CT Metal Artifact Reduction based on Virtual Generated Artifacts Using Modified pix2pix

Xie Kai
Second People's Hospital of Changzhou
Gao Liugang
Second People's Hospital of Changzhou
Lu Zhengda
Second People's Hospital of Changzhou
Li Chunying
Second People's Hospital of Changzhou
Lin Tao
Second People's Hospital of Changzhou
Sui Jianfeng
Seond People's Hospital of Changzhou
Bi Hui
Second People's Hospital of Changzhou
Ni Xinye (nxy2000@aliyun.com)
Nanjing Medical University
Dai Jianrong
Chinese Academy of Medical Science Cancer Institute

Research

Keywords: deep learning net, U-Net, pix2pix, metal artifact generation, metal artifact reduction, radiotherapy

Posted Date: October 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-96694/v1

CT Metal Artifact Reduction based on Virtual Generated Artifacts Using Modified pix2pix

Xie Kai[1,2#], Gao Liugang[1,2#], Lu Zhengda[1,2,3], Li Chunying[1,2], Lin Tao[1,2,]

Sui Jianfeng[1,2], Bi Hui[1,2], Ni Xinye[1,2*], Dai Jianrong[4]

[1]Department Department, Second People's Hospital of Changzhou, Nanjing Medical University, Changzhou, China

[2]Center for Medical Physics, Nanjing Medical University, Changzhou, China

[3]School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China

[4]Radiotherapy Department, Chinese Academy of Medical Science Cancer Institute, Beijing, China

[#]Equal contributors

*Correspondence: Ni Xinye (e-mail: nxy@njmu.edu.cn)

Abstract

Background: Metal artifacts introduce challenges in image-guided diagnosis or accurate dose calculations. This study aims to reduce metal artifacts from the spinal brace by using virtual generated artifacts through convolutional neural networks and to compare the performance of this approach with two other methods, namely, linear interpolation metal artifact reduction (LIMAR) and normalized metal artifact reduction (NMAR) .

Method: A total of 3,600-slice CT images of 60 vertebral metastases patients were selected. The spinal cord center was marked in each image, metal masks were added to two sides of the marker to generate artifact-insert CT images, and the CT values of the metal parts were copied to original CT images to obtain reference CT images. These images were divided into training (3,000 slices) and test (600) sets. The modified U-Net and pix2pix architecture was applied to understand the relationship between the reference and artifact-insert images. The mean absolute error (MAE), mean square error (MSE), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM) were calculated between the reference CT images and the predicted CT through LIMAR, NMAR, U-Net, and pix2pix. The CT values of organs from different images were compared. Radiotherapy treatment plans for vertebral metastases were designed, and dose calculation was performed. The dose distribution in different types of images was also compared.

Results: The MAE values between the reference images and those images generated by LIMAR, NMAR, U-Net, and pix2pix were 15.02, 16.16, 6.12, and 6.48 HU, respectively, and the corresponding PSNR values were 15.37, 152.70, 158.93, and 65.14 dB, respectively. Pix2pix restored more texture than U-Net according to the visual comparison. The average CT values in the artifact-insert images of the liver, spleen, and left and right kidneys were all significantly higher than those of the reference images ($p<0.05$). The average CT values of the organs in images processed by the four methods showed no significant differences from those of the organs in the reference images. The mean dose of planned target volume in the artifact-insert images was significantly lower than that in the reference CT images. The average γ passing rate (1%, 1 mm) of the artifact-insert images was significantly lower than that of the reference images ($95.9\pm1.4\%$ vs. $99.2\pm1.4\%$, $p<0.05$).

Conclusions: U-Net and pix2pix deep learning networks can remarkably reduce metal artifacts and improve critical structure visualization compared with LIMAR and NMAR according to the simulation data of artifact-insert images in the spinal brace. Pix2pix can restore more texture with the help of a discriminator. Metal artifacts increase the dose calculation uncertainty in radiotherapy. The dose calculated through images obtained by U-Net and pix2pix was identical with that calculated through reference images.

Keywords: deep learning net, U-Net, pix2pix, metal artifact generation, metal artifact reduction, radiotherapy

Background

Malignant bone tumors have a high postoperative recurrence rate and have been mainly treated before the 1960s by amputation with a survival rate of less than 20%. However, the application of chemoradiotherapy to malignant bone tumors after the 1990s increased this survival rate up to 50% to 80%. Sakaura et al. [1] found that although vertebral tumors have been successfully treated via total en bloc spondylectomy, the recurrence rate remains high. Therefore, chemoradiotherapy, immunotherapy, and other treatment means should be combined after the operation. Pekmezci et al. [2] also noted that a combination of surgical operation and radiotherapy is conducive to the treatment of spinal tumors.

Installing metal implants in the spine of a patient with a vertebral bone tumor can significantly affect CT and MR imaging quality, thereby preventing doctors from clearly distinguishing normal tissues from lesion ones [3,4]. In addition, for those patients who require follow-up radiotherapy, when the radiation field passes through a metal implant or when the implant is located in the target region, the CT artifact will affect the planned dose

[5].

Many studies on CT artifacts have been carried out, and various metal artifact reduction (MAR) methods have been proposed. Traditional MAR methods can be divided into (1) optimization of acquisition conditions, (2) model-based iterative reconstruction method, and (3) projection-correction-based method.

Optimization of acquisition conditions include adjusting scanning mAs, kV, and slice thickness [6], where dual-energy CT has a certain application effect on artifact reduction. Among the available model-based methods [7-9], X-ray generation, energy spectrum hardening, detector receiving, and system noise must be modeled. Afterward, CT images can be obtained through iterative reconstruction. This process requires a detailed understanding of the whole operations of a CT machine, which is difficult to realize in clinical practice. Meanwhile, according to projection-correction methods [10-13], those projections that are formed when X-rays pass through metal are inaccurate and should be regarded as missing data. Therefore, the interpolation or prior image forward projection method is generally applied as an alternative. Priori images are generally without artifacts and are obtained by segmenting and filtering the original CT images. The projection value obtained through prior image forward projection is used to replace the metal projection region in the original projection. Afterward, the corrected image is obtained through filtering, back projection, and reconstruction.

To reduce artifacts in the spinal brace, Wang et al. [4] argued that the spectral CT images with fast-kVp switching CT can be used as monochromatic images at energy levels to facilitate the reduction of artifacts caused by pedicle screws in the spinal CT image of a patient. Kotsenas et al. [5] applied a prototype of iterative metal artifact reduction algorithm in clinically evaluating the CT data of spinal fusion patients and in the anatomic visualization of critical soft-tissue structures in the postoperative spine. Afterward, they reduced metal artifacts through subjective and objective measurements to enhance the confidence of most spinal fusion patients in diagnosis.

However, the aforementioned methods require a large number of formulas and have individual characteristics. Deep learning has achieved great success in the image field, such as in image denoising [14], super-resolution image reconstruction [15], and image synthesis [16], and a series of studies have been carried out on metal artifact reduction via deep learning. Gjesteby et al. [17] applied a combination of deep learning and normalized metal artifact reduction (NMAR) on a reconstructed image to correct the region with serious artifacts. Park et al. [18] used U-Net to correct an inconsistent sinogram and eliminate beam-hardening

factors triggered by main metals along the metal trace in a sinogram. Zhang et al. [19] proposed an open MAR framework called CNN-MAR based on the deep learning model to reduce metal artifacts in CT images. Previous studies on artifact reduction via deep learning have achieved some progress in cone beam CT [20], prostatic CT [21], and cervical CT [7]. However, given the limited amount of directly usable and comparable data, a golden standard that can be directly compared with the artifact reduction results remains lacking.

In deep learning, U-Net is initially applied to image segmentation as a variant of a convolutional neural network [22]. The encoder in U-Net gradually reduces the spatial resolution of output characteristics through convolution and pooling operations, whereas the decoder restores the details and spatial resolution of the object step by step. A skipping-type connection is observed between the coder and decoder in order for the latter to recover the target details. Therefore, a paired training between the generated artifacts and original images can be implemented via CT artifact synthesis to solve the golden standard problem in artifact-insert images. Compared with U-Net, the pix2pix model [23] has achieved great success in image-to-image translation. In this study, U-Net and pix2pix were compared with linear interpolation metal artifact reduction (LIMAR) [24] and NMAR, planned radiotherapy doses were calculated based on the images generated through several methods, and the differences among these methods were analyzed.

