

Article

Not peer-reviewed version

---

# Hybrid-Frequency-Aware Mixture-of-Experts Method for CT Metal Artifact Reduction

---

[Pengju Liu](#)\*, [Hongzhi Zhang](#)\*, [Chuanhao Zhang](#), Feng Jiang

Posted Date: 16 December 2025

doi: 10.20944/preprints202512.1438.v1

Keywords: CT metal artifact reduction; Hybrid-Frequency-Aware; wavelet transform; Fourier transform; Mixture-of-Experts




Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

# Hybrid-Frequency-Aware Mixture-of-Experts Method for CT Metal Artifact Reduction

Pengju Liu <sup>1,2,\*</sup> , Hongzhi Zhang <sup>3,\*</sup>, Chuanhao Zhang <sup>4</sup> and Feng Jiang <sup>5</sup>

<sup>1</sup> School of Medicine and Health, Harbin Institute of Technology

<sup>2</sup> Harbin Institute of Technology Zhengzhou Research Institute

<sup>3</sup> Faculty of Computing, Harbin Institute of Technology

<sup>4</sup> School of Cyber Security and Smart Policing, Zhengzhou Police University

<sup>5</sup> Nanjing University of Information Science and Technology

\* Correspondence: lpj008@126.com (P.L.); zhanghz@hit.edu.cn (H.Z.)

## Abstract

In clinical CT imaging, high-density metallic implants often induce severe metal artifacts that obscure critical anatomical structures and degrade image quality, thereby hindering accurate diagnosis. Although deep learning has advanced CT metal artifact reduction (CT-MAR), many methods do not effectively use frequency information, which can limit the recovery of both fine details and overall image structure. To address this limitation, we propose a Hybrid-Frequency-Aware Mixture-of-Experts (HFMoE) network for CT-MAR. The proposed method synergizes the spatial-frequency localization of the wavelet transform with the global spectral representation of the Fourier transform to achieve precise multi-scale modeling of artifact characteristics. Specifically, we design a Hybrid-Frequency Interaction Encoder with three specialized branches, incorporating wavelet-domain, Fourier-domain, and cascaded wavelet–Fourier modulation, to distinctively refine local details, global structures, and complex cross-domain features. Then, they are fused via channel attention to yield a comprehensive representation. Furthermore, a frequency-aware Mixture-of-Experts (MoE) mechanism is introduced to dynamically route features to specific frequency experts based on the degradation severity, thereby adaptively assigning appropriate receptive fields to handle varying metal artifacts. Evaluations on synthetic (DeepLesion) and clinical (SpineWeb, CLINIC-metal) datasets show that HFMoE outperforms existing methods in both quantitative metrics and visual quality. Our method demonstrates the value of explicit frequency-domain adaptation for CT-MAR and could inform the design of other image restoration tasks.

**Keywords:** CT metal artifact reduction; Hybrid-Frequency-Aware; wavelet transform; Fourier transform; Mixture-of-Experts

## 1. Introduction

Computed Tomography (CT) is a fundamental tool in modern diagnostic radiology. However, high-density metallic implants, *e.g.*, dental fillings, hip prostheses, and surgical clips, can corrupt projection data, leading to severe metal artifacts in reconstructed images [1]. These artifacts manifest as dark bands and bright streaks, which not only degrade the perceptual image quality but also obscure critical anatomical details of surrounding tissues. Moreover, metal artifacts can compromise downstream clinical tasks, such as target delineation in radiotherapy and computer-aided diagnosis. Hence, developing effective CT metal artifact reduction (MAR) techniques to restore diagnostic information remains an important challenge in medical imaging.

Early solutions primarily relied on projection inpainting methods, such as Linear Interpolation (LI) [2] and Normalized MAR (NMAR) [3]. Although these methods can reduce strong streaks, they often introduce secondary artifacts and blur fine details because they do not account for the nonlinear effects of beam hardening and scattering. Recently, deep learning has significantly advanced this field

due to its powerful nonlinear representation capacity and sufficiently deep hierarchical architecture. In recent years, Deep Learning (DL) [4,5] has revolutionized this field owing to its powerful capability in non-linear feature modeling. Existing DL-based MAR methods are generally categorized into image domain [6,7], sinogram domain [8,9], and dual-domain collaboration approaches [10–14]. Among these, dual-domain methods have shown strong performance by combining information from both projection and image domains.

Despite progress, current deep learning methods remain limited in adapting to the varying frequency characteristics of artifacts. From the perspective of imaging physics, metal artifacts exhibit distinct spectral characteristics. Beam hardening typically induces low-frequency global shading (cupping effects), whereas photon starvation and scattering manifest as high-frequency radiating streaks and edge discontinuities. The distribution of these frequency components varies with the implant's geometry, material, and location. However, most existing models adopt a "static" inference paradigm, where the same convolutional kernels or attention mechanisms are applied uniformly across the entire input. This uniform approach struggles to handle complex artifact mixtures: standard CNNs exhibit spectral bias toward low frequencies, often leaving residual streaks or over-smoothing details [15]. As a result, such models often fail to simultaneously reduce large-scale shading and preserve fine details.

To address these issues, we propose that an effective MAR model should adapt to frequency content and employ dynamic computation. We propose a hybrid-frequency-aware mixture-of-experts (HFMoE) framework, which separates artifact processing into frequency bands and dynamically routes features to specialized modules. The core idea is two-fold. First, to overcome the spectral bias, we introduce a hybrid frequency analysis strategy. We combine the Wavelet Transform, which captures local high-frequency streaks, and the Fourier Transform, which models global low-frequency shading. Second, to address the limitation of static networks, we integrate the Mixture-of-Experts (MoE) philosophy into the decoder. Unlike static layers, our frequency-aware experts specialize in different artifact types. A dynamic gating mechanism learns to assign higher weights to the most relevant experts based on the local frequency content of the input, realizing a "divide-and-conquer" strategy for artifact removal. The encoder includes a hybrid-frequency module with parallel wavelet, Fourier, and cascaded branches, fused via channel attention. In the decoder, the MoE architecture takes both spatial and wavelet high-frequency features as input, dynamically allocating computational resources to suppress artifacts while preserving tissue details. Experiments on synthesized DeepLesion, Clinical SpineWeb, and CLINIC-metal datasets show that FAMoE improves upon current methods in artifact reduction and detail preservation.

In summary, our main contributions are as follows:

- We propose HFMoE, a unified network that integrates wavelet and Fourier transforms to fully exploit the complementary benefits of spatial-frequency and global spectral information for robust CT-MAR.
- A Hybrid-Frequency Interaction Encoder is developed to capture multi-scale features. It incorporates concurrent wavelet and Fourier modulation branches to simultaneously restore local textures and global structures.
- A Frequency-Aware MoE strategy is implemented in the decoder to dynamically allocate frequency-specific experts according to degradation severity of the input, enabling adaptive artifact correction.
- Extensive evaluations on multiple benchmarks, including synthesized and clinical datasets, validate that our method outperforms existing state-of-the-art methods in terms of both quantitative metrics and visual quality.

## 2. Related Works

### 2.1. CT Metal Artifact Reduction

Early MAR methods, such as linear interpolation [2] and normalized MAR [3], relied on heuristic inpainting or filtering, often introducing secondary artifacts or blurring. While foundational, they lack the adaptability for complex, patient-specific artifact patterns.

Deep learning [16,17] has revolutionized MAR, leading to single-domain, dual-domain, and multi-domain paradigms. Single-domain approaches operate either in the sinogram space (*e.g.*, CNN-MAR [8] and Deep-MAR [9]) or the image space (*e.g.*, OSCNet+[7]). Although effective, sinogram-domain methods can be sensitive to error propagation, while image-domain methods lack explicit physical data consistency [6]. Dual-domain methods, such as DuDoNet [10] and its successors [11,12,18], integrate both domains to enforce physical constraints, improving robustness. More recently, multi-domain frameworks like Quad-Net [19] have begun exploring complementary representations beyond the spatial and sinogram domains, hinting at the potential of frequency-aware processing.

