM Flash scanner. Among them, LDCT images are obtained by simulating noise pollution under 1/4 standard dose. In the experiment, 8 patients were selected as training data for the model, and the other 2 patients were selected as test data. Our approach is similar to other studies [1,62]. Table 2 lists the imaging conditions for each patient's original scans and the respective tube current intensities. It is worth noting that the noise in LDCT may no longer strictly follow the Poisson distribution, but the Poisson distribution is a good approximation when describing the statistical properties of X-rays, especially when the noise is relatively low. The benefit of using the Poisson noise model is that it simplifies the image reconstruction algorithm and interpretability, and it has been proven in many practical applications [63].

Table 2. Imaging conditions for the AAPM dataset. This table lists the imaging parameters for the AAPM dataset, including patient ID, number of slices, field of view (FOV) size, kilovolt peak (KVP), exposure time, and X-ray tube current (mA).

Patient ID	Numbers of Slices	Size of FOV	KVP	Exposure Time (ms)	X-ray Tube Current (mA)
L067	224	370	100	500	234.1
L096	330	430	120	500	327.6
L109	128	400	100	500	322.3
L143	234	440	120	500	416.9
L192	240	380	100	500	431.6
L286	210	380	120	500	328.9
L291	343	380	120	500	322.7
L310	214	380	120	500	300.0
L333	244	400	100	500	348.7
L506	211	380	100	500	277.7

2.2. Data Preprocessing

The data preprocessing part of this study aims to optimize the training process to better adapt to the input data requirements of the neural network model. For the original CT image size of 512×512, we took the following steps to process the data.

First, we introduce a key parameter patch-size, which defines the size of dividing small image patches. By dividing the image into smaller chunks, we are able to increase computational efficiency and allow the network to better learn local features. The actual size of the input image is 512×512, the size of the image block input to the network in the experiment is 64×64. Next, we preprocessed the image. First, we scaled the images to facilitate batch operations. We then convert the image data type to floating point to meet the input requirements of the neural network model. Regarding data shape conversion, we determine whether the patch-size parameter is defined based on conditions. If the patch-size parameter is set to a non-zero value, we perform a shape transformation operation on the image. By reshaping the image into patch-size patches, we can input each patch into the neural network as an independent sample. Through these data preprocessing steps, we effectively change the form of the original CT images to better suit the needs of the neural network model. This preprocessing can improve the effectiveness of network training and enable the network to better learn the local features of the image. Our data preprocessing process is key to improving model performance and training effectiveness.

2.3. Performance Metrics

Medical images contain more subtle structures and fewer channels than natural images, and appropriate evaluation metrics are crucial to evaluating LDCT images. We choose peak signal-to-noise ratio (*PSNR*), structural similarity (*SSIM*) and root mean square error (*RMSE*) as image quality evaluation metrics. In addition to these objective metrics, radiologist evaluations are also critical to the success of the denoising task. We will include actual radiologist evaluations in subsequent studies to support the validity of diagnosis based on denoised images.

2.3.1. PSNR

PSNR is an objective measure of the error between image pixels and is typically used for error-sensitive images. It is defined according to the mean square error (*MSE*), which is defined as

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [Y(i,j) - X(i,j)]^2 \tag{1}$$

where *MSE* represents the mean square error between the real image *Y* and the input noise image *X*, $X(i,j)$ and $Y(i,j)$ respectively correspond to the pixel values at the coordinates. m and n represent the height and width of the image, respectively. The smaller the *MSE*, the closer the two images are and the smaller the distortion. Correspondingly, *PSNR* is expressed as

$$PSNR = 10\log_{10}(\frac{(2^n-1)^2}{MSE}) \tag{2}$$

where n is the number of bits per pixel, which is generally 8. The larger the *PSNR* value, the smaller the distortion and the better the image effect [64].

2.3.2. SSIM

SSIM (structural similarity index) stands for structural similarity. It is an index used to measure the similarity of two images. It is better in line with human visual perception. The *SSIM* formula is based on three parameters between image *X* and *Y* : Luminance, Contrast, and Structure. The formula is as follows:

$$L(x,y) = \frac{2u_x u_y + c_1}{u_x^2 + u_y^2 + c_1} \tag{3}$$

$$C(x,y) = \frac{2\sigma_x \sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \tag{4}$$

$$S(x,y) = \frac{\sigma_x \sigma_y + c_3}{\sigma_x \sigma_y + c_3} \tag{5}$$

where U_x and U_y are the means of x and y respectively, σ_x and σ_y are the variances of x and y respectively, and σ_{xy} is the covariance of x and y, c_1, c_2 and c_3 are constants that make the result non-zero. The definition of *SSIM* can be obtained from the above three parameters:

$$SSIM(x,y) = [L(x,y)^\alpha * C(x,y)^\beta * S(x,y)^\gamma] \tag{6}$$

The structural similarity index defines structural information from the perspective of image composition, which reflects the structural information, brightness and contrast of the object. In the calculation of structural similarity, the mean is used as the estimate of brightness, the standard deviation is used as the estimate of contrast, and the covariance is used to measure the degree of structural similarity.

2.3.3. RMSE

Root mean square error (*RMSE*) is a common image quality evaluation metric, used to measure the degree of difference between the output image and the original image. It is defined by calculating the square root of the mean square error (*MSE*),

$$RMSE = \sqrt{MSE} \tag{7}$$

Mean squared error (*MSE*) calculates the squared difference between the pixels in two images and then averages all differences. *RMSE* is the square root of *MSE*. The smaller the value, the smaller the difference between the output image and the original image, indicating the better the processing effect.

2.4. Deep Learning Models

With the development of deep learning, medical image processing has attempted to use neural networks as a problem-solving tool. Deep learning has important roles including lesion detection and segmentation, disease prevention and diagnosis, etc. In these applications, clear medical images are crucial to solving problems. Deep learning methods have also shown good results on low-level tasks such as medical image denoising. In this paper, we evaluate seven deep learning models (REDCNN [34], EDCNN [57], QAE [58], OCTNet [59], UNet [59], WGAN [43], CTformer [60]), REDCNN, EDCNN, OCTNet, and UNet are CNN-based denoising methods, QAE s an autoencoder-based denoising method, WGAN is a GAN-based denoising method, and CTformer is a Transformer-based denoising method. Table 3 summarizes some features and parameters of the deep learning model we used.

Table 3. An overview of deep learning models. The table summarizes the deep learning models used in our study, detailing the reference numbers, the number of trainable parameters, and key features or remarks for each model.