Methods and material

Data acquisition

The chest and abdominal CT images of 60 patients accepting treatment at the Department of Radiotherapy of Changzhou Second People's Hospital were selected, and all patients were informed in advance. These patients were 56±7 years old on average with a median age of 58 years. This study was approved by the Research Ethics Board of the Second People's Hospital of Changzhou. The CT images were obtained by using a Siemens CT scanner (SOMATOM Force, Germany), and the scanning parameters were as follows: 120 kVp tube voltage, 300 mA tube current, 3 mm slice thickness, 512×512 reconstructed image size, and 0.73×0.73 mm^2 to 0.98×0.98 mm^2 spatial resolution. The CT images of 50 patients (3,000 slices) were used for the network training and those of 10 (600 slices) patients were used for testing.

Artifact generation

First, the treatment tables in CT images were removed. Some approximations were made to put the focus on MAR when simulating the spinal brace. The simulated metal was Ti with a

density of 4.54 g/cm3. The spinal cord center was marked, and binary masks were placed on the left, right, or both sides of the center in all CT images to represent the simulated metal positions. The mask was approximated by an ellipse with a major axis of 1.8 cm to 3.6 cm and a minor axis of 0.7 cm to 1.4 cm. The artifact-insert images were generated according to the masks. The metal artifacts were generated by using the method described in Zhang et al. [19].

The CT values of masks in the artifact-insert images were assigned to the original CT images, and then the original CT images containing metal information were obtained and called reference CT images. Fig. 1 shows the artifact-insert CT images of three different slices with a window width and level of 40 and 600 HU, respectively. The original CT images, binary masks, artifact-insert CT images, and reference CT images are successively presented from the left to the right sides of this figure. In the follow-up training, artifact-insert CT images were used as inputs, whereas reference CT images were treated as outputs.

Fig. 1 Comparison of **a** original CT images, **b** metal masks, **c** artifact-insert CT images, and **d** reference CT images in the same slices.

U-Net and pix2pix models

U-Net was initially applied in the biomedical image segmentation task, and context capturing and accurate positioning were combined through interconnected layers. Relative to the original U-Net, leaky ReLUs were used to replace ReLUs as the activation function. Batch normalization was applied before the activation of a function, and the training parameters were initialized by using the Xavier method [25]. Fig. 2 shows the architecture of this network. The sum of mean absolute error (MAE) and mean squared error (MSE) was used in the U-Net loss function.

$$MAE = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} \left| X(i, j) - Y(i, j) \right|$$

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

$$\text{loss}(UNet) = MAE + MSE$$

where X and Y are the CT images to be compared, and H and W are the height and width of the images, respectively.

The pix2pix model uses the previous U-Net architecture as a generator, and its discriminator is shown in Fig. 3. This model consists of five convolution layers, the first four

activation functions are leaky ReLUs, and the last function is sigmoid. The modified least-squared adversarial loss was used in the pix2pix loss function.

$$\text{loss}(pix2pix) = \text{loss}(UNet) + \log(D(X)^2) + \log(1 - D(G(Z))^2)$$

Peak signal-to-noise ratio (PSNR) [26] and SSIM [27] were used to compare the prediction results as defined below:

$$PSNR = 10\log_{10}\frac{(MAX)^2}{MSE}$$

where MAX is the range of images. PSNR is expressed in dB, and an increasing PSNR denotes an improving image quality.

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

where μ_x is the mean intensity of image x, σ_x is the variance, σ_{xy} is the correlation coefficient, $C1=(K1L)^2$, $C1=(K1L)^2$, L is the dynamic range of the pixel values, and the default values are K1=0.01 and K2=0.03.

The learning rate was set to 0.0001, the momentum factors of the Adam algorithm were 0.9 and 0.999, the batch size was 16, the size of the convolution kernel was 3×3, and the number of epochs was 1,000.

The tests were performed on an Intel(R) Core (TM) i9-9900K @ 3.60 GHz CPU with an NVIDIA GEFORCE RTX 2080 Ti graphics card and 12G video memory. Matlab R2016b was used to generate the artifact images, and Tensorflow 1.14 was used as the training platform.

Fig. 2 Architecture of the modified U-Net for MAR

Fig. 3 Pix2pix discriminator architecture

Treatment planning

The test data include 10 vertebral tumors who received radiotherapy. The volumetric modulated arc radiotherapy plan was designed in the reference CT images by a commercial treatment planning system (Monaco 5.11, Elekta, Sweden), and each planned target volume (PTV) was given at the prescribed dose of 40 Gy/20 fractions. The isocenter was the center of PTV, and the gantry rotated from −180° to 180° clockwise and then −180° anticlockwise [28]. 95% of the PTV was covered by the prescribed dose (40 Gy), the percent volume of the kidney covered by 20 Gy was <20%, and the maximum dose of the spinal cord was < 45 Gy.

Dose distributions were calculated by using the Monte Carlo method, where the dose grid and each control point accuracy were set to 3 mm and 2%, respectively. The irradiation field was copied to the artifact-insert CT images and images generated by LIMAR, NMAR, U-Net, and pix2pix. Dose was directly calculated without plan optimization. A non-parametric Wilcoxon signed-rank test was performed to compare the average CT and mean dose values. The level of statistical significance was set to p=0.05, and all computations were performed by using SPSS version 20.0 (IBM, Chicago, IL).

RESULTS

For the computer configuration used in this study, the training times of the whole training set were approximately 28 h and 30 h for U-Net and pix2pix, respectively. The loss function in the training process reached a plateau (Fig. 4).

Fig. 4 Training loss function in **a** U-Net, and **b** pix2pix

Fig. 5 presents visual examples of synthetic CT images obtained by different methods from a test subject. The U-Net and pix2pix model effectively reduces metal artifacts, whereas LIMAR and NMAR create new streaking artifacts and a blurred vertebral shape. This phenomenon was similar to that observed by Meyer et al. [11]. Compared with U-Net, pix2pix keeps obtains better textures with the help of a discriminator, such as in the liver, stomach, and kidney as denoted by the arrows in the figure.

Fig. 5 Visual comparison of MAR images obtained by using different methods. For one test subject, we show **a** the reference image and artifact-insert image and the MAR image and its difference image (compared with the reference image) generated by **b** LIMAR, **c** NMAR, **d** U-Net, and **e** pix2pix. The small text in each sub-image denotes the corresponding accuracy. The texture differences between U-Net and pix2pix results are presented by red arrows.

Fig. 6 shows the statistical histogram of the CT value difference between the reference images and other images. As shown in Fig. 6 (a) and (b), most of the CT value differences between the reference images and those images generated by LIMAR and NMAR were distributed within a positive interval. Meanwhile, in Fig. 6 (c) and (d), the CT value differences between the reference images and those images generated by U-Net and pix2pix were slightly around the concentration distribution of 0 HU in the histogram.

Fig. 6 Statistical histogram of the CT value differences between the reference images and those images generated by **a** LIMAR, **b** NMAR, **c** U-Net, and **d** pix2pix

Table 1 compares the results for the reference images and those images obtained by different methods. We used 1-SSIM because the SSIM result is extremely close to 1. Table 2 presents the average CT values for several organs in different types of CT images. The mean CT values of the liver, spleen, and left and right kidneys are significantly higher than those of the reference images ($p<0.05$). The six types of images exhibit no statistical difference from the mean CT values of the stomach due to the gastric filling differences among the patients. Table 3 shows the relative dose in 10 patients. The mean dose of PTV in the artifact-insert images is significantly lower than that in the reference images ($p=0.028$).

