End-to-End Simultaneous Combined Scatter Correction and Metal Projection Segmentation for CBCT Metal Artifact Reduction Network

Authors: Dr. Kai Chen, Xie, Shipeng, Dr. Yuan Gao, Gu, Dr. Jiabin, Xi, Dr. Yan, Ms. Hongyu Gao, Wang, Zihao, Wu, Dr. Zhan, Dr. Hui Tang, Chen, Prof. Yang, chen, Dr. kai

Date: 2025-03-27T00:00:00+00:00

Abstract

Cone-beam computed tomography (CBCT) is widely used for high-precision three-dimensional imaging in the field of nuclear science and technology (e.g., nuclear medicine radiation therapy, nondestructive testing of nuclear facilities). However, when patients with metal implants in their bodies are examined by CT, the projection data will be contaminated and the CT reconstruction will produce severe metal artifacts affecting the diagnostic accuracy. Therefore, developing high-performance metal artifact reduction (MAR) methods for CBCT is essential for clinical diagnosis. Some of the existing MAR methods ignore physical phenomena such as x-ray scattering during CT imaging. The remaining MAR methods utilize scattering correction as a preprocessing approach, and this non-end-to-end MAR method increases the complexity of the MAR process. In this article, we design an end-to-end MAR method that incorporates both scattering correction and metal segmentation in the projection domain in the same network, followed by morphological post-processing and metal artifact suppression. Both qualitative and quantitative results and analyses in simulated and real data sets illustrate the superiority of this end-to-end MAR method.

Full Text

Preamble

End-to-End Simultaneous Combined Scatter Correction and Metal Projection Segmentation for CBCT Metal Artifact Reduction Network

Kai Chen and Shi-peng Xie
School of Communications and Information Engineering, Nanjing University of

Posts and Telecommunications, Nanjing 210003, China

Yuan Gao
The Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), and Qilu University of Technology (Shandong Academy of Sciences), Jinan 250013, China

Jiabing Gu
The Affiliated Hospital of Qingdao University, Qingdao 266001, China

Yan Xi
The Shanghai First-Imaging Information Technology Co., Ltd., Shanghai 201315, China

Hong-yu Gao
Department of Pharmacy, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing 210096, China

Zi-hao Wang, Zhan Wu, and Hui Tang
Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing 210096, China

Yang Chen
School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China, and Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing 210096, China

Cone-beam computed tomography (CBCT) is widely used for high-precision three-dimensional imaging in nuclear science and technology applications, including nuclear medicine radiation therapy and nondestructive testing of nuclear facilities. However, when patients with metal implants undergo CT examination, the projection data becomes contaminated, leading to severe metal artifacts in CT reconstruction that compromise diagnostic accuracy. Consequently, developing high-performance metal artifact reduction (MAR) methods for CBCT is essential for clinical diagnosis. Existing MAR approaches either ignore physical phenomena such as X-ray scattering during CT imaging or treat scattering correction as a preprocessing step, which increases the complexity of the MAR pipeline through non-end-to-end processing.

In this article, we design an end-to-end MAR method that integrates both scattering correction and metal segmentation in the projection domain within a unified network, followed by morphological post-processing and metal artifact suppression. Both qualitative and quantitative results and analyses on simulated and real datasets demonstrate the superiority of this end-to-end MAR approach.

Introduction

Cone-beam CT (CBCT) represents one of the most prominent CT applications, offering high spatial resolution and low radiation dose while demonstrating significant clinical value as a medical measurement instrument in dentistry and interventional medicine. However, as illustrated in [Figure 1: see original paper], CT scans of patients with metal implants (such as dental fillings or orthopedic metallic medical consumables) produce severe metal artifacts that negatively impact subsequent clinical diagnosis. Various metal artifact correction methods have been employed to eliminate artifacts introduced by metallic substances, though the most effective approach—avoiding dense metallic components in the scanned object—is often impractical for patients with implanted medical devices such as metal dentures, vascular stents, artificial joints, and endobronchial devices. During the data acquisition phase, alternative strategies involve modifying CT acquisition parameters such as X-ray tube voltage, tube current, scan plane, and slice thickness. The most common approach involves algorithmic correction of metal artifacts from the collected CT data, which can be broadly categorized into three classes: projection domain interpolation-based MAR methods, iterative reconstruction-based MAR methods, and deep learning-based MAR methods.

Projection data affected by metal artifacts are considered damaged or incomplete. Interpolation-based methods in the projection domain leverage proximity to complete data to generate new projection values, reconstruct the repaired data, and obtain artifact-corrected CT images. Kalender et al. proposed a Linear Interpolation (LI) correction algorithm that manually segments metal objects and replaces the segmented regions using linear interpolation. While LI is simple and fast, secondary artifacts may arise due to segmentation inaccuracies. In 2004, Wei et al. addressed the limitations of manual segmentation through a threshold-based artifact correction algorithm that employed polynomial interpolation to suppress metal artifacts near bones by replacing metal pixels with polynomial interpolations of neighboring data. This method improved region delineation accuracy and more effectively resolved artifacts from metals and bones, though significant secondary artifacts persisted in the corrected results.

In 2010, Meyer et al. introduced Normalized Metal Artifact Reduction (NMAR), a correction method that normalizes a priori projection data. NMAR can recover tissue details at metal boundaries while suppressing secondary artifact generation. To further recover fine structures around metal implants, Meyer et al. subsequently fused the original projection with differently weighted NMAR projections, proposing the Adaptive Normalized Metal Artifact Reduction (ANMAR) algorithm to improve upon NMAR. Bannas et al. proposed a MAR algorithm based on image-constrained compressed sensing, which utilizes iterative compressed sensing reconstruction to obtain prior images and constrains the estimation of missing information through these priors to improve image quality. While projection domain interpolation-based MAR methods effectively suppress most metal artifacts and represent the most commonly used and widely developed approach, they suffer from inaccurate metal region segmentation and may introduce new secondary artifacts due to inherent interpolation algorithm limitations.