However, a critical gap persists: existing methods predominantly operate on spatially-localized features, either from convolutional operations or local window-based attention. They lack a dedicated mechanism to explicitly model, decompose, and adaptively process the distinct frequency components that characterize metal artifacts, which span both global low-frequency shadows and local high-frequency streaking. This oversight limits their ability to achieve optimal fidelity and structural preservation across diverse clinical scenarios.

### 2.2. Frequency-Domain Modeling for Image Restoration

Leveraging frequency-domain transformations (*e.g.*, Fourier, Wavelet) is a proven strategy for separating and processing different image components in general vision tasks [15]. Classical methods used thresholding in transform domains [20], while modern deep learning approaches integrate these transforms into networks. For instance, some works decompose features into high- and low-frequency subbands for targeted enhancement [21–23], or use Fourier convolutions to capture global context [24,25], which is similar to Transformer [26]. In medical imaging, wavelet transforms have been used for super-resolution to preserve detail [27,28], and hybrid spatial-Fourier models aim to balance local and global processing [29,30].

Despite these advances, most current methods are limited in two ways: 1) They often rely on a single type of frequency transform (*e.g.*, only Fourier or only Wavelet), failing to capture the hybrid frequency characteristics (global spectral shifts and multi-scale local discontinuities) that define CT metal artifacts. 2) Their processing is typically static; the same network path handles all frequency content, lacking the adaptability to specialize for the highly variable spectral signatures of different artifacts and anatomies.

### 2.3. Mixture-of-Experts in Low-Level Vision

The Mixture-of-Experts (MoE) paradigm provides a compelling framework for adaptive computation by employing multiple specialized sub-networks (experts) and a dynamic routing mechanism [31–33]. In low-level vision, MoE has been successfully applied to weather restoration [34], multi-task learning [35], and MRI denoising/super-resolution [36,37]. The core strength lies in its ability to conditionally activate different experts based on input features, allowing for task-aware or region-aware specialization. Recent innovations include gating mechanisms guided by semantic prompts [38] or designed for complexity-aware routing [39].

While promising, the application of MoE in CT-MAR remains nascent. Crucially, existing MoE designs in medical imaging do not explicitly route information based on frequency-domain characteristics. An MoE architecture where experts are specialized for distinct frequency bands or spectral patterns, and where routing is informed by a hybrid-frequency analysis of the input artifact, remains unexplored. This represents a significant opportunity to move beyond spatial-only adaptive processing.

### 3. Materials and Methods

#### 3.1. Frequency-Domain Foundations: DWT and FFT

Our method is grounded in the complementary analysis provided by the Discrete Wavelet Transform (DWT) and the Fast Fourier Transform (FFT). Metal artifacts in CT scans exhibit a complex, hybrid-frequency nature, comprising both localized high-frequency streaking and globally distributed low-frequency shading. Individually, DWT and FFT offer distinct advantages for modeling different aspects of this problem. Their combination forms the basis of our hybrid-frequency approach.

##### 3.1.1. Discrete Wavelet Transform for Localized Analysis

The 2D DWT performs a multi-resolution decomposition of an image, enabling joint analysis in space and frequency. For an input image or feature map  $\mathbf{X} \in \mathbb{R}^{H \times W}$ , a single decomposition level applies quadrature mirror filter banks, producing four subbands:

$$\{\mathbf{X}_{LL}, \mathbf{X}_{LH}, \mathbf{X}_{HL}, \mathbf{X}_{HH}\} = \text{DWT}(\mathbf{X}), \quad (1)$$

where  $\mathbf{X}_{LL}$  represents the low-frequency approximation component, capturing coarse-grained structural information. The subbands  $\mathbf{X}_{LH}$ ,  $\mathbf{X}_{HL}$ , and  $\mathbf{X}_{HH}$  correspond to high-frequency details along the horizontal, vertical, and diagonal directions, respectively. The original signal can be perfectly reconstructed via the Inverse DWT (IWT):  $\mathbf{X} = \text{IWT}(\mathbf{X}_{LL}, \mathbf{X}_{LH}, \mathbf{X}_{HL}, \mathbf{X}_{HH})$ .

In fact, DWT provides excellent space-frequency localization, which can effectively capture localized features, *i.e.*, edges and fine details, whereas corresponding to metal streaks. Thus, DWT is suitable for preserving local structural details.

##### 3.1.2. Fast Fourier Transform for Global Context

In contrast, the FFT provides a global representation of an image's frequency content. For the 2D feature map  $\mathbf{X}$ , its 2D FFT is defined as:

$$\mathcal{F}(\mathbf{X})[u, v] = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} \mathbf{X}[h, w] \cdot e^{-j2\pi(uh/H + vw/W)}, \quad (2)$$

with the inverse transform (iFFT) denoted by  $\mathcal{F}^{-1}(\cdot)$ . The magnitude spectrum  $|\mathcal{F}(\mathbf{X})|$  reveals the global energy distribution across all frequencies, while the phase spectrum encodes essential structural information. The key attribute of the FFT is its inherent global receptive field; each point in the frequency domain depends on every pixel in the spatial domain. This property makes the FFT exceptionally effective for modeling and suppressing globally correlated degradations, *i.e.*, the low-frequency shadows caused by beam hardening, a common source of metal artifacts. Furthermore, the FFT and iFFT operations are efficient, with a computational complexity of  $O(HW \log(HW))$ .

##### 3.1.3. Motivation for a Hybrid Frequency-Aware Design

DWT and FFT have complementary characteristics. While DWT offers precise localization ideal for high-frequency details, its receptive field for capturing long-range, global patterns is inherently limited by the filter window size. Conversely, the FFT is good at global frequency analysis but sacrifices all spatial locality, making it less suited for processing fine, localized structures without additional mechanisms.

Naturally, this complementarity matches the dual nature of CT metal artifacts. Therefore, we propose to construct a unified Wavelet-Fourier Hybrid Frequency-Aware Representation. This allows the network to leverage both transforms: DWT for local details and FFT for low-frequency shading. Integrating both domains provides a richer feature space for artifact reduction compared to single-domain methods.

### 3.2. Overview

Our proposed Hybrid-Frequency-Aware Mixture-of-Experts network (HF-MoE) follows an encoder-decoder architecture with skip connections. Given a metal-corrupted CT image  $\mathbf{Y} \in \mathbb{R}^{H \times W}$ , the network first extracts shallow features via a convolutional stem:  $\mathbf{F}_0 = \text{Conv}_{3 \times 3}(\mathbf{Y})$ .

The encoder comprises four stages, each performing feature downsampling via DWT. Each stage contains a Hybrid-Frequency Interaction Module (HFIM) (Section 3.3) that processes features in spatial and frequency domains. The bottleneck layer employs multiple HFIMs for deep feature refinement.

The decoder symmetrically upsamples features via IWT. Crucially, each decoder stage incorporates the proposed Frequency-Aware Mixture-of-Experts Module (FAME) (Section 3.4), which dynamically routes features to experts specialized for different frequency patterns and receptive fields. Skip connections aggregate multi-scale features from the encoder to the decoder.

The final restored image  $\hat{\mathbf{X}}$  is obtained through a reconstruction layer and a residual connection:

$$\hat{\mathbf{X}} = \text{Conv}_{3 \times 3}(\mathbf{F}_d) + \mathbf{Y}, \quad (3)$$

where  $\mathbf{F}_d$  denotes the output of the decoder. This residual design enables the network learn the artifact component.

### 3.3. Hybrid-Frequency Interaction Module (HFIM)

The proposed HFIM is designed to jointly leverage the complementary strengths of wavelet and Fourier transforms for local detail preservation and global artifact suppression. As shown in Figure 1, given an input feature  $\mathbf{F}_{in}^{HFIM} \in \mathbb{R}^{H \times W \times C}$ , it is first projected to  $\mathbf{F} = \text{Conv}_{1 \times 1}(\mathbf{F}_{in})$ .  $\mathbf{F}$  is then split into three parallel branches.

