

Article

Metal Artifact Correction in Industrial CT Images Based on a Dual-Domain Joint Deep Learning Framework

Shibo Jiang^{1,2}, Yuewen Sun^{1,2}, Shuo Xu^{1,2} , Zehuan Zhang^{1,2} and Zhifang Wu^{1,2,*}

¹ Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing 100084, China; jiangsb20@mails.tsinghua.edu.cn (S.J.); syw15@mails.tsinghua.edu.cn (Y.S.); xu-s19@mails.tsinghua.edu.cn (S.X.); zh-zhang21@mails.tsinghua.edu.cn (Z.Z.)

² Beijing Key Laboratory of Nuclear Detection Technology, Tsinghua University, Beijing 100084, China

* Correspondence: zhifang.wu@mail.tsinghua.edu.cn

Abstract: Industrial computed tomography (CT) images reconstructed directly from projection data using the filtered back projection (FBP) method exhibit strong metal artifacts due to factors such as beam hardening, scatter, statistical noise, and deficiencies in the reconstruction algorithms. Traditional correction approaches, confined to either the projection domain or the image domain, fail to fully utilize the rich information embedded in the data. To leverage information from both domains, we propose a joint deep learning framework that integrates UNet and ResNet architectures for the correction of metal artifacts in CT images. Initially, the UNet network is employed to correct the imperfect projection data (sinograms), the output of which serves as the input for the CT image reconstruction unit. Subsequently, the reconstructed CT images are fed into the ResNet, with both networks undergoing a joint training process to optimize image quality. We take the projection data obtained by analytical simulation as the data set. The resulting optimized industrial CT images show a significant reduction in metal artifacts, with the average Peak Signal-to-Noise Ratio (PSNR) reaching 36.13 and the average Structural Similarity Index (SSIM) achieving 0.953. By conducting simultaneous correction in both the projection and image domains, our method effectively harnesses the complementary information from both, exhibiting a marked improvement in correction results over the deep learning-based single-domain corrections. The generalization capability of our proposed method is further verified in ablation experiments and multi-material phantom CT artifact correction.

Keywords: industrial CT; metal artifact correction; dual-domain; UNet; ResNet; joint framework



Citation: Jiang, S.; Sun, Y.; Xu, S.; Zhang, Z.; Wu, Z. Metal Artifact Correction in Industrial CT Images Based on a Dual-Domain Joint Deep Learning Framework. *Appl. Sci.* **2024**, *14*, 3261. <https://doi.org/10.3390/app14083261>

Academic Editor: Giuseppe Lacidogna

Received: 23 February 2024

Revised: 9 April 2024

Accepted: 11 April 2024

Published: 12 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Industrial computed tomography (CT) technology plays an essential role in industrial inspection and quality control, and is particularly irreplaceable in non-destructive testing and the analysis of internal material defects [1]. However, when CT systems scan objects containing metallic materials, strong metal artifacts commonly emerge in the CT images [2]. These artifacts, resulting from uneven X-ray absorption due to materials with high atomic numbers, scatter, and shortcomings within the reconstruction algorithms themselves [3], can severely disrupt image quality, leading to the misinterpretation of image information. This poses significant challenges to subsequent processing and ultimately limits the effectiveness of CT in practical industrial applications.

Metal artifact reduction (MAR) has been an active research area within the field of CT image processing. Existing methodologies are broadly categorized into two groups: traditional image processing techniques and deep learning-based methods.

Traditional algorithms often concentrate on pre-processing the projection data, including interpolation techniques [4], iterative reconstruction methods [5], model-based artifact correction [6], and frequency domain filtering [7,8]. For instance, linear interpolation and sinogram interpolation methods attempt to compensate data in regions affected by artifacts.

Iterative reconstruction approaches, on the other hand, try to address the data loss caused by metal material by incorporating different mathematical models, such as statistical and geometrical models. However, these traditional methods may be constrained by their inherent model assumptions and may only exhibit effective correction in specific scenarios.

In recent years, the rise of deep learning technologies such as convolutional neural networks (CNNs) has led to rapid advancements in deep learning-based MAR methods [9–12]. Deep learning approaches demonstrate superior capabilities in image quality and processing speed by autonomously learning artifact characteristics and correction mappings. Although they significantly outperform traditional methods in performance metrics, they still rely heavily on a single source of information and do not fully exploit the rich information within the data. In the projection domain, deep learning models necessitate the direct handling of projection data, potentially leading to the neglect of spatial information in the image domain [13]. Deep learning-based corrections focused solely on the projection domain may be ineffective in addressing specific image domain artifacts due to the indirect nature of projection domain information, with changes that might be intricate in the image domain [14]. Such single-domain processing can miss correction errors caused by different factors:

- Beam Hardening: The nonlinear effects caused by varying attenuation characteristics of different materials are often not apparent in the projection domain.
- Scattering: Some scattering effects may be more readily identified and corrected in the image domain.
- Statistical Noise: The randomness of noise means that noise reduction within the projection domain might not cater to variations in local image characteristics.
- Reconstruction Algorithm Shortcomings [15]: For example, the filters used in filtered back projection (FBP) may not entirely eliminate streaks caused by metal.

Conversely, deep learning methods based solely on the image domain might overlook key physical characteristics of the projection data. This limitation arises due to incomplete information, with corrections in the image domain potentially losing part of the raw measurement data, especially when the information has been severely distorted by strong artifacts [16]. Moreover, local optimization in the image domain relies heavily on image features themselves, which could lead to overfitting on local features while ignoring global consistency issues [17]. Contextual information may be lost when focusing on local corrections in the image domain, thus affecting the correction quality [18].

Addressing this problem, we propose an innovative joint deep learning framework that synergistically harnesses the UNet and ResNet architectures through a dual-domain correction process [19,20]. Initially, UNet is utilized to directly handle the raw projection data containing noise and artifacts, obtaining pre-corrected data; then, these corrected data are used for image reconstruction with the FBP algorithm, and ResNet is responsible for further refining the CT images in the image domain to eliminate any remaining artifacts. A dual-domain approach that can utilize complementary information across multiple domains while simultaneously optimizing the entire system has stronger performance potential compared to single-domain correction methods. Notably, the combined UNet and ResNet architecture can extract multi-scale and in-depth features while validating and correcting the information across projection and image domains. This dual-domain method is more adept at processing complex artifacts as it considers the physical properties of the data acquisition process and utilizes rich image content information, providing a new pathway for obtaining high-quality, artifact-free industrial CT images with high efficiency and effectiveness.

2. Method

2.1. Overview of the Dual-Domain Joint Deep Learning Framework

In this study, we proposed a joint deep learning framework based on dual-domain (projection and image domains) for the correction of metal artifacts in industrial CT images. As illustrated in Figure 1, our framework incorporates two main components: initially

employing a UNet architecture in the projection domain to perform a pre-correction on the projection data, followed by generating images using a reconstruction algorithm, and then applying a ResNet architecture in the image domain to further refine the reconstructed images. The joint training process allows the two network modules to learn interactively, enabling a collective optimization of the overall correction performance.