Model	Ref	Trainable Parameters	Remarks	
REDCNN	[34]	1848865	(a)	Combine the autoencoder, deconvolution and shortcut connections into the ResNet.
EDCNN	[57]	80961	(a)	Design an edge enhancement module based on trainable Sobel convolution.
			(b)	Construct a dense connection to fuse edge features.
			(c)	Introduce the compound loss which integrates the MSE loss and multi-scale perceptual loss.
QAE	[58]	49818	(a)	Propose quadratic neurons by replacing the inner product.
			(b)	Encoder-decoder structure.
OCTNet	[59]	371073	(a)	Adopt multi-scale method to represent the CT denoising problem.
			(b)	Octave convolution proposed in CT image denoising.
UNet	[59]	7819201	(a)	Multiple residual connections are used for CT denoising.
WGAN	[43]	34071842	(a)	Introduce a new CT image denoising method based on GAN with Wasserstein distance.
			(b)	Use perceptual loss suppresses image noise.
CTFORMER	[60]	1448265	(a)	Propose a convolution-free Token2Token dilated vision Transformer.
			(b)	An overlapped inference mechanism effectively eliminate the boundary artifacts.

3. Experiment and Evaluation

This section shows the configuration of the experiments, presents the experimental results and brief analysis, and evaluates the LDCT image denoising performance of the deep learning models.

3.1. Experiment Design

We conducted four experiments aimed at performing denoising analysis on different deep learning models and evaluating their generalization capabilities as well as model complexity and inference speed.

In the first experiment, we performed denoising analysis using the AAPM dataset. We calculated the performance indicators of the region of interest and the enlarged image and evaluated their performance on the LDCT image denoising task. In the second experiment, to verify the generalization ability of the models on different datasets, we conducted a cross-experiment on the two datasets. The experiment allows comprehensive evaluation of the performance of the models on different datasets and verify their abilities to adapt to unseen data. In the third experiments, we conducted experiments based on the Piglet dataset. The experiment tested the generalization ability of the deep learning model to CT images with different noise levels. In the fourth experiment, we evaluated the complexity and inference speed of different models. The experiment analyzed indicators such as the number of parameters, computing resource requirements, and inference time of the models to find models with lower computational costs and efficient speed in practical applications. Through these four experiments, we can comprehensively evaluate the performance, generalization ability and computational efficiency of different deep learning models in LDCT image denoising, and explore the LDCT denoising method most suitable for this task.

3.2. Experiment Setup

We use the PyTorch 1.10 deep learning framework to implement all deep learning models, and the compilation environment for experimental is Python 3.8. We use NVIDIA RTX3090 24G GPU and Intel i9-10900X CPU to complete all model training and testing. During the optimization process, we use the Adam optimizer with default configuration and use 512×512 pixel size LDCT images as input. We set the learning rate to 0.00001, the batch size to 8, and conducted 200 rounds of iterative training to make the model converge. In our study, all models were trained from scratch, rather than being fine-tuned based on pre-trained models from other datasets. After training, we save the model with the best performance and evaluate it on the validation dataset. These settings were kept the same for all models.
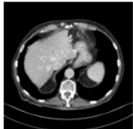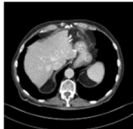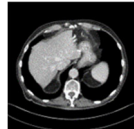
3.3. Experiment Result and Analysis

3.3.1. Performance Result of Deep Learning Models

To quantitatively analyze the denoising performance of the deep learning model, we use peak signal-to-noise ratio (*PSNR*), structural similarity (*SSIM*) and root-mean square error (*RMSE*) as objective metrics. Table 4 shows the denoising results of the deep learning model on the AAPM dataset, the test data comes from an abdominal image of patient L506. In this experiment, the size of the input image and output image of the AAPM dataset are both 512×512. During training, the image is divided into small image blocks. We extracted 123072 pairs of image blocks from 1923 CT images as training Input and label, size is 64×64. 29,120 pairs of image blocks were extracted from another 455 CT images for testing.

Table 4. Performance of different models on the AAPM dataset. The table compares the denoising performance of various deep learning models on the AAPM dataset using three metrics: PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), and RMSE (Root Mean Square Error). Higher PSNR and SSIM values indicate better denoising performance, while lower RMSE values are preferred. UNet achieved the highest PSNR score, suggesting its superior ability to retain overall image quality, while EDCNN obtained the best SSIM, highlighting its strength in preserving structural details. Visual examples of the denoised images are also provided to qualitatively compare the models' outputs. Bold indicates the best results.

Model	REDCNN	EDCNN	QAE	OCTNet	UNet	WGAN	CTformer
PSNR	31.6918	31.8518	28.1326	31.9020	**32.2510**	30.2021	31.4673
SSIM	0.8841	**0.8972**	0.8581	0.8853	0.8884	0.8359	0.8821
RMSE	10.6134	9.9714	15.9625	9.9126	**9.7624**	12.7829	10.6773
Image							
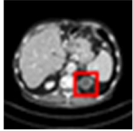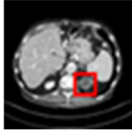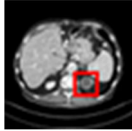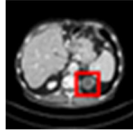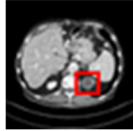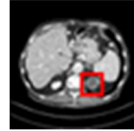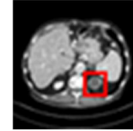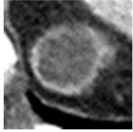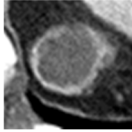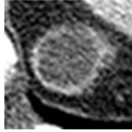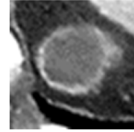
Quadratic Autoencoder is a special autoencoder network. The network introduces a quadratic term loss function, which improves image feature capabilities and makes network training more time-consuming. It's denoised image has more noise points, and its objective indicators such as PSNR and SSIM are the worst. WGAN's images can retain the texture information of the original image, but cannot completely remove stripe artifacts. Its PSNR and SSIM results are lower than those using MSE-loss as the loss function (REDCNN, EDCNN, UNet, OCTNet). At the same time, since REDCNN only uses the MSE loss function for training, the texture of the image is blurred. EDCNN adds a trainable Sobel operator for edge enhancement before training, so the edge details of its images are more prominent. It obtains better PSNR and the best SSIM on the AAPM dataset, with SSIM = 0.9009. The denoising effects of OCTNet and UNet are close, but the denoised images of OCTNet suffer from loss of details. CTformer uses the powerful feature extraction capability of the attention mechanism to remove noise in images and achieves excellent results. Most deep learning models use gradient loss as the loss function, which will pay more attention to the subtle structure of the image, but will smooth some areas of the LDCT image. UNet can effectively remove noise and stripe artifacts and maintain high structural simila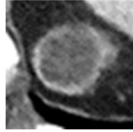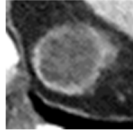rity with NDCT images. It has the highest peak signal-to-noise ratio and the smallest root mean square error, with PSNR = 32.2510 and RMSE = 9.7624.