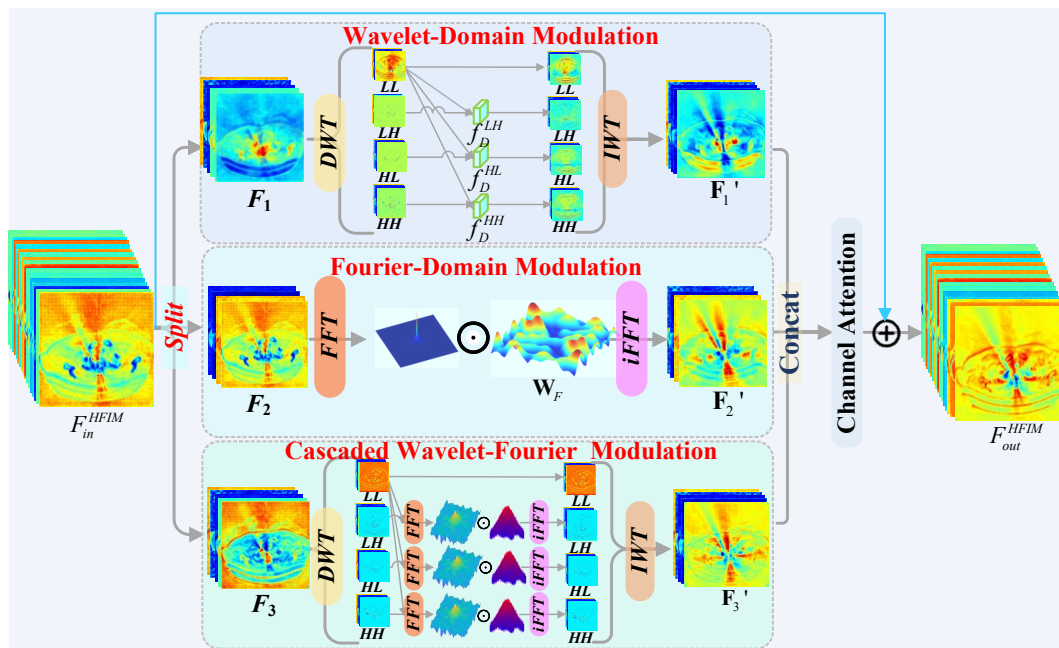


Figure 1. The architecture of Hybrid-Frequency Interaction Module.

**Branch 1: Wavelet-Domain Modulation.** This branch focuses on localized details.  $F_1$  is decomposed via DWT into four subbands:  $\{F_1^{LL}, F_1^{LH}, F_1^{HL}, F_1^{HH}\}$ . The high-frequency (HF) subbands are concatenated with the low-frequency (LF) component and processed by depthwise-separable convolution blocks  $f_D^{(.)}$ :

$$\mathbf{F}'_{1,H} = \text{Concat}\left(f_D^{LH}([F_1^{LL}, F_1^{LH}]), f_D^{HL}([F_1^{LL}, F_1^{HL}]), f_D^{HH}([F_1^{LL}, F_1^{HH}])\right). \quad (4)$$

The processed features are then reconstructed via IWT:  $\mathbf{F}'_1 = \text{IWT}(\mathbf{F}_1^{LL}, \mathbf{F}'_{1,H})$ .

**Branch 2: Fourier-Domain Modulation.** This branch targets globally distributed artifacts. The feature is transformed via 2D FFT:  $\mathcal{F}(F_2) \in \mathbb{C}^{H \times W \times C/3}$ . A learnable frequency filter  $\mathbf{W}_F \in \mathbb{C}^{H \times W \times C/3}$  performs element-wise modulation, followed by the inverse FFT (iFFT):

$$F'_2 = \mathcal{F}^{-1}(\mathbf{W}_F \odot \mathcal{F}(F_2)). \quad (5)$$

**Branch 3: Cascaded Wavelet-Fourier Modulation.** This branch models interactions across scales and frequencies. The HF subbands from an initial DWT of  $F_3$  are first transformed to the Fourier domain, filtered, and then transformed back:

$$F'_{3,i} = \mathcal{F}^{-1}\left(\mathbf{W}_{F,i} \odot \mathcal{F}\left(f_F^i([F_3^{LL}, F_3^i])\right)\right), \quad i \in \{LH, HL, HH\}. \quad (6)$$

The final output is obtained via IWT, i.e.,  $F'_3 = \text{IWT}(F_3^{LL}, F'_{3,LH}, F'_{3,HL}, F'_{3,HH})$ .

The outputs of the three branches are aggregated via channel attention and passed through a feed-forward network (FFN), resulting in the final output  $F_{out}^{HFIM}$  of the HFIM.

### 3.4. Frequency-Aware Mixture-of-Experts Module (FAME)

Building upon the MoE paradigm, the FAME module dynamically selects specialized experts based on the frequency characteristics of input features. As illustrated in Figure 2, the input feature  $F_{in}^{FAME} \in \mathbb{R}^{H \times W \times C}$  is first enriched with high-frequency cues extracted via DWT. This hybrid feature serves as input to the gating network and the expert pool.

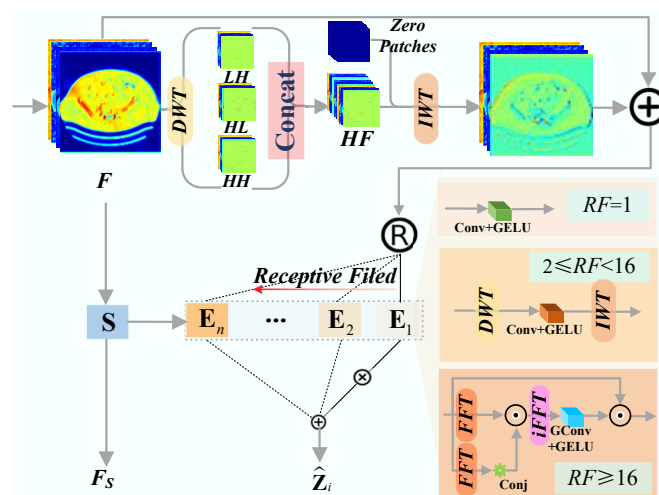


Figure 2. The architecture of Frequency-Aware Mixture-of-Experts.

**Expert Design.** Our expert pool  $\{E_i\}_{i=1}^N$  consists of three types of experts with increasing receptive field (RF) and parameter complexity, designed to handle different degradation patterns:

$$\hat{Z}_E^i = \begin{cases} \text{Conv}_{1 \times 1}(Z_E^1), RF = 1 \\ \text{IWT}(\text{Conv}_{k \times k}(\text{DWT}(Z_E^i))), 2 \leq RF < 16 \\ Z_E^i \odot \mathcal{F}^{-1}\left(\mathcal{F}(\mathbf{w}_Q(Z_E^i)) \odot (\mathcal{F}(\mathbf{w}_K(Z_E^i)))^H\right), RF \geq 16 \end{cases} \quad (7)$$

- **Local Expert** ( $RF = 1$ ): Employs pointwise convolution ( $1 \times 1$  Conv) for pixel-wise adjustment.
- **Wavelet Expert** ( $2 \leq RF < 16$ ): Applies a DWT-IWT sandwich with a  $k \times k$  depthwise convolution in the wavelet domain, effective for medium-range structural artifacts.
- **Fourier Expert** ( $RF \geq 16$ ): Approximates global self-attention via Fourier autocorrelation, efficient for suppressing globally correlated artifacts.

The  $i$ -th expert processes a routed feature  $Z_i$  and produces an output  $\hat{Z}_i$ .

**Frequency-Aware Routing.** A lightweight router  $\mathcal{G}$ , implemented as a linear layer followed by softmax, generates a routing weight vector  $\mathbf{g} = \mathcal{G}(\mathbf{F}) \in \mathbb{R}^N$ . For each spatial position (or token), only the expert with the highest weight is activated (Top-1 routing). To encourage balanced utilization and align expert choice with frequency content, we employ an auxiliary load-balancing loss  $\mathcal{L}_{aux}$  [39].