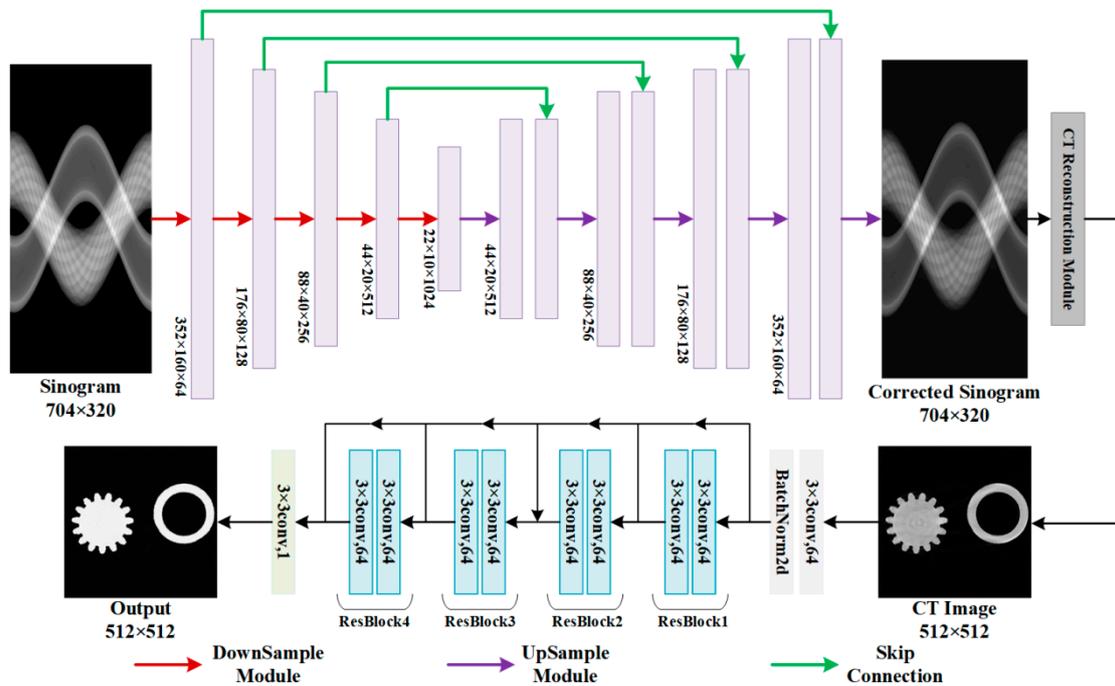


Figure 1. The dual-domain joint deep learning framework diagram we proposed in this study.

2.2. UNet Network for Projection Domain Correction

The UNet architecture is purposefully engineered to address artifact issues within projection data. This encoder–decoder network, which has a symmetric structure, enables feature reutilization via skip connections, thus contributing to the mitigation of information loss at deeper network levels [21]. The UNet architecture adopted in this study comprises five downsample layers and five upsample layers. The convolutional kernel size of the DownSample module is 3×3 , the stride is 2, and the padding is 1. The convolutional kernel size of the UpSample module is 1×1 , the stride is 1, and the padding is 0. With respect to the input projection data, that is, sinograms, UNet processes them through sequential convolution and pooling operations to encode and distill key features, followed by decoding through the corresponding transconvolution layers, reverting them to output projection data with preserved dimensions. This method proficiently reduces artifacts while maintaining the essential information in the projection data.

2.3. ResNet Network for Image Domain Correction

The CT image reconstruction module transforms the pre-corrected projection data into CT images, which are further refined using ResNet. ResNet, also known as a residual network [22], incorporates residual connections that facilitate the construction of deeper networks while preventing the gradient vanishing problem. The particular ResNet employed in this study includes four residual blocks. The size of the convolution kernel is 3×3 , as shown in Figure 1. Within our proposed framework, the input to the ResNet is CT images reconstructed via the FBP algorithm. The network's objective is to learn a residual mapping to amend metal artifacts and other defects, consequently delivering an output of an enhanced and clear CT image.

2.4. Dataset Acquisition

The dataset utilized in this study was obtained through analytical simulation, taking into account results from our previous research [23]. Multispectral X-rays with energy spectrum $\Omega(E)$ can be subdivided into smaller energy bins based on their energy levels, with photons within each bin having similar energies and approximately obeying the Lambert–Beer law. The material’s response to multispectral X-rays is the cumulative integration of responses from all energy bins. Given that the spectrum of the X-rays and the attenuation coefficient $\mu_{E,S}$ are known, the relationship between the incident X-ray intensity I_{input} and the transmitted intensity I_{output} can be represented by the following equation:

$$I_{output} = I_{input} \int \Omega(E) e^{-\int \mu_{E,S} dS} dE$$

In this way, a projection process of CT analytical simulation is transformed into a process of segmenting X-ray beam energy intervals, determining the types of substances, and calculating the attenuation coefficient of the substance in a given energy interval during the path attenuation. The X-ray energy spectrum of tungsten target material at an acceleration voltage of 260 KeV is shown in Figure 2, generated by SpekCalc (v.2009) software [24]. According to the simulation requirements, the X-ray energy spectrum is segmented into 30 energy zones. In the simulation calculation process, it is assumed that photons within an energy zone have the same energy, and the total number of photons in the energy zone can be calculated based on the proportion of that energy zone.

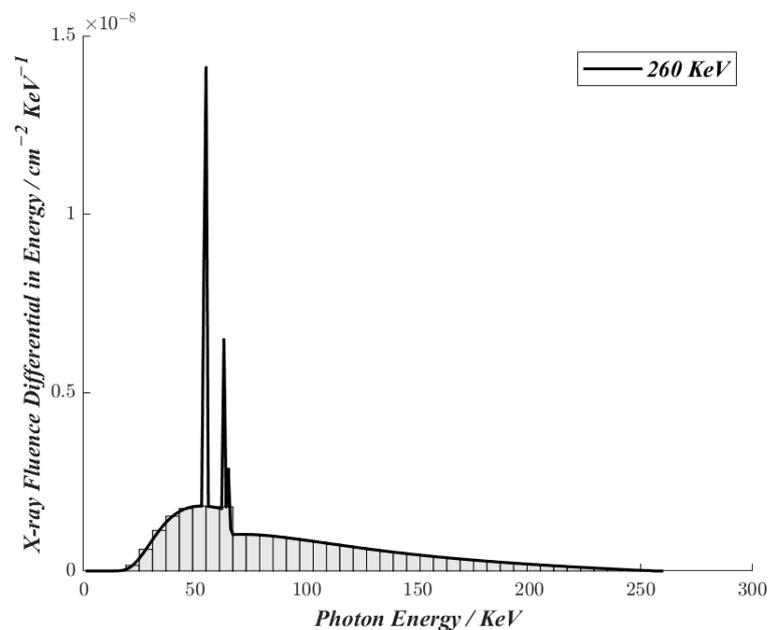


Figure 2. Energy spectrum of X-ray tube at 260 KeV. Assume that photons in each energy bin have the same energy during the simulation. Artifacts caused by beam hardening, scattering, and noise are approximately simulated based on this principle.