Furthermore, we evaluate the performance metrics of regions of interest (ROI). As shown in Table 5, we zoomed in on the aorta in the chest image, the red box represents the ROI, and we calculated the test results of the local ROI and drew the ROI image. The enlarged ROI image that all models show varying degrees of denoising

effects. REDCNN and EDCNN based on the MSE loss function perform well on values and images, but have edge blur in details. The ROI image of QAE is still not good in experiments. OCTNet and UNet achieved good results with their large number of dense cascades and residual connections. The area of interest (aorta) of the two still maintains good structural similarity after amplification, and the edge information is not blurred. Although the visual effects are not as good as CTformer, it achieves the best numerical results. The WGAN network based on Wasserstein distance has texture blur at both the macro level and the micro level, which may be related to the instability of its training. The CTformer enlarged image has significant noise and blur at the edges, indicating that the ability of this type of model to process single-channel CT images has some limitations.

Table 5. Test results of the AAPM dataset on abdominal images, the red box is the region of interest ROI. This table shows the denoising performance of various models on abdominal CT images from the AAPM dataset using PSNR, SSIM, and RMSE metrics. UNet achieves the highest PSNR, indicating better overall image quality, while EDCNN achieves the highest SSIM, highlighting better structural preservation. Visualized results include both the full image and the zoomed-in ROI for a detailed comparison. Bold indicates the best results.

Model	REDCNN	EDCNN	QAE	OCTNet	UNet	WGAN	CTformer
PSNR	26.1672	26.3421	22.0271	26.5634	**26.6253**	23.7721	25.4271
SSIM	0.5669	**0.5676**	0.4831	0.5661	0.5653	0.5659	0.5422
RMSE	21.3844	21.1312	28.9304	21.4931	**21.0823**	26.7823	22.2216
Pred-img							
ROI-img							

We calculated the average PSNR and SSIM of ROI images on the AAPM dataset and Piglet dataset. As shown in Figures 2 and 3, the results show that UNet has the best PSNR and EDCNN has the best SSIM. In general, all models show certain denoising capabilities. Different networks have different problems, the output image exhibits differences in the visual effects. REDCNN and EDCNN have texture blur but rich colors and contrast. UNet performs well in terms of structure preservation and texture details, and the test results at the macro and micro levels are excellent. UNet's excellent denoising performance deserves further study for designing better models.
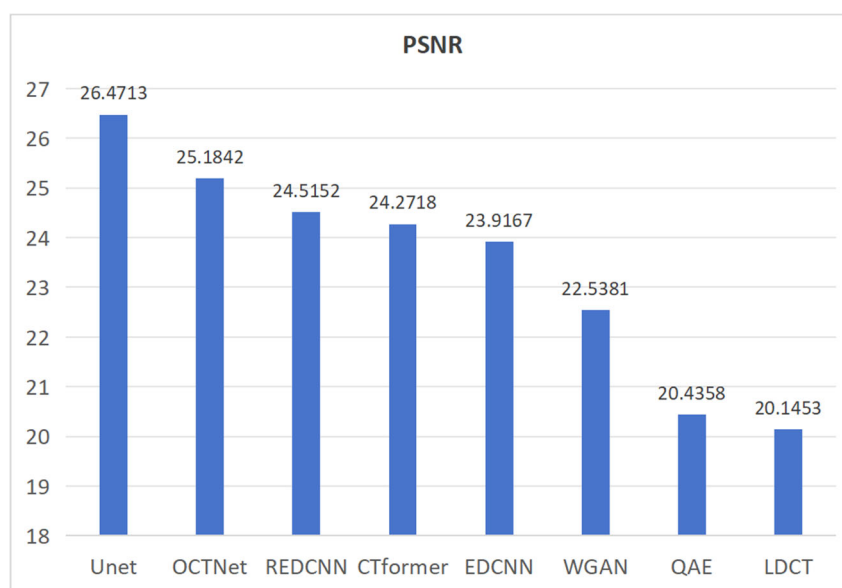

Figure 2. Average PSNR of ROI on two dataset. This figure compares the denoising performance of various models based on the average PSNR values for the ROI. Higher PSNR indicates better preservation of image quality after denoising, with UNet achieving the highest score.
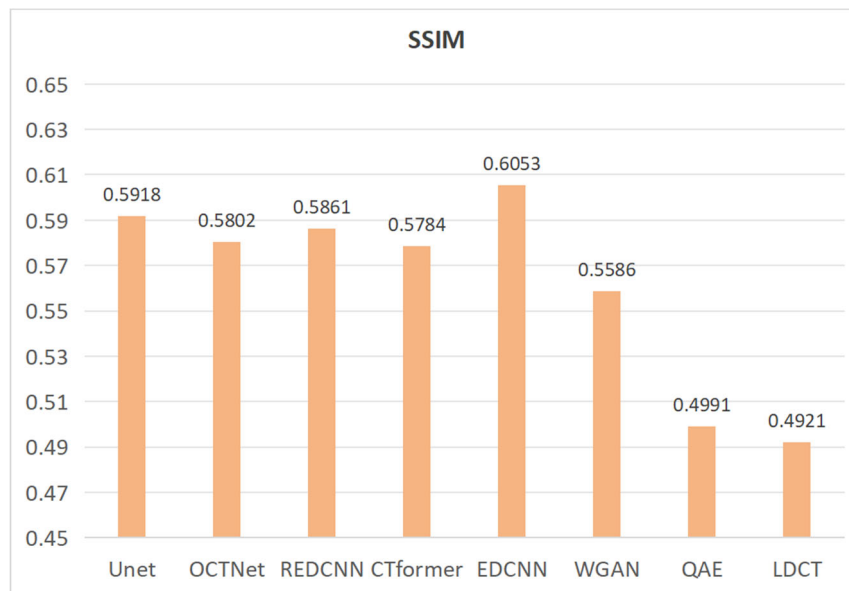
Figure 3. Average SSIM of ROI on two dataset. This figure shows the average SSIM values for the ROI, comparing the structural preservation capabilities of various models. Higher SSIM values indicate better structural similarity between the denoised and reference images, with EDCNN achieving the highest score.

3.3.2. Assessment of Generalization Performance

When using deep learning models to process LDCT denoising in actual clinical practice, their performance may be affected by acquisition parameters, equipment and other factors. To evaluate the generalization performance of deep learning models when processing new LDCT images, we tested their performance based on different datasets from two CT scanners. Specifically, we first train the model using CT images from one CT scanner, and then, we test the model using data from another CT scanner.

