

FIND-Net –基于字典核的傅里叶集成网络用于金属伪影消除

Farid Tasharofi¹[0009–0006–1822–7889], Fuxin Fan¹, Melika Qahqaie¹,
Mareike Thies¹, and Andreas Maier¹

Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen, Germany
farid.tasharofi@fau.de

Abstract. 在计算机断层扫描 (CT) 成像中, 由于高密度金属植入物导致的金属伪影严重降低了图像质量, 复杂化了诊断和治疗计划。虽然现有的深度学习算法在金属伪影消除 (MAR) 上取得了显著成功, 但它们往往在抑制伪影的同时难以保留结构细节。为了应对这一挑战, 我们提出了 FIND-Net (Fourier-Integrated Network with Dictionary Kernels), 这是一个新颖的 MAR 框架, 集成了频域和空间域处理, 以实现卓越的伪影抑制和结构保留。FIND-Net 结合了快速傅里叶卷积 (FFC) 层和可训练的高斯滤波, 将 MAR 视为同时在空间和频域中运行的混合任务。该方法增强了全局背景理解和频率选择性, 有效地减少了伪影, 同时保持了解剖结构。在合成数据集上的实验表明, FIND-Net 在最新的 MAR 方法上取得了统计学意义上的显著改进, MAE 减少了 3.07, SSIM 增加了 0.18, PSNR 提高了 0.90, 证实了在不同伪影复杂度下的稳健性。此外, 在实际临床 CT 扫描评估中, 证实了 FIND-Net 能够在有效抑制金属引起的失真同时最小化对清洁解剖区域的修改。这些发现强调了 FIND-Net 在提升 MAR 性能方面的潜力, 提供了卓越的结构保留和改善的临床适用性。代码可在 [这个链接](#)¹ 获得。

Keywords: CT Metal Artifact Reduction · Frequency-Domain Deep Learning · Fast Fourier Convolution

1 引言

计算机断层扫描 (CT) 在非侵入性诊断中起着至关重要的作用; 然而, 金属植入物会引入严重的伪影, 例如由束硬化、光子饥饿和散射引起的条纹和星形畸变 [1,2,3]。这些伪影模糊了解剖结构并降低了准确性。

传统的金属伪影消除 (MAR) 方法, 如线性插值 (LI), 在计算上是高效的, 但常常引入次级伪影, 破坏结构的真实度 [3,4]。基于深度学习的 MAR 方法, 特别是卷积神经网络 (CNNs), 已显示出优异的伪影抑制效果 [2,5,6,7]。然而, 其有限的感受野阻碍了有效处理大规模、全局分布变形的能力。在图像和正弦图域中有效捕捉长距离依赖关系的能力不足仍然是 CNN 方法的一个关键限制 [8]。

卷积神经网络 (CNNs) 主要依赖于小卷积核 (例如, 3×3 在 ResNet 中), 它们强调局部特征, 从而使全局伪影抑制具有挑战性 [8,9]。虽然正弦图域处理

¹ 代码可在 <https://github.com/Farid-Tasharofi/FIND-Net> 获得

旨在遵循成像系统约束，但常常引入次级失真或局部过度平滑。相反，从受金属污染的正弦图重建图像会将金属引起的伪影传播到整幅图像，导致全局分布的失真 [6]。频域技术为全局特征提取提供了一条有前途的途径 [9]；然而，它们与深度学习的结合仍然有限，因为大多数 MAR 方法主要在空间域中操作，并在处理长程依赖时表现出困难。

为了解决这些挑战，我们提出了 FIND-Net（傅里叶集成网络与字典核），这是一种新颖的 MAR 模型，它将快速傅里叶卷积（FFC）与可训练的高斯滤波相结合，以改善伪影抑制。通过结合空间和频域处理，FIND-Net 在捕获长程依赖关系的同时保留结构细节，减少金属引起的失真而不会过度平滑解剖结构。

2 方法论

问题表述 FIND-Net 扩展了来自 DICDNet [7] 的卷积字典模型，这一模型将被金属伪影污染的 CT 图像 $\mathbf{Y}$ 分解为无伪影的成分 $\mathbf{X}$ 和伪影项 $\mathbf{A}$ ：

$$\mathbf{I} \odot \mathbf{Y} = \mathbf{I} \odot \mathbf{X} + \mathbf{I} \odot \mathbf{A}, \quad (1)$$

