IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 4, 2019, accepted July 17, 2019, date of publication July 22, 2019, date of current version August 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2930302

Metal Artifact Reduction for X-Ray Computed Tomography Using U-Net in Image Domain

LINLIN ZHU, YU HAN, LEI LI, XIAOQI XI, MINGWAN ZHU, AND BIN YAN
National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450001, China

Corresponding author: Bin Yan (ybspace@hotmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61701089, and in part by the National Foundation of Henan Province of China under Grant 162300410333.

ABSTRACT Metal artifacts seriously degrade the quality of the CT data and bring great difficulties to subsequent image processing and analysis, which nowadays become a great concern in X-ray CT applications. In this paper, we introduce a U-net-based metal artifact reduction method into CT image domain. The proposed network reduces metal artifacts by learning an end-to-end mapping of images from metal-corrupted CT images to their corresponding artifact-free ground truth images. We design and optimize the network through the simulation experiments. The experimental results show that the proposed method can well reduce metal artifacts of CT images, and this method has higher computational efficiency and greatly shortens the processing time. It avoids complex image preprocessing and can accept input images of any size. Therefore, it can be a more automated way to handle large amounts of data, making it ideal for existing CT workflows.

INDEX TERMS Computed tomography, metal artifact reduction, deep learning, convolutional neural network, U-net.

I. INTRODUCTION

X-ray Computed Tomography (CT) is a non-destructive imaging technique which can reconstruct the internal image of the object in a non-invasive way. In recent year, X-ray CT has utilized in a vast range of areas include diagnostic [1], industrial fields [2] and security [3]. The quality of CT reconstructed images is critical to detecting analytical results. Most reconstruction algorithms, such as the filtered back projection (FBP) algorithm has a strict ideal assumption, such as single X-ray photon energy and uniform photon propagation direction. However, under realistic conditions, the ideal assumptions are difficult to satisfy. The premise hypothesis of the reconstruction algorithm is deviated from the actual physical imaging conditions, and is shown in the reconstructed image in various artifact forms [4]. Metal artifacts are a typical one of CT image artifact [5]. The specific performance is a large number of black strips and bright radial streak artifacts around the metal. The artifacts are common emerge in the scanned objects which with two or more constituent materials with large density differences between materials. These image artifacts degrade the image quality of the CT which makes the subsequent image analysis steps of object segmentation and identification inaccurate.

Although many different approaches for metal artifact reduction (MAR) [6]–[8] has been proposed during the past four decades, metal artifacts reduce is still remains a hot and difficult issue of CT technique. The proposed MAR algorithms can be divided mainly into three groups: physics-based correction, iterative reconstruction and interpolation methods. A direct way to reduce artifacts is to correct physical effects, e.g., beam hardening [9]–[11] and photon starvation [12]. Iterative reconstruction methods generally use a model-based approach to minimize a well-defined objective function subject to defined constraints. Man *et al.* [13] proposed a maximum-likelihood-based MAR approach by using poly-energetic model. Hamelin *et al.* [14] approximated data as Gaussian distribution and used the poly-energetic model to reduce metal artifacts. Zhang *et al.* [15] suggested a penalized least squares methods subject to non-negativity constraint and discard metal traces data. Interpolation methods are one of the most popular MAR techniques; the idea is to replace metal-corrupted data with estimated values in an effort to perform correction. The values for metal-corrupted sinogram regions are typically estimated using neighboring information. Kalender *et al.* [16] consider

The associate editor coordinating the review of this manuscript and approving it for publication was Guitao Cao.

metal-corrupted data as missing data and estimate the missing data using an one-dimensional linear interpolation method (LI-MAR). LI-MAR is considered a benchmark and is so far the most cited MAR technique. Morin and Raeside [17] use nearest neighbor search to estimate the missing data. Mahnken *et al.* [18] suggested a two dimensional interpolation that replaces the metal-corrupted data by a weighted sum of 16-nearest neighbors. Zhao *et al.* [19] performs sinogram interpolation using wavelet coefficients. Normalized MAR (NMAR) [20] uses a prior image to normalize the corrupted sinogram and a simple interpolation method to estimate metal corrupted data in the normalized sinogram. NMAR heavily relies on a prior image and is not suitable for security or other applications where a correct prior image cannot be robustly found. However, this method is still effective in practice. Zhang and Mou [21] calculated the image mutual correlation among the original reconstructed image, LI image and their difference image to build a combined prior image, which contains fewer artifacts. Schmidt *et al.* [22] selectively removes energy window measurements that are corrupted by metal, while extracting information from the remaining spectral measurements. The optimization framework incorporates constraints on the reconstructed basis material maps to mitigate the undersampling effects that occur when corrupted data is excluded. Jeon and Lee [23] using forward projection of the modified CT image and generate a sinogram containing information about the anatomical structure. Whether the metal artifacts are eliminated by interpolation or iterative method, some theoretical effects have been achieved, but for various reasons, the MAR methods that can be practically used in engineering, simple, effective, and easy to implement are still scarcity.

In CT, the most important cause of metal artifacts is beam hardening and scatter, which are essentially nonlinear changes in the projected value. This makes the reconstructed image have significant features. In recent years, deep learning has rapidly gained attention for employment on image processing [24]–[26]. High-level features can be efficiently derived from pixel-level data through a hierarchical multi-layer framework. Deep learning has been successfully applied in the field of image restoration and denoising, providing a new method for MAR in CT images [27]. Gjesteby *et al.* developed a few MAR methods based on deep learning that refine the performance of the state-of-the-art MAR method, NMAR, which were in the projection domain [28] and the image domain [29], [30], respectively. Their experiments prove that CNN can effectively improve NMAR, but the remaining artifacts are still considerable. Park *et al.* used a U-net to correct the beam hardening induced by metal in the projection domain [31] and image domain [32], respectively. They proposed a MAR method based on deep learning for specific implants. The method uses simulated metal artifacts to learn real metal artifacts. The proposed method does not require ground-truth image, insofar as it learns artifacts generated exclusively by metal, rather than artifact-free CT images. However, there is residual of artifacts in the process of acquiring real artifacts. Secondly, the method is only for the special case involving two metal implants, and the application problem is limited.

The size, material, number, and location of the implanted metal may cause variations in metal artifacts. Here, we introduce the modified U-net (U-MAR net) into the field of CT images, and use the strong fitting ability of the network to complete the MAR processing of the entire image. We use U-MAR net as an artifact reduction tool to remove metal artifacts directly from uncorrected reconstruction results. In the existing network model, the feature map is compressed using the pooling layer, which will result in a change in image size. After testing and comparison, to keep the output consistent with the original image size, we removed the pooling operation in the network. In order to make full use of the depth features, we removed the copy and crop operations. Through many practical experiments, we adjusted the network parameters to further optimize the training speed and accuracy of the network. Specifically, to generate training data for our method we built a database. Our method first conducted simulation experiments by adopting the anthropomorphic head phantom. Then we generate a dataset using a shared medical image dataset. Further, we conducted the actual experiment by constructing a cylindrical phantom containing aluminum column. The network reduces the metal artifacts from learning an end-to-end mapping of images from metal-corrupted CT images to their corresponding artifact-free ground truth. The reconstructed results are directly input into the network as training data. In network training, we can capture more original information to generate higher quality images. Second, this method is an open framework that can correct images of different sizes. On this basis, if we continue to increase training data, network capacity will be further enhanced.