As shown in Table 6, when the model is trained using the AAPM dataset, the best results on the AAPM dataset are PSNR = 33.0712, SSIM = 0.9221, and the best results on the Piglet dataset are PSNR = 27.9170, SSIM = 0.8616. When the model is trained using the Piglet dataset, the best results on the Piglet dataset are PSNR = 31.2192, SSIM = 0.8969, and the best results on the AAPM dataset are PSNR = 28.9191, SSIM = 0.8615. Among all models, OCTNet achieved better results in the model generalization performance test. The results show that the deep learning model performs better when the test and training data come from CT images from the same CT scanner. In summary, it is difficult to obtain the same performance when using a trained deep learning model to test new LDCT images.

Table 6. Image quality evaluation results of the model on two datasets. This table displays PSNR and SSIM results for each model when trained and tested on AAPM and Piglet datasets, highlighting their performance on both same-dataset and cross-dataset scenarios. UNet excels in same-dataset tests, while OCTNet shows better cross-dataset generalization. Bold indicates the best results.

Model	Train: AAPM Test: AAPM		Train: AAPM Test: PIGLET		Train: PIGLET Test: PIGLET		Train: PIGLET Test: AAPM	
INDEX	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
LDCT	29.2454	0.8732	25.0054	0.7234	25.0054	0.7234	29.2454	0.8732
REDCNN	32.3221	0.9103	27.3561	0.8371	30.6551	0.8965	28.5451	0.8472
EDCNN	32.9791	0.9037	27.0181	0.8231	31.1081	0.8899	28.0169	0.8349
QAE	29.2291	0.8759	25.0063	0.7762	25.0172	0.7881	26.0071	0.7876
OCTNet	32.7813	0.9082	**27.9170**	**0.8616**	31.1102	0.8971	**28.9191**	**0.8615**
UNet	**33.0712**	**0.9221**	27.6974	0.8421	**31.2192**	**0.8969**	28.6885	0.8571
WGAN	30.5192	0.8882	26.9159	0.8334	27.8292	0.8251	27.6601	0.8102
CTformer	32.2071	0.9092	27.0331	0.8264	30.3482	0.8020	28.0330	0.8351

3.3.3. Assessment of Training Strategies

To further explore the denoising ability of the deep learning model on CT images with different noise levels, we conducted the following experiments. Observe the performance of various deep learning models by varying the radiation dose. We use the Piglet dataset to train the model and test it on LDCT images with different noise

levels. Less radiation dose means more noise. Figure 4 shows CT images of the Piglet dataset denoised using different methods. Table 7 shows the objective indicators of various methods.
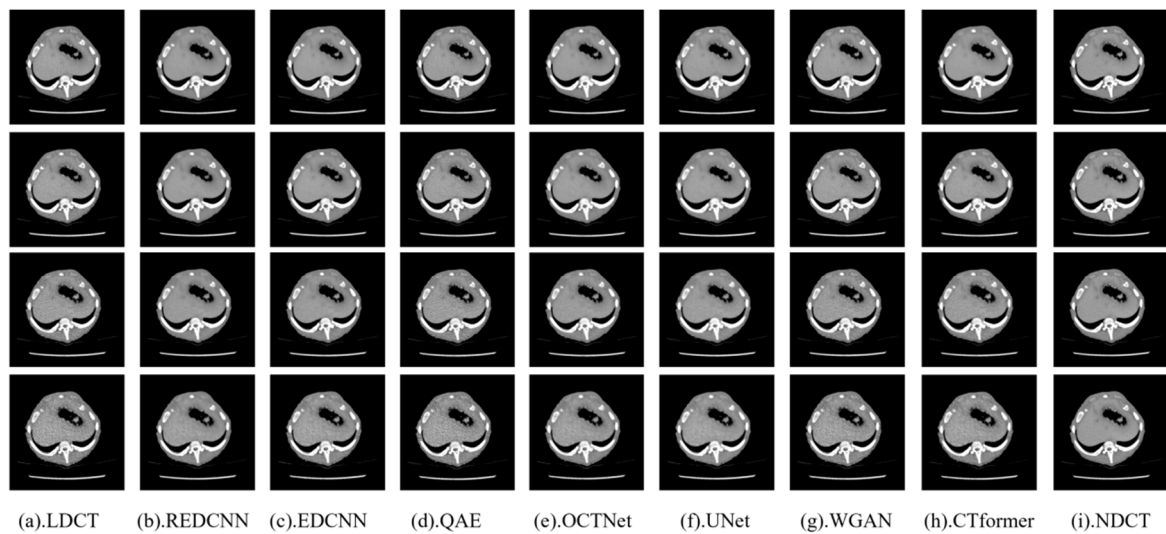

(a).LDCT   (b).REDCNN   (c).EDCNN   (d).QAE   (e).OCTNet   (f).UNet   (g).WGAN   (h).CTformer   (i).NDCT

Figure 4. The above figure is a visualization of four low doses CT images in the Piglet dataset using different methods to denoise. The first row is LDCT (50% of full dose reconstructed by FBP). The second row is LDCT (25% of full dose reconstructed by FBP). The third row is LDCT (10% of full dose reconstructed by FBP). The last row is LDCT (5% of full dose reconstructed by FBP). The last column NDCT is reconstructed by the FBP algorithm with a 100% dose.

Table 7. Test results when trained on the Piglet dataset with 100% dose images. This table shows the PSNR, SSIM, and RMSE results for models tested on CT images with different dose levels (50%, 25%, 10%, and 5%). It highlights how each model adapts to lower dose levels, with EDCNN and UNet showing superior results at various dose reductions. Bold indicates the best results.

Model	50% Dose			25% Dose			10% Dose			5% Dose		
	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM	RMSE
LDCT	31.0736	0.8771	11.1785	28.0292	0.8414	15.8708	26.7750	0.8114	18.3361	24.1211	0.7779	24.8889
REDCNN	28.5556	0.8940	14.9375	28.2353	0.8935	15.1827	31.4215	0.8963	10.7395	28.2437	0.8949	13.9151
EDCNN	**32.8634**	0.9004	**9.9734**	**30.7741**	0.8877	**12.1470**	31.5269	0.9077	10.5100	29.7512	**0.8962**	12.8965
QAE	31.0968	0.8769	11.1486	28.0512	0.8411	15.8307	26.7689	0.8113	18.3491	25.0084	0.7893	22.9941
OCTNet	31.5507	0.9041	12.1175	30.7046	0.8904	12.4709	31.5017	0.9017	10.6408	28.8929	0.8846	14.3683
UNet	32.2654	**0.9198**	10.6591	30.2083	**0.8999**	12.1786	**31.5763**	**0.9066**	10.5206	**29.9512**	0.8874	**12.6204**
WGAN	29.2374	0.8226	13.1043	29.0836	0.8812	14.0566	28.8378	0.8620	14.4600	26.0153	0.8141	20.0122
CTformer	29.8627	0.8720	12.1072	29.8460	0.8944	13.2094	29.3495	0.8605	13.6329	28.7136	0.8469	14.9342