Iterative reconstruction methods establish a difference function between artifact-free CT images and actually acquired CT images, then correct and evaluate iterative results by minimizing the objective function to obtain reconstructed images closest to the artifact-free ground truth. Wang et al. proposed two iterative reconstruction methods for metal artifact correction based on the maximum expectation formula and Algebraic Reconstruction Technique (ART). These methods exhibited similar computational complexity, with the former producing higher clarity images but with greater noise at the same number of iterations. Beekman et al. combined the Ordered Subsets acceleration algorithm with the EM algorithm, proposing Ordered Subsets Expectation Maximization (OSEM) for MAR, which reconstructed higher-quality images while improving algorithm convergence speed. Chang et al. proposed an a priori-based iterative metal artifact correction method that combines statistical methods with sinogram interpolation for estimating and correcting metal-induced bias, yielding improved image quality compared to conventional NMAR and MBIR methods. Zeng et al. proposed an iterative metal artifact correction method for the projection domain based on minimizing total variation and negative pixel energy. While iterative reconstruction methods offer strong noise immunity and can directly correct metal artifacts, they suffer from complex algorithmic steps, high computational cost, and long iteration times.

Recently, deep learning-based algorithms have attracted significant attention in medical image processing. Deep learning-based MAR methods can operate in the sinogram domain, image domain, or both simultaneously. Sinogram domain methods utilize neural networks to directly restore metal-affected projection data before reconstructing CT images. Image domain methods employ neural networks to map metal-damaged CT images directly to clean CT images. Dual-domain deep learning MAR networks simultaneously eliminate metal artifacts in both projection and image domains. Previous work suffers from three key limitations: (1) Most studies neglect the physical processes (photoelectric effect and scattering) occurring when X-ray photons interact with objects, treating MAR as mere image processing. The presence of metallic implants exacerbates the physical interaction between X-ray photons and matter, resulting in significant fragmentation of projection data at metal trajectory boundaries in the projection domain. (2) Slice-by-slice MAR methods exhibit limitations in three-dimensional spatial continuity, ignoring correlations between slices and allowing metal artifacts to persist in coronal and sagittal views despite achieving better suppression on individual slices. (3) Few MAR methods that incorporate scattering correction integrate it within the MAR framework; instead, they perform scattering removal as a separate preprocessing step before metal artifact suppression, increasing processing complexity through non-end-to-end operation.

This paper makes the following contributions: (1) We thoroughly investigate the physical process of CT imaging with metallic substances and X-ray photons, providing a detailed analysis of metal artifact causes—scattering and beam hardening—and propose a solution addressing the underlying mechanism. (2) We propose an integrated MAR framework with end-to-end scattering correction combined with metal projection segmentation to improve MAR performance while reducing computational complexity and improving efficiency. (3) We design fine-grained morphological processing after metal projection segmentation to address pixel misclassification and optimize metal mask boundaries for both precision and continuity. (4) The two-dimensional Delaunay triangulation interpolation restoration introduced by our method effectively preserves rich tissue structure details around metals while reducing residual metal artifacts in sagittal views. Qualitative and quantitative results on simulated and real datasets demonstrate the excellent MAR performance of the proposed method.

The remainder of this paper is organized as follows. Section II presents the mathematical model and neural network architecture details. Section III presents results and analysis. Section IV discusses relevant issues and presents conclusions.

II. Methodology

A. Principles and Mechanisms of Metal Artifact Generation

[Figure 2: see original paper] illustrates that scattering and beam hardening are critical factors contributing to metal artifacts. When X-rays emitted by the X-ray tube—the core element of medical measurement instrument CT—penetrate an object, the photons received by the detector include not only primary photons but also scattered ones, distorting measured values. This phenomenon, known as Compton scattering, represents the main cause of metal artifacts. Compton scattering can be described by the following equation:

$$\lambda - \lambda_0 = \frac{h}{mc}(1 - \cos\theta)$$

where λ is the scattering wavelength, λ_0 is the wavelength before Compton scattering occurs, h is Planck' s constant, m is electron mass, c is the speed of light, and θ is the photon direction rotation angle. Equation (1) shows that the probability of Compton scattering is directly related to electron density in the material, which is associated with the atomic number of the substance. Due to their high atomic numbers, metallic substances exhibit different scattering effects relative to tissue and bone. [Figure 3: see original paper] and [Figure 4: see original paper] visualize the impact of scattering on MAR. Without scatter correction, the overall image contrast was dark, some tissues and bones were difficult to observe, and certain tissue details were lost (slice 1 and slice 3 in [Figure 3: see original paper]). For slice 3 in [Figure 4: see original paper], blurring around metallic substances within the ROI before scatter correction severely affected image quality. Scatter correction enhanced bone contrast within the ROI and restored apparent artifacts. In summary, scattering phenomena during CT examination of objects containing metal inlays strongly exacerbate metal artifact generation, manifesting as reduced overall CT image contrast and causing artifacts.

As shown in [Figure 2: see original paper], CT—a clinically used medical measurement instrument—typically employs an X-ray beam energy spectrum containing photons of different energies to obtain high-quality CT images. As a multi-energy X-ray beam passes through the scanned object, lower-energy photons are preferentially attenuated while higher-energy photons are less easily attenuated. Consequently, the average beam energy (hardness) increases as rays penetrate more material, evidenced by a higher proportion of high-energy photons relative to low-energy photons—a phenomenon known as beam hardening. When reconstruction is performed using the algorithm, numerous black streak artifacts appear at metal locations in the reconstructed image ([Figure 5: see original paper], right), representing beam-hardening artifacts that constitute the most important reason for metal artifact appearance. Considering the presence of metallic substances in scanned objects, the scattering kernel widths of metallic substances differ in size from those of body tissues, motivating our design of a three-output scattering correction network for metallic substance projections.