II. METHODS

A. U-MAR NET

U-MAR net architecture is inspired by U-net [33]. U-net as a segmentation network requires an overall analysis of the multi-layer features. Due to the saliency of metal artifact features, we tried to deal with metal artifacts with a concise structure. Through fine tuning and multiple trials, we have achieved a better balance between network performance and simplicity. After testing, we removed the copy and crop operations as well as the pooling layer. In order to improve the effective receptive field, we use stride-2 convolutions. Our network is simpler and more stable, and we can get higher training speed under quality assurance. Fig.1 depicts the network structure of our method. All layers have a 3×3 kernel. We used the rectified linear unit (ReLU) nonlinearity with a 0.2 slope at the first 7 layers. Additionally, we use the batch normalization (BN) operation at the first three layers. The first three convolution layers contain 32 filters, while the fourth contains 16 filters so that only the most important features are obtaining. The deconvolution layers use the

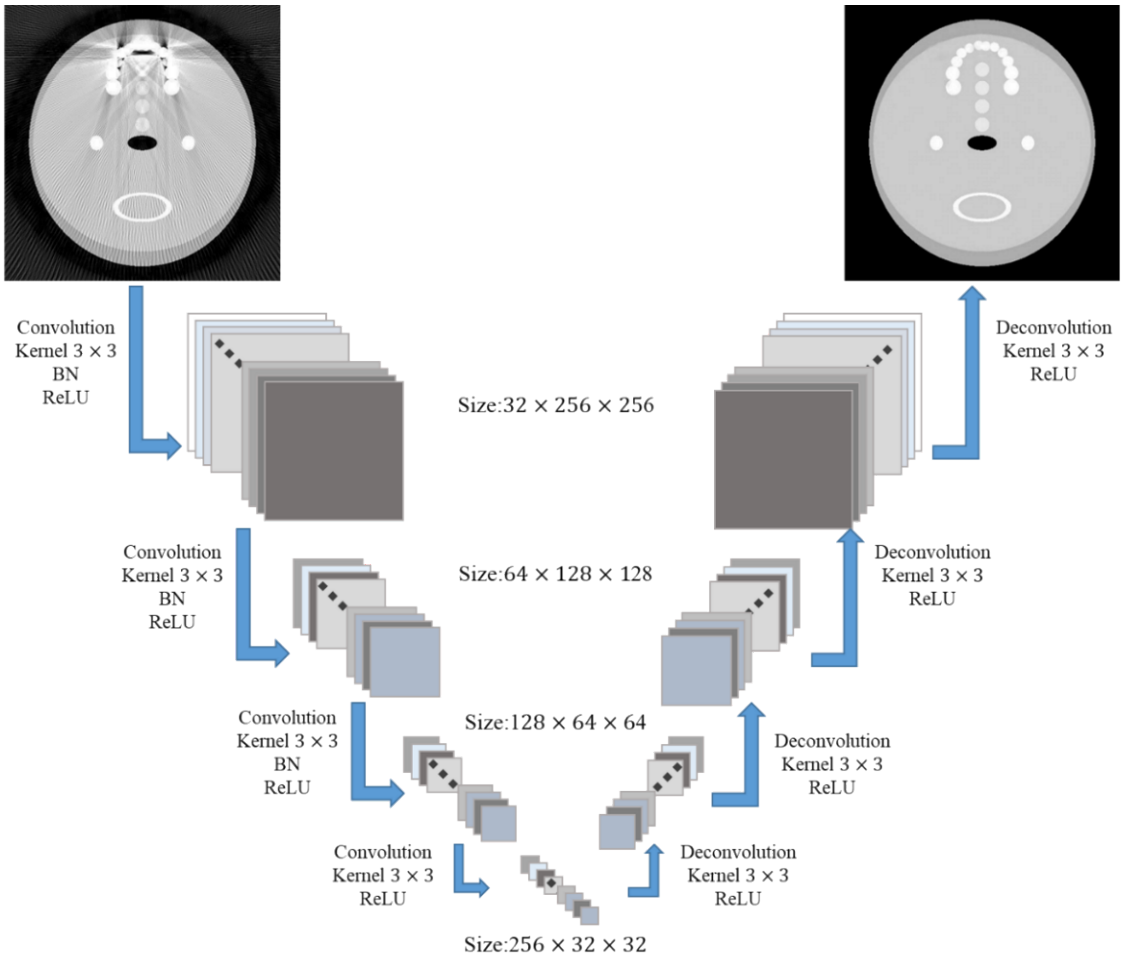


FIGURE 1. Configuration of U-MAR net for metal artifact reduction.

features to generate a feature map as the metal artifact distribution of the input images. The network reduces the metal artifacts by learning an end-to-end mapping of images from metal-corrupted CT images to their corresponding artifact-free ground truth. The input and output images are both 512 × 512 and it can be adjusted according to the specific image size. Output prediction from the extracted features is based on a mean squared error (MSE) loss function. The input to our method is the metal-corrupted image, and the corresponding target is the artifact-free image.

B. PARAMETER SELECTION

Deep network as a ''black box'' lacks sufficient theoretical analysis. In order to improve the performance of the network, we constantly adjust the parameters through experiments. In the experiment, the anthropomorphic head phantom is adopted. The phantom was designed by Catherine *et al.* [34] The attenuation coefficient distribution of this phantom is shown in Fig.2. Among them, the CT value of teeth and bones is 1000HU. Four details with a contrast of 150 HU with respect to the soft tissue background were inserted in between the teeth and the area marked with a red arrow in the figure is

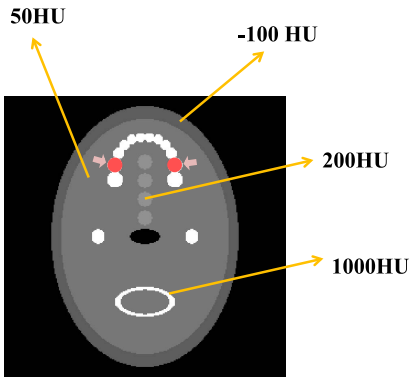


FIGURE 2. Digital jaw phantom for the simulation study. Inserting two metal into the teeth as indicated by the arrows.

an amalgam dentures. The simulation experiment uses a circular trajectory fan beam for scanning. 360-angle projection data are uniformly sampled in the 360-degree scanning range. The size of the object to be reconstructed is 512 × 512, and the number of detectors unit in the detector is 1024.