Fig. 7 shows the dose distribution from one patient, and the processed images show no significant differences in the reference images, except for the artifact-insert images. The dose distribution in the reference CT images is calculated as reference, and the γ passing rates of the dose in the other images are calculated under 1% absolute dose difference and 1 mm distance-to-agreement criteria. Fig. 8 and Fig. 9 show typical transverse-sectional and coronal-section γ passing rates within 60 mm from the center, and Table 4 presents the quantitative results. The average γ passing rate of the artifact-insert images is significantly lower than that of other types of images ($p<0.05$).

Fig. 10 shows the metal artifact reduction for a patient with a true metal artifact. The window width and window level are 40 and 1000 HU for these two figures. LIMAR and NMAR results blur the images while U-Net and pix2pix can reduce metal artifacts. However, the results of U-Net and pix2pix are not as satisfactory as in the simulated data because the simulated metal artifact differs from the real metal artifact due to the simulation parameters (e.g., mask size and shape and metal materials) and method. Therefore, some differences can be observed between the training and real patient datasets.

Table 1 MAE, MSE, and PSNR of CT images with LIMAR, NMAR, and U-Net and those of the reference CT images

Table 2 Average CT values of different organs in different types of CT images

Table 3 Relative dose differences in different types of CT images

Fig. 7 Dose distribution from one patient. **a** Reference image, **b** artifact-insert image, and MAR images generated by **c** LIMAR, **d** NMAR, **e** U-Net, and **f** pix2pix.

Table 4 Relative dose differences in different types of CT images

Fig. 8 γ passing rate distribution in different types of images compared with the dose of reference images used as the criterion in the transverse section. Those values greater than 1 denote failure. **a** Artifact-insert image, and MAR images generated by **b** LIMAR, **c** NMAR, **d** U-Net, and **e** pix2pix

Fig. 9 γ passing rate distribution in different types of images compared with the dose of reference images used as the criterion in the coronal section. Those values greater than 1 denote failure. **a** Artifact-insert image, and MAR images generated by **b** LIMAR, **c** NMAR, **d** U-Net, and **e** pix2pix

Fig. 10 Metal artifact reduction for two patients with a true metal artifact. **a** Original image, and MAR images generated by **b** LIMAR, **c** NMAR, **d** U-Net, and **e** pix2pix

Discussion

Metal implants in the vertebras of a patient will generate serious metal artifacts, thereby resulting in the deviation of CT values, which, in turn, will influence doctors in their judgment of tissue structures. Moreover, the rays passing through the artifact-influenced region in radiotherapy may lead to dose inaccuracy.

Liao et al. [29] divided CT artifacts into the image domain without artifacts and the artifact-affected domain to design a new artifact disentanglement network that removes artifacts from an artifact-affected image. Huang et al. [30] designed the RL-ARCNN residual network to produce cervical artifact-insert CT images and clinical real CT images and demonstrated its good performance in correcting cervical artifacts. Lee et al. [31] divided MAR into (1) segmentation of metal trace, (2) a fully convolutional network (FCN)-based restoration in the sinogram domain, and (3) FCN-based restoration in the image domain followed by a metal insertion to test its effects on physical teeth phantom and clinical data. Afterward, they used a modified U-Net and pix2pix model to directly construct the mapping relation between artifact-insert and reference CT images.

In this paper, a simulated dataset including reference and artifact-insert images in the vertebra was generated. Given the great success of the U-Net and pix2pix model in image translation tasks, we reduced the metal artifacts based on a modified U-Net and pix2pix model. Experimental results on a simulated dataset show that this model can remove metal artifacts more efficiently compared with LIMAR and NMAR, which create new streaking

artifacts. In TPS plans, the dose distribution difference between artifact-insert and reference images is significant. In terms of the transverse- and coronal-sectional γ passing rate distribution within 60 mm from the center, the U-Net and pix2pix model obtains better results than LIMAR and NMAR.

Nevertheless, this model can be further improved in two aspects. First, 2D data were used for training, and the training accuracy can be enhanced by refining the U-Net structure [32] by using 3D information [33] or a more advanced adversarial neural network [34]. Second, artifact-insert and reference CT images were produced through an artifact generation method that partially solves the problem of collecting artifacts and original images and eliminating the image synthesis error caused by image registration. However, the simulated artifact-insert CT images show certain differences from the real artifact CT. Therefore, while we obtain excellent results on the simulated datasets, our results on real patient images are not as good. Therefore, in our future study, we will reduce the differences between simulated and real images.

Conclusion

We have generated simulated data, including reference and artifact-insert images, and proposed a modified U-Net and pix2pix model for metal artifacts reduction. Results show that this model can reduce more metal artifacts compared with LIMAR and NMAR and that pix2pix can restore more texture than U-Net with the help of a discriminator. Metal artifacts can lead to dose calculation uncertainty in radiotherapy. Dose distribution can be accurately calculated based on the images generated by U-Net and pix2pix. The difference between the real and simulated patient data leads to unsatisfactory results and therefore require follow-up research.

Abbreviations

PTV: planned target volume; MAR: metal artifact reduction; LIMAR: linear interpolation metal artifact reduction; NMAR: normalized metal artifact reduction; MAE: mean absolute error; MSE: mean square error; PSNR: peak signal-to-noise ratio; SSIM: structural similarity.

Author contributions

Xie Kai and Gao Liugang contributed equally to this work, participated in the design of the study, carried out the study, performed the statistical analysis, and drafted the manuscript. Lu Zhengda, Li Chunying, Bi Hui, Lin Tao and Sui Jianfeng helped to carried out the study. Dai Jianrong reviewed and edited the manuscript. Ni Xinye conceived and designed the study,

edited and reviewed the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (No. 81871756), the Science and Technology Programs for Young Talents of Changzhou Health Commission (No. QN201932), Changzhou Science and Technology Program (No .CJ20200099）, Young Talent Development Plan of Changzhou Health Commission (NO. CZQM20200067 and CZQM20200075), and Changzhou Key Laboratory of Medical Physics (No. CM20193005).

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Ethics approval and consent to participate

This study was approved by the Research Ethics Board of the Second People's Hospital of Changzhou, Nanjing Medical University. Written informed consent was not required following national and institutional guidelines.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Sakaura H, Hosono N, Mukai Y, Ishii T, Yonenobu K, Yoshikawa H. Outcome of total en bloc spondylectomy for solitary metastasis of the thoracolumbar spine. J Spinal Disord Tec. 2004; 17:297-300.
2. Pekmezci M, Dirican B, Yapici B, Yazici M, Alanay A, Gurdalli S. Spinal Implants and Radiation Therapy: The Effect of Various Configurations of Titanium Implant Systems in a Single-Level Vertebral Metastasis Model. J Bone Joint Surg AM. 2006; 88:1093-100.
3. Stradiotti P, Curti A, Castellazzi G, Zerbi A. Metal-related artifacts in instrumented spine: echniques for reducing artifacts in CT and MRI-state of the art. Eur Spine J. 2009; 18:102-8.
4. Wang Y, Qian B, Li B, Qin G, Zhou Z, Qiu Y, et al. Metal artifacts reduction using monochromatic images from spectral CT: evaluation of pedicle screws in patients with

scoliosis. Eur J Radiol. 2013; 83:360-66.