The transverse slice of the phantom in Figure 3 is illustrated and the principle of industrial fan-beam CT is shown in Figure 4. In the simulation process, the size and spatial position of pixels determine the path length of X-ray penetration. Different grayscale values represent different materials. Data from the cross-section library are called based on material numbers to calculate the cumulative factor of the interaction between X-rays and substances in the attenuation path.

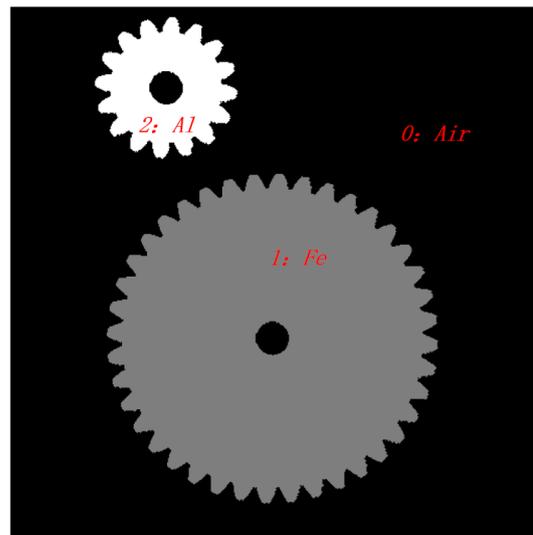


Figure 3. The slices of the scanned object correspond to different materials with varying grayscale values. During the simulation process, the cross-sectional data from the XCOM database are called based on the corresponding materials.

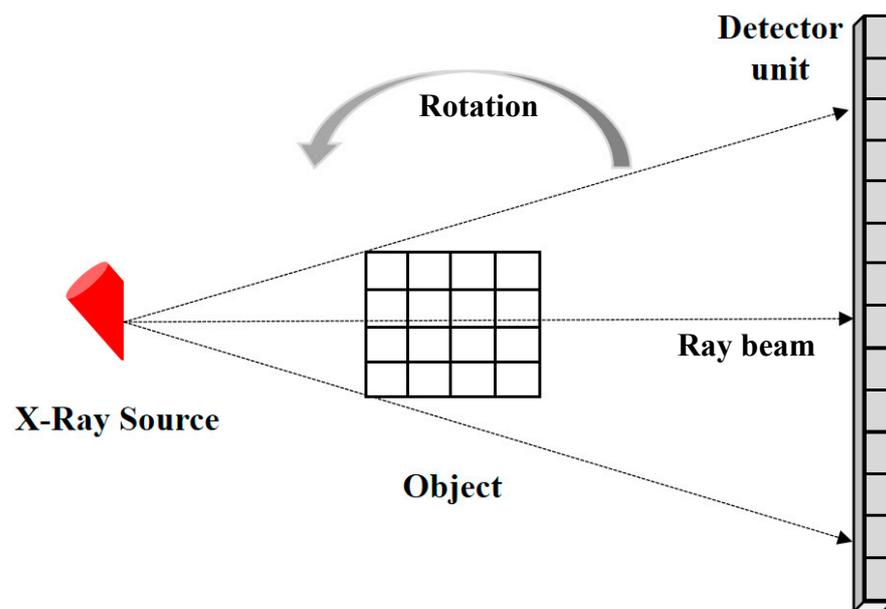


Figure 4. Schematic diagram of the working principle of fan-beam CT.

During the calculation of the cumulative factor, it is essential to have cross-sectional data for different material compositions and X-ray interactions at a given energy level. With the increasing depth of research on the interaction between X-rays and substances, more research units have access to accurate experimental and theoretical cross-sectional data. The cross-sectional database used originates from the X-ray Cross Sections Database (XCOM), developed by the National Institute of Standards and Technology of the United States. It covers cross-sectional data of X-ray interactions with substances up to 300 KeV, commonly used in our industrial micro-focus CT. This database provides a wide range of material types and their cross-sectional data, and also supports the calculation of compound cross-sections in different proportions. Our designed MATLAB(R2021b)-CUDA 11.3 hybrid program can rapidly retrieve cross-sectional data for specified materials and energies by accessing the XCOM database, along with the proportion of the total number of photons in that energy range to the overall photon share.

In the simulation process, the materials of the scanned objects included metals such as Fe, Al, alloy steel, and aluminum alloy. Fan-beam projection was employed to acquire a sequence of projection data, with the simulation parameters detailed in Table 1. It is worth noting that after obtaining the projection data, we added Poisson–Gaussian mixed noise corresponding to the order of magnitude of the projection data size to simulate the effects of scattering and statistical noise on the projection data in the actual CT imaging process, so as to be closer to the real situation. The number of detector units and the number of projection angles were determined by the requirements of actual industrial CT scan data and our network model inputs.

Table 1. Main parameters of simulation CT system.

Parameter	Numeric Value
Source-to-detector distance	1200 (mm)
Source-to-rotation distance	600 (mm)
Object size	512 × 512
Pixel size	0.1 × 0.1 (mm ²)
Number of detector units	704 × 1
Detector size	0.1 × 0.1 (mm ²)
Projection angles	320

Ideal CT image data were derived through direct linear integration. The simulation yielded a total of 2000 projection data sets, with 100 of these sets being allocated for the test dataset, while the remaining samples were the training dataset. The projection data from the training set were reconstructed using direct filtered back-projection, as demonstrated in Figure 5, resulting in CT images with characteristic, intense metal artifacts.

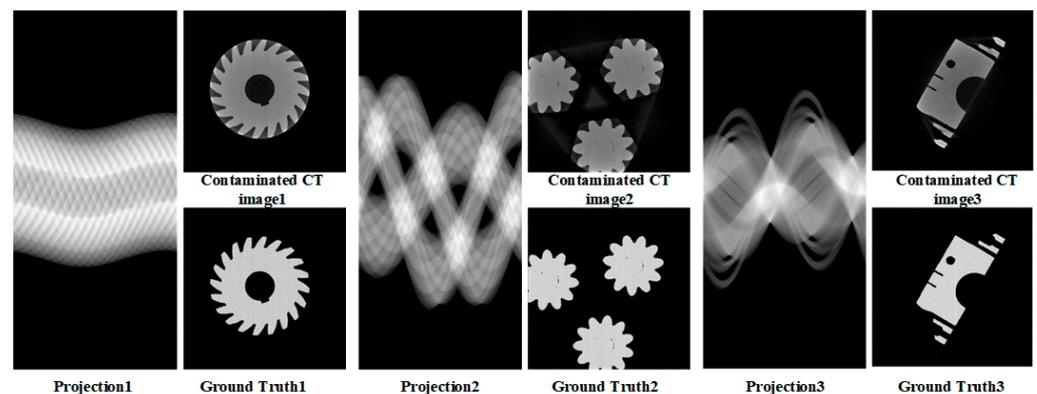


Figure 5. Partial projection data in the training set, reconstructed CT images containing metal artifacts, and corresponding ideal CT images. The materials corresponding to the CT images in the picture are iron, aluminum, and brass-zinc alloy.