In the simulation, we adopt the second-order nonlinear model proposed in [19]. It was assumed that metal artifacts

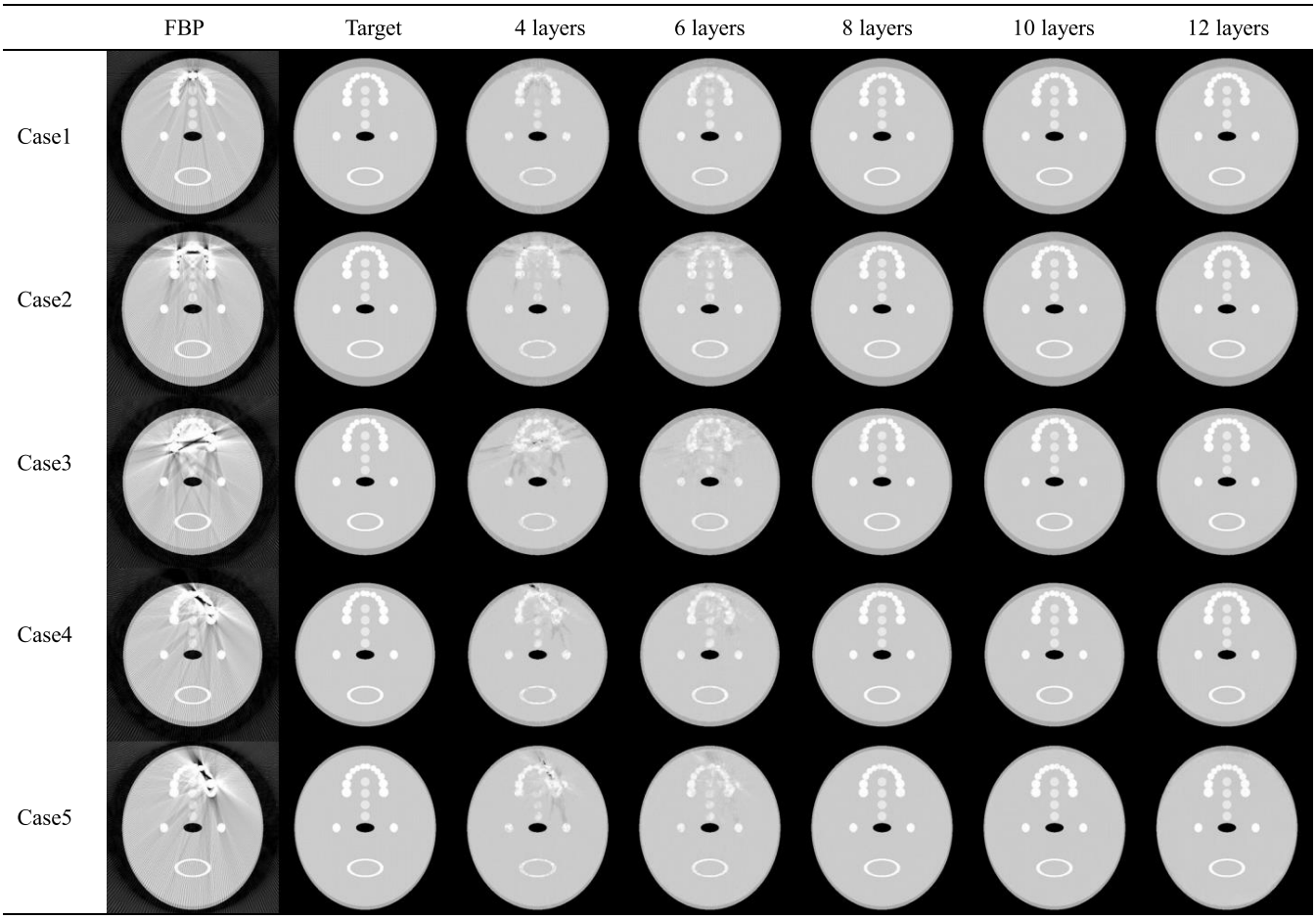


FIGURE 3. Five cases of jaw model CT images containing metal artifacts. Different case corresponds to the different position of metal implant. Window width WW = 750HU and window level WL = 0HU, respectively.

occurred only when the projection value exceeded a threshold $a > 0$ and had a magnitude $c > 0$. The projection correction formula is as follows:

$$g' = \begin{cases} g - \dfrac{2c(g-a)+1-\sqrt{4c(g-a)+1}}{2c}, & g \geq a \\ g, & g < a \end{cases} \quad (1)$$

where g' is the ray projection value after the deviation due to the influence of the metal, a is the attenuation threshold and c represents the amplitude of the nonlinear attenuation. The ray photons exceeding the threshold will produce a sharp attenuation in the high energy segment. We have set $a = 10$ and $c = 0.5$ in this paper.

In the simulations, we chose the geometry that represents a fan beam CT scanner setup. The imaging configuration are as follows: (1) The projection data includes 360 projection views at regular intervals on 512-bin linear detector arrays, (2) the distance from the detector to the light source is 600 mm, (3) the distance from the center of rotation to the source is 300 mm, and (4) the space of each detector bin is 1 mm. All reconstructed images include 512 × 512 square pixels. Each pixel has a size of 0.5mm × 0.5mm. Each projection data along the X-ray passing through the cross-sectional image is calculated as the cross-length of an X-ray with a pixel. Based on a second-order model similar to equation (1), only beam hardening is considered in the simulation. There is no simulated scattering in the simulation experiment.By changing the position and the number of metals, we generated 1,000 pairs of images for network training and generated 100 pairs of images for performance testing.

During the experiment, the learning rate started at 10^{-4} and decreased by a factor of 0.5 after every 10,000 iterations for a total of 400,000 training iterations. The network required approximately 40 hours for training. We first change the depth of the network. Through testing different layers of networks, we get the results as shown in Fig.3.

To quantitatively evaluate the results, root mean square error (RMSE) and normalized mean absolute distance (NMAD) were chosen as a measure of reconstruction quality. The definitions of RMSE and NMAD are as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f(i) - f_0(i))^2} \quad (2)$$

TABLE 1. Evaluation of results from different depth of the network.

RMSE						
	FBP	4 layers	6 layers	8 layers	10 layers	12 layers
Case1	19.60029	2.843116	2.325864	0.909089	1.180929	1.379455
Case2	21.80767	3.992375	3.27129	1.02789	1.239808	1.346155
Case3	25.5085	2.913439	2.679175	0.991559	1.276752	1.566139
Case4	25.02037	4.583205	3.175437	0.934104	1.197338	1.462531
Case5	21.66579	4.983296	3.611914	1.057285	1.153551	1.426854
NMAD						
Case1	0.168043	0.041521	0.028261	0.010008	0.011686	0.01571
Case2	0.173631	0.042576	0.03263	0.01122	0.012704	0.014452
Case3	0.210128	0.036085	0.026398	0.011688	0.01403	0.017197
Case4	0.199313	0.058276	0.030989	0.010648	0.012392	0.015315
Case5	0.174883	0.047293	0.03621	0.012206	0.013054	0.014442

TABLE 2. Evaluation of results from different kernel size of the network.

RMSE				
	FBP	3×3	5×5	7×7
Case1	19.60029	0.91842	0.93503	0.91618
Case2	21.80767	0.99390	0.94352	1.02557
Case3	25.50850	0.99731	1.13758	1.06565
Case4	21.05478	1.02819	0.98344	1.03303
Case5	20.28533	0.90279	0.96249	1.01367
NMAD				
Case1	0.16804	0.01083	0.01175	0.00965
Case2	0.17363	0.01064	0.01202	0.01094
Case3	0.21013	0.01146	0.01429	0.01284
Case4	0.17758	0.01175	0.01069	0.01174
Case5	0.17722	0.01093	0.01022	0.01063

$$NMAD = \sum_{i=1}^{N} |f(i) - f_0(i)| \Big/ \sum_{i=1}^{N} |f(i)| \tag{3}$$

where f and f_0 represent the reconstructed image and ideal phantom, respectively. i indexes the pixel in the image. N is the total number of pixels in the image. The values of RMSE and NMAD close to 0 indicate small differences between ideal phantom image and reconstructed results.

The RMSE and NMAD of each image were calculated, and the results are listed in Table 1.

Through the above indicators, we can see that the results of the eight-layer network are optimal. Convolutional networks extract features, and the more convolutional layers, the higher the extracted feature dimensions. In this research, the high-dimensional features correspond to the metal artifacts. In the process of deconvolution, after the high-dimensional metal artifact features are restored to the image, the original image and the resulting image are used to make the metal artifact reduce.