The final output of the FAME module is the aggregated expert outputs, modulated by a shared feature projection  $\mathbf{S}(\mathbf{F})$ , and processed by an FFN:

$$\mathbf{F}_{out}^{FAME} = \text{FFN} \left( \sum_{i=1}^N g_i \cdot (\hat{\mathbf{Z}}_i \odot \mathbf{S}(\mathbf{F})) \right). \quad (8)$$

### 3.5. Loss Functions

Our training objective combines three complementary losses:

**1. Reconstruction Loss ( $\mathcal{L}_{rec}$ ):** An  $\ell_1$  loss [40] ensures pixel-level fidelity, weighted inside ( $\mathcal{M}$ ) and outside the metal trace region:

$$\mathcal{L}_{rec} = \|(\hat{\mathbf{X}} - \mathbf{X}^{gt}) \odot \mathcal{M}\|_1 + \lambda_{bg} \|(\hat{\mathbf{X}} - \mathbf{X}^{gt}) \odot (1 - \mathcal{M})\|_1. \quad (9)$$

**2. Hybrid Frequency Loss ( $\mathcal{L}_{freq}$ ):** This loss enforces consistency in both wavelet and Fourier domains. Let  $\mathcal{D}_W(\cdot)$  and  $\mathcal{D}_F(\cdot)$  denote operations extracting high-frequency wavelet subbands and Fourier magnitude spectra, respectively.

$$\mathcal{L}_{freq} = \sum_{s \in \{LH, HL, HH\}} \alpha_s \| \mathcal{D}_W^s(\hat{\mathbf{X}}) - \mathcal{D}_W^s(\mathbf{X}^{gt}) \|_1 + \beta \| \mathcal{D}_F(\hat{\mathbf{X}}) - \mathcal{D}_F(\mathbf{X}^{gt}) \|_1. \quad (10)$$

**3. Routing Auxiliary Loss ( $\mathcal{L}_{aux}$ ):** Adopted from [39], this loss balances expert utilization and encourages alignment between routing decisions and task complexity.

The total loss is a weighted sum:  $\mathcal{L}_{total} = \mathcal{L}_{rec} + \lambda_f \mathcal{L}_{freq} + \lambda_a \mathcal{L}_{aux}$ . The hyperparameters  $\lambda_{bg}, \lambda_f, \lambda_a, \alpha_s, \beta$  are selected based on empirical observations.

**Table 1.** Performance evaluation in terms of PSNR/SSIM on the Synthesized DeepLesion benchmark. **Bold** numbers denote the highest scores.

Methods	Large Metal $\rightarrow$		Medium Metal $\rightarrow$	Small Metal		Average
Input	24.12 / 0.6761	26.13 / 0.7471	27.75 / 0.7659	28.53 / 0.7964	28.78 / 0.8076	27.06 / 0.7586
LI (1987) [2]	27.21 / 0.8920	28.31 / 0.9185	29.86 / 0.9464	30.40 / 0.9555	30.57 / 0.9608	29.27 / 0.9347
NMAR (2010) [3]	27.66 / 0.9114	28.81 / 0.9373	29.69 / 0.9465	30.44 / 0.9591	30.79 / 0.9669	29.48 / 0.9442
CNNMAR (2018) [8]	28.92 / 0.9433	29.89 / 0.9588	30.84 / 0.9706	31.11 / 0.9743	31.14 / 0.9752	30.38 / 0.9644
DuDoNet (2019) [10]	29.87 / 0.9723	30.60 / 0.9786	31.46 / 0.9839	31.85 / 0.9858	31.91 / 0.9862	31.14 / 0.9814
DuDoNet++ (2020) [18]	36.17 / 0.9784	38.34 / 0.9891	40.32 / 0.9913	41.56 / 0.9919	42.08 / 0.9921	39.69 / 0.9886
DSCMAR (2020) [41]	34.04 / 0.9343	33.10 / 0.9362	33.37 / 0.9384	32.75 / 0.9393	32.77 / 0.9395	33.21 / 0.9375
DAN-Net (2021) [42]	30.82 / 0.9750	31.30 / 0.9796	33.39 / 0.9852	35.02 / 0.9883	43.61 / 0.9950	34.83 / 0.9846
InDuDoNet (2021) [1]	36.74 / 0.9742	39.32 / 0.9893	41.86 / 0.9944	44.47 / 0.9948	45.01 / 0.9958	41.48 / 0.9897
DICDNet (2021) [43]	37.19 / 0.9853	39.53 / 0.9908	42.25 / 0.9941	44.91 / 0.9953	45.27 / 0.9958	41.83 / 0.9923
ACDNet (2022) [6]	37.91 / 0.9872	39.30 / 0.9920	41.14 / 0.9949	42.43 / 0.9961	42.64 / 0.9965	40.68 / 0.9933
OSNet (2022) [44]	37.70 / 0.9883	39.88 / 0.9902	42.92 / 0.9950	45.04 / 0.9958	45.45 / 0.9962	42.19 / 0.9931
InDuDoNet+ (2023) [45]	36.28 / 0.9736	39.23 / 0.9872	41.81 / 0.9937	45.03 / 0.9952	45.15 / 0.9959	41.50 / 0.9891
OSNet+ (2023) [7]	38.98 / 0.9897	40.72 / 0.9930	43.46 / 0.9956	45.51 / 0.9965	45.99 / 0.9968	42.93 / 0.9943
DuDoDp (2024) [46]	29.11 / 0.9580	29.30 / 0.9631	29.07 / 0.9668	28.33 / 0.9673	28.33 / 0.9673	28.83 / 0.9645
MARFormer (2024) [47]	40.56 / 0.9903	42.32 / 0.9933	43.90 / 0.9947	45.78 / 0.9956	45.70 / 0.9957	43.65 / 0.9939
MoCE-IR (2025) [39]	41.23 / 0.9889	42.44 / 0.9912	44.02 / 0.9923	45.48 / 0.9932	45.98 / 0.9931	43.86 / 0.9917
FAMoE(Ours)	<b>41.70 / 0.9916</b>	<b>43.09 / 0.9940</b>	<b>44.43 / 0.9951</b>	<b>46.20 / 0.9961</b>	<b>46.71 / 0.9963</b>	<b>44.43 / 0.9946</b>

## 4. Experimental Settings

### 4.1. Datasets

To comprehensively evaluate the proposed method, we utilize one synthesized dataset for training and quantitative benchmarking, and two clinical datasets for assessing cross-domain generalizability. **Synthesized DeepLesion Dataset.** Following the protocol established in OSCNet+ [7], we constructed a synthesized dataset based on the DeepLesion collection. The training set comprises 1,000 paired clean and metal-corrupted volumes, generated by simulating metal artifacts on 1,200 randomly sampled metal-free abdominal/thoracic CT images using 90 distinct implant types [8]. To simulate clinical scenarios, the pipeline includes beam-hardening and partial-volume effects, with randomized metal

mask size, orientation, and position. For evaluation, 200 pairs featuring 10 unseen implant types are reserved. These test implants cover a wide range of sizes, from 3 pixels (tiny fragments) to 2,061 pixels (large prostheses). Consistent with prior works [7], all volumes are re-sampled into 640 projection views and reconstructed at a resolution of  $416 \times 416$ .

**Clinical SpineWeb Dataset<sup>1</sup>.** To validate clinical generalization ability, we employ the SpineWeb dataset as an external test bed. This dataset consists of post-operative spinal CT scans with thoracolumbar instrumentation. We select a set of spine volumes with implants that are not used in training. Metal implants are segmented using a standard threshold of 2,500 HU. This dataset serves to evaluate the model’s performance on real-world metallic hardware distinct from the training distribution.

**Clinical CLINIC-metal Dataset.** We also utilize the CLINIC-metal pelvic CT dataset [48] to test generalization to different anatomies. This multi-center dataset contains pelvic CT scans with severe artifacts, challenging the model on different anatomies and acquisition protocols.