5. Kotsenas A L, Michalak G, Delone D R, Diehn F E, Grant K L, Halaweish A F, et al. CT Metal Artifact Reduction in the Spine: Can an Iterative Reconstruction Technique Improve Visualization? AM J Neuroradiol. 2015; 36:2184-90.
6. Katsura M, Sato J, Akahane M, Kunimatsu A, Abe O. Current and Novel Techniques for Metal Artifact Reduction at CT: Practical Guide for Radiologists. Radiographics. 2018; p.450-461.
7. Xia D, Roeske J C, Yu L, Pelizzari C A, Mundt A J, Pan X. A hybrid approach to reducing computed tomography metal artifacts in intracavitary brachytherapy. Brachytherapy. 2005; 4:18-23.
8. Boas FE, Fleischmann D. Evaluation of two iterative techniques for reducing metal artifacts in computed tomography. Int J Med Radiol. 2011; 259:894–902.
9. Aissa J, Boos J, Schleich C, Sedlmair M, Krzymyk K, Kröpil P, et al. Metal artifact reduction in computed tomography after deep brain stimulation electrode placement using iterative reconstructions. Invest Radiol. 2016; 52:18–22.
10. Kalender WA, Hebel R, Ebersberger J. Reduction of CT artifacts caused by metallic implants. Radiology. 1987; 164:576–7.
11. Meyer E, Raupach R, Lell M, Schmidt B, and Kachelries M. Normalized metal artifact reduction NMAR; in computed tomography. Med Phys. 2010; 37:5482-93.
12. Roeske J C, Lund C, Pelizzari C A, Pan X, Mundt A J. Reduction of computed tomography metal artifacts due to the Fletcher-Suit applicator in gynecology patients receiving intracavitary brachytherapy. Brachytherapy. 2003; 2:207-14.
13. Meyer E, Raupach R, Lell M, Schmidt B, Kachelries M. Frequency split metal artifact reduction FSMAR; in computed tomography. Med Phys. 2012; 39:1904-16.
14. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. IEEE T Image Process. 2017; 26:3142-55.
15. Dong C, Loy CC, He K, Tang X. Image Super-Resolution Using Deep Convolutional Networks. IEEE Trans Pattern Anal Mach Intell. 2014; 38:295-307.
16. Nie D, Trullo R, Lian J, Wang L, Petitjean C, Ruan S, et al. Medical Image Synthesis with Deep Convolutional Adversarial Networks. IEEE Trans. Biomed Eng. 2018; 65:2720-30.
17. Gjesteby L, Yang Q, Yan X, Claus B E H, Shan H. Deep learning methods for CT image-domain metal artifact reduction. Int Soc Optics Photon. 2017; 9:1-6.
18. Park H S, Lee S M, Kim H P, Seo J K, Chung Y E. Sinogram-consistency learning in CT

for metal artifact reduction. Med Phys. 2018; 45: 5376-5384.

19. Zhang Y, Yu H. Convolutional neural network based metal artifact reduction in X-ray computed tomography. IEEE Trans Med Imaging. 2018; 37:1370–82.
20. Xu X, Li L, Zhang L, Wang Q. A metal projection segmentation algorithm based on Random walks for dental CBCT metal artifacts correction. IEEE Nuclear Science Symposium and Medical Imaging Conference. 2013; p.1-4.
21. Yasaka K, Maeda E, Hanaoka S, Katsura M, Sato J, Ohtomo K. Single-energy metal artifact reduction for helical computed tomography of the pelvis in patients with metal hip prostheses. JPN J Radiol. 2016; 34:625-32.
22. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Computer Vision and Pattern Recognition. 2015; p.234-241.
23. Isola P, Zhu J, Zhou T, Efros A A. Image-to-Image Translation with Conditional Adversarial Networks. Computer vision and pattern recognition. 2017; p.5967-76.
24. Kalender W A, Hebel R, Ebersberger J. Reduction of CT artifacts caused by metallic implants. Radiology. 1987; 164: 576-577.
25. Glorot X, Bengio Y. Understanding the difficulty of training deep feed forward neural networks. International Conference on Artificial Intelligence and Statistics. 2010; 5:249-56.
26. Chen G, Yang C, Xie S. Gradient-based structural similarity for image quality assessment. International Conference on Image Processing. 2006; 1:2929-32.
27. Wang Z, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing. 2004; 13:600-12.
28. Dobler B, Weidner K, Koelbl O. Application of volumetric modulated arc therapy VMAT; in a dual-vendor environment. Radiat Oncol. 2010; 5:95-104.
29. Liao H, Lin W, Yuan J, Zhou S K, Luo J. ADN: Artifact Disentanglement Network for Unsupervised Metal Artifact Reduction. IEEE Trans Med Imaging. 2019; p.1-10.
30. Huang X, Wang J, Tang F, Zhong T, Zhang Y. Metal artifact reduction on cervical CT images by deep residual learning. Biomed Eng Online. 2018; 17: 175-89.
31. Lee D, Park C, Lim Y, Cho H. A Metal Artifact Reduction Method Using a Fully Convolutional Network in the Sinogram and Image Domains for Dental Computed Tomography. J Digit Imaging. 2019; p.1-9.

32. Ibtehaz N, Rahman M S. MultiRes UNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural Networks. 2020; 121: 74-87.
33. Cicek O, Abdulkadir A, Lienkamp S S, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. Medical image computing and computer assisted intervention. 2016; p.424-32.
34. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S. Generative adversarial nets. Advances in neural information processing systems. 2014; p.2672-80.

Figures

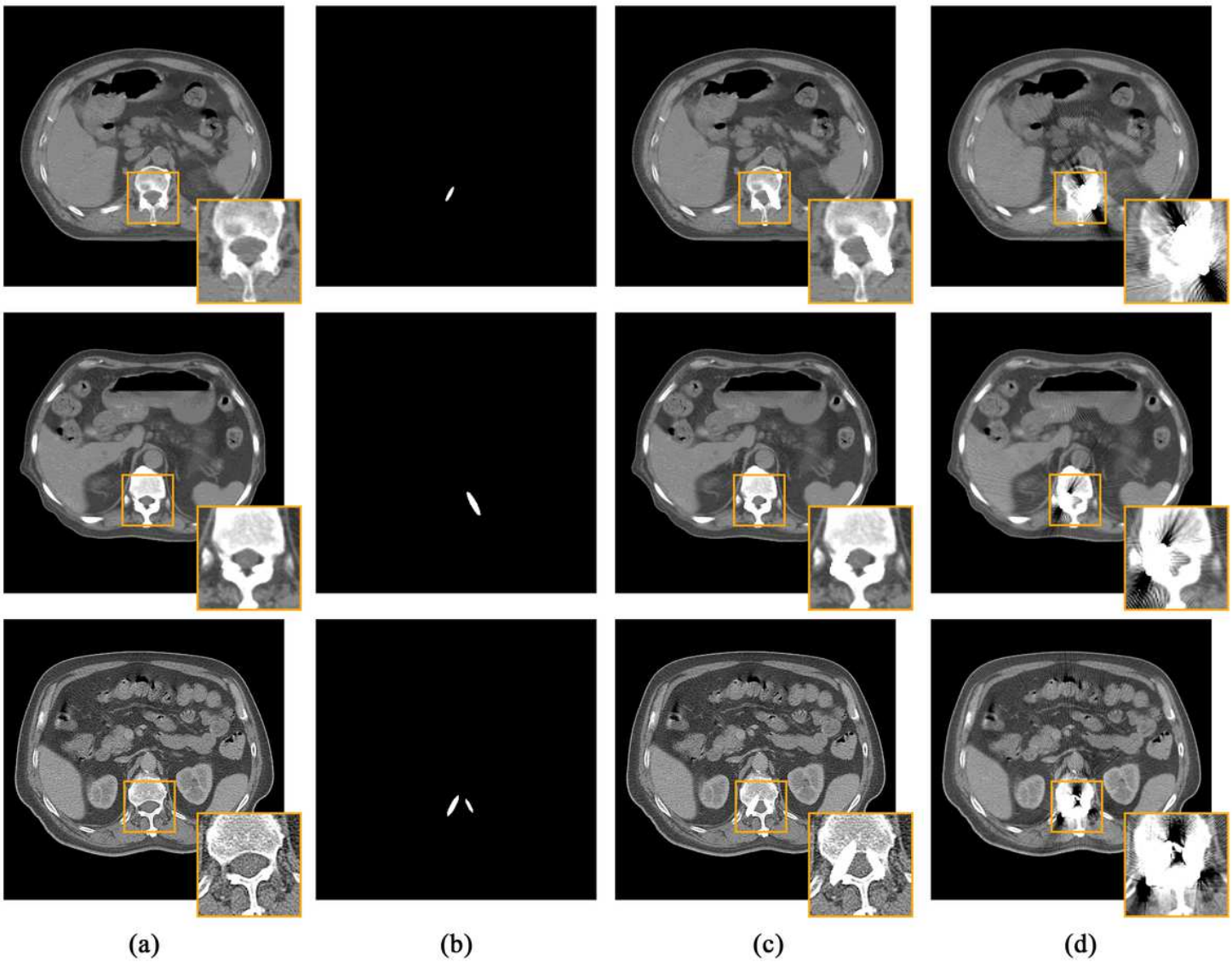

Figure 1

Comparison of a original CT images, b metal masks, c artifact-insert CT images, and d reference CT images in the same slices.