When the number of convolution layers is less, features of the lower dimensions, such as structural features of the tissue, are extracted, which increases the difficulty of resolution in the subsequent processing, resulting in a decrease in contrast of the results. When the number of convolution layers is more, the high-dimensional features are further extracted, which

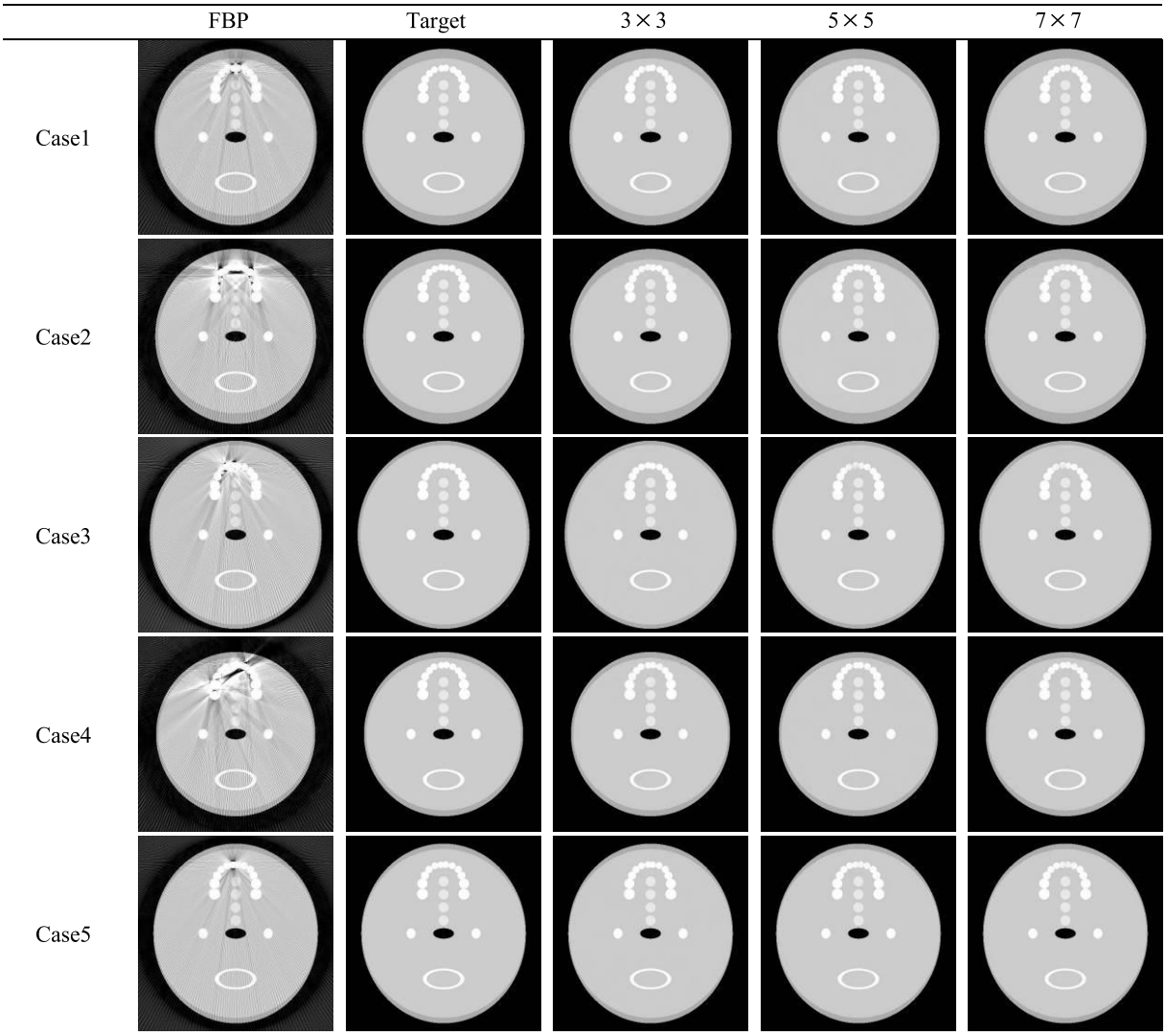


FIGURE 4. Five cases of jaw model CT images containing metal artifacts. Different case corresponds to the different position of metal implant. Window width WW = 750HU and window level WL = 0HU, respectively.

may not fully reflect the characteristics of metal artifacts, resulting in inaccurate artifact removal. At the same time, the increase in the number of convolution layers will greatly increase the parameter amount of the network. Increase the training time of the previous network. Based on the above results and analysis, this paper sets the number of network layers to eight.

After confirmed the depth of the network. We changed the size of the convolution kernel in network training. The convolution kernel is a weight matrix used to learn image-specific feature patterns. The smaller the relative difference of each element in the convolution kernel, the more the image can be subjected to fuzzy noise reduction processing; the larger the difference, the more the edge of the image can be extracted, or the sharpening effect can be obtained. In practical applications, the size of the convolution kernel is generally small, which is 3×3, 5×5 or 7×7, mainly because the convolution kernel is symmetric and has only one center, so the convolution kernel matrix is generally an odd matrix. In addition, the value of the convolution kernel matrix element affects the brightness of the processed picture. If the sum of all elements of the convolution kernel is 1, the output image is the same as the input image; If the sum of all elements of the convolution kernel is greater than 1, the output image will be brighter; if the sum of all elements is less than 1, the output image will be darkened. When the sum of all the elements of the convolution kernel is 0, most of the output image is black, and only part of the edge area has high brightness, so the convolution kernel can be used to extract image edge information. We get the results shown in Fig.4. The RMSE and NMAD of each image were also calculated, and the results are listed in Table 2.

The size of the convolution kernel has little effect on the final result of the network training, because the multiple rounds of training have fully extracted the feature parts. When the convolution kernel is small, the convergence speed can be

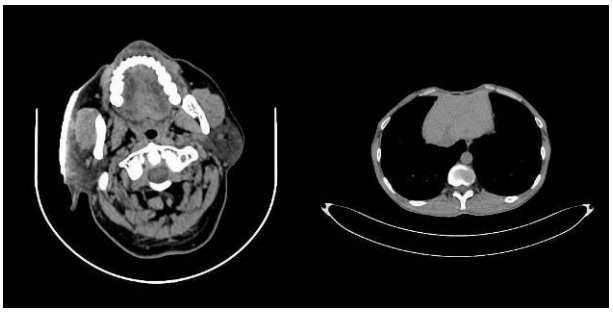


FIGURE 5. Two example scenes without metallic objects.

increased. This paper sets the size of the convolution kernel to 3×3.

III. RESULTS

A. SIMULATION STUDY

In the simulation study, we collected 500 metal-free CT images from the public medical image dataset. Two example scenes as shown in Fig.5. Based on these images we simulated metal implants images. The imaging configuration are as follows: (1) The projection data includes 360 projection views at regular intervals on 512-bin linear detector arrays, (2) the distance from the detector to the light source is 600 mm, (3) the distance from the center of rotation to the source is 300 mm, and (4) the space of each detector bin is 1 mm. All reconstructed images include 512×512 square pixels. Each pixel has a size of 0.5 mm$\times$0.5 mm. Each projection data along the X-ray passing through the cross-sectional image is calculated as the cross-length of an X-ray with a pixel. Based on the metal artifact simulation algorithm in [19], we carefully adjust the position, size and angle to simulate close clinical cases. In this work, a database is created in 1,000 cases. 360 degree projection data is evenly sampled over a 360 degree scan range.