，其中 $\mathbf{I}$ 是一个非金属掩膜，而 $\odot$ 表示点乘。伪影项被建模为：

$$\mathbf{A} = \sum_{n=1}^N \mathbf{K}_n \otimes \mathbf{M}_n. \quad (2)$$

这里， $\mathcal{K} = [\mathbf{K}_1, \dots, \mathbf{K}_N]$ 是捕捉金属伪影模式的字典核， $\mathcal{M} = [\mathbf{M}_1, \dots, \mathbf{M}_N]$ 是相应的特征图。运算符 $\otimes$ 表示二维卷积。这种分解使用局部化模式和空间先验对金属伪影结构进行编码，有助于在 MAR [7] 中的可解释性。然而，同时估计 $\mathbf{M}$ 和 $\mathbf{X}$ 本质上是病态的，需要一个优化框架来确保稳定性和有效的 MAR。

结合方程 1 和 2，我们可以将 MAR 重新表述为需要约束条件以实现稳定优化的逆问题。然而，由于非唯一性、不稳定性和欠定性，该问题是不适定的，因此需要额外的约束来获得有意义的解。为了解决这个问题，DICDNet [7] 引入了一个正则化优化框架：

$$\min_{\mathcal{M}, \mathbf{X}} \|\mathbf{I} \odot (\mathbf{Y} - \mathbf{X} - \mathbf{A})\|_F^2 + \lambda_1 P_1(\mathcal{M}) + \lambda_2 P_2(\mathbf{X}), \quad (3)$$

，其中 $P_1(\mathcal{M})$ 和 $P_2(\mathbf{X})$ 是先验项，用于约束 $\mathcal{M}$ 和 $\mathbf{X}$ ，以正则化解并解决不适定性。参数 λ_1 和 λ_2 是加权因子，控制正则化项对优化过程的影响。迭代更新过程被表述为：

$$\mathcal{M}^{(s)} = \text{prox}_{\lambda_1 \eta_1} \left(\mathcal{M}^{(s-1)} - \eta_1 \mathcal{K} \otimes^T \left(\mathbf{I} \odot \left(\mathcal{A}^{(s-1)} + \mathbf{X}^{(s-1)} - \mathbf{Y} \right) \right) \right), \quad (4)$$

在 $\mathcal{A}^{(s)} = \mathcal{K} \otimes \mathcal{M}^{(s)}$ 的情况下，

$$\mathbf{X}^{(s)} = \text{prox}_{\lambda_2 \eta_2} \left((1 - \eta_2 \mathbf{I}) \odot \mathbf{X}^{(s-1)} + \eta_2 \mathbf{I} \odot (\mathbf{Y} - \mathcal{A}^{(s)}) \right), \quad (5)$$

。近端算子保证了稳定性，同时施加约束。与 DICDNet [7] 中的方法类似， $\mathbf{X}^{(0)}$ 使用 LI 初始化，以提供迭代优化的合理起始点。DICDNet [7] 利用 ResNet [10]

作为近端算子来迭代地优化估计的金属伪影成分 ($\mathcal{M}$) 和重建图像 ($\mathbf{X}$)。具体来说, ResNet 用于模拟近端映射, 它有助于将估计的变量投射到一个更加结构化的解空间, 从而减轻噪声和伪影的影响。CNN 基于近端网络在保持结构一致性方面的有效性已经被 Fu 等人 [11] 进一步证明, 强调了其在施加领域特定先验知识和增强重建精度方面的能力。这种选择对于 MAR 任务特别有利, 因为基于 CNN 的架构利用空间相关性和分层特征提取来改善伪影抑制, 同时保持图像细节。

FIND-Net 通过在迭代框架中整合基于频率的处理扩展了这种方法。FIND-Net 并不仅依赖空间域学习, 而是结合了频域先验来完善伪影估计并增强图像重建。这种混合方法使 FIND-Net 能够捕捉到局部空间模式和全局频率特征。

网络架构 如图 1² 所示, FIND-Net 迭代运行: M-Net 使用频率增强特征来估计金属伪影, 而 X-Net 通过在空间和频率域中滤除残留伪影来逐步改进重建。两个网络都利用 ResNets 进行特征提取。为了增强 MAR 超越局部特征提取, FIND-Net 在 proxNet 的 ResBlocks 中用快速傅里叶卷积 (FFC) [9] 替代标准卷积, 实现空间和频率域的联合处理。FFC 将输入通道拆分为用于空间卷积的局部分支和应用基于 FFT 的光谱转换的全局分支, 跨分支交互改善特征传递。全局分支采用傅里叶单元 (FU) 和局部傅里叶单元 (LFU) 以捕获远程依赖关系, 同时保留局部纹理。

