A CBCT image noise reduction method based on cGAN

Yu Miao (custmiao@126.com)
changchun university of science and technology https://orcid.org/0000-0001-9513-3292
Yifei Ding
changchun university of science and technology
Xianchun Zhu
Jilin University Stomatology Hospital
Weili Shi
Changchun University of Science and Technology
Jiashi Zhao
Changchun University of Science and Technology
Ke Zhang
Changchun University of Science and Technology
Zhengang Jiang
Changchun University of Science and Technology
Huamin Yang
Changchun University of Science and Technology

Research

Keywords: Noise reduction, Gradient loss, GAN, LDCT

Posted Date: April 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-19128/v1

A CBCT image noise reduction method based on cGAN

Yu Miao [a], Yifei Ding [a], Xianchun Zhu [b], Weili Shi [a], Jiashi Zhao [a], Ke Zhang [a], Zhengang Jiang [a*], and Huamin Yang [a*]

[a] Changchun University of Science and Technology, [b] Jilin University Stomatology Hospital

*Correspondence: jiangzhengang@cust.edu.cn and yhm@cust.edu.cn, School of Computer Science and Technology, Changchun University of Science and Technology, No.7089, Weixing Road, Changchun, China

Abstract

Background: As an imaging modality, cone beam CT (CBCT) is widely used in dentistry, which can help dentists to observe tissues such as roots and jaws without confusion. CBCT has the advantages of convenience and low radiation dose, however low contrast and large noise are the serious points in the images. These disadvantages make it extremely difficult for doctors to accurately identify target tissues. Due to the differences in scanning methods and reconstruction algorithms between CBCT and multi-row detector spiral CT (MDCT)，the current CT noise reduction models have significant shortcomings when reduce the noise on CBCT images.

Methods: In this paper, we propose a method of image noise reduction based on conditional generative adversarial network to improve the quality of CBCT images. The normal-dose MDCT images are used as the ground truth images to train the model to generate the denoise images.

Results: In order to increase the model's sensitivity to the gradient information, a gradient loss function is involved in our proposed method. The verification experiments on the simulated data set and the real data set show that our model effectively generates the denoise images as well as preserves the quality of the images.

Conclusions: We compared the denoising effect between our model and other models with different loss functions. The scores by PSNR, MS-SSIM and GMSD showed that our model had better edge characteristics and denoising effect.

Keywords: Noise reduction, Gradient loss, GAN, LDCT

Background

With the development of CT equipment and technology, low dose CT (LDCT) imaging is becoming more and more popular [1]. CBCT (Cone beam CT), as a

representative LDCT, has been widely used in oral implants, orthodontics, etc. [2]. CBCT's advantages are lower dose, faster imaging speed and lower cost. However, due to its lower dose, the image quality is poor and the noise will be increased during the reconstruction.

In the recent years, machine learning has made the rapid progress in image processing. And it has been widely used in image noise reduction. For visible images, O.Kupyn et al. proposed DeblurGAN [3] to remove motion blur in images. Its model improved Generative Adversarial Networks (GAN) [4] and used the discriminator of patchGAN [5]. Other research groups used Generative Networks to reduce the noise for medical images. Yi et al. proposed Sharpness Aware Generative Adversarial Network (SAGAN), adding the detection network results into the loss function to train the network[6]. Wang et al. proposed a 3D conditional Generative Adversarial Network(cGAN), which can be directly trained by using 3D PET data [7]. Chen Hu et al. proposed RED-CNN to reduce the noise for LDCT images [8], which has greatly improved the training speed compared to other models. There are some disadvantages such as low radiation dose and high noise level of CBCT images. An additional difficulty is that CBCT and Multi-Detector row spiral CT (MDCT) have different hardware systems and reconstruction algorithms. The current models, however, have some limitations such as noise misidentification and incomplete noise reduction.

In this paper, we propose an image noise reduction method based on cGAN to improve the quality of CBCT images. We use the oral CBCT images as input, and the MDCT image as the ground truth image. We train the network to generate the corresponding MDCT image to reduce the noise in CBCT. And we propose a gradient loss function, which is a part of the generator G loss function, to make the model more sensitive to image gradient and improve the quality of the generated images.

The paper is organized as follows. Section 2 describes our model structure. Section 3 presents the experiment details and results. Section 4 discusses the models hyperparameter selection and its results. The conclusion is in Section 5.

Methods

cGAN has been widely used in many kinds of image processing. We improve the generator structure of cGAN by using the Unet-like structure and enhancing gradient propagation through skip connection. We design the gradient loss function which can better reflect the noise effect on the image gradient. Minimizing the gradient loss can retain the image edge information as well as reduce the noise.

cGAN

The original GAN consists of generator G and discriminator D. Both of them are trained through adversarial training process. G receives the noise signal as input and generated samples. D received real samples and generated samples. D tries to distinguish these two samples. Training G is to generate the highly similar samples with the real ones. Training D is to accurately identify whether the samples originate from the generated samples or the real samples. This training process is like min max game between G and D.

$$\min_G \max_D V\left(D,\ G\right) = E_{x\sim Pr}\left[\log\left(D(x)\right)\right] + E_{z\sim Pg}\left[\log\left(1 - D(G(z))\right)\right] \quad (1)$$

Where, the input x of D satisfies real data distribution P_r, and input z of G which is the random noise satisfies distribution Pg.

However, the original GAN may have problems such as model collapse and gradient disappearing. To solve these problems, Arjovsky proposed Wasserste GAN, which introduced quantified training indicators and largely avoided model collapse. Wasserste GAN is more easily trained than GAN.

Besides, cGAN was proposed to solve the randomness of the input and output of GAN. cGAN specifies the mapping from input to output by adding constraint y into generator G and discriminator D, respectively:

$$\min_G \max_D V\left(D,\ G\right) = E_{x\sim Pr}\left[\log\left(D(x|y)\right)\right] + E_{z\sim Pg}\left[\log\left(1 - D(G(z|y))\right)\right] \quad (2)$$

Constraint y is the ground truth image in noise reduction process. The images with noises are put into generator G, while both the images with noises and the ground truth images are put into discriminator D. Through adversarial training, generator G continuously generates images closer to the ground truth images and eventually to output the denoise images.

Models

We propose an improved cGAN, to solve noise misidentification and incomplete noise reduction problems existing in current cGAN models for CBCT images.

The generator G uses Unet-like structure in this paper which is similar to DeblurGAN[5]. Fig.1 shows the network structure. As an advantage over the direct convolution on the original image, Auto Encoder (AE) structure shortens training time by doing image transformation on a smaller size through downsampling and then restoring the original image size through deconvolution [11]. Long skip connections can help increase the number of feature channels, thus make the texture information of the original image easier propagation [12].

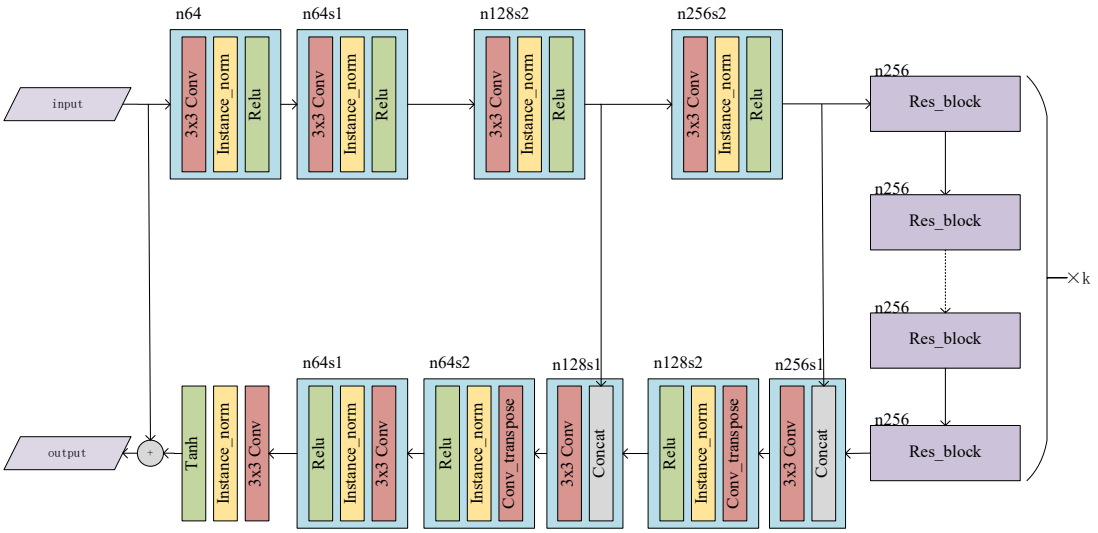


Figure 1 generator G structure

We use residual blocks in the residual net (ResNet) [13] to do the image transformation on the reduced size. Fig. 2 shows the residual block structure. The residual block helps the gradient propagate, and prevents the gradient from disappearing and avoids net degenerated when the network deepens. The number of residual blocks is 10 in our model.