As shown in Figure 4, in the first column, with the gradual reduction of radiation dose, the noise in CT images increases significantly. When the dose is 5% of the full dose, noise seriously affects the visual effect of CT images. Notably, the RED-CNN can remove noise to some extent, but its images inevitably exhibit a smoothing effect and lose some details. WGAN performs well when noise is low, but when it processes LDCT images with the highest noise levels, its denoised images will have many noise points. In contrast, EDCNN, OCTNet, UNet and CTformer denoise CT images of different doses, and the image quality obtained is significantly better than other algorithms. The the model based on MSE-loss and compound loss function performs well in obtaining LDCT denoised images at different radiation doses. The denoised image can retain the rich details and texture structure of the CT image well. Among all models, UNet has better visual effects on denoised images for all four dose levels.

We quantitatively analyze the denoising performance of different algorithms for different LDCT images. We calculated three objective indicators of the experimental results. The summary data are in Table 7. Of note, the LDCT image with a radiation dose of 50%, PSNR and SSIM are higher, which indicates that the 50% reduction of radiation dose has little effect on image quality. However, for images with a radiation dose of 5%, the values of the three metrics decreased significantly. That is, the image quality of Figure 4 becomes worse as the dose

decreases. The data show that UNet and EDCNN have better results in most cases, and both of them are ranked first and second in four rounds of tests (bold numbers represent the best, italics numbers represent the second best.). UNet has been ranked first in three rounds of tests many times, indicating that UNet has some robustness. Therefore, UNet performs well in denoising experiments on CT images with different noise levels.

3.3.4. Model Complexity Evaluation

Model efficiency is an important issue in deep learning. An excellent deep learning model should have both excellent denoising capabilities and fast inference speed. Based on the above criteria, we compared the number of trainable parameters (params), memory usage (MACs), and inference speed based on different devices (CPU and GPU) of the seven models. The experiment completes all experiments using Intel i9-10900X CPU and NVIDIA RTX3090 24G GPU.

QAE uses 15 3 × 3 kernels in each convolutional layer, while REDCNN has 32 5 × 5 kernels, which means that REDCNN has 4 times more parameters than QAE. As shown in Table 8, WGAN occupies the largest real-time memory. WGAN includes a generator and a discriminator and uses perceptual loss as the loss function. Its trainable parameter amount and memory usage are the highest, which makes it difficult to deploy the whole model. On the contrary, EDCNN has a relatively small number of parameters but a high memory usage, indicating that the network can effectively fuse image information, which can also explain its better PSNR. In addition to this, models process images with a CPU takes much longer than with a GPU.

Table 8. CPU computation speed and GPU computation speed for the two datasets on seven models. This table presents the parameter count, MACs, and computation times on both CPU and GPU for seven models, along with their average PSNR scores. It highlights the efficiency and speed differences between models when processing the AAPM and Piglet datasets, with CTformer achieving the fastest CPU computation time on both datasets. Bold indicates the best results.

Model	Params	MACs (G)	AAPM-Dataset		Piglet-Dataset		Avg-PSNR
			CPU Times (ms)	GPU Times (ms)	CPU Times (ms)	GPU Times (ms)	
REDCNN	1848865	4.3	3182.1	12.1	109.2	5.2	31.4277
EDCNN	80961	5.2	571.6	5.9	159.1	3.7	32.0321
QAE	49818	2.5	1024.4	**4.8**	315.6	**2.8**	27.1272
OCTNet	371073	3.1	684.2	13.4	229.3	6.5	31.8922
UNet	7819201	4.9	726.3	11.4	238.2	4.6	**32.4521**
WGAN	34071842	6.4	3682.1	28.2	1317.2	15.6	29.1281
CTformer	1448265	6.2	**531.8**	9.4	**134.6**	3.5	31.3180

In our experiments, the number of images in the AAPM and piglet test datasets were 1923 and 720, respectively. We calculated the processing speed of a single 512 × 512 LDCT image. QAE has the fastest inference speed, with a single image taking 3.9 ms. WGAN is the slowest, which also means that training WGAN takes more time. Therefore, we plotted a scatter plot of the model's inference time versus PSNR, as shown in Figure 5, with the best results in the upper left. The results show that only UNet and EDCNN have PSNRs exceeding 32. Among them, UNet has the highest PSNR, but UNet's inference time is slower. EDCNN achieves a balance between inference time and de-noising effect. Therefore, in terms of calculation speed and denoising performance, EDCNN is the strongest competitor compared to other models.
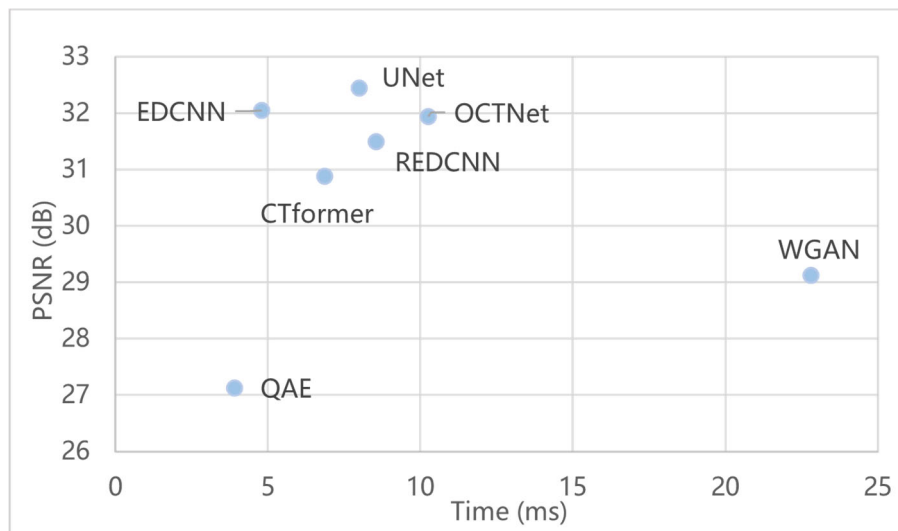
Figure 5. Performance results and inference speed of different deep learning models. This figure illustrates the trade-off between denoising performance (measured by PSNR) and inference speed (time in ms) for various models, highlighting the balance between image quality and computational efficiency.

4. Discussion

Deep learning already occupies a significant position in medical image processing. In LDCT image denoising research, many studies have different experimental conditions and training strategies. To accurately judge and compare the denoising performance of networks, in our study, we trained and evaluated seven deep learning models under the same conditions and studied their training strategies.