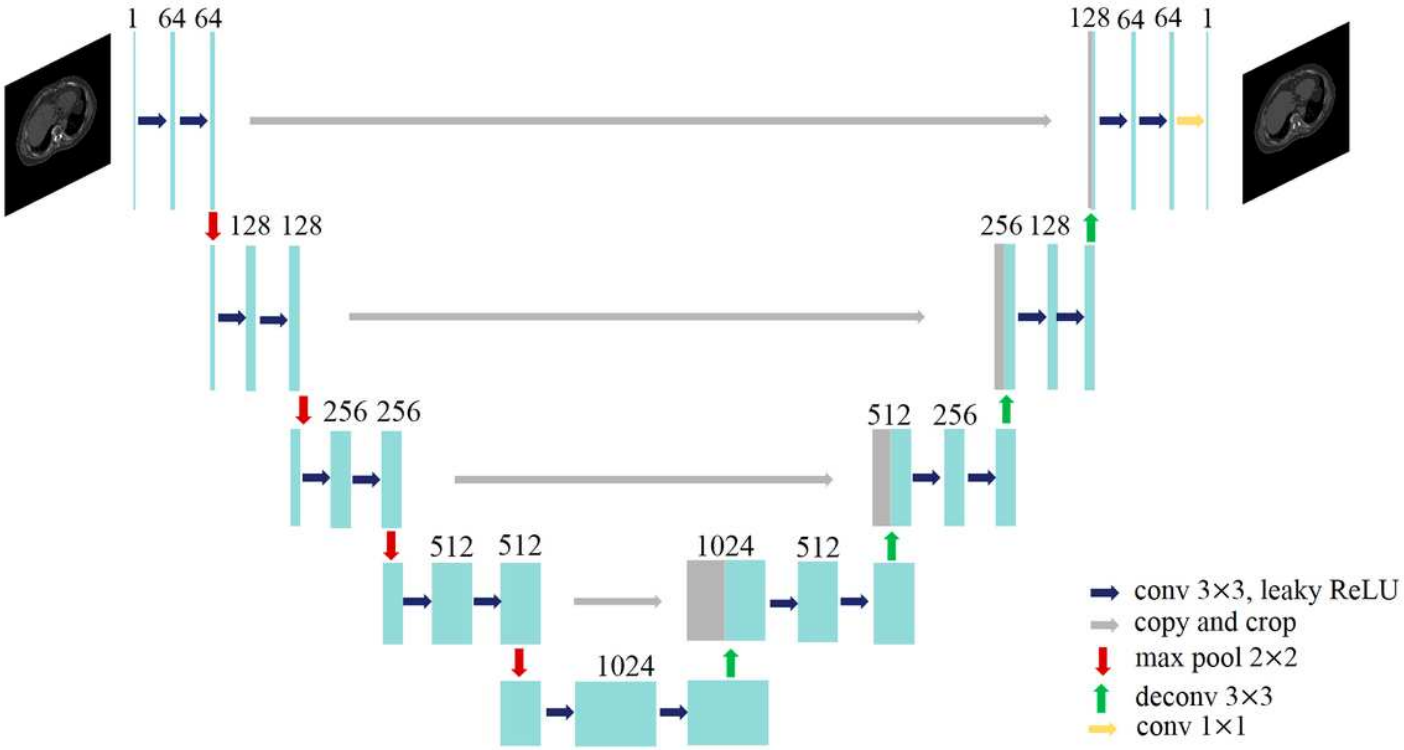

Figure 2

Architecture of the modified U-Net for MAR

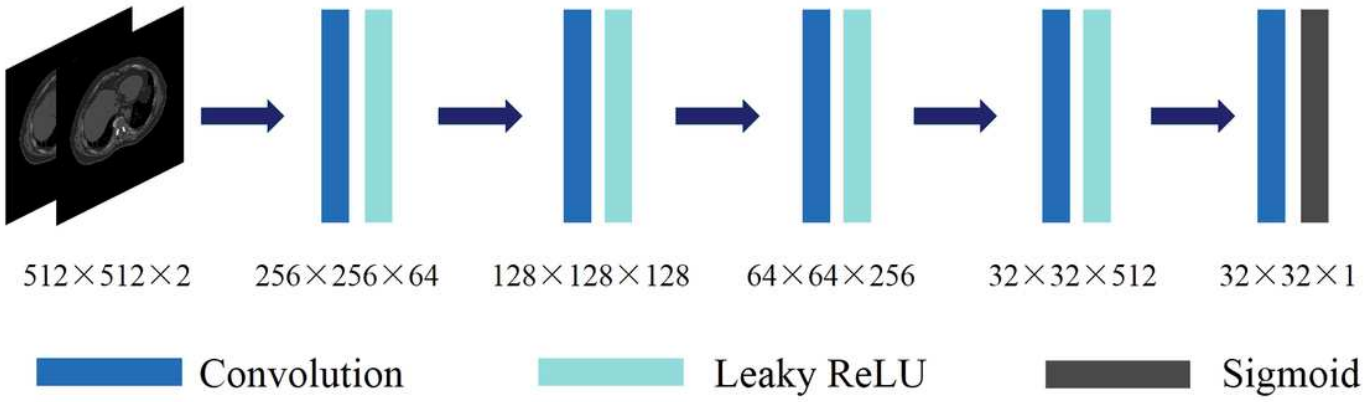

Figure 3

Pix2pix discriminator architecture

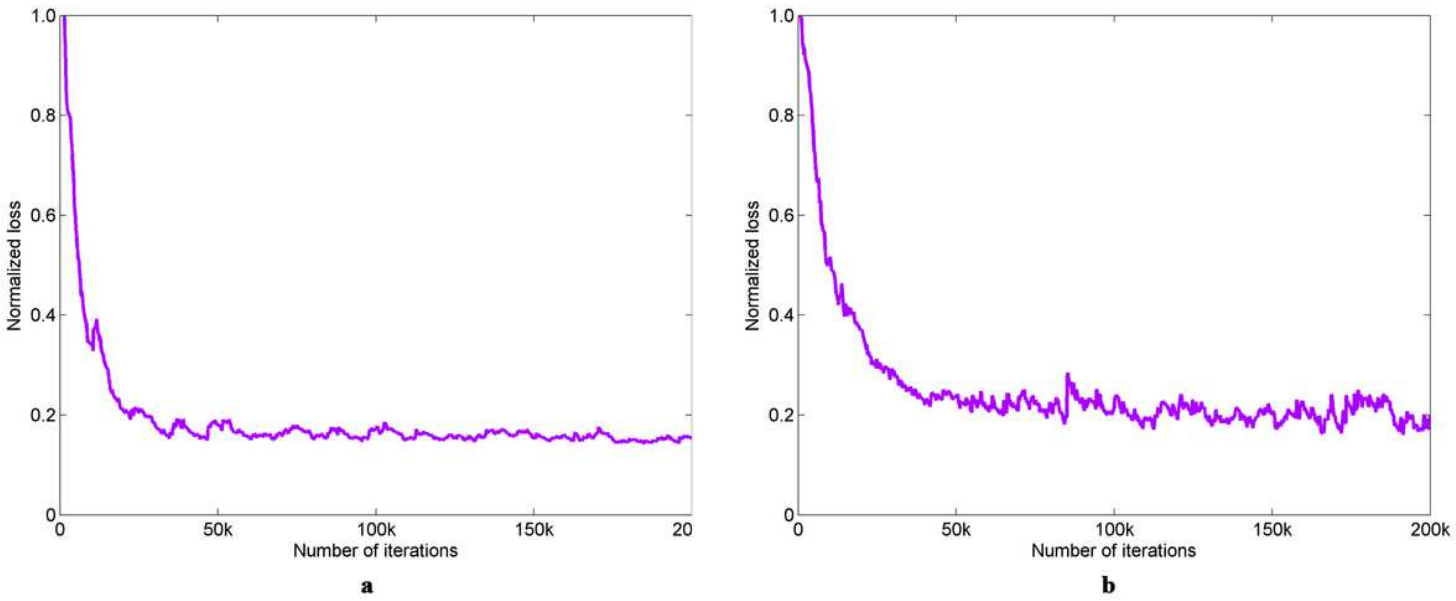

Figure 4

Training loss function in a U-Net, and b pix2pix

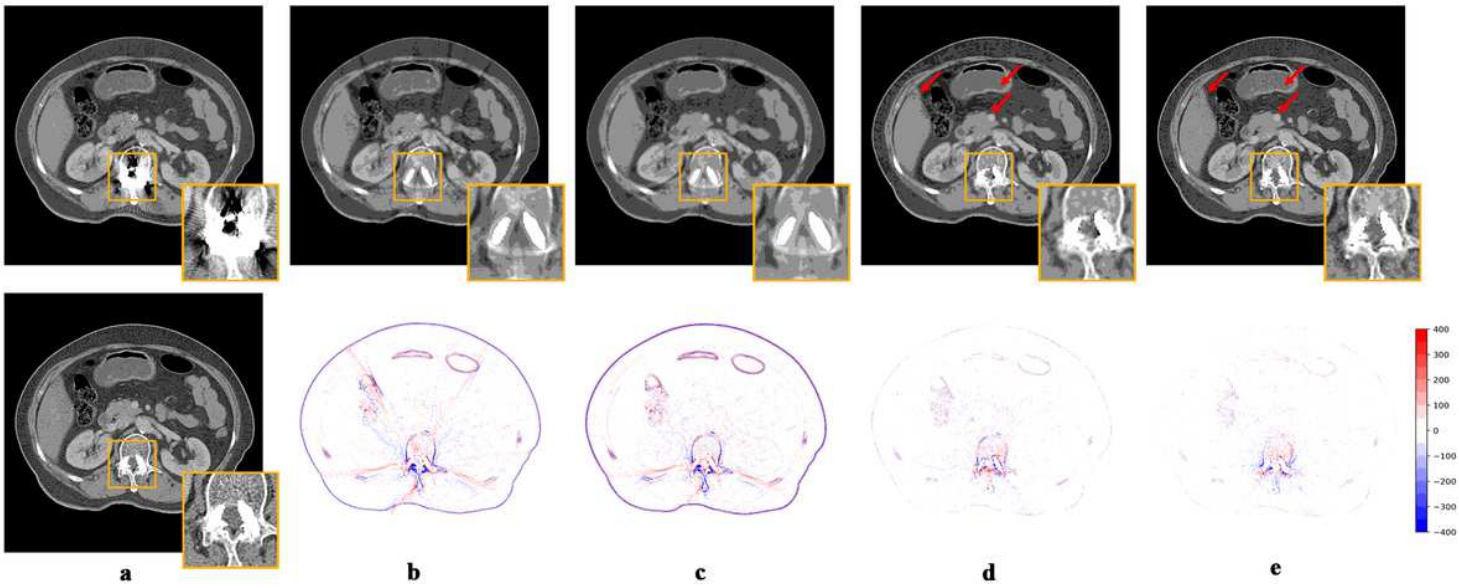

Figure 5

Visual comparison of MAR images obtained by using different methods. For one test subject, we show a the reference image and artifact-insert image and