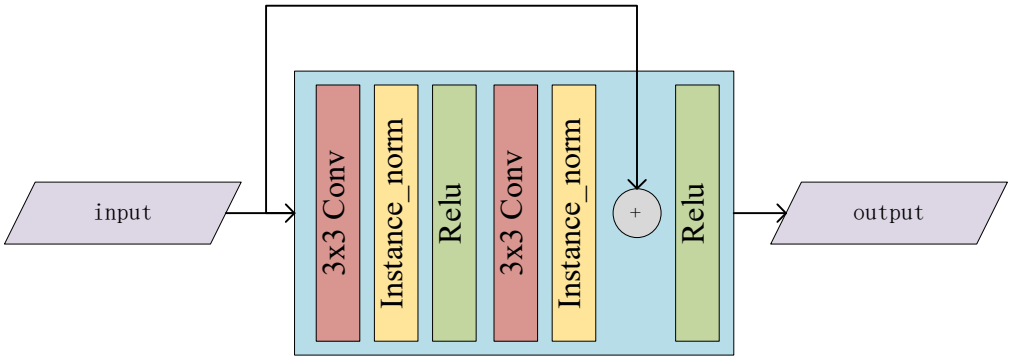


Figure 2 Residual Block structure

The discriminator D structure of our model is consistent with that of patchGAN proposed by P.Isola et al. [4]

The generators loss function is the sum of content and adversarial loss:

$$\mathcal{L} = \mathcal{L}_{adv} + \mathcal{L}_{content} \tag{3}$$

The adversarial loss is a loss function in the original GAN. It aims at generating images that approximate the ground truth image with high precision.

The content loss has three parts: the perceptual loss, the L2 loss (mean square error, MSE), and the gradient loss.

$$\mathcal{L}_{content} = \lambda_1 * \mathcal{L}_{perceptual} + \lambda_2 * \mathcal{L}_{L2} + \lambda_3 * \mathcal{L}_{gradient} \quad (4)$$

In recent researches, the output of deepen network expressed the abstract features [14]. The perceptual loss focused on how to restore the general content of the generated image [15]. In our paper, we calculate the perceptual loss from the output of third activated function on the third layer of VGG.

Although the L2 loss which is a classical choice for the content loss, has obvious merits. It is smooth. Its derivatives are easy to find. And it has stable solutions. But when it is the only optimization target the generated images are blurred.

We propose the gradient loss, which makes the model more sensitive to image gradient and edge information. The gradient vector of image function f (x, y) at the point (x, y) is defined as:

$$\nabla f(x,y) = \left[G_x, G_y\right]^T = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]^T \quad (5)$$

The digital image can be expressed by a two-dimensional discrete function. We use the approximation differential derivatives in horizontal and vertical directions. For the gradient G of the digital image at the point (i, j), we have:

$$grad = \left[grad_x, grad_y\right]^T \quad (6)$$

$$grad_x = f(i+1,j) - f(i-1,j) \quad (7)$$

$$grad_y = f(i,j+1) - f(i,j-1) \quad (8)$$

The difference vector measures how far the generated images are from the ground truth images. The gradient loss, defined as the mean of the squares of the difference vectors, is:

$$L_{\text{gradient}} = \frac{1}{W*H}\sum\left\|grad\left(G_{ij}\right) - grad\left(N_{ij}\right)\right\|^2 \quad (9)$$

Where $G_{x_{ij}}$ and $G_{y_{ij}}$ represent the gradient of the generated image，$\tilde{G}_{x_{ij}}$ and $\tilde{G}_{y_{ij}}$ represent the gradient of the ground truth image at point (i, j) , with W, H being the length and height of the image respectively.

In medical images, both edge and noise are high-frequency components. It depends on the gradient information when doing the edge detecting and segmentation. The noise greatly effects the accuracy [16]. L2 loss only concerns the difference value of each pixel. It can't reflect the effect of noise on the gradient. Gradient loss solves this problem.

For example, for the sequence $a_n = [1,2,3,4,5,6]$, if we define the gradient as $a_i = a_{i+1} - a_i$, the gradient sequence is $\Delta a_n = [1,1,1,1,1]$.Inserting a noise of

amplitude 1 at the places 3 and 4 in the sequence a_n, the possible results are shown in Table 1.

Table 1 the different noise impact

No.	sequence	MSE	Gradient sequence	Gradient loss
1	[1,2,2,3,5,6]	2	[1,0,1,2,1]	2
2	[1,2,4,5,5,6]	2	[1,2,1,0,1]	2
3	[1,2,2,5,5,6]	2	[1,0,3,0,1]	6
4	[1,2,4,3,5,6]	2	[1,2,-1,2,1]	6

It can be observed that No.1 and No. 2 roughly maintain the overall gradient, while No. 3 appears a large change, and No. 4 appears the opposite gradient. These four sets of data have the same L2 loss. The gradient loss better reflects how noises impact the gradient information. The bigger the gradient changes, the greater the loss value.

The gradient loss is the mean square for gradient difference vector length. It is more sensitive to the points with larger gradient error in the image, which have bigger effect on edge detecting and segmentation, while the points with smaller gradient errors do not. The drawback of gradient loss is that it cannot perceive the brightness change of the whole image, and it is not sensitive to the synchronous change of the continuous areas. we combined the L2 loss and the gradient loss. It showed that the results were better than using these two loss functions separately through the experiments shown in the discussion section.

Results

We discuss our model performance on noise reduction for CBCT data in this section. Against RED-CNN [8] and SAGAN [6], we compare the noise reduction results on the simulation data with Poisson noise and the real data. We also discuss the hyperparameter selection and model adjustment.

Experiments environment

The experiment data included CBCT and the head MDCT from 5 volunteers. CBCTs tube potential was 90kV, current was 10mA, and slice thickness was 0.3mm; MDCTs tube potential was 120kV, current was 232mA, and slice thickness was 0.9mm. The direction and position of the data were deviated because it came from different scanning equipment. We preprocessed the data firstly. We used iso-surface to 3D reconstruct the data, used ICP to register the 3D model, and resampled data again to obtain the experiment data of the same direction and the same slice.

We used the CBCT image as the generator G input and the MDCT image as the ground truth image to reduce the noise in the CBCT image.

The experiment used 620 images as training data and 150 ones as test data. The batch size was 1, the patch was 256, the number of training was 300, the first 150 epoch learning rates remained 10^{-4}, and the last 150 ones training rates linearly decreased to 0. The experiment system was Ubuntu 16.04, the CPU was 20-core Intel Xeon E5-2698, and the model was trained on a tesla V100.

Image quality evaluation method

We use the peak signal-to-noise ratio (PSNR), multi-level structural similarity (MS-SSIM), and gradient amplitude similarity deviation (GMSD) to evaluate the quality of the generated image.

PSNR is a full reference image quality evaluation standard. It is calculated as follow:

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

Where, MSE represents the mean square error between the current image X and the reference image Y, MSE is calculated as follow:

$$MSE = \frac{1}{H*W}\sum_{i=1}^{H}\sum_{j=1}^{W}\big(X(i,j)-Y(i,j)\big)^2 \tag{11}$$

Where W represents the length and H represents height of the image, n represents the number of pixel bit, and the unit of PSNR is dB. The larger the value, the smaller the distortion.