2.5. Joint Training Process

In order to achieve collaborative optimization between the projection domain and the image domain, we devised a joint training strategy. This strategy involves optimizing the parameters of the UNet network and the ResNet network through a unified loss function. The loss function consists of multiple components aimed at evaluating several key aspects of artifact correction, including the following:

Mean Square Error (MSE) loss, denoted as $Loss_{MSE}$.

$$Loss_{MSE} = \frac{\sum_{i=1}^n (y_i - y_i^p)^2}{n}$$

where y_i represents the ground truth value and y_i^p represents the predicted value. $Loss_{MSE}$ evaluates and minimizes the overall pixel density error between the corrected image and

the ground truth image, helping adjust the global intensity values and suppress global distortions caused by metal artifacts.

L_1 loss, denoted as $Loss_{L_1}$.

$$Loss_{L_1} = \sum_{i=1}^n |y_i - y_i^p|$$

$Loss_{L_1}$ is more sensitive to edges in the image, which is beneficial for preserving image details and textures, especially when it is necessary to retain object edges and details during the correction process.

Structural Similarity Index (SSIM) loss [25], denoted as $Loss_{SSIM}$.

$$Loss_{SSIM} = 1 - SSIM(x, y)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

$SSIM(x, y)$ measures the similarity between the test image x and the reference image y , where higher values indicate better structural similarity. μ_x, μ_y are the mean luminance values, σ_x, σ_y are the standard deviations with respect to μ_x, μ_y , and C_1 and C_2 are constants representing contrast factors. By comparing the brightness, contrast, and structure of the images, $Loss_{SSIM}$ helps ensure visual realism and attractiveness in the corrected images. Additionally, it preserves the structural information of non-metallic regions in the images, preventing excessive smoothing or the loss of details while removing metal-induced artifacts.

The total loss function, denoted as $Loss_{Total}$, is a weighted combination of the aforementioned losses.

$$Loss_{Total} = \alpha Loss_{MSE} + \beta Loss_{L_1} + \lambda Loss_{SSIM}$$

The joint loss function enables the model to reduce pixel-level errors while maintaining image structure and details, leading to comprehensive performance improvements. In this study, the MSE loss is assigned the highest weight of 0.8, while the L_1 loss and SSIM loss are assigned weights of 0.1 each. The MSE loss focuses on the global average error at the pixel level. In the early stage of metal artifact correction, it is typically more important to reduce the large-scale deviations caused by artifacts rather than handle details or structure. Assigning a higher weight to the MSE loss and lower weights to the L_1 and SSIM losses prevents the model from excessively focusing on image details and structural fidelity during the early training stage, ensuring that the model does not overlook large-scale correction due to details.

The optimizer is Adam, with an initial learning rate of 0.0001. The Adam optimizer is widely used in deep learning training due to its fast convergence and insensitivity to initial values, making it suitable for handling large-scale datasets. During the optimization process, the loss function computes the difference between the predictions from the joint network and the desired outputs, and then adjusts the network parameters of UNet and ResNet through backpropagation. The joint optimization of the two networks allows error propagation from the projection domain to the image domain, enabling mutual learning to optimize the correction effects.

3. Experiment and Results

3.1. Experiment

3.1.1. Training Details

The network training was executed using the PyTorch toolkit, with a total of 500 iterations for the training process. The dataset, acquired via numerical simulation, includes 1900 sets of data. The dimension of the projection data is 704×320 and the size of the reconstructed CT images is 512×512 . The batch size was configured to 1 for the training. During the training,

validation was performed after every 100 iterations to assess the model's performance. The experiments were conducted in an environment equipped with an NVIDIA GeForce RTX 3060 GPU and a 12th Gen Intel® Core™ i7-10700F CPU, running at 2.10 GHz.

3.1.2. Evaluation Metrics

In order to quantitatively assess the performance of the proposed joint deep learning framework for metal artifact reduction in industrial CT images, two widely recognized image quality metrics were employed: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [26]. The PSNR is an estimation of the reconstruction quality relative to the maximum possible pixel values. It is calculated in the log scale of the Mean Squared Error (MSE) between the reconstructed and the original image, and is typically expressed in decibels (dB). The PSNR is defined as:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAXValue^2}{MSE} \right)$$

where *MAXValue* represents the maximum possible pixel value of the image (for instance, 255 for an 8-bit grayscale image), and the MSE is the mean of the squares of the pixel-wise differences between the original and the reconstructed image.

The SSIM index is a more sophisticated metric that measures the similarity between two images, addressing structural information, luminance, and contrast. SSIM has been described in Section 2 and will not be described here.

3.1.3. Ablation Study

In order to further validate the performance of the proposed joint deep learning framework, we conducted ablation experiments by comparing the use of only the UNet network for correcting projection data and the use of only the ResNet for correcting metal artifacts in CT images. Specifically, we utilized the UNet network alone to correct the projection data, and the resulting output was directly input into the CT image reconstruction module to obtain the corrected CT images, omitting the ResNet network for image artifact correction. We also performed filtered back-projection reconstruction on the projection data in the dataset to obtain CT images contaminated with metal artifacts. These images were then input into the ResNet network for metal artifact correction, omitting the UNet network for projection data correction. The training details of UNet and ResNet are consistent with those of the joint network model. The Adam optimizer is used with an initial learning rate of 0.0001, the batch size is 1, and the loss function is the same. The models are trained for 500 epochs. The correction results of both approaches were compared with those achieved using the proposed joint deep learning framework presented in this paper.

3.2. Results

After completing the training process, the performance of our model was tested using the test set, and partial output results are shown in Figure 6. To compare the effects of using different networks, testing was conducted separately using only the UNet network for projection data correction and only the ResNet for image domain artifact correction. The output results for these scenarios are shown in Figures 7 and 8.

To quantitatively evaluate the correction results, the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) were calculated, as shown in Table 2.

In order to further validate the generalization ability of the model, we modified the CT scan parameters for partial phantom models. Specifically, the source-to-rotation distance parameter was set to 500 mm, and the source-to-detector distance was set to 1500 mm. The obtained data were then used to re-assess the performance of the metal artifact correction for each model, as depicted in Figures 9 and 10. The quantitative evaluation metrics, including PSNR and SSIM, are presented in Table 3. We additionally acquired projection data of multi-material phantoms not present in the training set and input them into our network model, as shown in Figure 11 for calibration. Due to the varying attenuation characteristics

of different materials and rays, the complexity of the projection data of multi-material phantoms and CT image artifacts is significantly increased compared to single-material phantoms. The correction results further validate the generalization capability of our proposed joint deep learning framework.

Table 2. Comparison of quantitative evaluation indicators for metal artifact correction in CT images of different models.