为了提高频率适应性, Global Branch 中的可训练高斯滤波在训练过程中动态调整均值和方差, 有选择地增强重要的频率, 同时抑制无关的频率。与 FFC 的统一频率映射 [9] 不同, 这种自适应机制在 FU 和 LFU 中的 1×1 卷积之前应用高斯滤波, 确保稳健的抑制同时保留解剖细节。标准的 FU 变换为:

$$X^{FU_{out}} = \mathcal{F}^{-1}(\text{ReLU}(\text{BN}(\text{Conv}_{1 \times 1}(\mathcal{F}(X)))))$$

校正发现网络通过可训练高斯滤波扩展了这一点:

$$X^{FU_{out}} = \mathcal{F}^{-1}(\text{ReLU}(\text{BN}(\text{Conv}_{1 \times 1}(G_{\sigma,c}(u,v) \cdot \mathcal{F}(X)))))$$

其中 $G_{\sigma,c}(u,v) = \exp\left(-\left(\frac{D^2 - c^2}{D \cdot \sigma + \epsilon}\right)^2\right)$, D 为归一化的频率距离。可学习的 σ 和 c 动态控制滤波带宽和中心频率, 从而能够灵活适应变化的伪影模式。

在各个阶段中, 对全局分支的通道分配逐渐增加, 以平衡空间和光谱学习, 从初始阶段的 $\alpha_{in}^{(s)} = \alpha_{out}^{(s)} = 0.0$ 开始, 到后期阶段达到 0.8。早期阶段强调空间处理, 中期阶段结合频率成分, 后期阶段则专注于全局特征提取, 以实现最佳的伪影抑制。

训练损失: 训练在每个阶段 s 中对复原的 CT 图像 $\mathbf{X}^{(s)}$ 和提取的伪影 $\mathbf{A}^{(s)}$ 采用多项损失:

$$\begin{aligned} \mathcal{L} = & \sum_{s=0}^S \omega_s \mathbf{I} \odot \left\| \mathbf{X} - \mathbf{X}^{(s)} \right\|_F^2 + \gamma_1 \sum_{s=0}^S \omega_s \mathbf{I} \odot \left\| \mathbf{X} - \mathbf{X}^{(s)} \right\|_1 \\ & + \gamma_2 \sum_{s=1}^S \omega_s \mathbf{I} \odot \left\| \mathbf{Y} - \mathbf{X} - \mathbf{A}^{(s)} \right\|_1, \end{aligned} \quad (6)$$