MS-SSIM [17] is the improved structural similarity proposed by Wang et al., Which is closer to the subjective quality evaluation results than SSIM. SSIM is calculated as:

$$SSIM(X,Y) = L_M(X,Y)C_J(X,Y)S_J(X,Y) \tag{12}$$

MS-SSIM adds scale information when calculating SSIM, reduces the original image with various extent and calculates the contrast factor and structure factor on each size image. The calculation method is shown as follow:

$$MSSSIM(X,Y) = [L_M(X,Y)]^{\alpha_M}\prod_{J=1}^{M}\big[C_J(X,Y)\big]^{\beta_M}\big[S_J(X,Y)\big]^{\gamma_M} \tag{13}$$

X is current image and Y is the reference image, L（X,Y）is the brightness contrast factor, C (X, Y) is the contrast factor of contrast, and S (X, Y) is the structure contrast factor, β_1=γ_1=0.0448，β_2=γ_2=0.2856，β_3=γ_3=0.3001，β_4=γ_4=0.2363，

α_5=β_5=γ_5=0.1333， M represents the scale factor of image reduction. L, C and S are calculated as follows:

$$L(X,Y) = \frac{2u_X u_Y + C_1}{u_X^2 + u_Y^2 + C_1} \tag{14}$$

$$C(X,Y) = \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2} \tag{15}$$

$$S(X,Y) = \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3} \tag{16}$$

$u_X \; and \; u_Y$ are the image means, $\sigma_X \; and \; \sigma_Y$ are the image variances, and σ_{XY} is the covariance. The higher the MS-SSIM value, the more similar the two image structures are.

GMSD was proposed by Zhang et al. in 2014. It evaluates the structural distortion of local images by the local gradient amplitude, and uses the standard deviation of local image quality to measure the global image quality [18], the calculation method is:

$$GMSD = \sqrt{\frac{1}{N}\sum_{\text{i-1}}^{N}(GMS(i) - GMSM)^2} \tag{17}$$

GMS is the similarity of the gradient amplitude, and the calculation formula is as follow, GMSM is the mean of the regional gradient field.

$$GMS(i) = \frac{2m_r(i)m_d(i) + c}{m_r^2(i) + m_d^2(i) + c} \tag{18}$$

m_r represents the gradient magnitude of the current image and m_d represents the gradient magnitude of the ground truth image.

$$m_r(i) = \sqrt{(r \otimes h_x)^2(i) + \left(r \otimes h_y\right)^2(i)} \tag{19}$$

$$m_d(i) = \sqrt{(d \otimes h_x)^2(i) + \left(d \otimes h_y\right)^2(i)} \tag{20}$$

$h_x \; and \; h_y$ are the operators of Prewitt. The smaller the GMSD value, the closer the current image is to the ground truth image.

Simulation dataset

We added Poisson noise to the MDCT to simulate LDCT data. The number of images, size, and network parameters remained the same with the real data set. Table 2 shows the objective evaluation scores of each model in the test data. The denoise effect of the RED-CNN in the simulation data set was slightly better than our model. It showed that SAGAN had a lower overall score because it misidentified the low-density area

as error. But the main tissues such as teeth were clear. Figure 3 shows the results of each model.

Table 2 the evaluation score by RED-CNN、SAGAM and our model in the simulation data

classification	model	PSNR	MS-SSIM	GMSD
all	Noise image	39.531	0.996	0.000
	RED-CNN	**45.886**	**0.998**	0.000
	SAGAN	28.279	0.936	0.011
	Our model	43.835	**0.998**	0.000

(a) (b) (c) (d) (e)

Figure 3 the results of the different model on simulated data. (a) CBCT data (b) RED-CNN result (c) SAGAN result (d) our model result (e) MDCT data

Real dataset

The real data set used the CBCT image as the LDCT image and its corresponding MDCT image as the ground truth image. The data was divided into three groups. We discuss the denoise results of each model on the crown, root and jaw bone. Figure 4 shows the results in the test set. It showed that the images trained by RED-CNN were blurred. The images trained by SAGAN model were clearer, which was not sensitive to the noise. Some noise had not been removed, and the part of the low-density bone was misidentified as noise.

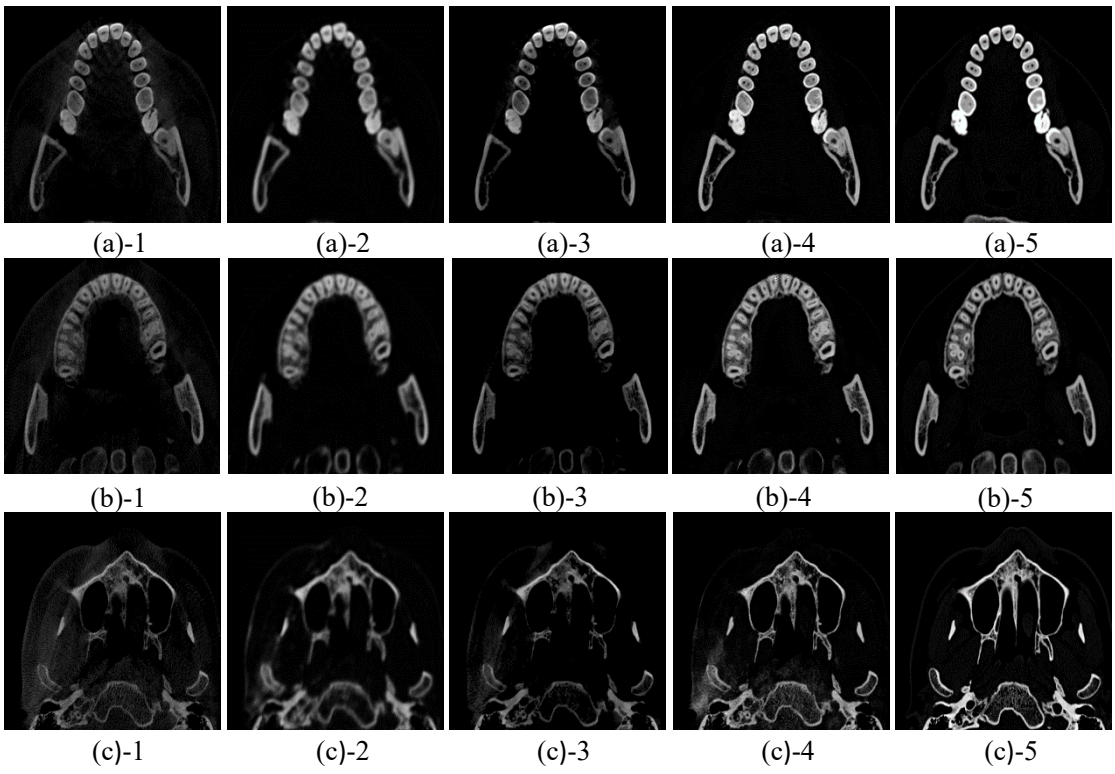


Figure 4 Qualitative comparison of each model on the test set. (a) crown part (b) root part (c) jaw part (number 1 was the original CBCT image, number 2 was the denoise image by RED-CNN, number 3 was the denoise image by SAGAN, and number 4 was denoise image by our model, number 5 was the MDCT image.)

We use PSNR, MS-SSIM, and GMSD as evaluation indicators. Table 3 shows the quantitative results. The scores of our model are the best.

Table 3 Evaluation scores of RED-CNN, SAGAM and our model in real data

classification	model	PSNR	MS-SSIM	GMSD
crown part	CBCT	18.536	0.742	0.033
	RED-CNN	21.249	0.867	0.027
	SAGAN	18.241	0.792	0.039
	Our model	**22.834**	**0.895**	**0.022**
root part	CBCT	22.366	0.795	0.024
	RED-CNN	24.477	0.897	0.020
	SAGAN	21.561	0.83	0.031
	Our model	**25.957**	**0.920**	**0.016**
jaw part	CBCT	22.435	0.776	0.031
	RED-CNN	23.328	0.851	0.031
	SAGAN	21.899	0.791	0.041

	Our model	**24.034**	**0.872**	**0.026**
all parts	CBCT	21.112	0.771	0.029
	RED-CNN	23.018	0.872	0.026
	SAGAN	20.567	0.804	0.037
	Our-model	**24.275**	**0.896**	**0.021**

We constructed the error images according to the average of absolute difference between all generated images and ground truth images pixel by pixel in the test set. Contrast with the error image, it can be seen directly that the error of our model was significantly smaller than the contrast model in Figure 5.

(a) (b) (c)

Figure 5 (a) error image by RED-CNN (b) error image by SAGAN (c) error image by our model

Discussion

Hyperparameter selection and model adjustment

In this section we demonstrate the effectiveness of the long skip link through the contrast experiments. We discuss the hyperparameter selection and the effect of the preprocessing.