#### 4.2. Implementation Details

Table 2 outlines the architectural specifications of FAMoE. We employ progressive multi-resolution training, increasing input size from  $256 \times 256$  to  $384 \times 384$ , then to  $416 \times 416$ . The model is optimized using the AdamW optimizer with a cosine annealing schedule, where the learning rate decays from an initial  $1 \times 10^{-3}$  to  $1 \times 10^{-6}$  following [49]. Data augmentation techniques, including random flipping, rotation, cropping, and MixUp, are applied to prevent overfitting. The framework is implemented in PyTorch. All experiments are conducted on a single NVIDIA RTX 4080S GPU. The model is trained for 2,000 epochs, requiring approximately 72 hours.

**Table 2.** HFMoE network configurations.

Task	Features	$L_{E1}$	$L_{E2}$	$L_{E3}$	$L_{E4}$	$L_{D4}$	$L_{D3}$	$L_{D2}$	$L_{D1}$
CT-MAR	16	1	2	4	6	6	4	2	1

#### 4.3. Experimental Results and Discussion

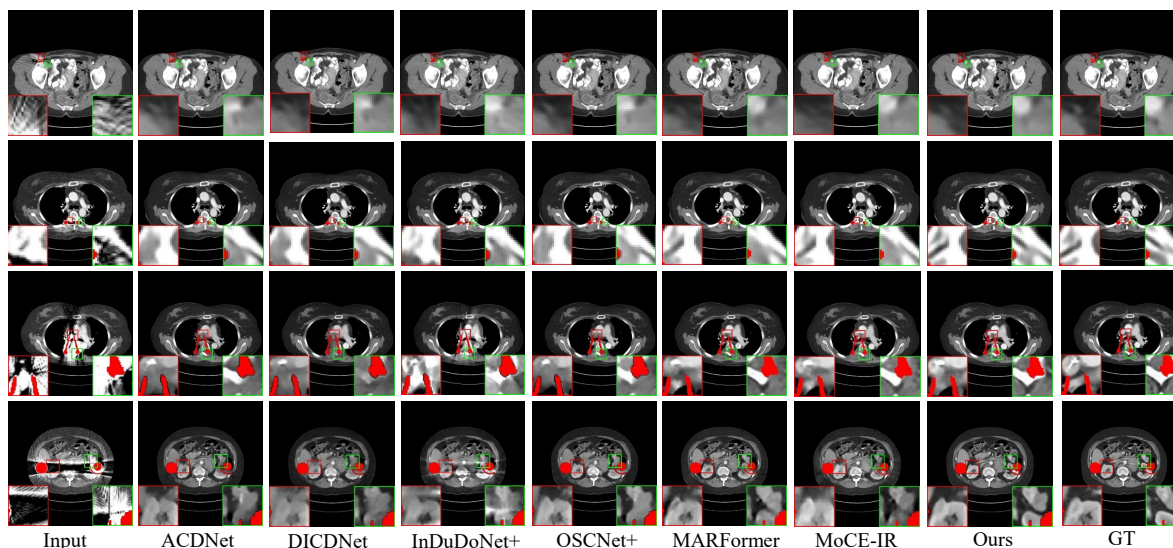
##### 4.3.1. Quantitative Evaluation

Table 1 compares our proposed HFMoE with 16 existing MAR methods. For fair comparison, models with open-source codes (*e.g.*, OSCNet+, InDuDoNet) were evaluated using official pre-trained weights, while others were retrained following their original protocols. HFMoE achieves the best results on all metrics. As detailed in Table 1, our method achieves the highest PSNR across all implant sizes. Notably, HFMoE surpasses the best-performing MoE-based competitor, MoCE-IR, by 0.57 dB in PSNR, and outperforms the leading dual-domain method, InDuDoNet+, by a substantial margin of 2.5 dB. The improved SSIM scores also indicate better structural preservation.

##### 4.3.2. Qualitative Evaluation

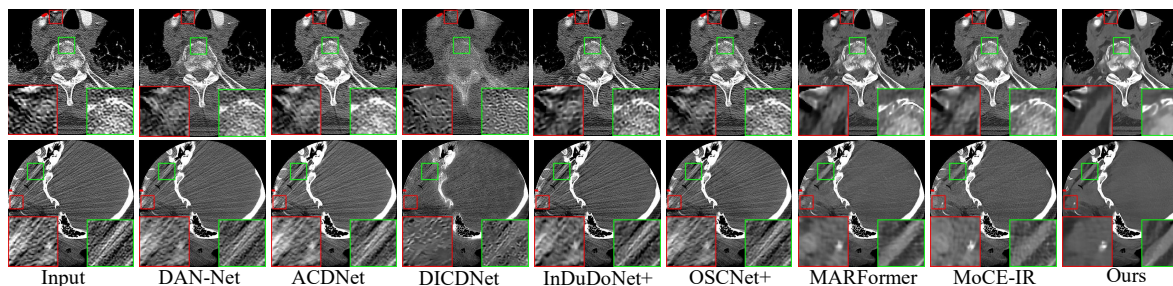
**Synthesized DeepLesion.** Figure 3 illustrates the reconstruction results on cases with varying implant sizes. HFMoE exhibits superior artifact suppression capabilities, particularly in regions surrounding large implants. While competing methods often leave prominent residual streaks (highlighted in circles), our approach reduces these streaks and better recovers anatomical boundaries. For small implants, HFMoE also produces more realistic textures.

<sup>1</sup> Available at: <http://spineweb.digitalimaginggroup.ca>

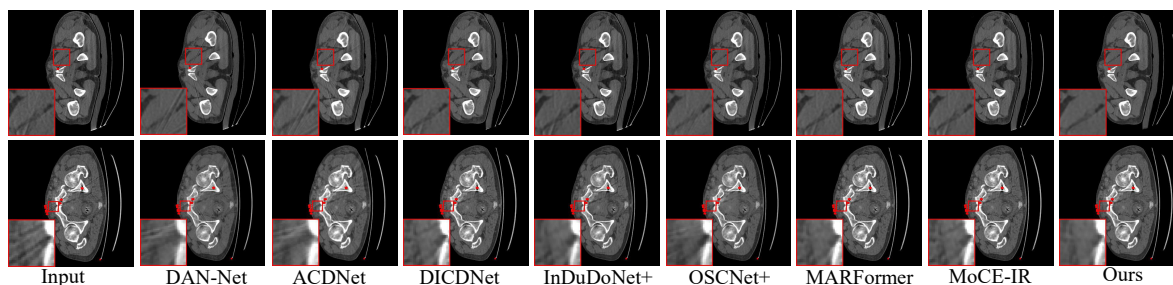


**Figure 3.** Visual comparison of metal artifact reduction methods on the Synthesized DeepLesion dataset (cases 100, 1175, 1362, 1618). The proposed HFMoE method recovers superior anatomical detail. Display:  $[-175, 275]$  HU; metal (blue). ROIs are enlarged. [Zoom in](#) for a better view.

**Clinical Datasets.** Qualitative results on the clinical SpineWeb and CLINIC-metal datasets are presented in Figure 4 and Figure 5, respectively. On SpineWeb, HFMoE significantly reduces beam-hardening streaks from spinal hardware, making vertebral structures more visible compared to baseline methods. On the CLINIC-metal dataset with complex pelvic anatomy, our method generalizes well textures which demonstrates robust generalization. HFMoE better preserves tissue boundaries and sharpness compared to other methods, which often cause blurring or distortion.