² 代码可以在 <https://github.com/Farid-Tasharofi/FIND-Net> 获得

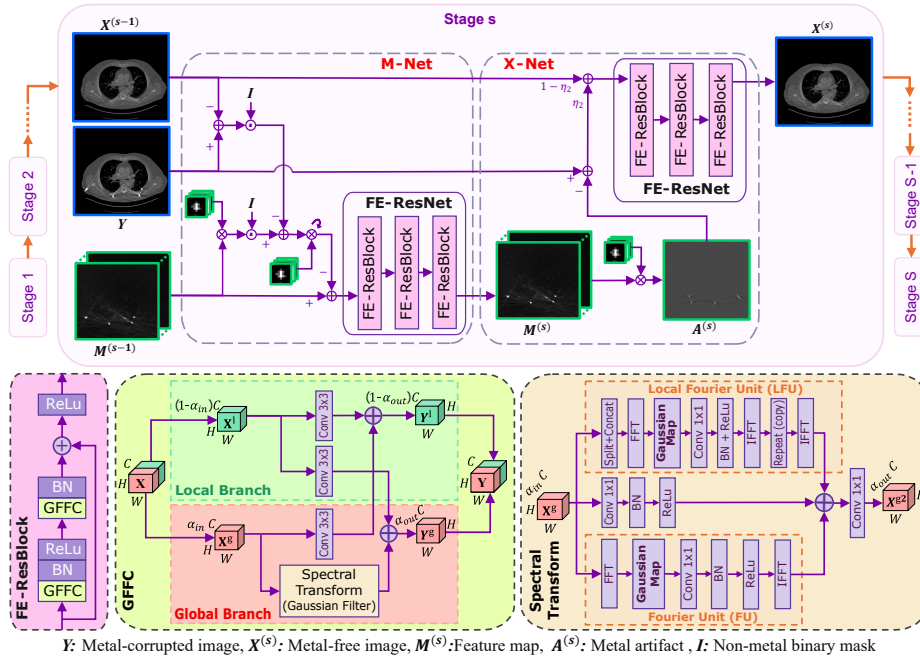


Fig. 1. FIND-Net 在迭代阶段中运行: M-Net 估计金属伪影, 而 X-Net 精炼重建。每个阶段使用频率增强 ResNet (FE-ResNet) 和频率增强 Res 块 (FE-ResBlocks), 集成了 GFFC, 一种带有可训练高斯滤波的改进快速傅里叶卷积 (FFC)。GFFC 使用 α 将通道分为局部分支和全局分支。频谱变换模块 (LFU, FU) 应用傅里叶变换 (FFT), 高斯滤波和反傅里叶变换 (IFFT) 进行频率增强。符号: $\otimes$ (卷积), $\otimes^T$ (转置卷积), $\oplus$ (逐元素求和), $\odot$ (逐点乘法), BN (批量归一化)。

, 其中 I 是去金属掩膜, ω_s 表示阶段权重。第一项通过惩罚 Frobenius 范数来最小化重建误差, 而第二和第三项基于 L_1 范数, 鼓励稀疏性及在无伪影和重建图像之间加强一致性 [7]。

我们利用 AAPM CT-MAR Grand Challenge 数据集 [12,13], 这些数据集使用 XCIST [14], 通过 NIH DeepLesion (肺, 腹部, 肝脏, 骨盆) [15] 以及 UCLH Stroke EIT (头部) [16] 生成, 并结合了合成金属物体。该数据集包括金属污染和无伪影的 CT 图像, 相应的正弦图和金属掩膜, 每个的形状为 512×512 像素。用于训练的共有 5500 个案例 (1300 个头部, 4200 个身体), 验证和测试各有 700 个案例 (200 个头部, 500 个身体)。通过量子噪声, 束硬化和散射辐射引入了逼真的金属伪影。最终评估集包含 29 个案例, 包括正弦图和金属污染的重建图像, 但无真实值, 反映了现实临床的约束。该工作使用了 AAPM CT-MAR Grand Challenge 数据集 [12,13]。AAPM CT-MAR Grand Challenge 数据集使用开源的 CT 仿真环境 XCIST [14] 生成, 采用混合数据仿真框架结合了公开可用的临床图像 [15,16] 和虚拟金属物体。

这些模型使用 PyTorch 实现, 并在 NVIDIA A100 GPU (80 GB HBM2) 上训练, 训练周期最多达 200 个 epoch, 批次大小为 6。基于验证损失应用了早停。所有模型都在相同的数据集上训练。DICDNet 和 FIND-Net 遵循 DICDNet 的

配置 [7]，包含 10 个阶段，AdamW ($\beta_1 = 0.9, \beta_2 = 0.999$)，学习率为 1×10^{-4} ，权重衰减为 1×10^{-5} 。学习率的适应使用了余弦退火的线性热启动方法。

我们将 FIND-Net 与 DICDNet、OSCNet 和 OSCNet+ 进行比较。OSCNet 向 DICDNet 添加了旋转感知卷积，而 OSCNet+ 为不同的伪影模式引入了动态卷积。表展示了所有测试用例的综合测试结果，表明 FIND-Net 在平均绝对误差 (MAE)、结构相似性指数 (SSIM) 和峰值信噪比 (PSNR) 上优于所有方法。所有深度学习模型的性能均优于传统的 LI 方法。OSCNet 的变体表现不如 DICDNet，表明其旋转对称性编码的泛化能力有限。