Long skip link

We respectively tested the quality of the images generated by the model without skip connections and with one skip connection. Table 4 shows the scores that show the quality of the image generated by the model added the skip-in link is improved.

Table 4 impact of skip connections on image quality

method	PSNR	MS-SSIM	GMSD
our model	24.275	0.896	0.021
one skip link model after the first down sampling	24.041	0.894	0.022
no skip link model	23.978	0.893	0.022

The number of Residual Blocks

The number of residual blocks determines the depth of the model. We compared the different depth models with the residual blocks number of 6, 8, 10, 12, and 14.Figure 6 shows the results. It can be observed that with the depth of the model increasing within a certain range, the quality of the generated image was improved. But as the depth continued to increase, the quality of the generated image decreased instead. At the same time, as the model depth increased, the training time linearly increased.

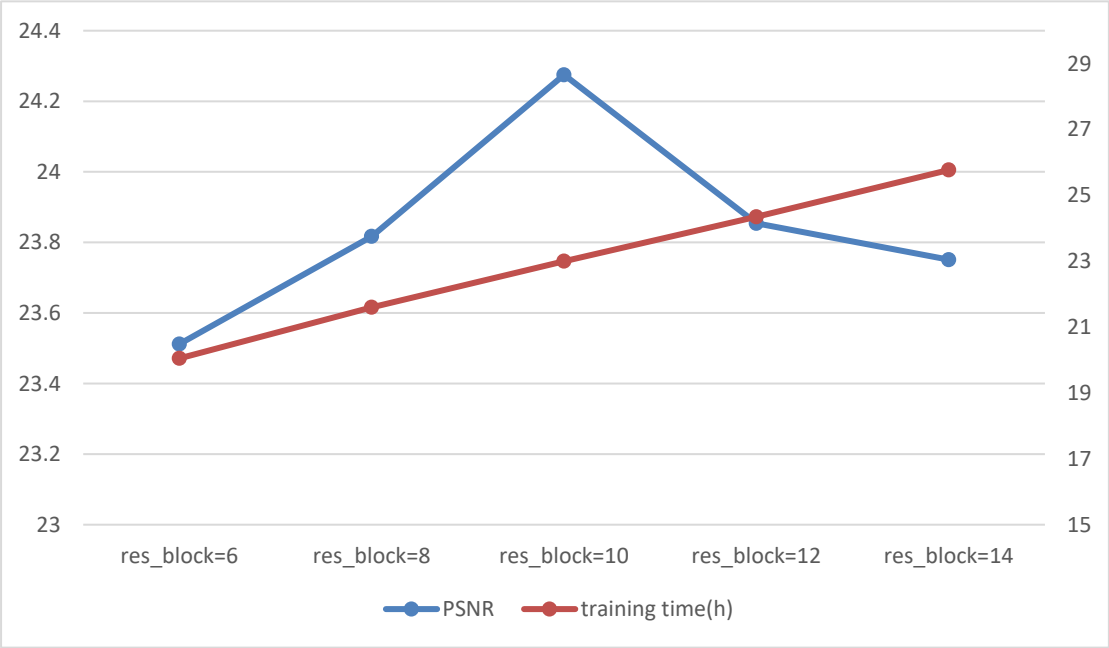


Figure 6 Impact of Residual Blocks number on image quality

The effects of different noise on noise reduction

The denoising by contrast model is poor in oral CBCT images because both the tissue type and the CT imaging methods are different.

The current LDCT denoising model is mainly used for soft tissues such as the abdominal cavity and the brain. Such tissues have similar CT values at adjacent points. And their tissues are continuous. The stomatology tests include examinations of bones, teeth, and gums. The density of bone cancellous and bone dense is uneven, and the CT values at the edge of bones change suddenly. These characteristics are obvious different from soft tissues, which effect the denoising results.

MDCT scan method is multi-row detector scanning layer-by-layer. CBCT is quite different. CBCT uses a cone beam emitter to make a circular projection around the projection object. The reconstruction algorithms are also different. It can be concluded CBCT contains more kinds of noise. The radial noise shown in Figure 7(a) was related to its position in the scanning area. For this kind of noise, the SAGAN

model had the best denoise images, while there was still some noise denoised by RED-CNN and our model. Figure 7(b) Mainly compared the impact of CT partial volume effect on the denoised images. It showed that there was an obvious volume effect on the edge of the teeth in the CBCT image. Our model outperforms SAGAN and REDCNN in terms of noise reduction. The noise had not been completely removed by SAGAN and RED-CNN. Figure 7(c) was the strong noise aera. In this aera, the denoised images by our model were better than other models.

In summary, our model could identify noise more accurately. It couldn't misidentify the noise as tissues compared with other models while its limitation was the incomplete denoising for radial noise.

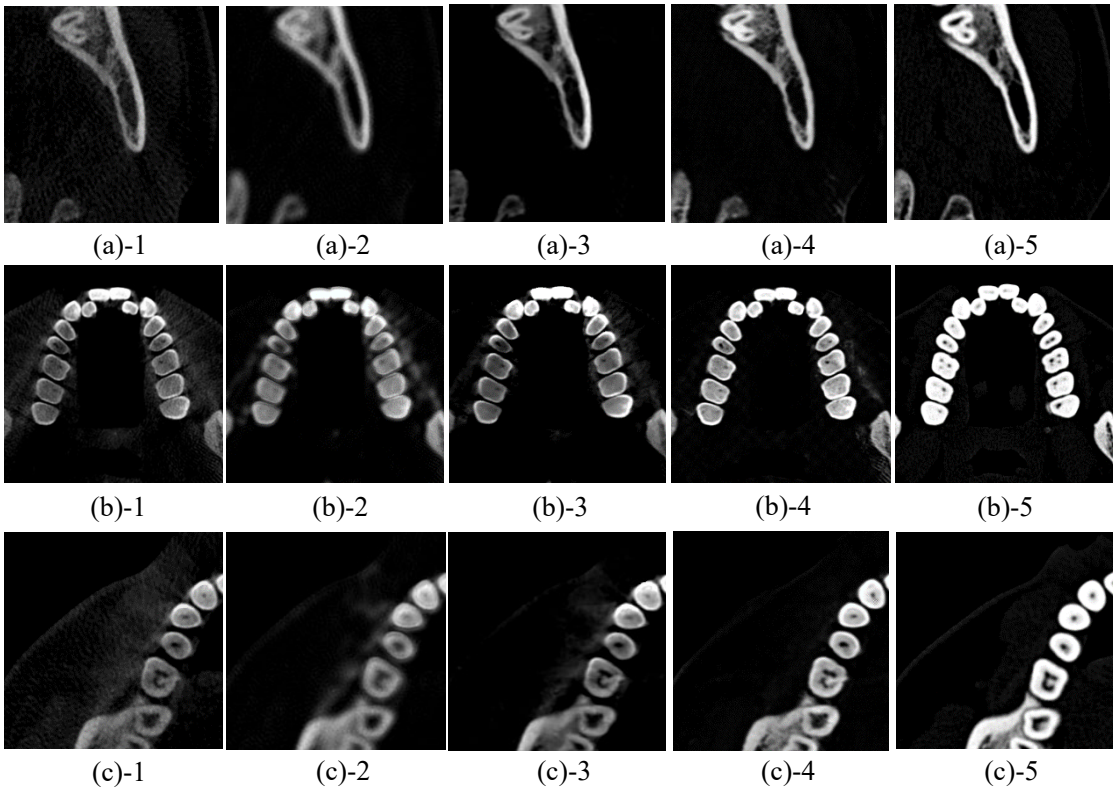


Figure 7 the impact of various noises on CBCT and the denoise effect of each model. (a)radial noise (b)volume effect (c) strong noise area (number 1 was the original CBCT image, number 2 was denoised by RED-CNN number 3 was denoised by SAGAN, and number 4 was denoised by our model, No. 5 was the MDCT image.)

Through the experiments, it can be concluded that although the RED-CNN performed well on objective evaluation indicators such as PSNR, its visual effect was poor that was more blurred than other models. We did Gaussian filtering and mean filtering with a window size of 3 on the MDCT images with Poisson noise, and

calculated the image quality of the results. Table 5 shows the experiment results. It showed that for Poisson noise, above filtering had improved the objective evaluation score of the image in some extent. when training the model, it was necessary to avoid image blurred when denoising.