B. End-to-End Combined Scattering Correction and Metal Projection Segmentation in Integrated MAR Scheme

We propose a framework for simultaneous scattering correction and metal projection segmentation for MAR, as shown in [Figure 6: see original paper]. The framework inputs are CBCT projection and air projection after logarithmic operation. On one path, the projection is corrected by a scattering correction network, then processed by Empirical Beam Hardening Correction (EBHC), followed by FDK reconstruction and threshold segmentation (TS). These metal regions are then forward-projected from the image domain to the sinogram domain using the forward projection algorithm (FP). On the other path, the corrected projection data and air projection are fed into a metal projection segmentation network to segment metal regions in the projection domain. Morphological fine-tuning of segmented metal projections employs forward projection. The projections are then inpainted by Delaunay triangulation interpolation to eliminate metal artifacts and reconstruct high-quality CBCT images, followed by metal backfilling (MB).

The scattering correction network employs a classical encoder-decoder architecture that simultaneously outputs the scattering amplitude matrix, convolutional Gaussian kernel parameters, and convolutional kernel selection matrix. The convolutional scattering amplitude matrix selects scattering kernel parameters adapted to different metal region properties, accurately estimates scattering signals in different regions, and effectively captures scattering signal diversity due to material differences. The convolutional kernel selection matrix outputs a matrix matching the input image size to facilitate selection of Gaussian kernel parameters that determine each pixel point, automatically learning and determining the optimal Gaussian kernel parameter for each pixel to optimize scattering signal processing. The scattered signal estimation and correction results are based on the learned scattering amplitude matrix $c(x, y)$, eight pairs of Gaussian kernel parameters σ_1 and σ_2, the categorization indicator function $M_i(x, y)$ determined by the Gaussian kernel selection matrix, and the total detected signal $I(x, y)$. The total computational procedure is illustrated in the following equation:

$$\text{Corrected Result} = \mathcal{F}^{-1}\left\{\mathcal{F}\{I(x,y)c(x,y)\}\sum_i M_i(x,y)\mathcal{F}\{g_i(x,y)\}\right\}$$

The computational process of scattering correction is graphically demonstrated in [Figure 7: see original paper]. The metal projection segmentation network adopts the U-Net architecture as shown in [Figure 6: see original paper], consisting primarily of upsampling and downsampling operations. The loss function for the simultaneous scattering correction and metal projection segmentation network is given by:

$$\text{Loss} = -\sum_i [y_i \log(p_i) + (1 - y_i)\log(1 - p_i)]$$

where N is the number of pixels in the image, y_i is the true label of the i-th pixel (taking value 0 or 1), and p_i is the probability predicted by the network that the i-th pixel belongs to the positive class.

C. Refinement of Metal Projection After Segmentation

After obtaining metal projection segmentation results, pixel misclassification may occur during the segmentation process, as shown in [Figure 8: see original paper]. Some regions within the projection data are incorrectly labeled as metals. Therefore, obtaining metal segmentation results from the projection domain network alone is insufficient; these results must be fine-tuned.

Algorithm 1: Metal Artifact Reduction via Dual-Threshold Projection Refinement

Require: Reconstructed volume $V \in \mathbb{R}^{N_x \times N_y \times N_z}$
Ensure: Refined segmentation mask $M_{\text{final}} \in \{0, 1\}^{N_x \times N_y}$

1. Initialize projection index: $p \leftarrow 1$
2. **repeat**
3. **COARSE SEGMENTATION:**
4. $M_1^{(p)} \leftarrow T_{\theta_1}(P_p)$ // High sensitivity threshold θ_1
5. $M_2^{(p)} \leftarrow T_{\theta_2}(P_p)$ // High specificity threshold θ_2

6. **NEURAL PRIOR INTEGRATION:**
7. $\tilde{M}^{(p)} \leftarrow \mathcal{N}(P_p) \vee M_1^{(p)}$ // Neural network operator $\mathcal{N}$
8. **CONSENSUS REFINEMENT:**
9. $M_{\text{refined}}^{(p)} \leftarrow \tilde{M}^{(p)} \wedge M_2^{(p)}$
10. **GEODESIC PROPAGATION:**
11. $M_{\text{final}}^{(p)} \leftarrow \mathcal{G}(M_{\text{refined}}^{(p)}, P_p)$
12. $p \leftarrow p + 1$
13. **until** $p > N_{\text{proj}}$ // Total projections: N_{proj}

In the original uncorrected reconstructed images, two critical thresholds are established, denoted as T_{hres1} and T_{hres2}, where T_{hres1} is set as a higher threshold and T_{hres2} as a lower threshold. These thresholds effectively segment metal masks in the image domain, serving as a foundation for subsequent processing. Forward projection techniques are then applied to the segmented metal masks, converting segmentation results from the image domain to the projection domain, with resulting projections denoted as P_{mt1} and P_{mt2}. The network segmentation result (Net_{mt}) is combined with P_{mt1} using a logical OR operation, producing an intermediate result (Tmp_1) that expands the metal region to ensure all potential metal components are included. The expanded result (Tmp_1) is refined through a logical AND operation with P_{mt2}, filtering out falsely identified metal regions and improving segmentation precision. Finally, a morphological region-growing operation is applied to the resulting Tmp_2, further optimizing metal mask boundaries to ensure both accuracy and continuity. By iteratively applying this entire workflow, an efficient network-threshold-based segmentation method is implemented for CBCT projections.