In the denoising performance analysis, UNet has the best effect. Through multi-layer residual connections, UNet can extract more image information and obtain the best results, with a PSNR of 32.25. The successful performance of the UNet architecture is due in part to its features specifically designed for biomedical image segmentation, including efficient utilization of small amounts of training data. However, it is worth noting that regardless of the size of the dataset, UNet is also likely to improve performance due to its efficient architecture. In our first experiment (3.1.1) and third experiment (3.1.3), the two experiments are based on different datasets and the number of images in the datasets is different. In both experiments, UNet achieved excellent results. Therefore, UNet's own efficient architecture is the main reason for its success in small datasets. Similarly, REDCNN also achieves good performance using symmetric encoders and decoders. EDCNN introduces a trainable sobel operator before the residual connection to enhance the edge information of the output image, thus achieving better results. The LDCT denoising network based on CNN showed better performance, while the denoising network based on GAN and transformer was overall inferior to the denoising network based on CNN. Additionally, we zoom in on the region of interest to focus on detail recovery and edge information. Regarding detail recovery, EDCNN's denoised image shows no obvious noise after enlargement, and the edge details of blood vessels are not blurred, indicating that its denoising effect is better than other models. Overall, the top-performing models did not exhibit significant differences in structural similarity.

In deep learning, generalization performance is one of the important metrics for evaluating the stability of the model. In our study, we conducted cross-experiments to explore the model's generalization ability on different datasets. UNet shows excellent denoising performance on the same dataset, while OCTNet has better denoising performance on different datasets. Therefore, OCTNet has stronger generalization ability than other models.

In LDCT image processing, model running speed is also an important metric. Experimental results show that QAE has the fastest computing speed, which is consistent with its size, but its denoising performance is poor. WGAN consists of a generator and a discriminator and uses the VGG network as a feature extractor and complex loss function, and its calculation speed is the slowest. Overall, EDCNN balances computational speed and denoising performance.

Our work has some shortcomings that we hope to address in the future. First, in this study, we employed the AAPM dataset and the Piglet dataset. The AAPM dataset primarily contains contrast-enhanced abdominal CT images, while the Piglet dataset consists of low-dose CT images of experimental pigs obtained by reducing the tube current. These two datasets represent different anatomical regions—human abdomen and experimental pig—providing an opportunity to evaluate the applicability of the models across varying anatomical areas. Further validation of the models' performance on CT data from different anatomical regions allows comprehensive

assessment of the denoising methods' generalizability and applicability. In future research, we plan to test these denoising methods on CT datasets from other anatomical regions, such as the brain and chest, to more thoroughly evaluate their effectiveness and explore their potential in various clinical scenarios. Secondly, generalizability is a key issue. When evaluating the generalization ability of denoising methods, we did not consider their versatility. For example, a method trained on brain CT scans should also be applicable to chest CT scans. In future research, we plan to test these denoising methods on CT datasets from more anatomical regions (such as the brain, chest, etc.) to further explore their transferability and applicability across different anatomical regions. Moreover, transfer learning can utilize models pre-trained on large-scale datasets to provide better initial feature representations for other related tasks, thereby improving model training efficiency and generalization ability. This is especially recognized in medical image analysis [65–67]. In our study, although all models were trained from scratch, we acknowledge the potential advantages of transfer learning in enhancing model performance. In future research, we plan to evaluate the effectiveness of transfer learning in LDCT denoising, such as using models pre-trained on other CT image tasks as initial models and observing their impact on LDCT denoising. We believe that this exploration could further improve the denoising performance of the models and validate their broader applicability across different datasets and tasks. Finally, although we have conducted many sets of experiments, considering many top deep learning models, We need to continue updating the latest neural network models to take advantage of the new deep learning advancement.

5. Conclusion

In our study, we implemented and evaluated the performance and efficiency of seven LDCT denoising models. The results show that UNet has the best performance in terms of PSNR, due to its multi-layer residual connected encoder. The output image of EDCNN is most similar to the original image and has the highest structural similarity. UNet has better denoising effect, but the calculation time is longer, which will increase the time consumption in actual clinical processing. In contrast, EDCNN can balance performance and efficiency, which has potential for practical applications. In addition, to assess the model's performance on new data, we evaluated its generalization performance, providing a benchmark for future research.

Author Contributions

All authors contributed to the study's conception and design. Material preparation and data collection were performed by H.Z., Y.Z., L.Q., and D.T. The first draft of the manuscript was written by H.Z. and D.T. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

We are using a public dataset, and all the data used is publicly available.

Conflicts of Interest

The authors have no relevant financial or non-financial interests to disclose.

References

1. Zhang, Z.; Yu, L.; Liang, X.; Zhao, W.; Xing, L. TransCT: Dual-path transformer for low dose computed tomography. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, 27 September–1 October 2021; Part VI 24, pp. 55–64.
2. Jiang, H. *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*; SPIE: Bellingham, WA, USA, 2009.

3.   Brenner, D.J.; Hall, E.J. Computed Tomography — An Increasing Source of Radiation Exposure. *N. Engl. J. Med.* **2007**, *357*, 2277–2284.

4.   de Gonzalez, A.B.; Darby, S. Risk of cancer from diagnostic X-rays: Estimates for the UK and 14 other countries. *Lancet* **2004**, *363*, 345–351.

5.   Naidich, D.P.; Marshall, C.H.; Gribbin, C.; Arams, R.S.; McCauley, D.I. Low-dose CT of the lungs: preliminary observations. *Radiology* **1990**, *175*, 729–731.

6.   Yin, X.; Coatrieux, J.-L.; Zhao, Q.; Liu, J.; Yang, W.; Yang, J.; Quan, G.; Chen, Y.; Shu, H.; Luo, L. Domain Progressive 3D Residual Convolution Network to Improve Low-Dose CT Imaging. *IEEE Trans. Med Imaging* **2019**, *38*, 2903–2913.

7.   Han, Y.S.; Yoo, J.; Ye, J.C. Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis. *arXiv* 2016, arXiv:1611.06391.

8.   Chen, Y.; Yin, X.; Shi, L.; Shu, H.; Luo, L.; Coatrieux, J.-L.; Toumoulin, C. Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing. *Phys. Med. Biol.* **2013**, *58*, 5803–5820.

9.   Thanh, D.; Surya, P.; Hieu, L.M. A Review on CT and X-Ray Images Denoising Methods. *Informatica* **2019**, *43*, 151–159.

10. Diwakar, M.; Kumar, M. A review on CT image noise and its denoising. *Biomed. Signal Process. Control.* **2018**, *42*, 73–88.