Table 5 image quality after Gaussian filtering and mean filtering

method	PSNR	MS-SSIM	GMSD
MDCT with Poisson noise	39.531	0.996	0.000
mean filtering	42.424	0.996	0.000
Gaussian filtering	43.961	0.997	0.000

Effectiveness of gradient loss

We discuss the effect of the generator G loss function on the generated image quality. We used different loss function and their combination as loss function. Figure 8 shows the comparison results.

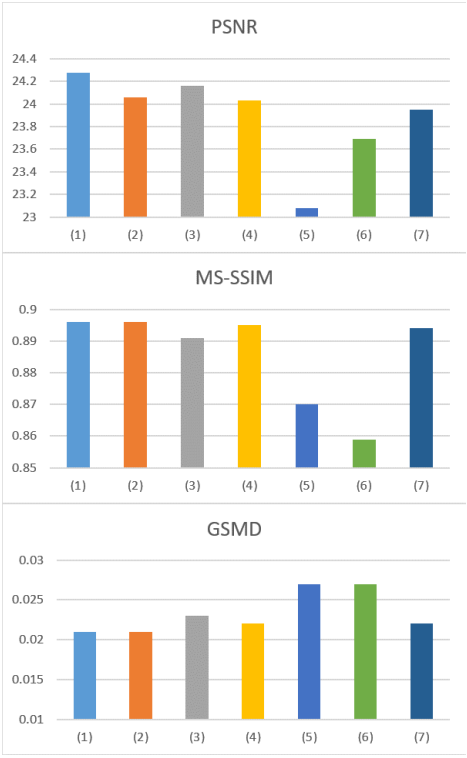
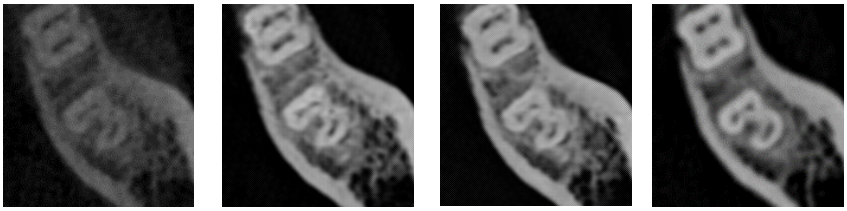


Figure 8 (1) our model (2) perceived loss combined with adversarial loss (3) gradient loss combined with MSE(4) gradient loss combined with MSE and perceived loss (5) gradient loss combined with MSE and adversarial loss (6) perceived loss combined with adversarial loss and MSE (7) perceived loss combined with adversarial loss and gradient loss

Comparing the results of (2) and (3), we can see that compared with “perceived loss combined with adversarial loss”, the model using “gradient loss combined with

MSE" did better on PSNR, but it did worse on GMSD and MS-SSIM. We combined perceived loss with gradient loss and MSE and we combined adversarial loss with gradient loss and MSE. The results are shown in Figure 8(4) and (5), respectively. The former improved the quality of the generated image. For the model with adversarial loss, the generated image quality was poor and it appeared fake details. We also compared the "perceived loss combined with adversarial loss and MSE" model with "perceived loss combined with adversarial and gradient loss" model. Figure 8(6) and (7) show the results. It can be observed that our model obtained the best results which used the combination of the four loss functions as the final function.

We compared the effect between the models without gradient loss and the model with gradient loss on the edge information of the generated image. Figure 8(1) and (6) show the quantitative results. In the test data, the model with gradient loss had better edge characteristics, as shown in Figure 9.



(a) (b) (c) (d)

Figure 9 (a)CBCT image, (b) the model with gradient loss, (c) the model without gradient loss , (d)MDCT image

Figure 10 shows the difference of objective evaluation score between the model with gradient loss and the model without gradient loss. We compared the difference between the PSNR and MS-SSIM scores of each image in the test set. And formula (21), (22) define ΔPSNR and ΔMS-SSIM respectively:

$$\Delta PSNR = \mathrm{PSNR}_{\text{with gradient loss model}} - \mathrm{PSNR}_{\text{without gradient loss model}} \quad (21)$$

$$\Delta MS-SSIM = \mathrm{MS\text{-}SSIM}_{\text{with gradient loss model}} - \mathrm{MS\text{-}SSIM}_{\text{without gradient loss model}} \quad (22)$$

It can be observed that the median of ΔPSNR is above 0, and the first quartile is slightly lower than 0 in Figure 10, while both the median and lower quartile of ΔMS-SSIM are above 0. It can be concluded that the test image gradient loss improved the quality of the generated image. But there are also some images, the model without gradient loss denoised better yet.

Figure 10 impact of gradient loss on the quality of images

Conclusion

We propose an improved cGAN for denoising in oral CBCT data. we use adversarial and content loss with gradient loss to train the model. Through the experiments, the results of our model were the best compared with other models.

Our model had a higher evaluation score and better visual quality. It is satisfied to retain bone cancellous. And it can show the edge of the image more obviously. Our model is more competitive for the clinical application. Compared with other models which only reduce simulated noise, we directly input CBCT images to reduce the noise.

However, our model still has some limitations. First, it also exists noise in MDCT. We need to consider its effect on the training and evaluation. Second, the scores of PSNR and other indicators may be inconsistent with the visual effect in the evaluation for medical images. It is necessary to further optimize the quantitative evaluation method of CT image quality. Last, we use the data came from the different scanning equipment. We preprocessed the image such as the resample. It is necessary to further analyze the registration errors effect on the experiment results.

For all methods based on deep learning, it needs to be trained for specific dose levels, window widths and window levels. Our model only did the noise reduction for high-density bones and teeth in oral CBCT data. The denoise performance of other parts and other tissues remains to be evaluated.

Abbreviations

CBCT: cone beam computed tomography; MDCT: multi-row detector spiral CT; LDCT: low dose; GAN: Generative Adversarial Networks; cGAN: conditional Generative Adversarial Network; PSNR: signal-to-noise ratio; MS-SSIM: multi-level structural similarity; GMSD: gradient amplitude similarity deviation.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This work is supported by the Science & Technology Development Program of Jilin Province, China (Nos. 20160204048GX, 20170204031GX, 20180201037SF, 20190201196JC and 20190302112GX), the National Key Research and Development Program of China under Grant No. 2017YFC0108303 and the Science & Technology Development Program of Jilin Province, China (No. 2018C039-1).

Authors' contributions

HMY and ZGJ: design the research, supervise all the process, and provide valuable guidance and research grant. YFD and YM: responsible for coding the algorithms, conduct all experiments and the data analysis, and write the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the editor and the anonymous reviewers for their helpful suggestions and comments which will provide a great contribution to this manuscript.

Availability of data and materials

Not applicable.

References

[1] Wolterink J M , Leiner T , Viergever M A , et al. Generative Adversarial Networks for Noise Reduction in Low-Dose CT[J]. IEEE Transactions on Medical Imaging, 2017:1-1.

[2] JU Hao, ZHU Hong-hua, DUAN Tao, LI Zhi-min, SUN Hong-chen, Farman AG, Scar fe WC. The research progress of basic principle and application of CBCT [J]. Journal of Medical Imaging, 2015, 25(05):907-909+942.

[3] O. Kupyn , V. Budzan , M. Mykhailych , D. Mishkin , J. Matas , DeblurGAN: blind motion deblurring using conditional adversarial networks, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 8183–8192.

[4] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu,D.Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. GenerativeAdversarial Networks. June 2014. 1, 3.

[5] P. Isola , J.-.Y. Zhu , T. Zhou , A .A . Efros , Image-to-image translation with condi- tional adversarial networks, in: IEEE Conference on Computer Vision and Pat- tern Recognition (CVPR), 2017, pp. 1125–1134 .

[6] Yi X , Babyn P . Sharpness-Aware Low-Dose CT Denoising Using Conditional Generative Adversarial Network[J]. Journal of Digital Imaging, 2018.

[7] Wang Y , Yu B , Wang L , et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose[J]. NeuroImage, 2018, 174.

[8] Chen H , Zhang Y , Kalra M K , et al. Low-Dose CT with a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN)[J]. IEEE Transactions on Medical Imaging, 2017:1-1.

[9] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein GAN.ArXiv e-prints, Jan. 2017. 1, 3, 4, 5, 7.