The learning rate started at 10^{-4} and decreased by a factor of 0.5 after every 10,000 iterations for a total of 400,000 training iterations. We implemented our models using Tensorflow deep learning framework on a CPU (Intel(R) Xeon E5-2630, 2.20GHz) and four GPU (NVIDIA GeForce GTX-1080 Ti, 12GB) system. The network required approximately 40 hours for training.

The comparison results are shown in Fig.6. The input shows significant metal artifacts and is processed using a U-MAR net model. The experimental results validate the effectiveness of the proposed method for metal artifact reduction of the reconstructed slices. This is mainly because the metal artifacts in images present some characteristics, and the network trained with numerous data can learn the relationship between images with artifacts and the corresponding artifact-free images. Then metal artifacts can be suppressed. This shows that the training of the U-MAR net model is effective. To quantify the reconstruction accuracy, we still use RMSE and NMAD as a measure of reconstruction error. The FBP reconstruction produced by the ideal sinogram is considered to be the ground truth. Subsequently, RMSE and NMAD reconstructed by different methods were calculated. The calculation results are shown in Fig.7. Obviously, by applying the proposed method, RMSE and NMAD are greatly reduced.

B. REAL DATA STUDY

The metal artifacts obtained by the simulation are difficult to express consistent with the real artifacts. There are many

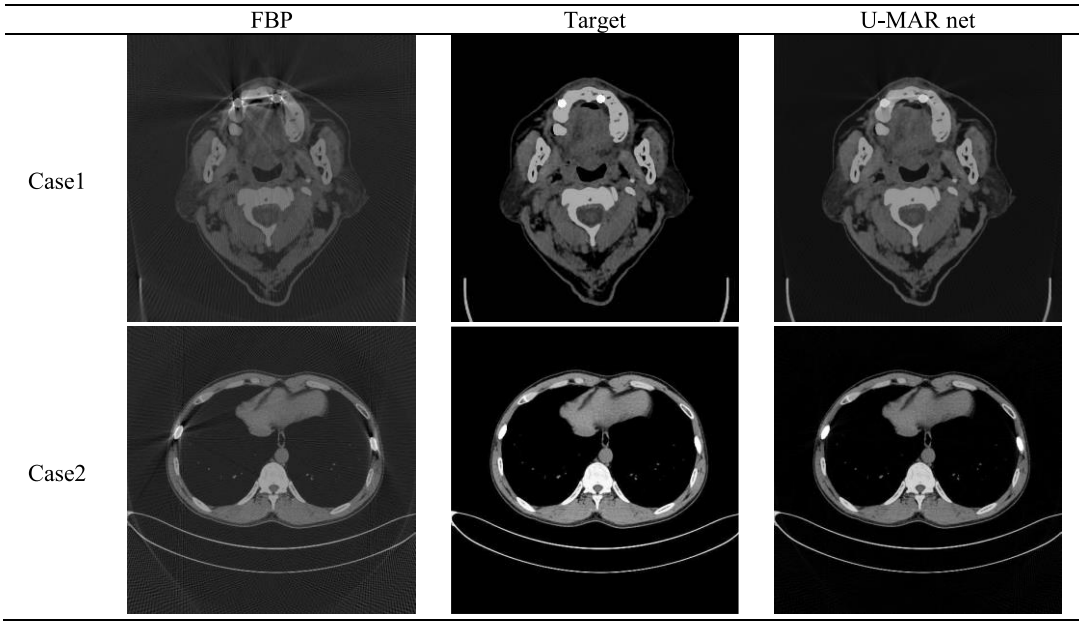


FIGURE 6. Two cases of head model CT images containing metal artifacts. Different case corresponds to the different position of metal implant. Window width WW = 750HU and window level WL = 0HU, respectively.

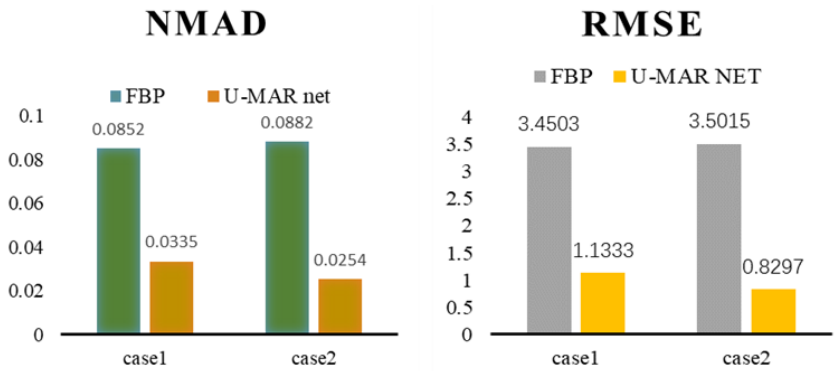


FIGURE 7. NMAD and RMSE for different cases in simulation experiments.

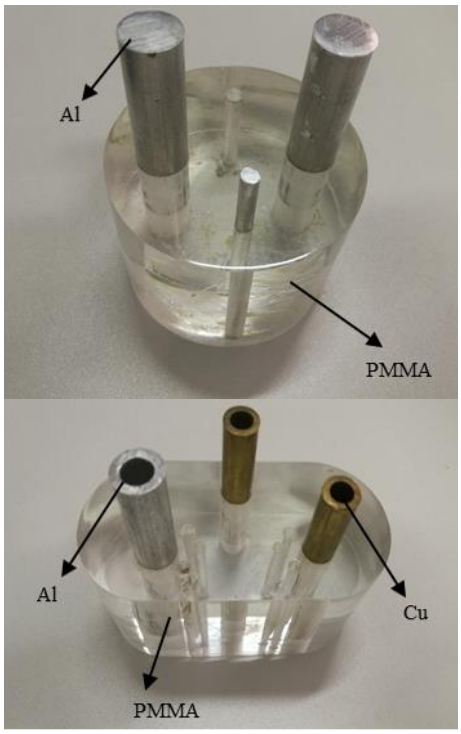


FIGURE 8. The PMMA phantom containing aluminum and copper columns.

reasons for the imaging complexity of metal artifacts. The simulation can only be applied to one aspect. Therefore, to validate the effectiveness of our method for real data, we built a belly model containing metal aluminum column.

In order to further evaluate the performance of the proposed method under actual conditions, we performed two PMMA phantom containing aluminum and copper columns study of actual data. The phantom is shown in Fig.8.

In this study, CT projection data were obtained using a CT scanner including an X-ray source (Hawkeye130, Thales, France) and a flat panel detector (4343F, Thales, France). The tube voltage was set to 100 kVp. The X-ray tube current was set to 300μA. The extracted projection data slices were used for 2D studies. The imaging configuration are as follows: (1) 360 projection views were acquired evenly for a 360° rotation on a circular orbit, (2) the distance from the detector to the light source is 155 mm, (3) the distance from the center of rotation to the source is 44 mm, and (4) the space of each detector bin is 1 mm. All reconstructed images include 512×512 square pixels.