D. Triangulation-Based Metal Artifact Suppression for CBCT MAR

In CBCT metal artifact suppression, linear interpolation remains widely utilized due to its simplicity and ease of implementation. However, this approach often results in significant detail loss, particularly in regions surrounding metallic objects. By estimating pixel values in metal regions using averages of adjacent non-metal areas, linear interpolation fails to preserve fine textures and intricate details, potentially introducing blurring artifacts that degrade overall image quality. To overcome these limitations, this study proposes a triangulation-based interpolation technique that effectively repairs projection-domain images, suppresses metal artifacts, and reconstructs high-quality CBCT images. This method corrects affected regions in the projection domain and subsequently restores metal regions in the reconstructed image. The mathematical formulation is presented in Equation (4):

$$P(x) = \lambda_A V_A + \lambda_B V_B + \lambda_C V_C$$

where V_A, V_B, and V_C denote intensity values at vertices A, B, and C of a triangle in the 2D projection domain, respectively, and $P(x)$ represents the interpolated intensity value at point P within the triangle. The weights λ_A, λ_B, and λ_C are barycentric coordinates of P, satisfying:

$$\lambda_A + \lambda_B + \lambda_C = 1$$

The barycentric coordinates are derived based on geometric relationships as described in Equation (6):

$$\lambda_A = \frac{S(\triangle PBC)}{S(\triangle ABC)}, \quad \lambda_B = \frac{S(\triangle PCA)}{S(\triangle ABC)}, \quad \lambda_C = \frac{S(\triangle PAB)}{S(\triangle ABC)}$$

where $S(\triangle ABC)$ represents the area of the triangle formed by vertices A, B, and C, and $S(\triangle PBC)$, $S(\triangle PCA)$, and $S(\triangle PAB)$ denote areas of sub-triangles formed by point P and respective vertex pairs. The proposed method employs 2D Delaunay triangulation-based interpolation to repair metal-corrupted regions in the projection domain, ensuring that interpolated values accurately reflect the spatial and intensity distribution of surrounding non-metal regions. After processing all projection images, the Feldkamp-Davis-Kress (FDK) algorithm performs 3D image reconstruction. In the final step, restored metal regions are seamlessly backfilled into the reconstructed image, completing metal artifact suppression. Compared to traditional linear interpolation, the proposed triangulation-based method offers significant advantages in preserving structural details and textures near metal objects. By leveraging barycentric coordinate-based interpolation, this approach ensures more precise and realistic reconstruction of projection-domain data, leading to improved image quality in CBCT applications.

III. Experiment Results and Analysis

A. Experimental Setup

The projection domain correction network and metal segmentation network were implemented using the PyTorch framework. Prior to full network training, the scatter correction network was pre-trained to obtain an initial scatter correction model. Subsequently, the complete network was trained end-to-end by jointly optimizing both projection correction and metal segmentation networks. Training was conducted on a 64-bit Windows 10 system equipped with 32 GB RAM and an NVIDIA GeForce GTX-3090 GPU with 24 GB video memory, using an Intel i5-650 processor. The Adam optimizer was utilized with parameters $\beta_1 = 0.5$ and $\beta_2 = 0.99$. A batch size of 6 was used across 400 training epochs. The initial learning rate was set to 0.0001 and decayed linearly from epoch 160 to 320 with a decay factor $\gamma = 0.5$. The triangulation-based interpolation method for suppressing metal artifacts was implemented in MATLAB R2021a, with computational efficiency enhanced via Mex compilation.

B. Datasets and Training Protocols

To comprehensively evaluate the proposed CBCT MAR method, extensive simulated and clinical datasets were utilized, meticulously designed to replicate realistic imaging scenarios while ensuring experimental consistency. and summarize key acquisition parameters for simulation-based and clinical data, respectively.

The SPEKTR 3.0 toolkit generated the polychromatic energy spectrum, simulating a realistic X-ray source. Detector dimensions were set to 960×960 pixels with a pixel size of 0.616 mm, corresponding to a field of view (FOV) of 480×480 mm. Key geometric parameters included a source-to-object distance (SOD) of 612 mm and source-to-detector distance (SDD) of 1138 mm. A total scanning angle of 360° was covered with 400 evenly spaced projections.

TABLE 1: Simulation Parameters for Data Generation

Parameter	Value
Source-to-Object Distance (SOD)	612 mm
Source-to-Detector Distance (SDD)	1138 mm
Total scan angle	360°
Total projection number	400
Detector size	960×960
Pixel spacing	0.616 mm
Field of View (FOV)	480×480 mm

To ensure sufficient training data, this study employed the publicly available DeepLesion dataset from the NIH Clinical Center. Twenty patients' abdominal and thoracic CT scans were selected, with soft tissue and bone structures carefully segmented to ensure diversity. Twenty metal masks were manually designed and inserted into anatomical regions of distinct patients within the dataset. Tissue densities and corresponding coordinates were filled into files required for MC-GPU simulation, yielding projection data with and without scattered signals. The photon count for simulation was set to 10^5 incident photons per detector pixel. Model voxel size was set to 0.08 cm $\times\, 0.08$ cm $\times\, 0.16$ cm, with 400 projections per model. [Figure 9: see original paper] illustrates a sample projection from the simulated dataset.

The clinical data acquisition system used for all experiments featured a detector pixel matrix of 960×960 and pixel size of 0.45 mm $\times\, 0.45$ mm. CT scan parameters included a tube voltage of 80 kV, tube current of 3 mA, scanning angle of 360°, 400 scans, and total scanning time of 80 seconds. Reconstructed volume dimensions were $512 \times 512 \times 512$ for the arch nail model and $394 \times 394 \times 300$ for the human spine. All projection data were downsampled $2\times$ to 480×480 via bilinear interpolation before network input and upsampled to original size after scatter correction. [Figure 10: see original paper] shows a sample projection from the clinical dataset.