[10] Mirza M, Osindero S. Conditional Generative Adversarial Nets[J]. Computer Science, 2014:2672-2680.

[11] Si Yang, Xiao Qinkun, Li Xing. Human action recognition using auto-encoder and neural network[J]. Foreign Electronic Measurement Technology, 2018,37(1):78-84.

[12] Ronneberger O , Fischer P , Brox T . U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. 2015.

[13] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015.

[14] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham,A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. ArXiv e-prints, Sept. 2016.

[15] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. arxiv, 2016.

[16] WANG Zhenhua, HU Fuyuan, Lü Fan, XIA Zhenping. An adaptive image enhancement algorithm for edge detection [J]. Journal of Suzhou University of Science and Technology(Natural Science),2018,35(02):68-72.

[17] Wang Z , Simoncelli E P , Bovik A C . Multi-scale structural similarity for image quality assessment[C]// Proc IEEE Asilomar Conference on Signals. 2003.

[18] Xue W , Zhang L , Mou X , et al. Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index[J]. IEEE TRANSACTIONS ON IMAGE PROCESSING, 2014, 23(2):684-695.

Figures

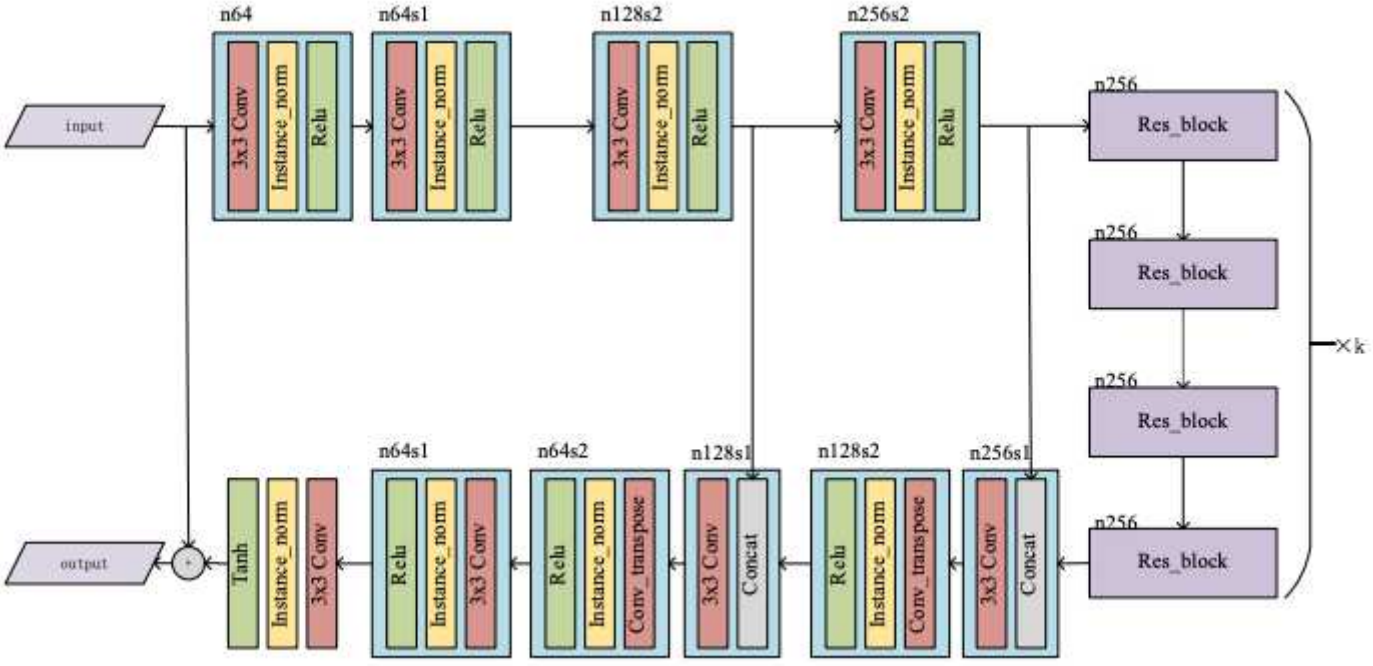


Figure 1

generator G structure

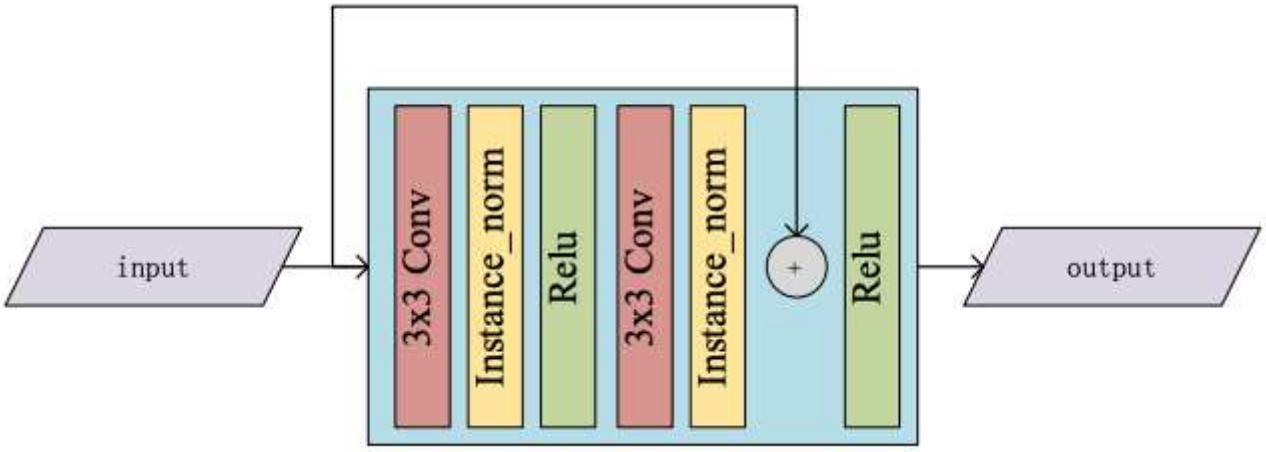


Figure 2

Residual Block structure

classification	model	PSNR	MS-SSIM	GMSD
all	Noise image	39.531	0.996	0.000
	RED-CNN	**45.886**	**0.998**	0.000
	SAGAN	28.279	0.936	0.011
	Our model	43.835	**0.998**	0.000

(a) (b) (c) (d) (e)

Figure 3

the results of the different model on simulated data. (a) CBCT data (b) RED-CNN result (c) SAGAN result (d) our model result (e) MDCT data