the MAR image and its difference image (compared with the reference image) generated by b LIMAR, c NMAR, d U-Net, and e pix2pix. The small text in each sub-image denotes the corresponding accuracy. The texture differences between U-Net and pix2pix results are presented by red arrows.

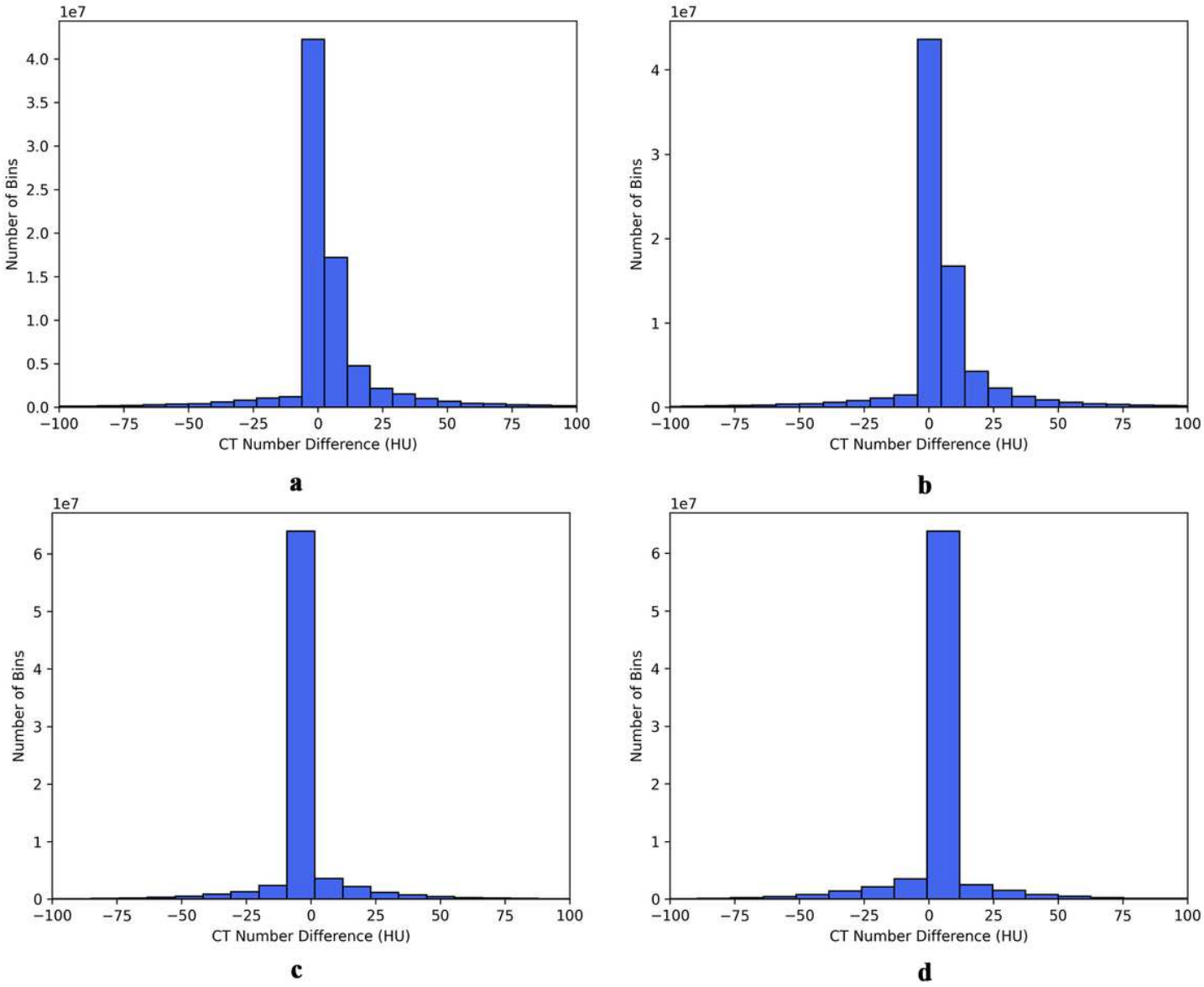

Figure 6

Statistical histogram of the CT value differences between the reference images and those images generated by a LIMAR, b NMAR, c U-Net, and d pix2pix