Metrics	Joint Network		UNet		ResNet	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Image 1	39.07	0.960	24.50	0.816	36.74	0.956
Image 2	36.97	0.971	28.05	0.829	34.69	0.946
Image 3	35.67	0.966	27.65	0.782	32.44	0.952
Image 4	33.90	0.970	28.82	0.851	29.28	0.944
Image 5	29.94	0.857	22.39	0.657	26.57	0.874
Image 6	32.99	0.962	27.80	0.801	28.95	0.948
Image 7	41.23	0.985	27.12	0.875	37.69	0.977
Image 8	39.25	0.952	25.21	0.758	36.17	0.949
Average	36.13	0.953	26.44	0.796	32.82	0.943

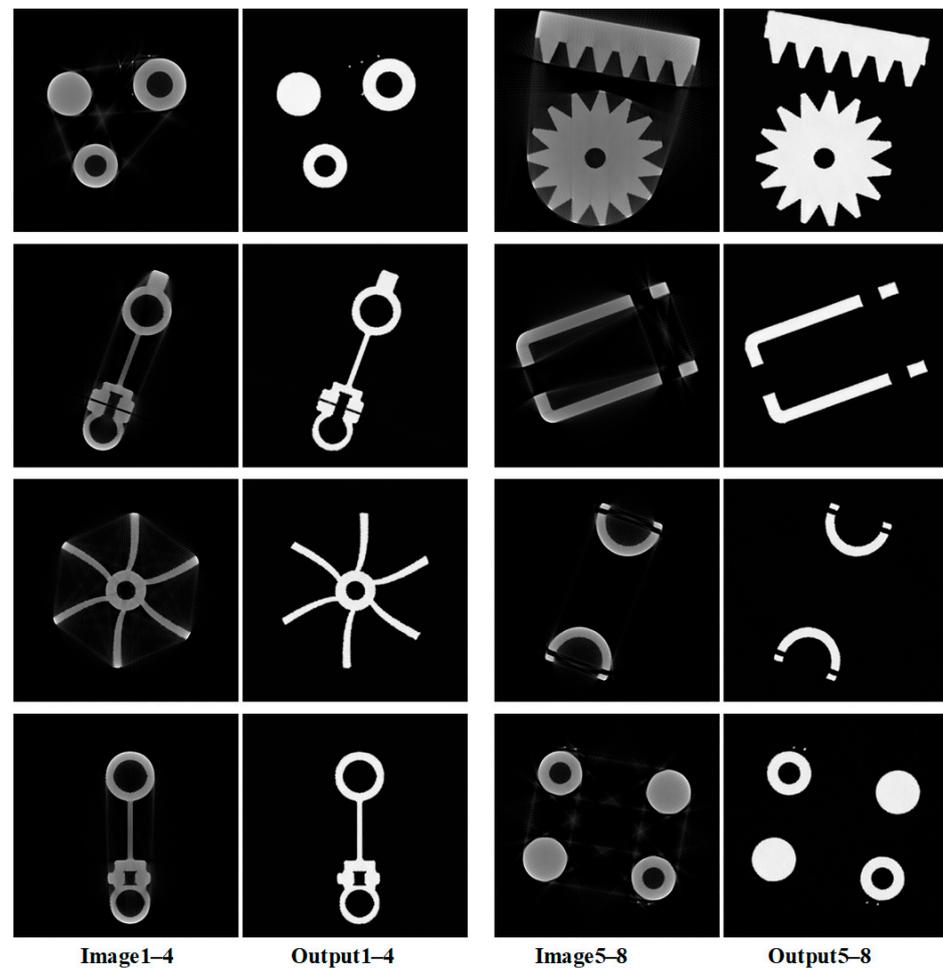


Figure 6. The CT images directly reconstructed from the projection data in the test set, as well as the corresponding output results of our dual-domain joint deep learning framework.

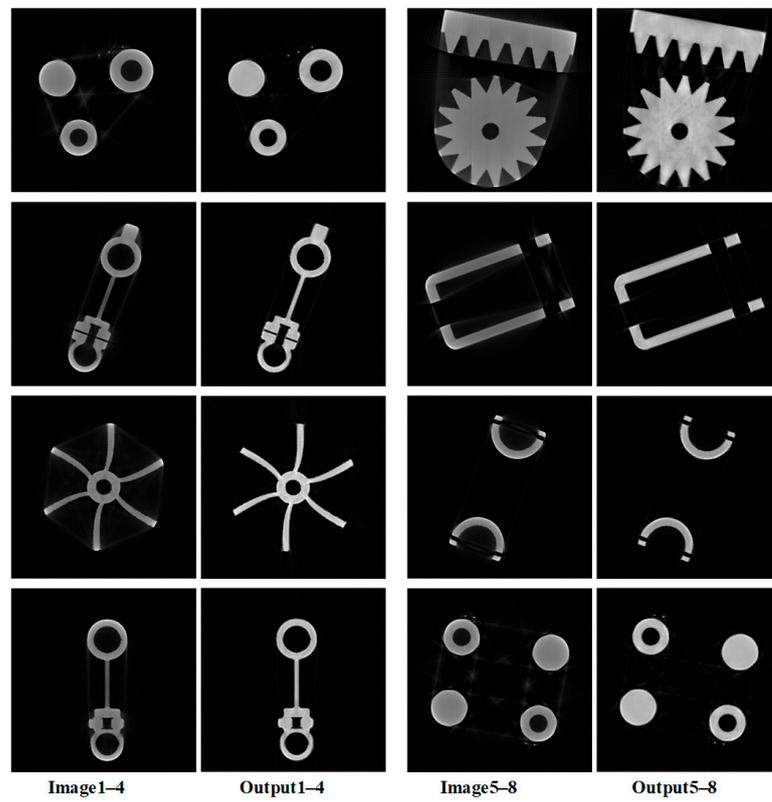


Figure 7. The CT images directly reconstructed from the projection data in the test set, as well as the corresponding output results of UNet.

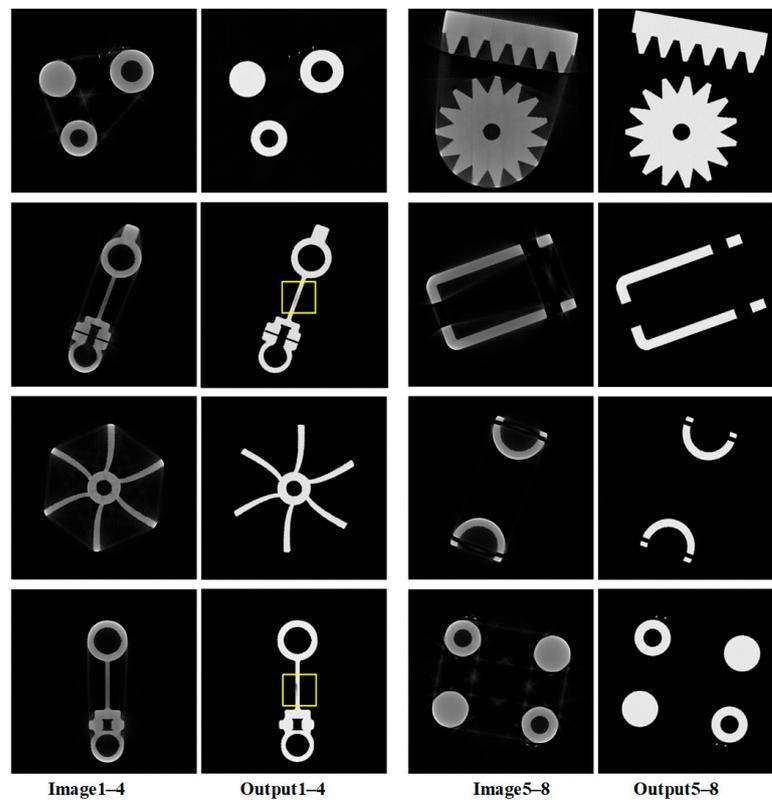


Figure 8. The CT images directly reconstructed from the projection data in the test set, as well as the corresponding output results of ResNet. Some output results show feature loss, as indicated by the yellow box in the figure.