TABLE 2: Clinical Data Acquisition Protocol

Parameter	Value
Source-to-Object Distance (SOD)	612 mm
Source-to-Detector Distance (SDD)	1138 mm
Total scan angle	360°
Total projection number	400
Detector size	960×960
Pixel spacing	0.45 mm
Reconstruction volume	$512 \times 512 \times 512$
Physical coverage	$394 \times 394 \times 400$ mm
Scan time	80 s

C. Results and Analysis of Comparison Experiments

1. Analysis of Metal Artifact Suppression Results We compared our method against linear interpolation, VMS, and PDS to demonstrate superiority. [Figure 11: see original paper] shows MAR results for two clinical cases across different methods, with selected ROIs zoomed for detailed observation. Without any metal artifact correction, artifacts from metal implants dominated large tissue regions, obscuring crucial anatomical details and producing severe streaking artifacts along angular views. Linear interpolation proved ineffective for clinical data, failing to reconstruct sufficient tissue details—most vertebral structural details were lost or heavily distorted while severe secondary artifacts were introduced, with residual artifacts observable in angular views. The VMS method exhibited notable improvement in suppressing metal artifacts, yielding results comparable to PDS and our approach in certain aspects, though performance was inconsistent with residual metal artifacts in Slice 1 that were prominent in certain regions or angular directions, limiting VMS effectiveness. The PDS method demonstrated relatively good artifact suppression in angular views but failed to preserve fine structural details in ROIs. For instance, in Slice 1, bone structures and surrounding soft tissues within the ROI were not as well-preserved compared to our method, and in Slice 2, noticeable bright spots appeared in the vertebral region post-processing, indicating incomplete artifact suppression. In contrast, our method achieved the most effective metal artifact suppression among evaluated techniques while demonstrating superior preservation of tissue details surrounding metal implants. Results in Slice 1 highlight our method' s ability to maintain ROI integrity, particularly around bone structures, while Slice 2 results confirm its capacity to address bright-spot issues in the vertebral region caused by residual artifacts. Overall, our method achieved the best balance between metal artifact suppression and anatomical structure preservation.

[Figure 12: see original paper] presents a comparative analysis of different MAR results applied to a CT slice of a spinal vertebra model. The uncorrected image displayed significant metal artifacts that severely obstructed visualization of surrounding bony structures, with pronounced streak patterns creating distinct disruptions in anatomical details. Linear interpolation achieved only partial artifact correction, leaving considerable residual streaks around metallic regions while introducing secondary artifacts that caused substantial loss of soft tissue details near implants, underscoring its inadequacy for clinical applications requiring high anatomical fidelity. The VMS method demonstrated improvement through noticeable artifact intensity reduction while preserving more surrounding tissue information, though residual artifacts remained visible, particularly in high attenuation gradient regions. However, the processed image still revealed evident loss of critical details in fine anatomical structures, restricting its clinical utility. The PDS method achieved further improvements, effectively reducing residual streak artifacts and preserving greater anatomical detail compared to VMS, especially in highlighted regions where residual artifacts were less prominent. Despite these advancements, subtle inaccuracies persisted in low-contrast regions, indicating challenges in handling complex attenuation variations. In contrast, our method applied to the simulated dataset demonstrated superior performance in both artifact suppression and preservation of critical anatomical details, achieving nearly complete metal artifact removal with absence of visible streaks around implants and preservation of fine tissue details. In summary, conclusions from both real and simulated datasets indicate significant improvement over other techniques, underscoring the robustness and effectiveness of our approach and emphasizing its potential to address longstanding MAR challenges as a promising candidate for clinical integration.

2. Comparison with Slice-by-Slice MAR Processing Our method directly suppresses metal artifacts in the three-dimensional projection domain. To evaluate performance, we conducted comparative analysis against image-domain MAR methods based on slice-by-slice processing. [Figure 13: see original paper] illustrates results from the simulated dataset using both approaches on two CBCT reconstructed slices from the same volume. Severe beam-hardening streak artifacts persisted between two metal implants in the uncorrected image, causing significant loss of structural information around the spinal region. The image-domain MAR method partially mitigated these artifacts but exhibited notably insufficient ability to recover fine anatomical details in certain regions, such as vertebral edge areas in slice 1. Quantitative evaluation further highlighted our method' s superiority in structural similarity and absolute percentage error, achieving lower error indices than the slice-by-slice MAR method.

[Figure 14: see original paper] presents experimental results from clinical data comparing our method and the slice-by-slice MAR method on two reconstructed slices from the same CBCT volume. Due to the lack of ground-truth reference for clinical datasets, evaluation relied on visual assessment and indirect comparison to baseline images. Severe metal artifacts in uncorrected images significantly disrupted visualization of fine anatomical details, particularly in the region indicated by the arrow in slice 2. The slice-by-slice MAR method suppressed artifacts to some extent but introduced simulated structural deformations, such as loss of spinal detail in the same region. In contrast, our method achieved effective artifact suppression while preserving anatomical structures and soft tissue details, as demonstrated in corrected images. In summary, our method consistently outperformed the image-domain MAR method in both simulated and clinical datasets, exhibiting robust capabilities in achieving superior artifact suppression, maintaining structural integrity, and minimizing quantitative evaluation errors. These findings demonstrate our approach' s potential for reliable and accurate diagnostic analysis while addressing metal artifact suppression challenges in clinical CBCT imaging.