In the cylindrical model, three typical structures containing metal artifacts were constructed by adding different numbers of aluminum columns, and the reconstruction results under three conditions were obtained. In order to obtain a corresponding artifact-free image, we only insert the aluminum column into the upper half of the cylinder mold during the scanning process. The lower half of the cylinder mold does not contain metal, because the material of the cylinder mold is uniform and there is no metal artifact, so it can be used as a corresponding label image.

In order to increase the number of samples, we save a pair of slice images every 0.5 degrees of rotation, and obtain a total of 700 pairs of training images and 20 pairs of test images. After a total of 500,000 training iterations. The learning rate started at 10^{-4} and decreased by a factor of 0.5 after every 10,000 iterations for a total of one million training iterations. The network required approximately 80 hours for training. After completing training, the time for image processing is no more than one second. Fig.9 shows the convergence curve of network training and Fig.10 shows the images reconstructed by different MAR methods. We found that NMAR and the U-MAR net approach successfully reduced the original metal artifacts. Subsequently, the RMSE and NMAD of the reconstruction results using different methods were calculated. The calculation results are shown in Fig.11.

The experimental results using the proposed correction method are shown in Fig.10. The trained network reduces the metal artifacts of the slices. The image obtained by reconstructing an ideal sinogram is considered to be a ground truth. Subsequently, the RMSE and NMAD of the reconstruction results using different methods were calculated. The calculation results are shown in Fig.11. Among the two measurement methods, the quantitative results of the proposed U-MAR

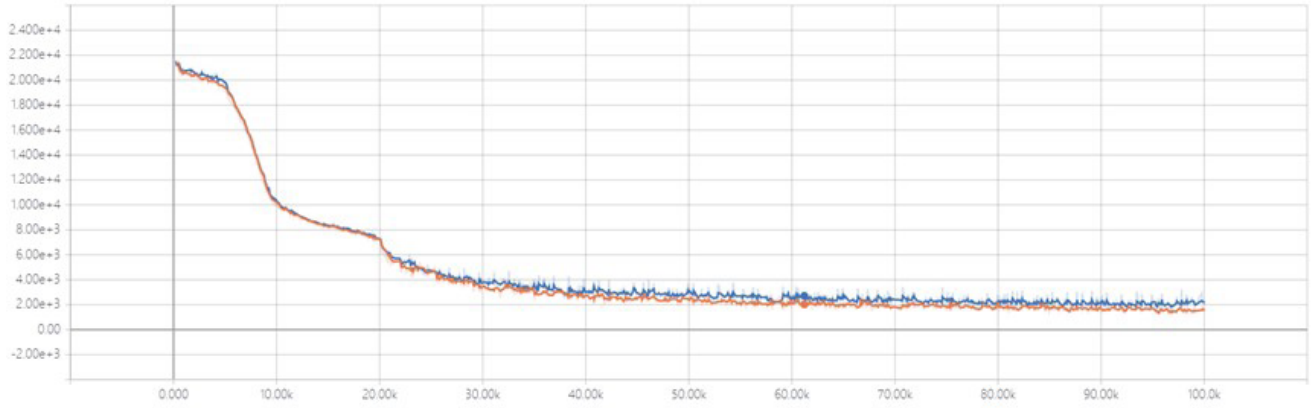


FIGURE 9. The convergence curve of network training. The orange line is the training set loss and the blue line is the validation set loss.

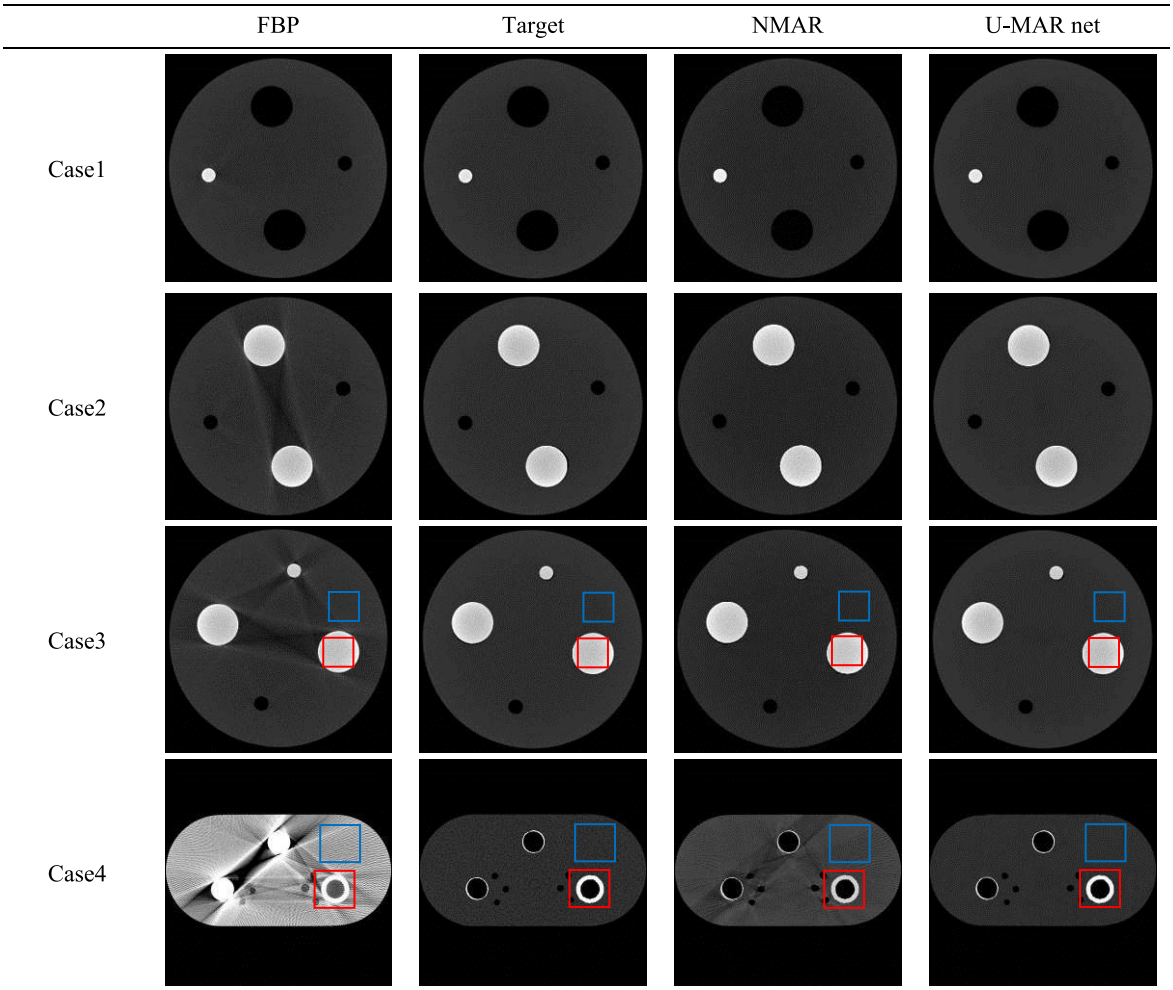


FIGURE 10. The result of actual experiments. Window width WW = 750HU and window level WL = 50HU in case 1 to case 3 and WW = 1500HU and WL = 500HU in case 4. Rectangle ROI within an insert for CNR analysis.

net inpainting exhibited show a superior over the results of other algorithms. The experimental results using the proposed correction method are shown From Fig.10, we can see that the uncorrected image directly reconstructed by FBP displays severe artifacts, and the image corrected with the proposed method shows less artifacts.

In the actual experiments, the results contain noises, so we add the contrast-to-noise ratio (CNR) evaluation index [35].