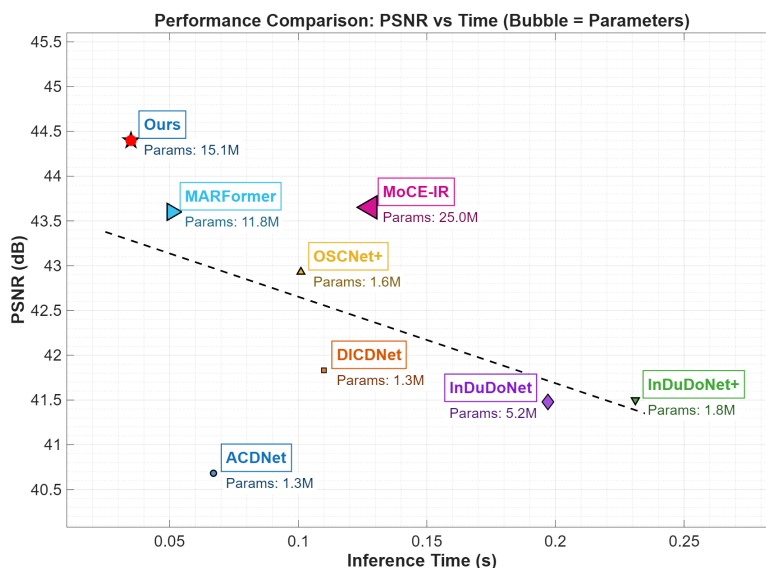
**Figure 4.** Qualitative comparison on clinical spine CT scans. The proposed HFMoE provides superior artifact suppression with better structural preservation, whereas competing methods show residual streaking. Display window:  $[50, 500]$  HU. ROIs are enlarged. [Zoom in](#) for a better view.



**Figure 5.** Qualitative comparison on the CLINIC-Metal dataset. The proposed HFMoE recovers anatomical structures with superior clarity and fidelity compared to competing methods. Display window:  $[50, 500]$  HU. ROIs are enlarged. [Zoom in](#) for a better view.

#### 4.4. Computational Efficiency

Figure 6 contrasts the inference efficiency of leading MAR methods. HFMoE has an inference time of 38 ms per slice ( $416 \times 416$ ). This is about 40% faster than ACDNet (65 ms) and also achieves over  $\sim 0.8$  dB higher PSNR than MARFormer and MoCE-IR. The model has a small memory footprint: training uses  $\sim 32$  GB VRAM (batch size 4), and inference requires  $\sim 180$  MB per slice. The efficiency and performance make it a practical candidate for clinical use.



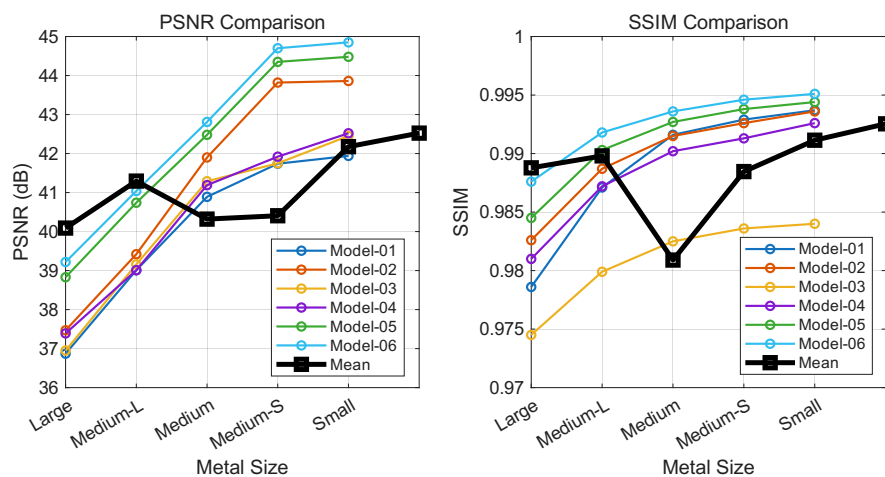
**Figure 6.** Runtime *v.s.* PSNR with Model Complexity. Horizontal axis: running time (s); Vertical axis: PSNR (dB); Parameter count is noted below each method.

#### 4.5. Ablation Study

We conduct a comprehensive ablation study on the Synthesized DeepLesion dataset to validate the contribution of three core components: the Hybrid-Frequency Interaction Module (HFIM), the Frequency-Aware Mixture-of-Experts (FAME), and the multi-scale loss function.

##### 4.5.1. Efficacy of HFIM

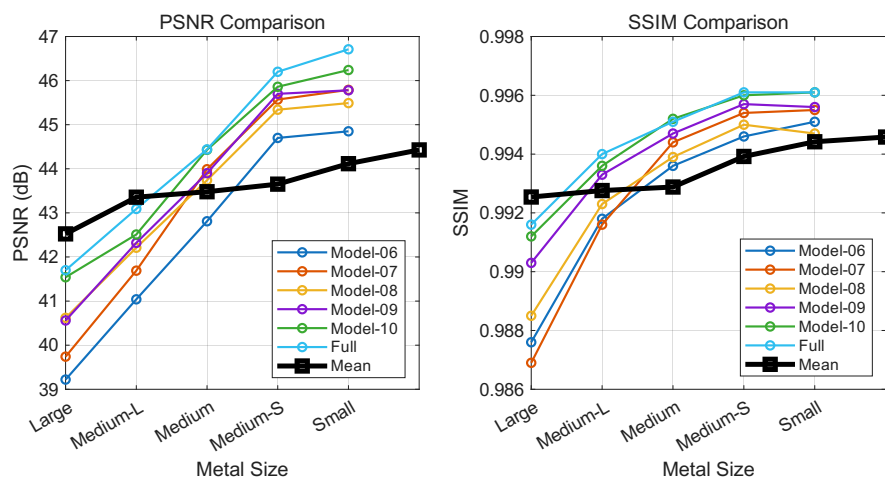
The HFIM leverage complementary information from Wavelet and Fourier domains. To analyze its impact, we constructed a baseline Transformer model (Model-01) and progressively integrated frequency modulation blocks. As shown in Figure 7, introducing wavelet-domain modulation (Model-02) yields performance gains by capturing local high-frequency details. While Fourier modulation alone (Model-03) proves less stable due to spectral sensitivity, its parallel combination with wavelet modulation (Model-05) produces a marked improvement, suggesting that the two domains offer complementary representations. The final configuration (Model-06), which incorporates a cascaded wavelet-Fourier strategy, achieves the best performance, confirming the synergy of the proposed triple-branch mechanism.



**Figure 7.** Contribution analysis of HFIM variants: PSNR (left) and SSIM (right) versus implant size (large to small). Curves represent different model configurations. *Note: The black curve sequences Model-1 to Model-6 (left to right), not implant size.*

#### 4.5.2. Role of FAME

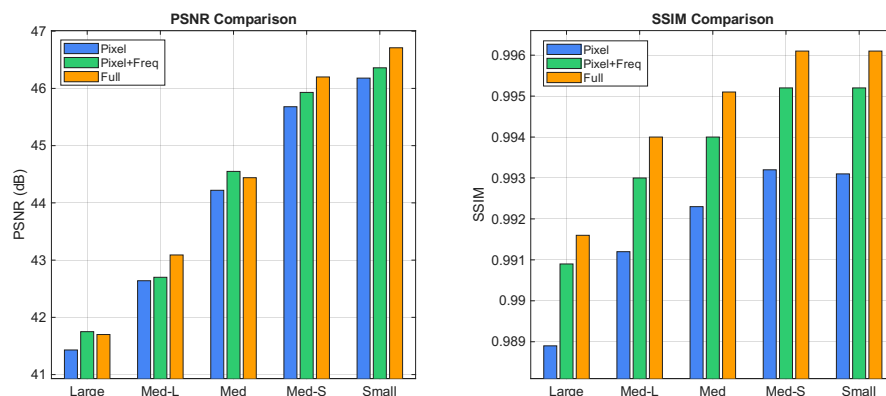
The FAME module utilizes a gating mechanism to dynamically route features based on local frequency content. We compared our frequency-aware design against standard MoE and single-domain variants (Figure 8). Replacing the standard MoE experts (Model-07) with multi-scale convolutions (Model-08) offers limited gains. However, employing domain-specific experts, either Wavelet (Model-09) or Fourier (Model-10), leads to noticeable improvements. The full FAME module, which adaptively selects between efficient wavelet experts for simple regions and global Fourier experts for complex textures, achieves optimal performance. This indicates that our proposed FAME module can adaptively assign appropriate receptive fields to handle varying metal artifacts.