11. Wang, H.; Chi, J.; Wu, C.; Yu, X.; Wu, H. Degradation adaption localto-global transformer for low-dose CT image denoising. *J. Digit. Imaging* **2023**, *36*, 1894–1909.

12. Chen, Z.; Gao, Q.; Zhang, Y.; Shan, H. Ascon: Anatomy-aware supervised contrastive learning framework for low-dose CT denoising. In Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention, Vancouver, BC, Canada, 8–12 October 2023; pp. 355–365.

13. Manduca, A.; Yu, L.; Trzasko, J.D.; Khaylova, N.; Kofler, J.M.; McCollough, C.M.; Fletcher, J.G. Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT. *Med. Phys.* **2009**, *36*, 4911–4919.

14. Kachelriess, M.; Watzke, O.; Kalender, W.A. Generalized multidimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT. *Med. Phys.* **2001**, *28*, 475–490.

15. Hsieh, J. Adaptive streak artifact reduction in computed tomography resulting from excessive X-ray photon noise. *Med. Phys.* **1998**, *25*, 2139–2147.

16. Wang, J.; Li, T.; Lu, H.; Liang, Z. Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography. *IEEE Trans. Med. Imaging* **2006**, *25*, 1272–1283.

17. Zeng, D.; Huang, J.; Bian, Z.; Niu, S.; Zhang, H.; Feng, Q.; Liang, Z.; Ma, J. A Simple Low-Dose X-Ray CT Simulation from High-Dose Scan. *IEEE Trans. Nucl. Sci.* **2015**, *62*, 2226–2233.

18. Fletcher, J.G.; Grant, K.L.; Fidler, J.L.; Shiung, M.; Yu, L.; Wang, J.; Schmidt, B.; Allmendinger, T.; McCollough, C.H. Validation of dual source single-tube reconstruction as a method to obtain half-dose images to evaluate radiation dose and noise reduction: Phantom and human assessment using CT colonography and sinogram-affirmed iterative reconstruction (safire). *J. Comput. Assist. Tomogr.* **2012**, *36*, 560–569.

19. Pickhardt, P.J.; Lubner, M.G.; Kim, D.H.; Tang, J.; Ruma, J.A.; del Rio, A.M.; Chen, G.-H. Abdominal CT with Model-Based Iterative Reconstruction (MBIR): Initial Results of a Prospective Trial Comparing Ultralow-Dose with Standard Dose Imaging. *Am. J. Roentgenol.* **2012**, *199*, 1266–1274.

20. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88.

21. Kaur, P.; Singh, G.; Kaur, P. A review of denoising medical images using machine learning approaches. *Curr. Med. Imaging* **2018**, *14*, 675–685.

22. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; Volume 2, pp. 60–65.

23. Balda, M.; Hornegger, J.; Heismann, B. Ray Contribution Masks for Structure Adaptive Sinogram Filtering. *IEEE Trans. Med Imaging* **2012**, *31*, 1228–1239.

24. Mallat, S.G. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693.

25. Yu, F.; Chen, Y.; Luo, L. CT image denoising based on sparse representation using global dictionary. In Proceedings of the 2013 ICME International Conference on Complex Medical Engineering, Beijing, China, 25–28 May 2013; pp. 408–411.

26. Chen, Y.; Yang, Z.; Hu, Y.; Yang, G.; Zhu, Y.; Li, Y.; Luo, L.; Chen, W.; Toumoulin, C. Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means. *Phys. Med. Biol.* **2012**, *57*, 2667–2688.

27. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095.

28. Hashemi, S.; Paul, N.S.; Beheshti, S.; Cobbold, R.S.C. Adaptively Tuned Iterative Low Dose CT Image Denoising. *Comput. Math. Methods Med.* **2015**, *2015*, 638568.

29. Ha, S.; Mueller, K. Low dose CT image restoration using a database of image patches. *Phys. Med. Biol.* **2015**, *60*, 869–882.

30. Zhang, Z.; Han, X.; Pearson, E.; Pelizzari, C.; Sidky, E.Y.; Pan, X. Artifact reduction in short-scan CBCT by use of optimization-based reconstruction. *Phys. Med. Biol.* **2016**, *61*, 3387–3406.

31. Chen, H.; Zhang, Y.; Kalra, M.K.; Lin, F.; Chen, Y.; Liao, P.; Zhou, J.; Wang, G. Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network. *IEEE Trans. Med. Imaging* **2017**, *36*, 2524–2535.

32. Shan, H.; Padole, A.; Homayounieh, F.; Kruger, U.; Khera, R.D.; Nitiwarangkul, C.; Kalra, M.K.; Wang, G. Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. *Nat. Mach. Intell.* **2019**, *1*, 269–276.

33. Kang, E.; Chang, W.; Yoo, J.; Ye, J.C. Deep Convolutional Framelet Denosing for Low-Dose CT via Wavelet Residual Network. *IEEE Trans. Med Imaging* **2018**, *37*, 1358–1369.

34. Chen, H.; Zhang, Y.; Zhang, W.; Liao, P.; Li, K.; Zhou, J.; Wang, G. Low dose CT denoising with convolutional neural network. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 143–146.

35. Zhang, Y.; Tian, Y.; Kong, Y.; Zhong, B.; Fu, Y. Residual dense network for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2472–2481.

36. Rai, S.; Bhatt, J.S.; Patra, S.K. Augmented Noise Learning Framework for Enhancing Medical Image Denoising. *IEEE Access* **2021**, *9*, 117153–117168.

37. Rai, S.; Bhatt, J.S.; Patra, S.K. Accessible, affordable and low-risk lungs health monitoring in COVID-19: Deep cascade reconstruction from degraded lr-uldct. In Proceedings of the 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), Kolkata, India, 28–31 March 2022; pp. 1–5.

38. Choi, K.; Vania, M.; Kim, S. Semi-supervised learning for lowdose CT image restoration with hierarchical deep generative adversarial network (hd-gan). In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 2683–2686.

39. Wang, L.; Gao, Q.; Meng, M.; Li, S.; Zhu, M.; Li, D.; Chen, G.; Zeng, D.; Xie, Q.; Zhao, Q.; et al. Semi-supervised noise distribution learning for low-dose CT restoration. *Med. Imaging 2020 Phys. Med. Imaging* **2020**, *11312*, 1026–1030.

40. Bizopoulos, P.; Vretos, N.; Daras, P. Comprehensive comparison of deep learning models for lung and COVID-19 lesion segmentation in CT scans. *arXiv* 2020, arXiv:2009.06412, 2020.

41. Shahidi, F.; Daud, S.M.; Abas, H.; Ahmad, N.A.; Maarop, N. Breast Cancer Classification Using Deep Learning Approaches and Histopathology Image: A Comparison Study. *IEEE Access* **2020**, *8*, 187531–187552.