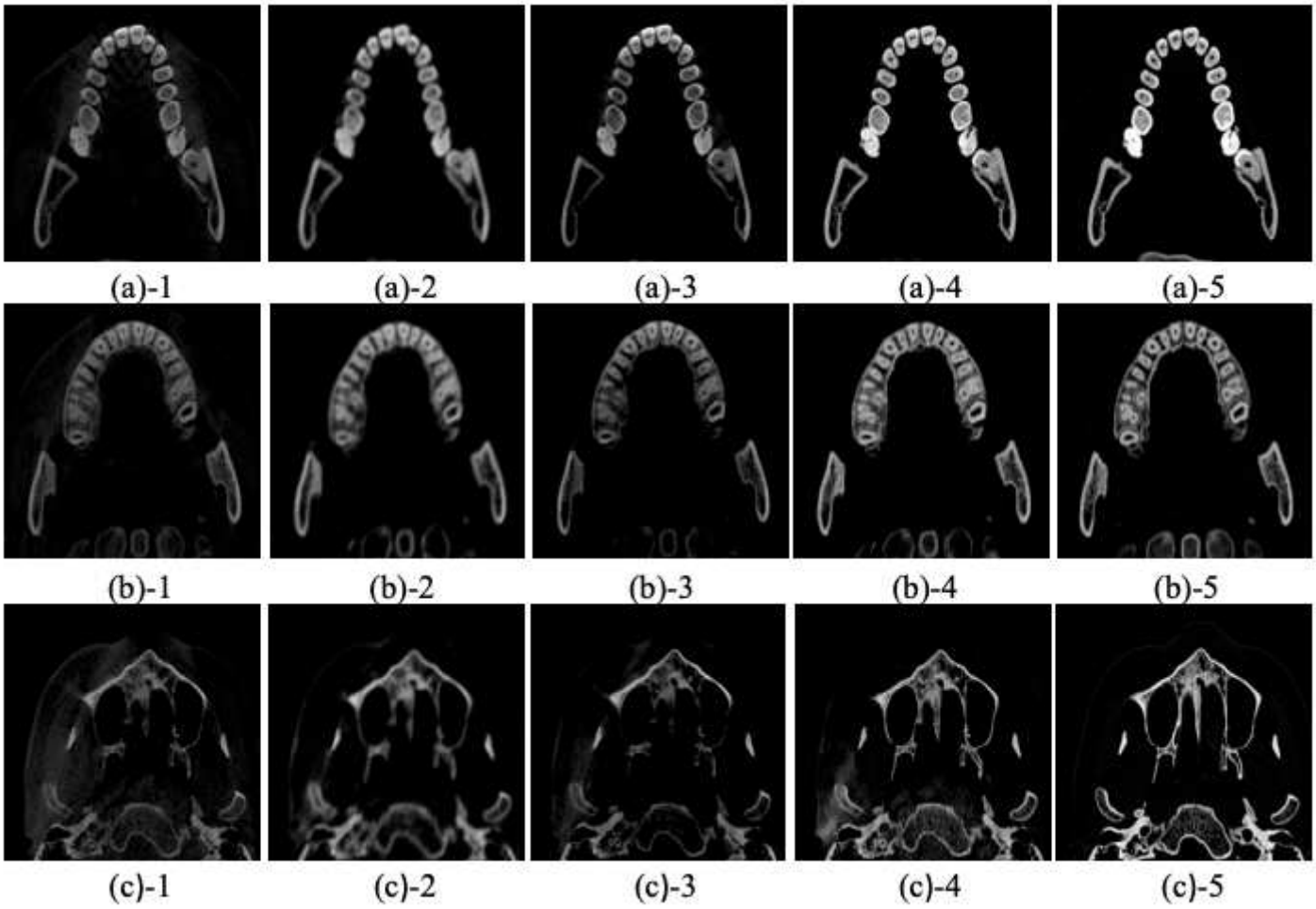


Figure 4

Qualitative comparison of each model on the test set. (a) crown part (b) root part (c) jaw part (number 1 was the original CBCT image, number 2 was the denoise image by RED-CNN, number 3 was the denoise image by SAGAN, and number 4 was denoise image by our model, number 5 was the MDCT image.)