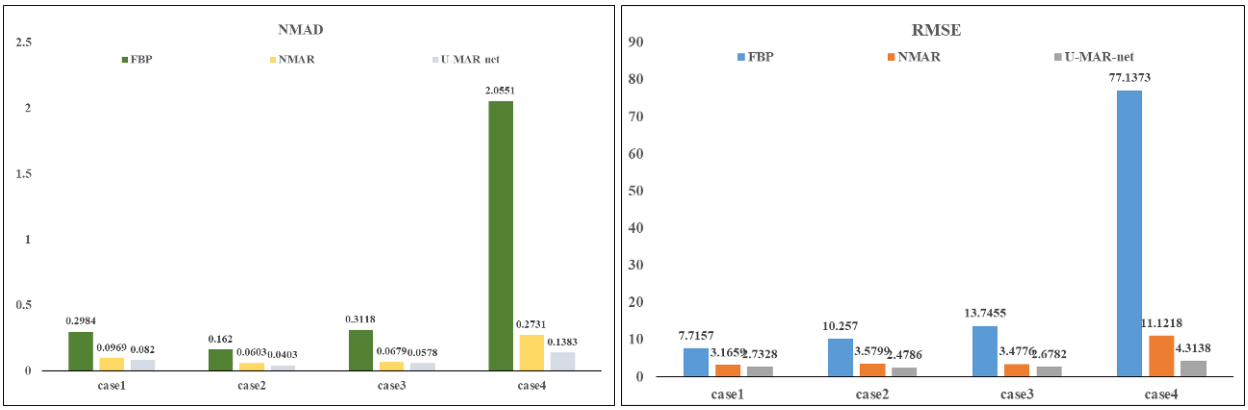


FIGURE 11. NMAD and RMSE for different cases in actual experiments.

TABLE 3. CNR for different cases in actual experiments.

	FBP	Target	NMAR	U-MAR net
PMMA(Al)	5.8385	6.8293	6.3188	6.6312
PMMA(Al+Cu)	0.5581	4.5178	1.1876	4.9879

The CNR is calculated from two regions of interest (ROIs), one is within an insert material for analysis and the other as a reference in an adjacent normal phantom material region with the similar distance to the phantom center. It's calculation formula is as follows:

$$CNR = |\mu_{M1} - \mu_{M2}| \Big/ \sqrt{\delta_{M1} + \delta_{M2}} \tag{4}$$

where: μ_{M1} and μ_{M2} represent the mean CT values of the two selected regions, respectively; δ_{M1} and δ_{M2} represents the standard deviation of the CT values of the two selected regions, respectively. In order to better illustrate the problem, we calculate the CNR for each phantom with the most serious artifact case. The PMMA(Al) stands for the case 3 in Fig.10 and The PMMA(Al+Cu) stands for the case 4 in Fig.10. In this two case, the red rectangular is ROIS of the insert material and the blue rectangular is ROIs of the reference. The calculation results are shown in Table 3.

The CNR results show that the proposed method can improve the CNR well. Compared with NMAR, especially in the case of serious artifacts, the correction effect of this method is better. Due to the convolution characteristics, the proposed method also has some denoising ability.

The results demonstrate the effectiveness of the method proposed in this paper. The experimental results in Fig.10 illustrate that the method can effectively reduce metal artifacts. It can be clearly seen from Fig.11 that by applying NMAR and our proposed method, the RMSE and NMAD of the reconstruction results can be significantly reduced, indicating that both methods have achieved good results in removing metal artifacts. Compared with NMAR, the results of the proposed method further reduce RMSE and NMAD, indicating that the proposed method is further improved in MAR compared to the NMAR method. The processing time of an image using U-MAR net is less than one second. Compared with the NMAR algorithm, the method does not need to re-project to acquire the prior image when processing the data, so the processing time is greatly shortened, and there is no need to adjust the parameters after network training.

IV. DISCUSSION AND CONCLUSION

The method uses the convolution to extract different feature layers and restores the image by up sampling. In the whole process, the metal artifacts are removed by adjusting the corresponding system parameters. Metal artifacts have different characteristics between different layers. Through multiple convolution operations, metal artifact features at different levels are extracted, and image correction is performed by removing this part of the data.

U-MAR net uses deconvolution layer, for example, the feature map of the last convolution layer and restore it to the same size of the input image, thus generating a prediction for each pixel, while preserving the original spatial information of the input image. The results demonstrate that the proposed method can well reduce metal artifacts of CT images. The convolution layer utilized in the U-MAR net can better extract detailed information about the image. U-MAR network avoids complex image preprocessing and accepts input images of any size so that the original image can be directly input.

Evidently, U-MAR net can be a more automatic approach to process the massive amounts of similar data. The trained U-MAR net model can be saved and directly applied for future MAR. For the sake of feature extraction, we suggest that the training data should contain the similar kind of feature, which contributes to extract the accurate features and also contribute to the efficiency of training.

After validated with the synthetic and real data, the present U-MAR net proved to be an effective approach to reduce metal artifacts. Besides, various featured datasets have been utilized to evaluate the proposed U-MAR net which shows satisfied performance. Above all, U-MAR net can be a more intelligent approach for MAR in X-ray CT. And this approach

can be applied in actual CT system to reduce metal artifact, which benefits to automation and massive data processing.

In future work, we will further enhance the applicability of the network. Combine artifact removal with denoise and apply the network to 3D data MAR to further improve U-MAR net performance.

ACKNOWLEDGMENT

This research used the resource of X-ray CT reconstructed data from National Digital Switching System Engineering and Technological Research Center.

REFERENCES

[1] C. H. McCollough, S. Leng, L. Yu, and J. G. Fletcher, "Dual-and multi-energy CT: Principles, technical approaches, and clinical applications," *Radiology*, vol. 276, no. 3, pp. 637–653, 2015.

[2] M. B. Bauza, J. Tenboer, M. Li, A. Lisovich, J. Zhou, D. Pratt, J. Edwards, H. Zhang, C. Turch, and R. Knebel, "Realization of industry 4.0 with high speed CT in high volume production," *CIRP J. Manuf. Sci. Technol.*, vol. 22, pp. 121–125, Aug. 2018.

[3] A. Mouton and T. P. Breckon, "Materials-based 3D segmentation of unknown objects from dual-energy computed tomography imagery in baggage security screening," *Pattern Recognit.*, vol. 48, no. 6, pp. 1961–1978, 2015.

[4] H. Jiang, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*, 2nd ed. Bellingham, WA, USA: SPIE, 2009.

[5] F. E. Boas and D. Fleischmann, "CT artifacts: Causes and reduction techniques," *Imag. Med.*, vol. 4, no. 2, pp. 229–240, 2012.

[6] L. Gjesteby, B. De Man, Y. Jin, H. Paganetti, J. Verburg, D. Giantsoudi, and G. Wang, "Metal artifact reduction in CT: Where are we after four decades?" *IEEE Access*, vol. 4, pp. 5826–5849, 2016.

[7] J. Y. Huang, J. R. Kerns, J. L. Nute, X. Liu, and P. A. Balter, "An evaluation of three commercially available metal artifact reduction methods for CT imaging," *Phys. Med. Biol.*, vol. 60, no. 3, pp. 1047–1067, 2015.

[8] M. Andre, M. Najla, V. S. Katrien, N. Johan, and T. P. Breckon, "An experimental survey of metal artefact reduction in computed tomography," *J. X-Ray Sci. Technol.*, vol. 21, no. 2, pp. 193–226, 2013.