42. Yi, X.; Babyn, P. Sharpness-Aware Low-Dose CT Denoising Using Conditional Generative Adversarial Network. *J. Digit. Imaging* **2018**, *31*, 655–669.

43. Yang, Q.; Yan, P.; Zhang, Y.; Yu, H.; Shi, Y.; Mou, X.; Kalra, M.K.; Zhang, Y.; Sun, L.; Wang, G. Low-Dose CT Image Denoising Using a Generative Adversarial Network with Wasserstein Distance and Perceptual Loss. *IEEE Trans. Med. Imaging* **2018**, *37*, 1348–1357.

44. Nishio, M.; Nagashima, C.; Hirabayashi, S.; Ohnishi, A.; Sasaki, K.; Sagawa, T.; Hamada, M.; Yamashita, T. Convolutional auto-encoder for image denoising of ultra-low-dose CT. *Heliyon* **2017**, *3*, e00393.

45. Liu, Y.; Zhang, Y. Low-dose CT restoration via stacked sparse denoising autoencoders. *Neurocomputing* **2018**, *284*, 80–89.

46. Liu, H.; Liao, P.; Chen, H.; Zhang, Y. ERA-WGAT: Edge-enhanced residual autoencoder with a window-based graph attention convolutional network for low-dose CT denoising. *Biomed. Opt. Express* **2022**, *13*, 5775–5793.

47. Wang, D.; Xu, Y.; Han, S.; Yu, H. Masked autoencoders for low-dose CT denoising. In Proceedings of the 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), Cartagena, Colombia, 18–21 April 2023; pp. 1–4.

48. Li, M.; Hsu, W.; Xie, X.; Cong, J.; Gao, W. SACNN: Self-Attention Convolutional Neural Network for Low-Dose CT Denoising With Self-Supervised Perceptual Loss Network. *IEEE Trans. Med. Imaging* **2020**, *39*, 2289–2301.

49. Karimi, D.; Dou, H.; Warfield, S.K.; Gholipour, A. Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis. *Med. Image Anal.* **2020**, *65*, 101759.

50. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–11.

51. Luthra, A.; Sulakhe, H.; Mittal, T.; Iyer, A.; Yadav, S. Eformer: Edge enhancement based transformer for medical image denoising. *arXiv* 2021, arXiv:2109.08044.

52. Yuan, J.; Zhou, F.; Guo, Z.; Li, X.; Yu, H. HCformer: Hybrid CNN-Transformer for LDCT Image Denoising. *J. Digit. Imaging* **2023**, *36*, 2290–2305.

53. Chyophel Lepcha, D.; Goyal, B.; Dogra, A. Low-dose CT image denoising using sparse 3dD transformation with probabilistic non-local means for clinical applications. *Imaging Sci. J.* **2023**, *71*, 97–109.

54. Othman, A.E.; Brockmann, C.; Yang, Z.; Kim, C.; Afat, S.; Pjontek, R.; Nikoubashman, O.; Brockmann, M.A.; Nikolaou, K.; Wiesmann, M.; et al. Impact of image denoising on image quality, quantitative parameters and sensitivity of ultra-low-dose volume perfusion CT imaging. *Eur. Radiol.* **2015**, *26*, 167–174.

55. Kulathilake, K.A.S.H.; Abdullah, N.A.; Sabri, A.Q.M.; Lai, K.W. A review on Deep Learning approaches for low-dose Computed Tomography restoration. *Complex Intell. Syst.* **2021**, *9*, 2713–2745.

56. Mück, J.; Reiter, E.; Klingert, W.; Bertolani, E.; Schenk, M.; Nikolaou, K.; Afat, S.; Brendlin, A.S. Towards safer imaging: A comparative study of deep learning-based denoising and iterative reconstruction in intraindividual low-dose CT scans using an in-vivo large animal model. *Eur. J. Radiol.* **2023**, *171*, 111267.

57. Liang, T.; Jin, Y.; Li, Y.; Wang, T. Edcnn: Edge enhancement-based densely connected network with compound loss for low-dose CT denoising. In Proceedings of the 2020 15th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 6–9 December 2020; Volume 1, pp. 193–198.

58. Fan, F.; Shan, H.; Kalra, M.K.; Singh, R.; Qian, G.; Getzin, M.; Teng, Y.; Hahn, J.; Wang, G. Quadratic Autoencoder (Q-AE) for Low-Dose CT Denoising. *IEEE Trans. Med. Imaging* **2019**, *39*, 2035–2050.

59. Won, D.K.; An, S.; Park, S.H.; Ye, D.H. Low-dose CT denoising using octave convolution with high and low frequency bands. In *Predictive Intelligence in Medicine: Third International Workshop, PRIME 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, 8 October 2020*; Springer: Cham, Switzerland, 2020; pp. 68–78.

60. Wang, D.; Fan, F.; Wu, Z.; Liu, R.; Wang, F.; Yu, H. CTformer: convolution-free Token2Token dilated vision transformer for low-dose CT denoising. *Phys. Med. Biol.* **2023**, *68*, 065012.

61. AAPM. Low Dose CT Grand Challenge. 2017. Available online: http://www.aapm.org/grandchallenge/lowdosect/ (accessed on 2 August 2024).

62. Yang, L.; Shangguan, H.; Zhang, X.; Wang, A.; Han, Z. High-Frequency Sensitive Generative Adversarial Network for Low-Dose CT Image Denoising. *IEEE Access* **2019**, *8*, 930–943.

63. Lee, S.; Lee, M.S.; Kang, M.G. Poisson–Gaussian Noise Analysis and Estimation for Low-Dose X-ray Images in the NSCT Domain. *Sensors* **2018**, *18*, 1019.

64. Liu, H.; Jin, X.; Liu, L. Low-Dose CT Image Denoising Based on Improved DD-Net and Local Filtered Mechanism. *Comput. Intell. Neurosci.* **2022**, *2022*, 2692301.

65. Yu, X.; Wang, J.; Hong, Q.Q.; Teku, R.; Wang, S.H.; Zhang, Y.D. Transfer learning for medical images analyses: A survey. *Neurocomputing* **2022**, *489*, 230–254.

66. Huang, C.; Wang, J.; Wang, S.H.; Zhang, Y.D. Applicable artificial intelligence for brain disease: A survey. *Neurocomputing* **2022**, *504*, 223–239.

67. Tian, D.; Zhu, B.; Wang, J.; Kong, L.; Gao, B.; Wang, Y.; Xu, D.; Zhang, R.; Yao, Y. Brachial plexus nerve trunk recognition from ultrasound images: A comparative study of deep learning models. *IEEE Access* **2022**, *10*, 82003–82014.