**Figure 8.** Performance of FAME variants across implant sizes (large to small). Left: PSNR; Right: SSIM. Each curve denotes a model configuration. *Note: The black curve sequences Model-6 to Model-11 (left to right), not implant size.*

#### 4.5.3. Impact of Loss Function

We finally analyze the composite loss function, which consists of reconstruction, frequency-aware, and routing-assisted terms. Ablation results in Figure 9 indicate that reliance solely on reconstruction loss is insufficient for recovering fine details. Adding the frequency-aware loss improves performance by explicitly constraining artifact removal in spectral domains. The routing-assisted loss further refines the model by ensuring balanced expert utilization, leading to the best perceptual quality.



**Figure 9.** Performance of loss function variants evaluated by PSNR (left) and SSIM (right). Implant size decreases along the horizontal axis; colored bars denote the metric values of each variant.

## 5. Discussion

In medical image processing, wavelet and Fourier transforms offer complementary frequency-domain representations: wavelets capture localized time-frequency features and edges, while Fourier analysis characterizes global periodic structures and spectral distributions. While each transform has been used separately in CNNs or Transformers for tasks like denoising, few methods combine them to leverage their complementary strengths. We propose a hybrid-frequency-aware mixture-of-experts (HFMoE) model for CT-MAR. It uses a gating mechanism to dynamically combine wavelet and Fourier features, adapting the artifact suppression process to different frequency content.

Our work has several limitations that suggest directions for future research. First, the model is trained only on synthetic data, which may not fully capture the artifact patterns in real clinical scans, affecting its generalization. Second, our method processes 2D slices independently, without exploiting 3D spatial context. Third, imaging parameters (*e.g.*, scanner settings, X-ray energy) that influence artifact formation are not incorporated. Future work could: i) use more realistic training data or domain adaptation techniques to improve generalization to clinical data; ii) extend HFMoE to 3D to utilize volumetric context; and iii) incorporate imaging parameters (*e.g.*, scanner settings) to enhance robustness.

## 6. Conclusions

This paper presents a Hybrid-Frequency-aware Mixture-of-Experts (HFMoE) model for metal artifact reduction in CT imaging. The core innovation lies in the synergistic integration of wavelet and Fourier transforms to jointly exploit their complementary strengths: wavelet transforms capture localized time-frequency details, while Fourier transforms characterize global spectral structures. The proposed method has two key components. First, a hybrid module processes features in parallel through wavelet, Fourier, and combined wavelet-Fourier branches, followed by adaptive fusion. Second, a routing mechanism dynamically selects specialized experts, each tailored to handle different artifact patterns based on local frequency content. Experiments on synthetic and clinical datasets show that HFMoE improves artifact reduction and preserves anatomical details better than existing methods, as measured by quantitative metrics and visual inspection. Future work will extend the framework to 3D reconstruction and explore its transferability to other multimodal medical imaging tasks.

**Author Contributions:** Conceptualization, Pengju Liu and Hongzhi Zhang; methodology, Pengju Liu; software, Pengju Liu; validation, Pengju Liu, Hongzhi Zhang and Feng Jiang; formal analysis, Pengju Liu; investigation, Pengju Liu; resources, Pengju Liu and Hongzhi Zhang; data curation, Pengju Liu, Chuanhao Zhang; writing—original draft preparation, Pengju Liu; writing—review and editing, Hongzhi Zhang, Chuanhao Zhang; visualization, Hongzhi Zhang and Feng Jiang; supervision, Hongzhi Zhang and Feng Jiang; project administration, Hongzhi Zhang and Feng Jiang; funding acquisition, Feng Jiang, Chuanhao Zhang. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the China Postdoctoral Science Foundation (Grant No. 2024M754208), the National Natural Science Foundation of China (Grant Nos. 62371164 and 62576123), and the Henan Provincial Science and Technology Research Project (Grant No. 252102210147).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created.

**Acknowledgments:** All authors declare that they have no known conflicts of interest in terms of competing financial interests or personal relationships that could have an influence or are relevant to the work reported in this paper.

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

## References

1. Wang, H.; Li, Y.; Zhang, H.; Chen, J.; Ma, K.; Meng, D.; Zheng, Y. InDuDoNet: An interpretable dual domain network for CT metal artifact reduction. In Proceedings of the MICCAI, 2021, pp. 107–118.
2. Kalender, W.A.; Hebel, R.; Ebersberger, J. Reduction of CT artifacts caused by metallic implants. *Radiology* **1987**, *164*, 576–577.
3. Meyer, E.; Raupach, R.; Lell, M.; Schmidt, B.; Kachelrieß, M. Normalized metal artifact reduction (NMAR) in computed tomography. *Med Phys* **2010**, *37*, 5482–5493.
4. Tian, C.; Zheng, M.; Lin, C.W.; Li, Z.; Zhang, D. Heterogeneous window transformer for image denoising. *IEEE T-SMC* **2024**.
5. Tian, C.; Cheng, T.; Peng, Z.; Zuo, W.; Tian, Y.; Zhang, Q.; Wang, F.Y.; Zhang, D. A survey on deep learning fundamentals. *Artificial Intelligence Review* **2025**, *58*, 381.
6. Wang, H.; Li, Y.; Meng, D.; Zheng, Y. Adaptive Convolutional Dictionary Network for CT Metal Artifact Reduction. In Proceedings of the IJCAI. IEEE, 2022, pp. 1401–1407.
7. Wang, H.; Xie, Q.; Zeng, D.; Ma, J.; Meng, D.; Zheng, Y. OSCNet: Orientation-Shared Convolutional Network for CT Metal Artifact Learning. *IEEE T-MI* **2023**, *43*, 489–502.
8. Zhang, Y.; Yu, H. Convolutional neural network based metal artifact reduction in X-ray computed tomography. *IEEE T-MI* **2018**, *37*, 1370–1381.
9. Ghani, M.U.; Karl, W.C. Fast enhanced CT metal artifact reduction using data domain deep learning. *IEEE T-MI* **2019**, *6*, 181–193.
10. 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 CVPR, 2019, pp. 10512–10521.
11. Zhou, B.; Chen, X.; Zhou, S.K.; Duncan, J.S.; Liu, C. DuDoDR-Net: Dual-domain data consistent recurrent network for simultaneous sparse view and metal artifact reduction in computed tomography. *Med Image Anal* **2022**, *75*, 102289.
12. Wang, H.; Yang, S.; Bai, X.; Wang, Z.; Wu, J.; Lv, Y.; Cao, G. IRDNet: Iterative Relation-Based Dual-Domain Network via Metal Artifact Feature Guidance for CT Metal Artifact Reduction. *IEEE T-RPMC* **2024**, *8*, 959–972.
13. Zheng, S.; Zhang, D.; Yu, C.; Jia, L.; Zhu, L.; Huang, Z.; Zhu, D.; Yu, H. MAReraser: Metal Artifact Reduction with Image Prior Using CNN and Transformer Together. In Proceedings of the BIBM. IEEE, 2024, pp. 4060–4065.
14. Yao, X.; Tan, J.; Deng, Z.; Xiong, D.; Zhao, Q.; Wu, M. MUPO-Net: A Multilevel Dual-domain Progressive Enhancement Network with Embedded Attention for CT Metal Artifact Reduction. In Proceedings of the ICASSP. IEEE, 2025, pp. 1–5.
15. Tian, C.; Zheng, M.; Li, B.; Zhang, Y.; Zhang, S.; Zhang, D. Perceptive self-supervised learning network for noisy image watermark removal. *IEEE TCSVT* **2024**, *34*, 7069–7079.
16. Tian, C.; Song, M.; Zuo, W.; Du, B.; Zhang, Y.; Zhang, S. Application of convolutional neural networks in image super-resolution. *CAAI-TIS* **2025**.
17. Tian, C.; Xie, J.; Li, L.; Zuo, W.; Zhang, Y.; Zhang, D. A Perception CNN for Facial Expression Recognition. *IEEE T-IP*.
18. Lyu, Y.; Lin, W.A.; Liao, H.; Lu, J.; Zhou, S.K. Encoding metal mask projection for metal artifact reduction in computed tomography. In Proceedings of the MICCAI, 2020, pp. 147–157.