[9] H. S. Park, D. Hwang, and J. K. Seo, "Metal artifact reduction for polychromatic X-ray CT based on a beam-hardening corrector," *IEEE Trans. Med. Imag.*, vol. 35, no. 2, pp. 480–487, Feb. 2016.

[10] J. Hsieh, R. C. Molthen, C. A. Dawson, and R. H. Johnson, "An iterative approach to the beam hardening correction in cone beam CT," *Med. Phys.*, vol. 27, no. 1, pp. 23–29, 2000.

[11] Y. Zhang, X. Mou, and S. Tang, "Beam hardening correction for fan-beam CT imaging with multiple materials," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, Oct./Nov. 2010, pp. 3566–3570.

[12] M. Kachelrieß, O. Watzke, and W. A. Kalender, "Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT," *Med. Phys.*, vol. 28, no. 4, pp. 475–490, 2001.

[13] B. D. Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "An iterative maximum-likelihood polychromatic algorithm for CT," *IEEE Trans. Med. Imag.*, vol. 20, no. 10, pp. 999–1008, Oct. 2001.

[14] B. Hamelin, Y. Goussard, D. Gendron, J.-P. Dussault, G. Cloutier, G. Beaudoin, and G. Soulez, "Iterative CT reconstruction of real data with metal artifact reduction," in *Proc. 5th IEEE Int. Symp. Biomed. Imag. From Nano Macro*, May 2008, pp. 1453–1456.

[15] X. Zhang, J. Wang, and L. Xing, "Metal artifact reduction in X-ray computed tomography (CT) by constrained optimization," *Med. Phys.*, vol. 38, no. 2, pp. 701–711, 2011.

[16] W. A. Kalender, R. Hebel, and J. Ebersberger, "Reduction of CT artifacts caused by metallic implants," *Radiology*, vol. 164, pp. 576–577, Aug. 1987.

[17] R. L. Morin and D. E. Raeside, "A pattern recognition method for the removal of streaking artifact in computed tomography," *Radiology*, vol. 141, no. 1, pp. 229–233, 1981.

[18] A. H. Mahnken, R. Raupach, J. E. Wildberger, B. Jung, N. Heussen, T. G. Flohr, R. W. Günther, and S. Schaller, "A new algorithm for metal artifact reduction in computed tomography: *In vitro* and *in vivo* evaluation after total hip replacement," *Investigat. Radiol.*, vol. 38, no. 12, pp. 769–775, 2003.

[19] S. Zhao, D. D. Robeltson, G. Wang, B. Whiting, and K. T. Bae, "X-ray CT metal artifact reduction using wavelets: An application for imaging total hip prostheses," *IEEE Trans. Med. Imag.*, vol. 19, no. 12, pp. 1238–1247, Dec. 2000.

[20] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrie, "Normalized metal artifact reduction (NMAR) in computed tomography," *Med. Phys.*, vol. 37, no. 10, pp. 5482–5493, 2010.

[21] Y. Zhang and X. Mou, "Metal artifact reduction based on the combined prior image," 2014, *arXiv:1408.5198*. [Online]. Available: https://arxiv.org/abs/1408.5198

[22] T. G. Schmidt, R. F. Barber, and E. Y. Sidky, "Spectral CT metal artifact reduction with an optimization-based reconstruction algorithm," *Proc. SPIE*, vol. 10132, Mar. 2017, Art. no. 101321B.

[23] S. Jeon and C.-O. Lee, "A CT metal artifact reduction algorithm based on sinogram surgery," *J. X-Ray Sci. Technol.*, vol. 26, no. 3, pp. 413–434, 2018.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[27] G. Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.

[28] L. Gjesteby, Q. Yang, Y. Xi, Y. Zhou, J. Zhang, and G. Wang, "Deep learning methods to guide CT image reconstruction and reduce metal artifacts," *Proc. SPIE*, vol. 10132, Mar. 2017, Art. no. 101322W.

[29] L. Gjesteby, Q. Yang, Y. Xi, B. Claus, Y. Jin, B. De Man, G. Wang, "Reducing metal streak artifacts in CT images via deep learning: Pilot results," in *Proc. 14th Int. Meeting Fully Three-Dimensional Image Reconstruct. Radiol. Nucl. Med.*, 2017, vol. 14, no. 6, pp. 611–614.

[30] L. Gjesteby, Q. Yang, Y. Xi, H. Shan, B. Claus, Y. Jin, B. De Man, and G. Wang, "Deep learning methods for CT image-domain metal artifact reduction," *Proc. SPIE*, vol. 10391, Sep. 2017, Art. no. 103910W.

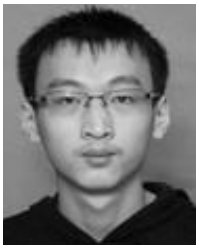
[31] H. S. Park, S. M. Lee, H. P. Kim, and J. K. Seo, "CT sinogram-consistency learning for metal-induced beam hardening correction," 2017, *arXiv:1708.00607*. [Online]. Available: https://arxiv.org/abs/1708.00607

[32] H. S. Park, S. M. Lee, H. P. Kim, and J. K. Seo, "Machine-learning-based nonlinear decomposition of CT images for metal artifact reduction," 2017, *arXiv:1708.00244*. [Online]. Available: https://arxiv.org/abs/1708.00244

[33] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[34] C. Lemmens, D. Faul, and J. Nuyts, "Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 250–260, Feb. 2009.

[35] X. Shi-Peng and L. Li-Min, "Scatter correction for cone-beam computed tomography using self-adaptive scatter kernel superposition," *Chin. Phys. C*, vol. 36, no. 6, pp. 566–572, 2012.



LINLIN ZHU was born in Anhui, China, in 1996. He received the B.E. degree from the Harbin Institute of Technology, in 2017. He is currently pursuing the master's degree with the National Digital Switching System Engineering and Technology Research Center. His current research interest includes the metal artifact reduce of X-ray CT.

YU HAN was born in Hebei, China, in 1987. He received the Ph.D. degree from the National Digital Switching System Engineering and Technology Research Center. He is currently a Lectorate with the National Digital Switching System Engineering and Technology Research Center. His current research interests include X-ray CT system design, reconstruction, and calibration algorithm in cone-beam CT.

LEI LI was born in Hebei, China, in 1981. He received the Ph.D. degree from the National Digital Switching System Engineering and Technology Research Center. He is currently a Professor with the National Digital Switching System Engineering and Technology Research Center. His current research interests include dual-energy reconstruction technique and the design of CT systems.

XIAOQI XI was born in Henan, China, in 1984. She received the Ph.D. degree from the Institute of High Energy Physics, Chinese Academy of Sciences. She is currently a Professor with the National Digital Switching System Engineering and Technology Research Center. Her current research interests include X-ray CT system design and CT tomography.

MINGWAN ZHU was born in Anhui, China, in 1996. She received the B.E. degree from Ningbo University, in 2018. She is currently pursuing the master's degree with the National Digital Switching System Engineering and Technology Research Center. Her current research interest includes geometric calibration of X-ray cone-beam CT.

BIN YAN was born in Henan, China. He received the Ph.D. degree from the Institute of High Energy Physics, Chinese Academy of Sciences. He is currently a Professor with the National Digital Switching System Engineering and Technology Research Center. His current research interests include X-ray CT and intelligent information processing.