D. Results and Analysis of Ablation Experiments

1. Contribution of Morphological Refinement Post-Processing to Metal Artifact Suppression This study highlights the contribution of segmentation-based post-processing refinement in suppressing metal artifacts in projection data. [Figure 15: see original paper] demonstrates comparative analysis involving three image sets derived from clinical datasets. Uncorrected images in the first column exhibited severe metal artifacts that obscured surrounding anatomical structures, particularly around metallic implants, severely compromising clinical diagnostics by impeding evaluation of adjacent tissues. Images in the second column, which relied solely on interpolation-based correction informed by segmentation outcomes without further refinement, revealed partial artifact intensity reduction but with persistent residual distortions and inaccuracies, especially in critical regions such as vertebral edges. These residual artifacts continued to impede accurate tissue characterization and diagnostic reliability. The third column demonstrates results obtained by incorporating our proposed segmentation-based post-processing refinement, which effectively addressed prior method limitations by leveraging refined segmentation to guide projection-domain corrections, resulting in substantial metal artifact suppression. Compared to interpolation-only correction, the refinement process significantly reduced residual streaks and distortions, as illustrated by visually improved anatomical structures and more distinguishable soft tissue regions. Quantitative assessments further reinforced these observations, showcasing superior performance metrics for our method compared to its predecessors and clearly underlining the critical role of segmentation-based refinement in suppressing projection-domain metal artifacts.

2. Contribution of Two-Dimensional Delaunay Triangulation Interpolation to Metal Artifact Suppression We performed comparison of MAR methods based on different interpolation techniques using a clinical dataset of two cases, analyzing one cross-sectional image per defect site for each processing method. [Figure 16: see original paper] illustrates results from various interpolation-based approaches on clinical datasets. The NMAR method failed to adequately suppress metal artifacts, with residual streak artifacts surrounding metallic implants introducing significant diagnostic errors as projection data contained residual streaking and secondary artifacts. As shown in the axial slice, NMAR processing introduced additional secondary streaks, yielding degraded image quality only marginally improved compared to uncorrected images—attributable to NMAR limitations in accounting for spatial continuity in three-dimensional volumetric imaging. Consequently, NMAR exhibited suboptimal artifact suppression, failing to preserve detailed structural information around metallic objects and leaving residual streaks visible in critical anatomical regions. In contrast, our interpolation-based approach demonstrated superior performance in suppressing metal artifacts, conserving rich details around metals after processing while leaving few residual artifacts in sagittal views.

[Figure 17: see original paper] presents MAR processing results from two simulated datasets. NMAR-processed slices retained prominent streak artifacts, failing to recover essential structural details within high-density regions due to absence of iterative corrections for spatial inconsistencies in projection data, consequently introducing significant visual distortions unsuitable for accurate diagnostic interpretation. In contrast, our method achieved more comprehensive metal artifact suppression in simulated datasets. As evident in [Figure 17: see original paper], streak artifacts were almost entirely eliminated with clearer visualization of bony structures and soft tissues surrounding metallic implants. The interpolation refinement contributed to superior anatomical fidelity, effectively resolving residual streaks prominent in NMAR-processed images. Cross-sectional analysis further corroborated these findings, with negligible residual artifacts observed at defect sites, demonstrating the robustness and generalizability of our method in addressing projection-domain inconsistencies while preserving structural integrity. In summary, our MAR method exhibited significant advantages over NMAR by achieving enhanced metal artifact suppression and preserving essential structural information, making it well-suited for both clinical and simulated datasets and highlighting its potential to provide artifact-free imaging for accurate diagnostic purposes.

E. Time Cost Analysis

Computational performance of our method was evaluated and compared with independent scattering correction and metallic segmentation networks, as shown in . Evaluation metrics included computational time, parameter count, and floating-point operations (FLOPs). The scattering correction network required 0.0486 s per projection image, with 8.751 M parameters and 26.421 G FLOPs. The metallic segmentation network demonstrated significantly faster computation at 0.0051 s per projection, with 0.863 M parameters and 30.413 G FLOPs. The hybrid method integrating scattering correction and metallic segmentation achieved an inference time of 0.0492 s, parameter count of 9.513 M, and required 23.371 G FLOPs, demonstrating notable trade-offs between computational cost and efficiency.

TABLE 3: Performance Evaluation of Scattering and Segmentation Combined Processing

Method	Computation Time (s)	Parameters (M)	FLOPs (G)
Scattering Estimation Network	0.0486	8.751	26.421
Metallic Segmentation Network	0.0051	0.863	30.413
Step-by-Step	0.0537	9.614	56.834
The proposed	**0.0492**	**9.513**	**23.371**

Comparing computational performance between individual networks and the hybrid approach revealed several hybrid model advantages. While independent networks demonstrated fast computation and low parameter counts, hybrid model integration reduced redundant computations, achieving greater efficiency. The slight parameter count increase from 8.751 M (scattering correction) to 9.513 M (hybrid model) was justified by reduced FLOPs and enhanced integration capabilities. The hybrid model also demonstrated notable computational overhead reduction. While independent networks required 0.0537 s combined processing time per projection image, the hybrid approach completed the task in 0.0492 s, representing significant performance improvement attributed to unified architecture minimizing data transfer and eliminating redundant computations between separate tasks. Furthermore, FLOP reduction from 30.413 G (metallic segmentation) and 26.421 G (scattering correction) to 23.371 G in the hybrid model highlighted computational advantages. By merging networks, the hybrid approach reduced excessive memory and computational resource requirements while maintaining high processing accuracy. These results underscore the hybrid model' s suitability for clinical applications requiring real-time metal artifact reduction in CT imaging. The hybrid design effectively preserved scattering correction and metallic segmentation functionalities, enabling improved artifact suppression with reduced computational burden. Its optimized parameter count and computation time position it as a reliable solution for large-scale data processing and real-time imaging workflows. In summary, the hybrid network architecture provides a computationally efficient and accurate approach for metal artifact reduction, representing significant advancement through its ability to integrate tasks while minimizing costs and processing time.