19. Li, Z.; Gao, Q.; Wu, Y.; Niu, C.; Zhang, J.; Wang, M.; Wang, G.; Shan, H. Quad-Net: Quad-domain network for CT metal artifact reduction. *IEEE T-MI* **2024**, *43*, 1866–1879.
20. Peng, C.; Qiu, B.; Li, M.; Yang, Y.; Zhang, C.; Gong, L.; Zheng, J. GPU-accelerated dynamic wavelet thresholding algorithm for X-ray CT metal artifact reduction. *IEEE T-RPMC* **2017**, *2*, 17–26.
21. Cui, Y.; Ren, W.; Cao, X.; Knoll, A. Image restoration via frequency selection. *IEEE T-PAMI* **2023**, *46*, 1093–1108.
22. Dai, T.; Wang, J.; Guo, H.; Li, J.; Wang, J.; Zhu, Z. FreqFormer: Frequency-aware transformer for lightweight image super-resolution. In Proceedings of the IJCAI, 2024, pp. 731–739.
23. Tian, C.; Zhang, X.; Zhang, Q.; Yang, M.; Ju, Z. Image super-resolution via dynamic network. *CAAI-TIS* **2024**, *9*, 837–849.
24. Suvorov, R.; Logacheva, E.; Mashikhin, A.; Remizova, A.; Ashukha, A.; Silvestrov, A.; Kong, N.; Goka, H.; Park, K.; Lempitsky, V. Resolution-Robust Large Mask Inpainting With Fourier Convolutions. In Proceedings of the IEEE/CVF WACV, January 2022, pp. 2149–2159.
25. Liu, T.; Li, B.; Du, X.; Jiang, B.; Geng, L.; Wang, F.; Zhao, Z. Fair: Frequency-aware image restoration for industrial visual anomaly detection. *arXiv preprint arXiv:2309.07068* **2023**.
26. Tian, C.; Liu, K.; Zhang, B.; Huang, Z.; Lin, C.W.; Zhang, D. A Dynamic Transformer Network for Vehicle Detection. *IEEE T-CE* **2025**.
27. Deeba, F.; Kun, S.; Ali Dharejo, F.; Zhou, Y. Wavelet-Based Enhanced Medical Image Super Resolution. *IEEE Access* **2020**, *8*, 37035–37044. <https://doi.org/10.1109/ACCESS.2020.2974278>.
28. Dharejo, F.A.; Zawish, M.; Deeba, F.; Zhou, Y.; Dev, K.; Khowaja, S.A.; Qureshi, N.M.F. Multimodal-boost: Multimodal medical image super-resolution using multi-attention network with wavelet transform. *IEEE ACM T COMPUT BI* **2022**, *20*, 2420–2433.
29. Jiang, X.; Zhang, X.; Gao, N.; Deng, Y. When fast fourier transform meets transformer for image restoration. In Proceedings of the ECCV. Springer, 2024, pp. 381–402.
30. Zhao, C.; Cai, W.; Dong, C.; Hu, C. Wavelet-based fourier information interaction with frequency diffusion adjustment for underwater image restoration. In Proceedings of the CVPR, 2024, pp. 8281–8291.
31. Ren, Y.; Li, X.; Li, B.; Wang, X.; Guo, M.; Zhao, S.; Zhang, L.; Chen, Z. MoE-DiffIR: Task-Customized Diffusion Priors for Universal Compressed Image Restoration. In Proceedings of the Proc. Eur. Conf. Comput. Vis. (ECCV); Leonardis, A.; Ricci, E.; Roth, S.; Russakovsky, O.; Sattler, T.; Varol, G., Eds., 2024, Vol. 15067, pp. 116–134.
32. Mandal, D.; Chattopadhyay, S.; Tong, G.; Chakravarthula, P. UniCoRN: Latent Diffusion-based Unified Controllable Image Restoration Network across Multiple Degradations. *arXiv preprint arXiv:2503.15868* **2025**.
33. Lin, J.; Zhang, Z.; Li, W.; Pei, R.; Xu, H.; Zhang, H.; Zuo, W. UniRestorer: Universal Image Restoration via Adaptively Estimating Image Degradation at Proper Granularity. *arXiv preprint arXiv:2412.20157* **2024**.
34. An, T.; Gao, H.; Liu, R.; Dai, K.; Xie, T.; Li, R.; Wang, K.; Zhao, L. An MoE-Driven Unified Image Restoration Framework for Adverse Weather Conditions. *IEEE T-CSVT* **2025**.
35. Yang, Z.; Chen, H.; Qian, Z.; Yi, Y.; Zhang, H.; Zhao, D.; Wei, B.; Xu, Y. All-in-one medical image restoration via task-adaptive routing. In Proceedings of the MICCAI. Springer, 2024, pp. 67–77.
36. Deng, Z.; Campbell, J. Sparse mixture-of-experts for non-uniform noise reduction in MRI images. In Proceedings of the WACV, 2025, pp. 297–305.
37. Wang, Z.; Ru, Y.; Chetouani, A.; Chen, F.; Bauer, F.; Zhang, L.; Hans, D.; Jennane, R.; Jarraya, M.; Chen, Y.H. MoEDiff-SR: Mixture of Experts-Guided Diffusion Model for Region-Adaptive MRI Super-Resolution. *arXiv preprint arXiv:2504.07308* **2025**.
38. Wang, Y.; Li, Y.; Zheng, Z.; Zhang, X.P.; Wei, M. M2Restore: Mixture-of-Experts-based Mamba-CNN Fusion Framework for All-in-One Image Restoration. *arXiv preprint arXiv:2506.07814* **2025**.
39. Zamfir, E.; Wu, Z.; Mehta, N.; Tan, Y.; Paudel, D.P.; Zhang, Y.; Timofte, R. Complexity experts are task-discriminative learners for any image restoration. In Proceedings of the Proceedings of CVPR, 2025, pp. 12753–12763.
40. Tian, C.; Song, M.; Fan, X.; Zheng, X.; Zhang, B.; Zhang, D. A Tree-guided CNN for image super-resolution. *IEEE T-CE* **2025**.
41. Yu, L.; Zhang, Z.; Li, X.; Xing, L. Deep Sinogram Completion with Image Prior for Metal Artifact Reduction in CT Images. *IEEE T-MI* **2020**, *40*, 228–238.
42. Wang, T.; Xia, W.; Huang, Y.; Sun, H.; Liu, Y.; Chen, H.; Zhou, J.; Zhang, Y. Dual-domain adaptive-scaling non-local network for CT metal artifact reduction. In Proceedings of the MICCAI. Springer, 2021, pp. 243–253.

43. Wang, H.; Li, Y.; He, N.; Ma, K.; Meng, D.; Zheng, Y. DICDNet: Deep Interpretable Convolutional Dictionary Network for Metal Artifact Reduction in CT Images. *IEEE T-MI* **2021**, *41*, 869–880.
44. Wang, H.; Xie, Q.; Li, Y.; Huang, Y.; Meng, D.; Zheng, Y. Orientation-Shared Convolution Representation for CT Metal Artifact Learning. In Proceedings of the MICCAI, 2022, pp. 665–675.
45. 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.
46. Liu, X.; Xie, Y.; Diao, S.; Tan, S.; Liang, X. Unsupervised CT metal artifact reduction by plugging diffusion priors in dual domains. *IEEE T-MI* **2024**, *43*, 3533–3545.
47. BaoShun, S.; ShaoLei, Z.; ZhaoRan, F. Artifact Region-Aware Transformer: Global Context Helps CT Metal Artifact Reduction. *IEEE SPL* **2024**, *31*, 1249–1253.
48. Liu, P.; Han, H.; Du, Y.; Zhu, H.; Li, Y.; Gu, F.; Xiao, H.; Li, J.; Zhao, C.; Xiao, L.; et al. Deep learning to segment pelvic bones: Large-scale CT datasets and baseline models. *IJ CARS* **2021**, *16*, 749–756.
49. Cheng, T.; Bi, T.; Ji, W.; Tian, C. Graph convolutional network for image restoration: A survey. *Mathematics* **2024**, *12*, 2020.

**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.