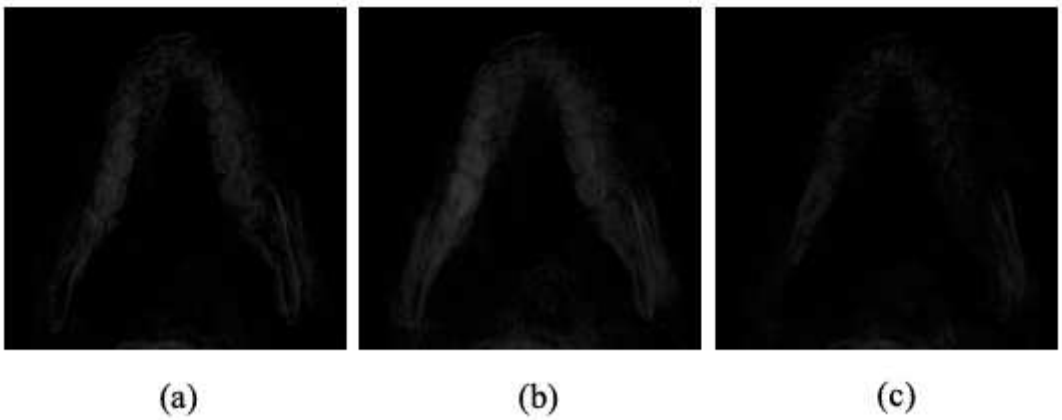


Figure 5

(a) error image by RED-CNN (b) error image by SAGAN (c) error image by our model

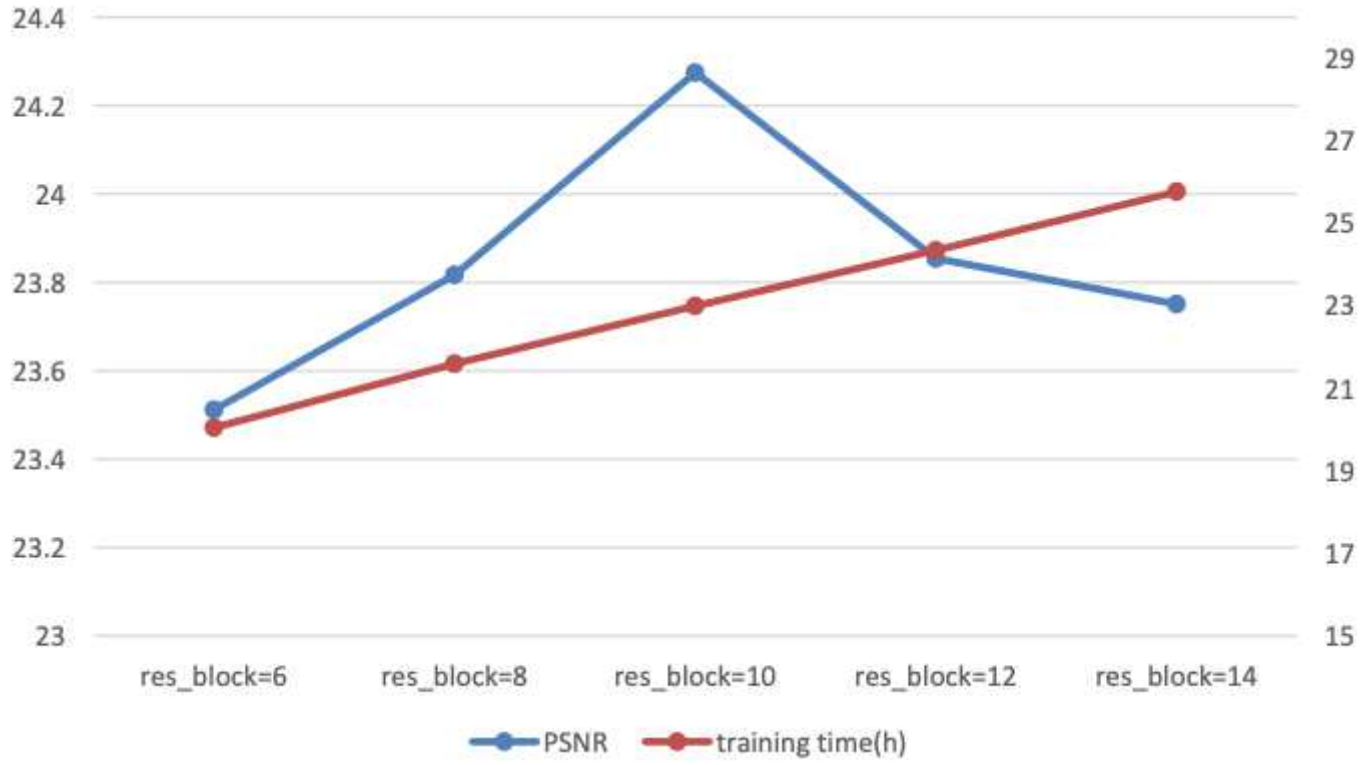


Figure 6

Impact of Residual Blocks number on image quality

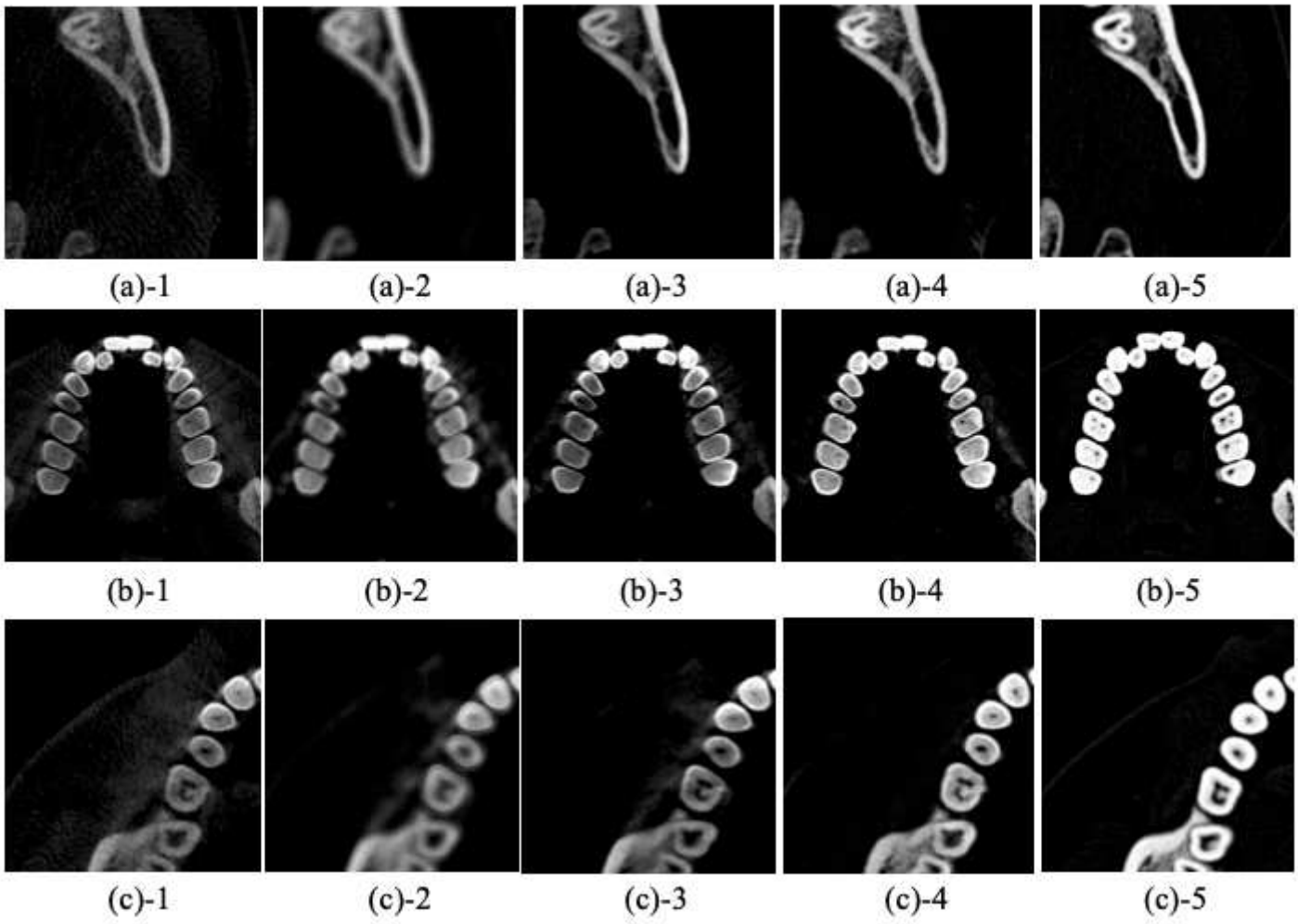


Figure 7

the impact of various noises on CBCT and the denoise effect of each model. (a)radial noise (b)volume effect (c) strong noise area (number 1 was the original CBCT image, number 2 was denoised by RED-CNN number 3 was denoised by SAGAN, and number 4 was denoised by our model, No. 5 was the MDCT image.)

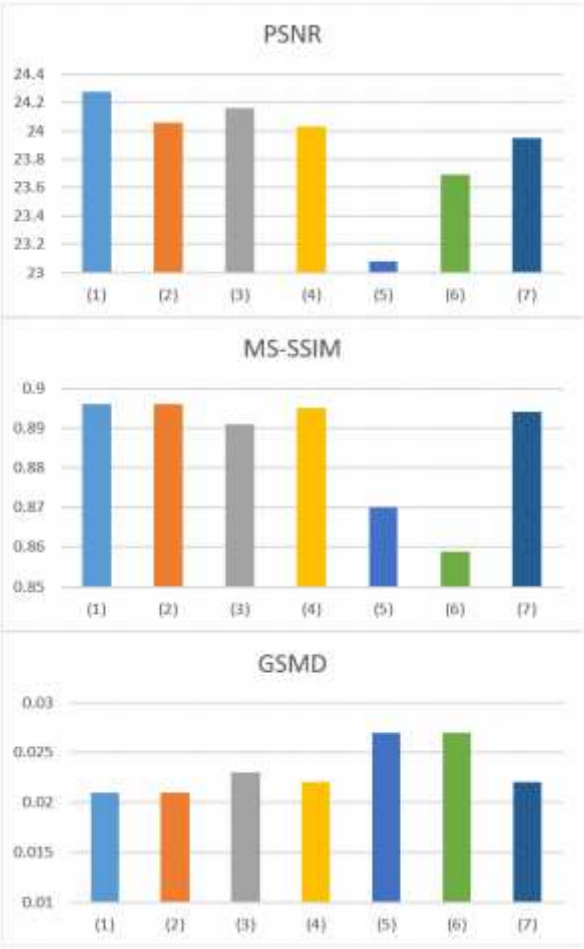


Figure 8

(1) our model (2) perceived loss combined with adversarial loss (3) gradient loss combined with MSE(4) gradient loss combined with MSE and perceived loss (5) gradient loss combined with MSE and adversarial loss (6) perceived loss combined with adversarial loss and MSE (7) perceived loss combined with adversarial loss and gradient loss

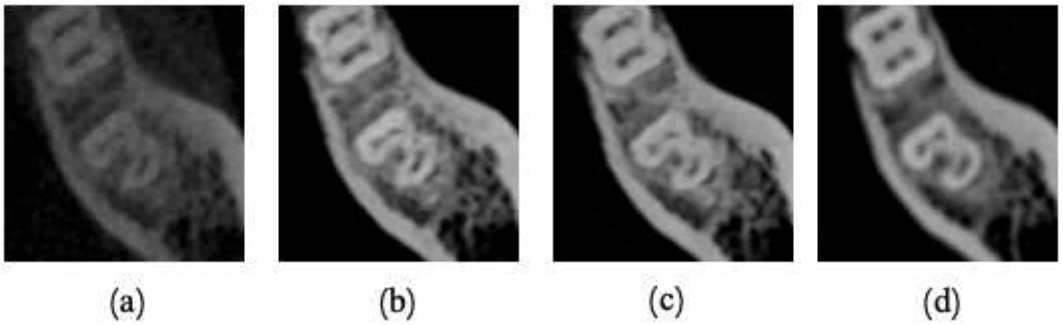


Figure 9

(a)CBCT image, (b) the model with gradient loss, (c) the model without gradient loss , (d)MDCT image

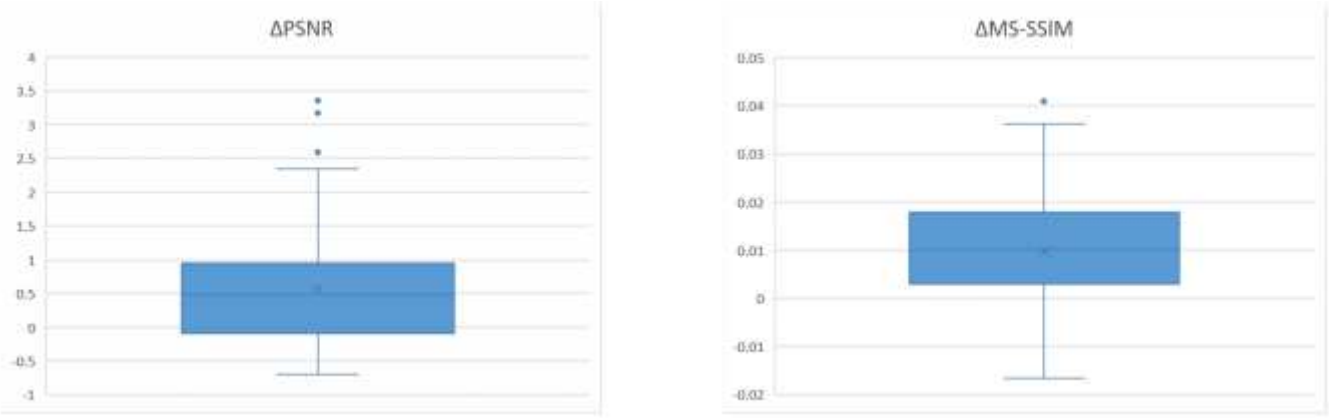


Figure 10

impact of gradient loss on the quality of images