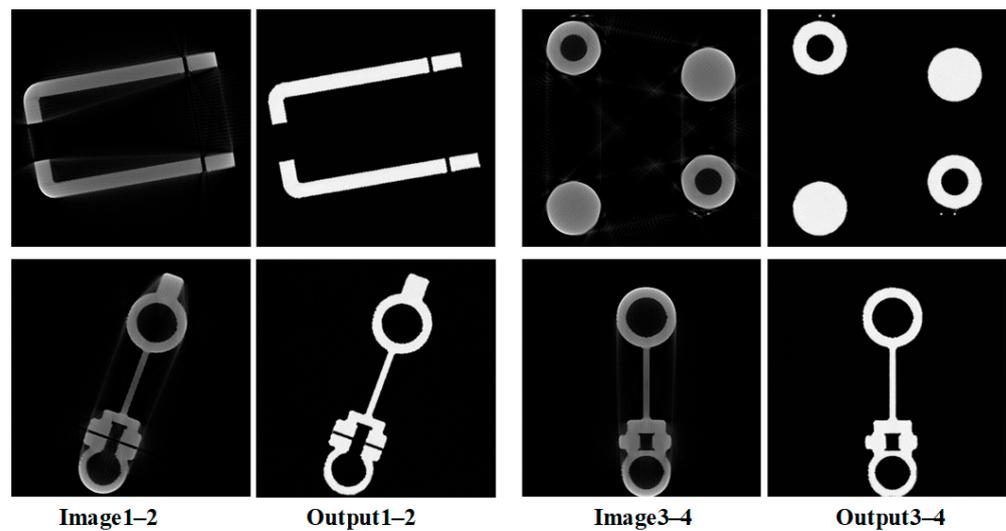


Figure 9. The correction results of our proposed dual-domain joint deep learning architecture after changing the CT scanning parameters. From the graph, it can be intuitively seen that the correction result remains relatively ideal.

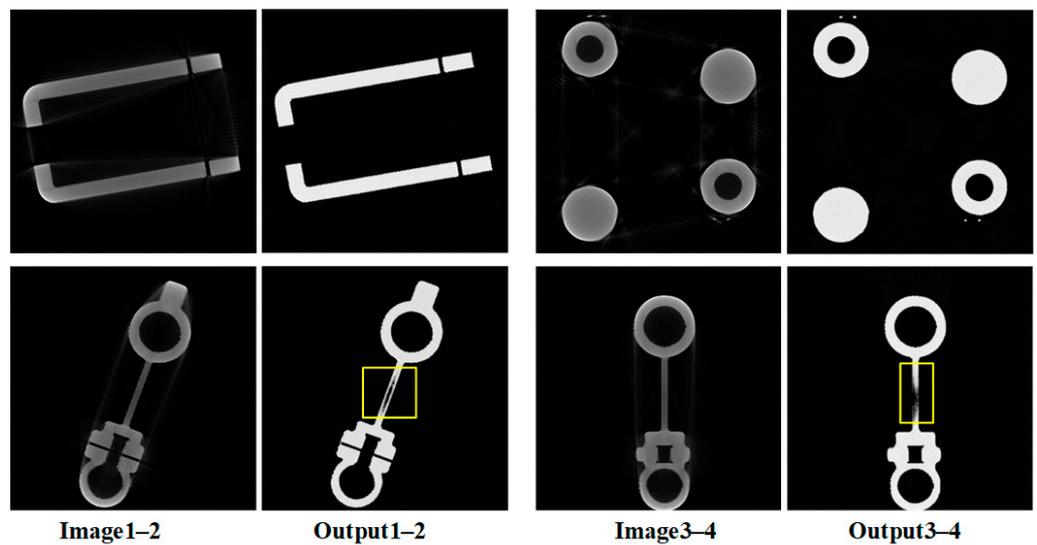


Figure 10. After changing the CT scanning parameters, the calibration results of the ResNet network. As indicated by the yellow box in the figure, when the grayscale difference between metal artifacts and the phantom decreases, the CT image correction results show more severe feature loss.

Table 3. Quantitative evaluation indicators for the calibration results of the two models after changing CT scanning parameters.

Metrics	Joint Network		ResNet	
	PSNR	SSIM	PSNR	SSIM
Image 1	33.75	0.956	31.83	0.944
Image 2	34.59	0.960	26.90	0.885
Image 3	35.25	0.951	32.81	0.933
Image 4	34.32	0.967	25.32	0.852
Average	34.48	0.959	29.22	0.904

To verify the artifact correction effect of our proposed joint deep learning framework on real CT images, we obtained usable fan-beam projection data by controlling the gantry angle of the industrial cone-beam CT, and the corresponding image slices were the bronze

sword and gear hull. The actual projection data were input into the trained network model, and the correction results are shown in Figure 12.

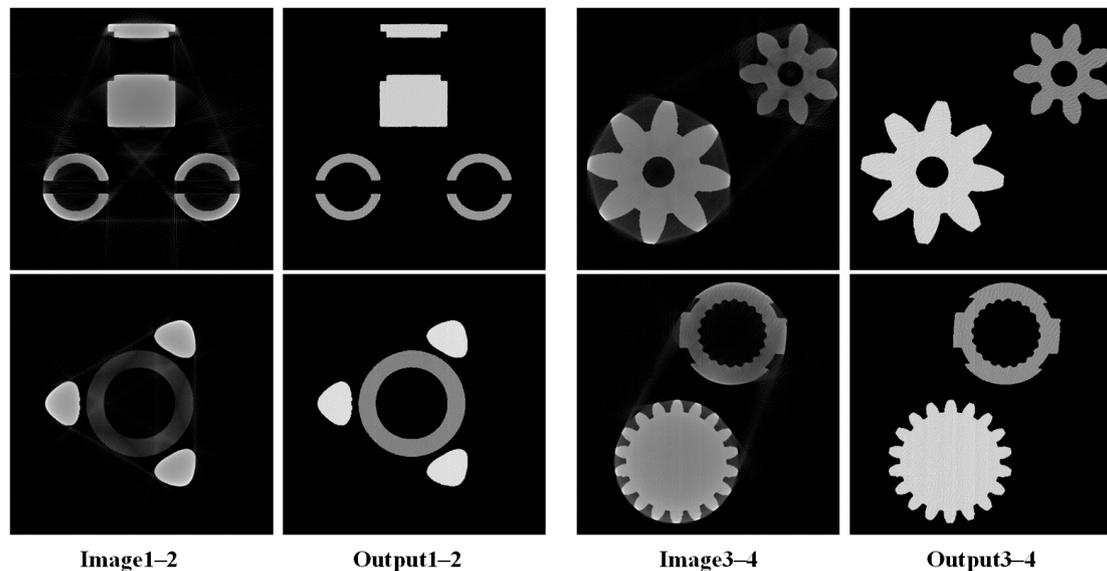


Figure 11. The calibration results of the additional projection data of multi-material phantoms, with the corresponding materials being copper–iron combination, aluminum–titanium alloy combination, iron–titanium alloy combination, and iron–zinc alloy combination. From the images, it can be observed that the joint deep learning framework has achieved good calibration results.