IV. Discussion and Conclusion

In this article, we developed an end-to-end MAR method incorporating both scattering correction and metal segmentation in the projection domain within a unified network. After obtaining metal projection segmentation results, refinement addresses accurately and continuously labeled metal masks for subsequent interpolation algorithms, and finally, reconstructed metal artifacts are eliminated using Delaunay triangular interpolation. Using actual scanned pedicle nail models and clinical spinal nail data from Jiangsu First Image' s equipment, our method considers inter-slice correlations and maximizes elimination of metal artifacts observable in coronal and sagittal views, outperforming slice-by-slice metal artifact removal approaches. Ablation experiment results demonstrated that improved scattering correction accuracy enhances metal artifact suppression while improving tissue detail fidelity and overall image contrast, and that Delaunay triangular interpolation after metal projection segmentation improves metal artifact suppression both qualitatively and quantitatively.

Although the proposed method demonstrates encouraging MAR improvements, several issues require attention. **Limitations:** (1) The scatter correction network is not computationally efficient enough due to using PyTorch' s native Fourier package for transform computations, which may reduce computational efficiency. (2) Dataset diversity is insufficient. Due to real imaging system complexity and metal shape variety, a rich collection of metal CBCT projections (varying metal sizes, materials, etc.) is lacking to further optimize the network for different scenarios. **Future research:** (1) **Multi-modal Imaging Integration:** Combining CBCT data with other imaging modalities such as MRI or PET to further improve metal artifact suppression and diagnostic accuracy. (2) **Real-Time Processing:** Optimizing the framework for real-time applications through model compression and hardware acceleration. (3) **Generalizability:** Expanding the dataset to include diverse clinical scenarios, enabling robust performance across various implant types and anatomical regions. (4) **Patient-Specific Models:** Developing adaptive models that account for individual patient anatomy and implant characteristics, enhancing clinical applicability.

V. REFERENCES

[1] Jia Hao, Li Zhang, Liang Li, Zhiqiang Chen, and Kejun Kang. A practical image reconstruction and processing method for symmetrically off-center detector cbct system. *Nuclear Science and Techniques*, 24(4), 2013.

[2] Kai Chen, Li-Bo Zhang, Jia-Shun Liu, Yuan Gao, Zhan Wu, Hai-Chen Zhu, Chang-Ping Du, Xiao-Li Mai, Chun-Feng Yang, and Yang Chen. Robust restoration of low-dose cerebral perfusion ct images using ncs-unet. *Nuclear Science and Techniques*, 33(3):30, 2022.

[3] Cui Zhang, Xiao Dong Pan, and Gong Ping Shang. Improvements to conventional x-ray tube-based cone-beam computed tomography system. *Nuclear Science and Techniques*, 29(3), 2018.

[4] Hong Kai Yang, Kai Chao Liang, Ke Jun Kang, and Yu Xiang Xing. Slice-wise reconstruction for low-dose cone-beam ct using a deep residual convolutional neural network. *Nuclear Science and Techniques*, 30(04):28–36, 2019.

[5] Xing Yue Ruan, Xiu Fang Li, Meng Ya Guo, Mei Chen, Ming Lv, Rui Li, and Zhi Ling Chen. Cone-beam computed tomography noise reduction method based on u-net with convolutional block attention module in proton therapy. *Nuclear Science and Techniques*, 35(7):89–100, 2024.

[6] Wenjun Zhang, Haining Ding, Hongchun Xu, MingMing Jin, and Gang Huang. An unsupervised deep learning network model for artifact correction

of cone-beam computed tomography images. *Biomedical Signal Processing and Control*, 95:106362, 2024.

[7] Chen Wang, Jingyu Yang, Baoyu Wu, Ruijun Liu, and Peng Yu. Trans-vnet: Transformer-based tooth semantic segmentation in cbct images. *Biomedical Signal Processing and Control*, 97:106666, 2024.

[8] Min Yuan, Yufei Xie, Rongchang Zhao, Ningning Lv, Zijian Zhang, Liye Zhu, and Xiaolin Wu. Generating synthesized computed tomography from cbct/ldct using a novel generative-transformer adversarial-cnn. *Biomedical Signal Processing and Control*, 96:106660, 2024.

[9] Fuqiang Yang, Dinghua Zhang, Hua Zhang, and Kuidong Huang. Cupping artifacts correction for polychromatic x-ray cone-beam computed tomography based on projection compensation and hardening behavior. *Biomedical Signal Processing and Control*, 57:101823, 2020.

[10] Andre Mouton, Najla Megherbi, Katrien Van Slambrouck, Johan Nuyts, and Toby P Breckon. An experimental survey of metal artefact reduction in computed tomography. *Journal of X-ray Science and Technology*, 21(2):193–226, 2013.

[11] Willi A Kalender, Robert Hebel, and Johannes Ebersberger. Reduction of ct artifacts caused by metallic implants. *Radiology*, 164(2):576–577, 1987.

[12] Jikun Wei, Laigao Chen, George A Sandison, Yun Liang, and Lisa X Xu. X-ray ct high-density artefact suppression in the presence of bones. *Physics in Medicine & Biology*, 49(24):5407, 2004.

[13] Esther Meyer, Rainer Raupach, Michael Lell, Bernhard Schmidt, and Marc Kachelrieß. Normalized metal artifact reduction (nmar) in computed tomography. *Medical Physics*, 37(10):5482–5493, 2010.

[14] Esther Meyer, Rainer Raupach, Bernhard Schmidt, Andreas H Mahnken, and Marc Kachelrieß. Adaptive normalized metal artifact reduction (anmar) in computed tomography. In *2011 IEEE Nuclear Science Symposium Conference Record*, pages 2560–2565. IEEE, 2011.

[15] Peter Bannas, Yinsheng Li, Utaroh Motosugi, Ke Li, Meghan Lubner, Guang-Hong Chen, and Perry J Pickhardt. Prior image constrained compressed sensing metal artifact reduction (piccs-mar): 2d and 3d image quality improvement with hip prostheses at ct colonography. *European Radiology*, 26:2039–2046, 2016.

[16] Ge Wang, Donald L Snyder, Joseph A O' Sullivan, and Michael W Vannier. Iterative deblurring for ct metal artifact reduction. *IEEE Transactions on Medical Imaging*, 15(5):657–664, 1996.