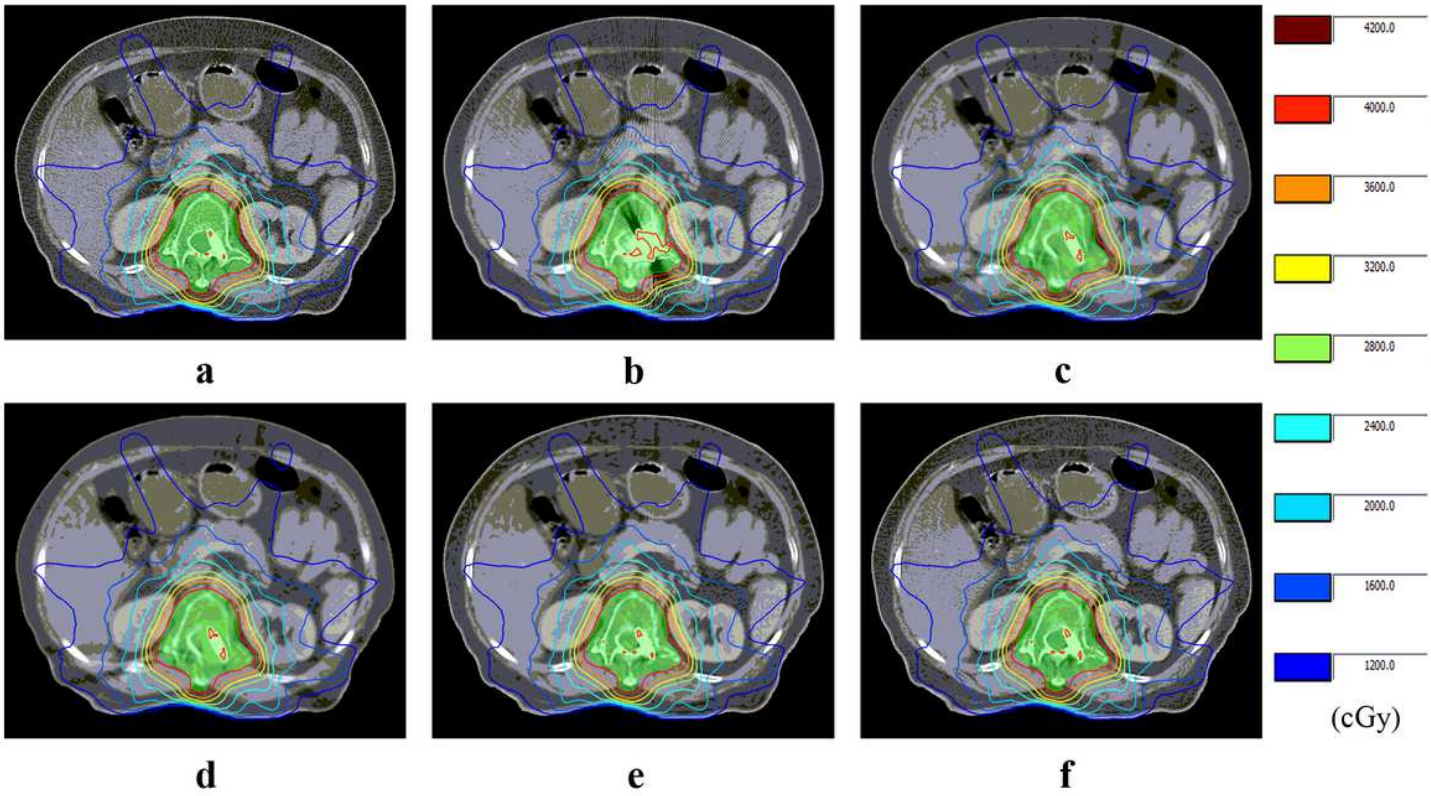

Figure 7

Dose distribution from one patient. a Reference image, b artifact-insert image, and MAR images generated by c LIMAR, d NMAR, e U-Net, and f pix2pix.

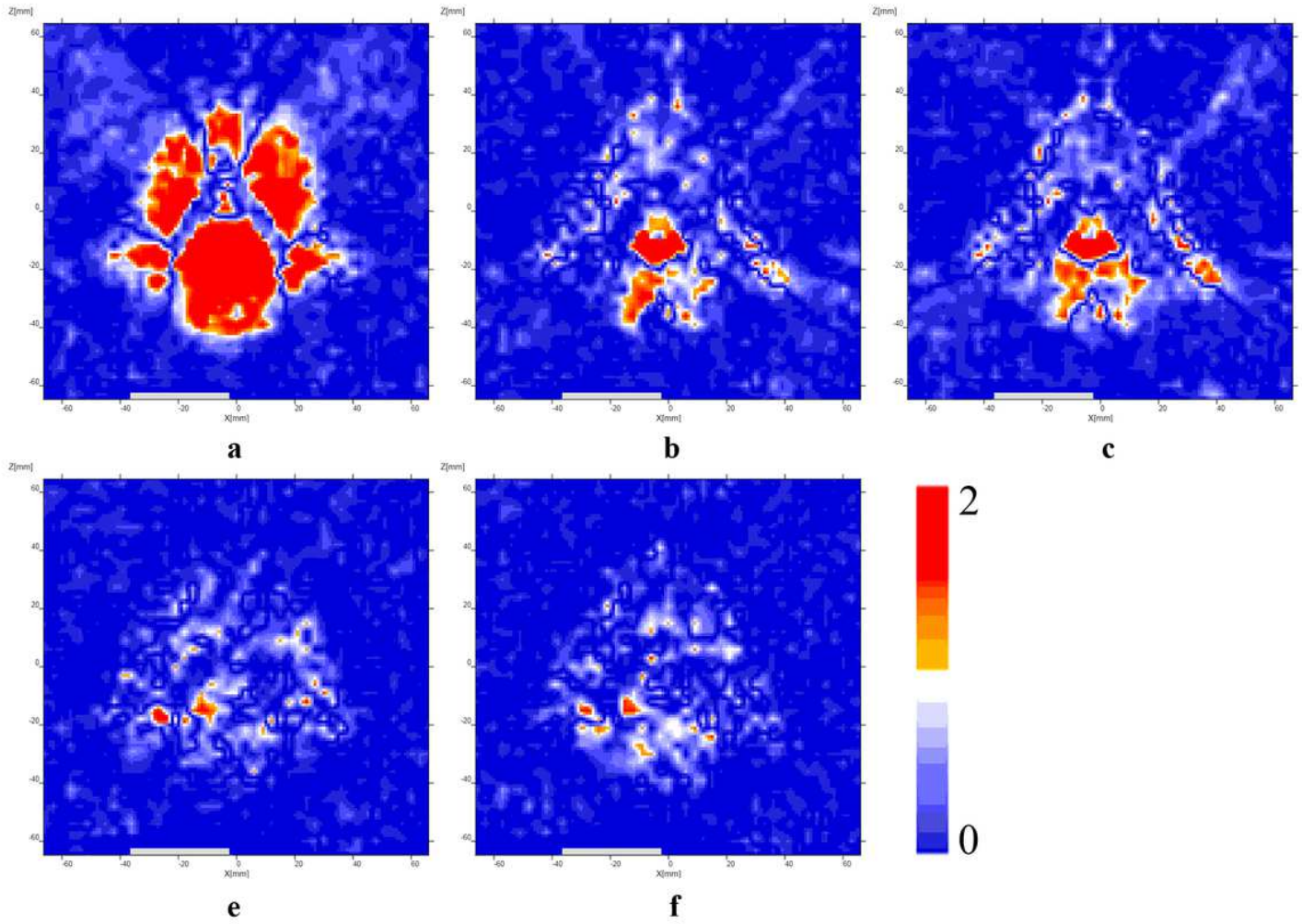

Figure 8

γ passing rate distribution in different types of images compared with the dose of reference images used as the criterion in the transverse section. Those values greater than 1 denote failure. a Artifact-insert image, and MAR images generated by b LIMAR, c NMAR, d U-Net, and e pix2pix