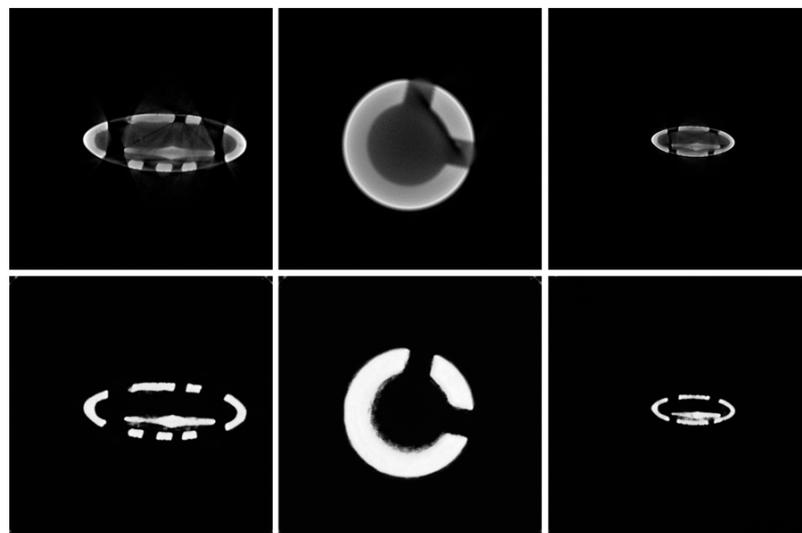


Figure 12. Artifact correction effect of our proposed network model on real industrial CT images. It can be seen that a relatively ideal artifact correction result is achieved, but it is limited by the interference between different layers in obtaining fan-beam projection data during the actual operation, so the projection data are not absolutely ideal fan-beam data. Thus, there is still room for improvement in the correction effect of some edges.

4. Discussion

From Figures 6–8, it can be observed that the dual-domain joint deep learning framework effectively removes metal artifacts from industrial CT images while preserving the characteristic image details. In contrast, employing only the UNet network for processing projection data does not completely eliminate the metal artifacts when reconstructed from the processed projection data. This is primarily attributed to the non-linear effects resulting

from the distinct attenuation characteristics of different materials, which are not prominently manifested in the projection domain. Thus, relying solely on the UNet network for projection data correction does not adequately address the non-linear effects, leading to residual strong metal artifacts in the reconstructed CT images. It is precisely for this reason that in the comparison experiment of correction results after changing CT scanning parameters, we directly abandoned the operation of Unet correction projection data and only compared our proposed network architecture with single-domain ResNet.

Utilizing only the ResNet for metal artifact correction in the image domain typically yields favorable correction results. However, in cases where the metal artifacts and the grayscale of the phantom are closely matched, ResNet fails to effectively preserve image details and features during artifact correction, as evident in the highlighted section in Figure 6. This deficiency arises from the exclusive correction in the image domain, which overlooks pertinent information from the original measurements. Particularly in the presence of severe artifacts, the image domain information undergoes substantial distortion.

The quantitative comparison presented in Table 2 demonstrates that the CT image metal artifact correction outcomes achieved via the dual-domain joint deep learning framework surpass those obtained through single-domain correction methods, with average PSNR and SSIM values reaching 36.13 and 0.953, respectively. This quantitative analysis further substantiates the superior performance of the proposed dual-domain joint deep learning framework in CT image metal artifact correction.

Furthermore, modifying the CT scan parameters results in conspicuous changes in the reconstructed CT image features. When inputting data acquired with altered parameters into our trained dual-domain joint deep learning architecture and the ResNet network model, the output results are depicted in Figures 9 and 10, respectively. It is evident that the dual-domain joint deep learning network maintains a consistently ideal output, achieving commendable metal artifact correction effects. Conversely, the ResNet network, which exhibited satisfactory performance before the CT scan parameter modifications, exhibits significant distortion in its output results. The quantitative evaluation presented in Table 3 further demonstrates the stability of the metal artifact correction performance of the joint deep learning network, with average PSNR and SSIM values remaining consistent, while a substantial decline is observed in the correction results produced by the ResNet network.

The performance comparison of each model before and after changes in CT scan parameters, as well as the calibration results of CT image artifacts of multi-material phantoms not present in the training set shown in Figure 11, further validate the generalization capability of the proposed dual-domain joint deep learning architecture, laying a solid foundation for practical applications. From the correction results of real industrial CT images in Figure 12, our proposed joint network model achieves a relatively ideal correction effect, which preliminarily verifies the practical application potential of our proposed joint deep learning model.

5. Conclusions

In this study, we employed an innovative dual-domain processing strategy to correct metal artifacts. By integrating information obtained from both the projection domain and the image domain, using a joint deep learning framework that combines the UNet and ResNet architectures, we aimed to better address complex artifact issues.

Firstly, projection domain correction allows the direct manipulation of the measured raw data, enabling the early correction of physical artifacts generated during the data acquisition process, such as non-linear artifacts caused by X-ray spectrum hardening and scattering effects. This direct correction at the raw data level is crucial for the final image quality as it helps reduce artifacts that may be amplified during subsequent reconstruction steps. Furthermore, the image domain correction focuses on artifacts induced by data reconstruction, including limitations of the reconstruction algorithm (such as deficiencies in filtered back-projection) or residual artifacts caused by imperfect correction in the projection domain. Image domain correction not only provides a complementary mechanism to

further enhance the quality of the reconstructed images but also refines the final output by leveraging spatial priors (e.g., image continuity and edge information). By combining these two correction methods, our framework capitalizes on the complementary advantages of different domains and adapts to variations in artifact types and intensities, robustly providing correction effects under various conditions. Additionally, dual-domain processing facilitates network self-calibration as systematic biases can be detected and corrected during domain transformations, significantly improving the robustness of the correction strategy.

The results demonstrate that the joint deep learning framework effectively utilizes the rich information present in both domains, greatly improving the quality of the images. The average PSNR of the corrected CT images reached 36.13, with an average SSIM value of 0.953. In comparison to single-domain correction methods, the proposed approach demonstrates a notable improvement in correction effectiveness. Furthermore, it exhibits stable performance even after modifying the CT scanning parameters, highlighting the remarkable effectiveness of the proposed method. In summary, the success of our approach underscores the potential of the joint deep learning framework in addressing the complex challenges in industrial CT imaging, paving the way for future optimizations in leveraging multi-domain information to enhance image reconstruction and reduce artifacts.