[17] Pieter EB Vaissier, Marlies C Goorden, Aaron B Taylor, and Freek J Beekman. Fast count-regulated osem reconstruction with adaptive resolution recovery. *IEEE Transactions on Medical Imaging*, 32(12):2250–2261, 2013.

[18] Zhiqian Chang, Dong Hye Ye, Somesh Srivastava, Jean-Baptiste Thibault, Ken Sauer, and Charles Bouman. Prior-guided metal artifact reduction for iterative x-ray computed tomography. *IEEE Transactions on Medical Imaging*, 38(6):1532–1542, 2018.

[19] Gengsheng L Zeng. A projection-domain iterative algorithm for metal artifact reduction by minimizing the total-variation norm and the negative-pixel energy. *Visual Computing for Industry, Biomedicine, and Art*, 5(1):1, 2022.

[20] Kai Chen, Zhengyuan Zhou, Yuchen Li, Xu Ji, Jiasong Wu, Jean-Louis Coatrieux, Yang Chen, and Gouenou Coatrieux. Red-net: Residual and enhanced discriminative network for image steganalysis in the internet of medical things and telemedicine. *IEEE Journal of Biomedical and Health Informatics*, 28(3):1611–1622, 2023.

[21] Kai Chen, Yang Chen, Gouenou Coatrieux, Jiasong Wu, and Jean Louis Coatrieux. Restoration-enhanced reversible information steganography network for ct images in the internet of medical things. *IEEE Transactions on Instrumentation and Measurement*, 2023.

[22] Kai Chen, Guohui Ji, Chenrui Wang, Zhiguang Peng, Xu Ji, Hui Tang, Chunfeng Yang, and Yang Chen. Ct-net: Cascaded t-shape network using spectral redundancy for dual-energy ct limited-angle reconstruction. *Biomedical Signal Processing and Control*, 2023.

[23] Kai Chen, Tianling Lyu, Jean Louis Coatrieux, and Yang Chen. Dual-energy ct metal artifact reduction by combined material decomposition and projection domain threshold segmentation. In *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.

[24] Muhammad Usman Ghani and W Clem Karl. Fast enhanced ct metal artifact reduction using data domain deep learning. *IEEE Transactions on Computational Imaging*, 6:181–193, 2019.

[25] Haofu Liao, Wei-An Lin, Zhimin Huo, Levon Vogelsang, William J Sehnert, S Kevin Zhou, and Jiebo Luo. Generative mask pyramid network for ct/cbct metal artifact reduction with joint projection-sinogram correction. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22*, pages 77–85. Springer, 2019.

[26] Yulin Zhu, Hanqing Zhao, Tangsheng Wang, Lei Deng, Yupeng Yang, Yuming Jiang, Na Li, Yinping Chan, Jingjing Dai, Chulong Zhang, et al. Sinogram domain metal artifact correction of ct via deep learning. *Computers in Biology and Medicine*, 155:106710, 2023.

[27] Haofu Liao, Wei-An Lin, S Kevin Zhou, and Jiebo Luo. Adn: artifact disentanglement network for unsupervised metal artifact reduction. *IEEE Transactions on Medical Imaging*, 39(3):634–643, 2019.

[28] Ge Wang, Jong Chu Ye, Klaus Mueller, and Jeffrey A Fessler. Image reconstruction is a new frontier of machine learning. *IEEE Transactions on Medical Imaging*, 37(6):1289–1296, 2018.

[29] Xia Huang, Jian Wang, Fan Tang, Tao Zhong, and Yu Zhang. Metal artifact reduction on cervical ct images by deep residual learning. *Biomedical Engineering Online*, 17:1–15, 2018.

[30] Yanbo Zhang and Hengyong Yu. Convolutional neural network based metal artifact reduction in x-ray computed tomography. *IEEE Transactions on Medical Imaging*, 37(6):1370–1381, 2018.

[31] Hong Wang, Yuexiang Li, Haimiao Zhang, Deyu Meng, and Yefeng Zheng. Indudonet+: A deep unfolding dual domain network for metal artifact reduction in ct images. *Medical Image Analysis*, 85:102729, 2023.

[32] Bo Zhou, Xiongchao Chen, S Kevin Zhou, James S Duncan, and Chi Liu. Dudodr-net: Dual-domain data consistent recurrent network for simultaneous sparse view and metal artifact reduction in computed tomography. *Medical Image Analysis*, 75:102289, 2022.

[33] Yuanyuan Lyu, Wei-An Lin, Haofu Liao, Jingjing Lu, and S Kevin Zhou. Encoding metal mask projection for metal artifact reduction in computed tomography. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23*, pages 147–157. Springer, 2020.

[34] Lequan Yu, Zhicheng Zhang, Xiaomeng Li, and Lei Xing. Deep sinogram completion with image prior for metal artifact reduction in ct images. *IEEE Transactions on Medical Imaging*, 40(1):228–238, 2020.

[35] Wei-An Lin, Haofu Liao, Cheng Peng, Xiaohang Sun, Jingdan Zhang, Jiebo Luo, Rama Chellappa, and Shaohua Kevin Zhou. Dudonet: Dual domain network for ct metal artifact reduction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10512–10521, 2019.

[36] Yiannis Kyriakou, Esther Meyer, Daniel Prell, and Marc Kachelrieß. Empirical beam hardening correction (ebhc) for ct. *Medical Physics*, 37(10):5179–5187, 2010.

[37] Tristan M Gottschalk, Andreas Maier, Florian Kordon, and Björn W Kreher. View-consistent metal segmentation in the projection domain for metal artifact reduction in cbct–an investigation of potential improvement. *arXiv preprint arXiv:2112.02101*, 2021.

Note: Figure translations are in progress. See original paper for figures.