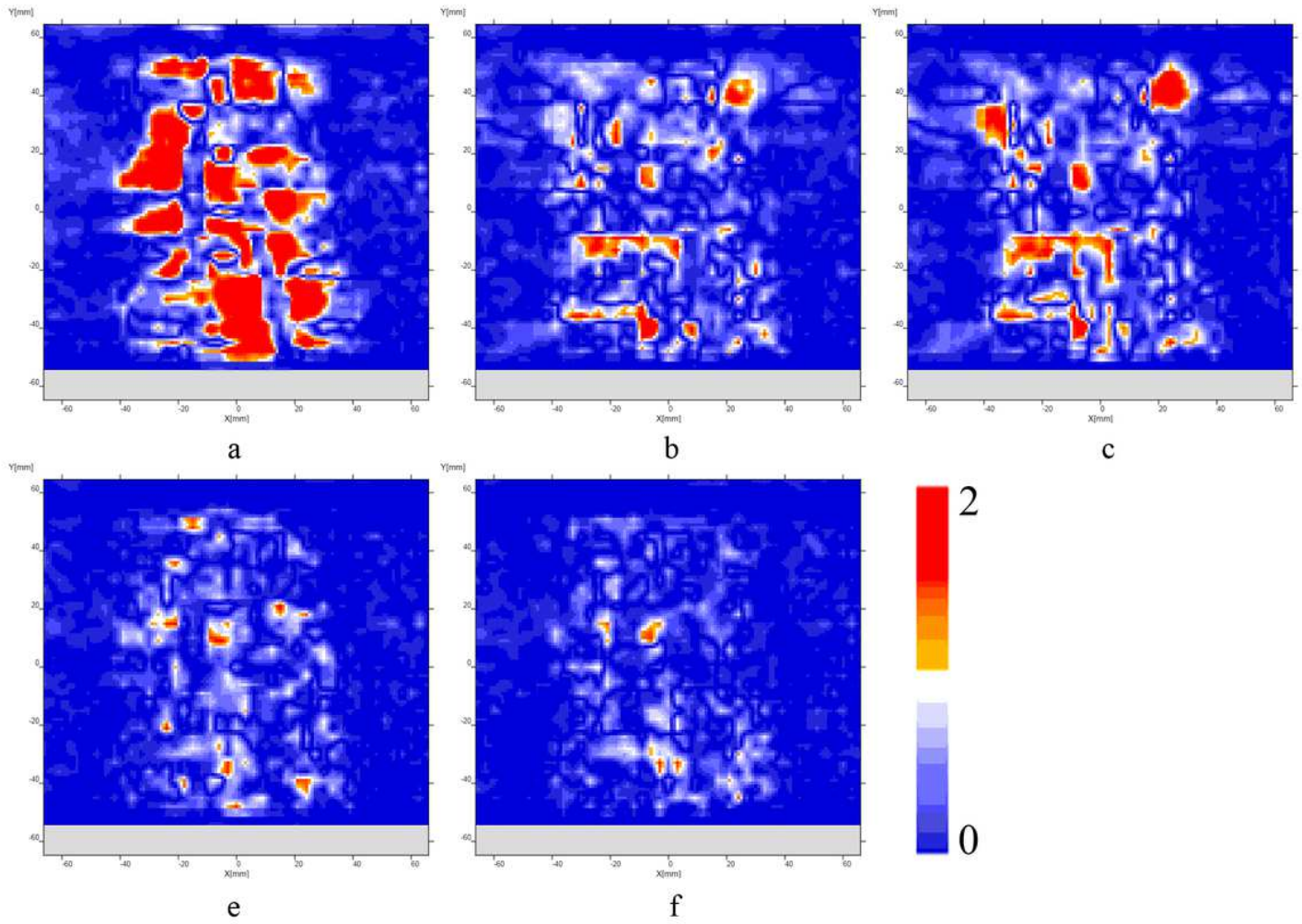

Figure 9

γ passing rate distribution in different types of images compared with the dose of reference images used as the criterion in the coronal section. Those values greater than 1 denote failure. a Artifact-insert image, and MAR images generated by b LIMAR, c NMAR, d U-Net, and e pix2pix

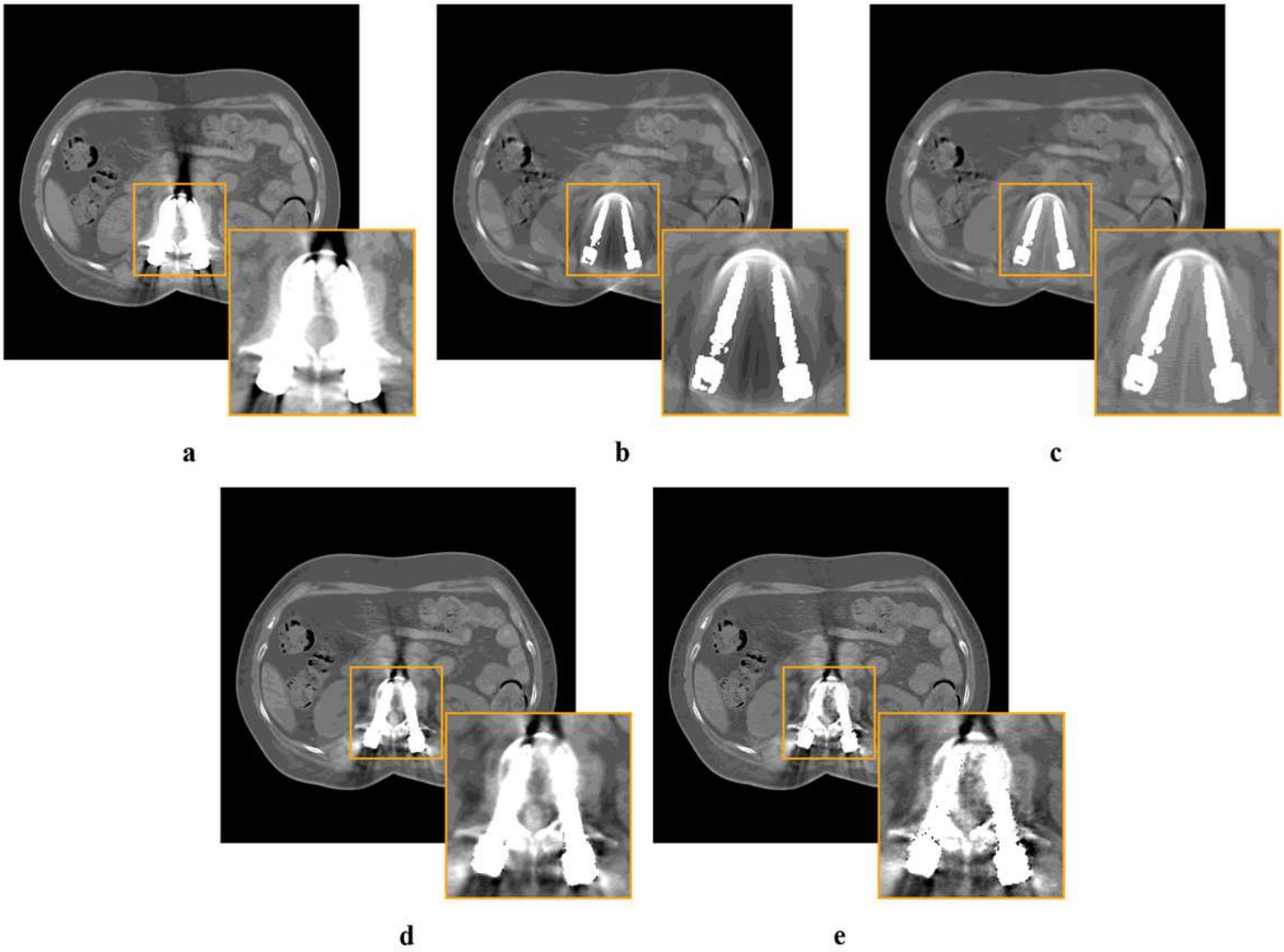

Figure 10

Metal artifact reduction for two patients with a true metal artifact. a Original image, and MAR images generated by b LIMAR, c NMAR, d U-Net, and e pix2pix