Author Contributions: Methodology, S.J., Y.S., Z.Z. and Z.W.; Software, S.J. and Y.S.; Validation, S.J.; Formal analysis, S.X.; Investigation, S.X.; Data curation, Z.Z.; Writing—original draft, S.J.; Writing—review and editing, S.J.; Supervision, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The code and model are available online (https://github.com/jiang-16/gan_jiang, accessed on 16 September 2023).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Withers, P.J.; Bouman, C.A.; Carmignato, S.; Cnudde, V.; Grimaldi, D.; Hagen, C.K.; Maire, E.; Manley, M.; Du Plessis, A.; Stock, S.R. Stock. X-ray computed tomography. *Nat. Rev. Methods Primers* **2021**, *1*, 18. [CrossRef]
2. Hampel, U. X-ray computed tomography. In *Industrial Tomography*; Woodhead Publishing: Sawston, UK, 2022; pp. 207–229.
3. Boas, F.E.; Fleischmann, D. CT artifacts: Causes and reduction techniques. *Imaging Med.* **2012**, *4*, 229–240. [CrossRef]
4. Gu, J.; Zhang, L.; Chen, Z.; Xing, Y.; Huang, Z. A method based on interpolation for metal artifacts reduction in CT images. *J. X-ray Sci. Technol.* **2006**, *14*, 11–19.
5. Acharya, R.; Kumar, U.; Patankar, V.H.; Kar, S.; Dash, A. Reducing Metal Artifact using Iterative Reconstruction in Industrial CT. In Proceedings of the 2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India, 15–16 January 2021; pp. 1–6.
6. Paudel, M.R.; Mackenzie, M.; Fallone, B.G.; Rathee, S. Evaluation of metal artifacts in MVCT systems using a model based correction method. *Med. Phys.* **2012**, *39*, 6297–6308. [CrossRef] [PubMed]
7. Hokamp, N.G.; Eck, B.; Siedek, F.; Dos Santos, D.P.; Holz, J.A.; Maintz, D.; Haneder, S. Quantification of metal artifacts in computed tomography: Methodological considerations. *Quant. Imaging Med. Surg.* **2020**, *10*, 1033. [CrossRef] [PubMed]
8. Anhaus, J.A.; Killermann, P.; Sedlmair, M.; Winter, J.; Mahnken, A.H.; Hofmann, C. Nonlinearly scaled prior image-controlled frequency split for high-frequency metal artifact reduction in computed tomography. *Med. Phys.* **2022**, *49*, 5870–5885. [CrossRef] [PubMed]
9. Arabi, H.; Zaidi, H. Deep learning-based metal artefact reduction in PET/CT imaging. *Eur. Radiol.* **2021**, *31*, 6384–6396. [CrossRef] [PubMed]
10. Zhang, Y.; Yu, H. Convolutional neural network based metal artifact reduction in x-ray computed tomography. *IEEE Trans. Med. Imaging* **2018**, *37*, 1370–1381. [CrossRef] [PubMed]
11. Zhang, Y.; Chu, Y.; Yu, H. Reduction of metal artifacts in x-ray CT images using a convolutional neural network. In Proceedings of the Developments in X-ray Tomography XI. SPIE, San Diego, CA, USA, 6–10 August 2017; Volume 10391, pp. 136–146.
12. 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*, 1–15. [CrossRef] [PubMed]
13. Ghani, M.U.; Karl, W.C. Deep learning based sinogram correction for metal artifact reduction. *Electron. Imaging* **2018**, *2018*, 472-1–472-8. [CrossRef]

14. Lyu, Y.; Fu, J.; Peng, C.; Zhou, S.K. U-DuDoNet: Unpaired dual-domain network for CT metal artifact reduction. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, 27 September–1 October 2021; Proceedings, Part VI 24; Springer International Publishing: Cham, Switzerland, 2021; pp. 296–306.
15. Zhang, X.; Wang, J.; Xing, L. Metal artifact reduction in X-ray computed tomography (CT) by constrained optimization. *Med. Phys.* **2011**, *38*, 701–711. [[CrossRef](#)] [[PubMed](#)]
16. Wang, H.; Li, Y.; Zhang, H.; Meng, D.; Zheng, Y. InDuDoNet+: A deep unfolding dual domain network for metal artifact reduction in CT images. *Med. Image Anal.* **2023**, *85*, 102729. [[CrossRef](#)] [[PubMed](#)]
17. Busi, M.; Kehl, C.; Frisvad, J.R.; Olsen, U.L. Metal artifact reduction in spectral X-ray CT using spectral deep learning. *J. Imaging* **2022**, *8*, 77. [[CrossRef](#)] [[PubMed](#)]
18. Yu, L.; Zhang, Z.; Li, X.; Xing, L. Deep sinogram completion with image prior for metal artifact reduction in CT images. *IEEE Trans. Med. Imaging* **2020**, *40*, 228–238. [[CrossRef](#)] [[PubMed](#)]
19. Hegazy, M.A.A.; Cho, M.H.; Cho, M.H.; Lee, S.Y. U-net based metal segmentation on projection domain for metal artifact reduction in dental CT. *Biomed. Eng. Lett.* **2019**, *9*, 375–385. [[CrossRef](#)] [[PubMed](#)]
20. Wang, H.; Xie, Q.; Zeng, D.; Ma, J.; Meng, D.; Zheng, Y. OSCNet: Orientation-Shared Convolutional Network for CT Metal Artifact Learning. *IEEE Trans. Med. Imaging* **2023**, *43*, 489–502. [[CrossRef](#)] [[PubMed](#)]
21. Alom, M.Z.; Hasan, M.; Yakopcic, C.; Taha, T.M.; Asari, V.K. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv* **2018**, arXiv:1802.06955.
22. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
23. Jiang, S.; Sun, Y.; Xu, S.; Wu, Z. Metal artifact correction of CT images based on Generative Adversarial Networks. *J. Harbin Eng. Univ.* **2022**, *43*, 1766–1771.
24. Zhong, X.Y.; Wang, Y.Z.; Cai, A.L.; Liang, N.N.; Li, L.; Yan, B. Dual-Energy CT Image Super-resolution via Generative Adversarial Network. In Proceedings of the 2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), Guangzhou, China, 14–16 May 2021; pp. 343–347.
25. Yang, H.H.; Yang, C.H.H.; Tsai, Y.C.J. Y-net: Multi-scale feature aggregation network with wavelet structure similarity loss function for single image dehazing. In Proceedings of the ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 2628–2632.
26. Lin, W.A.; Liao, H.; Peng, C.; Sun, X.; Zhang, J.; Luo, J.; Chellappa, R.; Zhou, S.K. Dudonet: Dual domain network for ct metal artifact reduction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–17 June 2019; pp. 10512–10521.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.