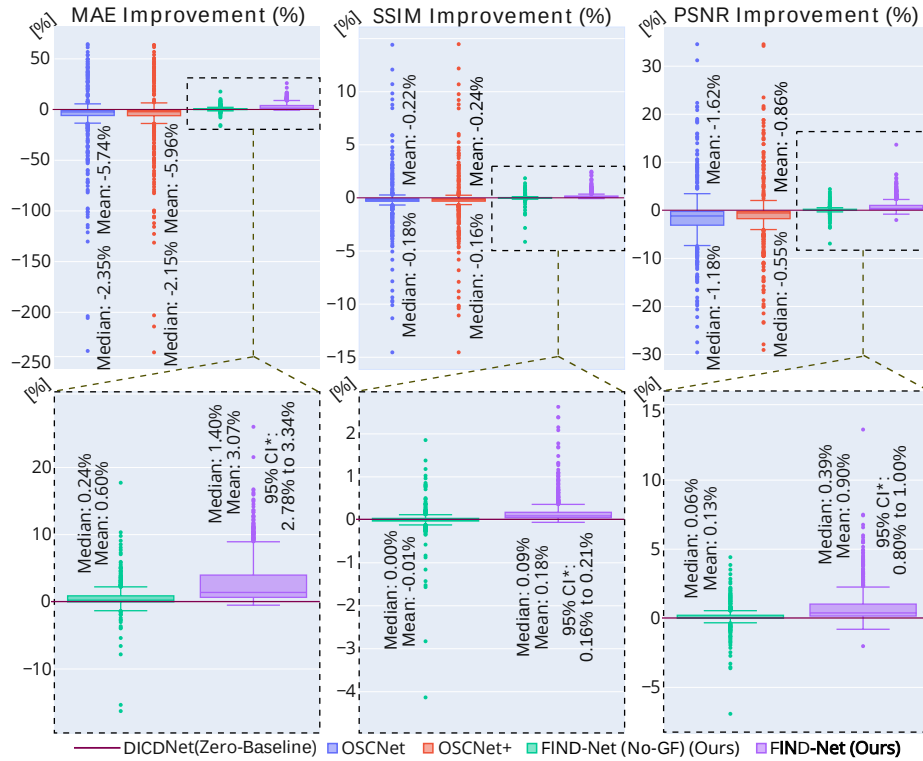


Fig. 2. 箱线图比较了各 MAR 模型相对于 DICDNet 的 MAE、SSIM 和 PSNR 改进 (%)。总体分布显示，并提供了放大视图以突出显示 FIND-Net 和未经过高斯滤波的 FIND-Net (FIND-Net (No-GF))。标注了均值、中位数和 95 % 置信区间 (*)。零基线对应于 DICDNet，负值表示相对于其表现更差。

我们进一步从与基线模型 DICDNet 的百分比性能提升来评估 FIND-Net 的表现，以及与 OSCNet、OSCNet+ 和 FIND-Net (No-GF) (不使用高斯滤波的 FIND-Net) 的比较。如图 2 所示，FIND-Net 始终实现性能提升，与 DICDNet 相比，MAE 减少 3.07%，SSIM 增加 0.18%，PSNR 提高 0.90%。标注的 95 % 置信区间确认了这些改进的统计显著性。箱线图进一步突出了 FIND-Net 的稳健性能，放大视图强调了高斯滤波在增强频率域学习中的贡献。特别是 MAE

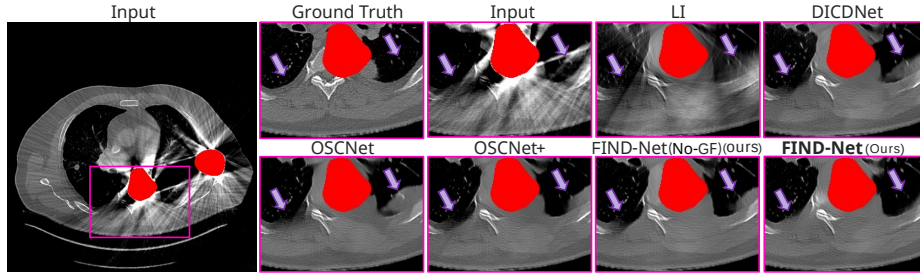


Fig. 3. 关于金属污染的合成 CT 图像的 MAR 结果的定性比较。突出显示的 ROI 展示了 FIND-Net 在伪影抑制方面的优越性，特别是在细条纹去除和边缘保留方面，相较于其他方法。红色遮罩突出显示了金属遮罩。(No-GF 表示没有高斯滤波)

的减少表明了更优秀的伪影抑制，进一步证实了 FIND-Net 在减少金属伪影方面的有效性。

计算复杂性： FIND-Net 通过利用基于 FFT 的优化，将计算复杂度减少了大约 16 % (334.45 GFLOPs 279.55 GFLOPs)。然而，由于频域转换的开销，推理时间从每张图像 0.15 秒增加到 1.07 秒。

定性结果 图 3 比较了 FIND-Net 与其他模型在一个合成测试案例中的表现。虽然 LI 能部分抑制条纹，DICDNet 在文饰还原方面有改进，特别是在右箭头区域，优于 OSCNet 的变体。FIND-Net 进一步优化了重建效果，因为高斯滤波消除了其他模型中持续存在的低对比度伪影。左箭头区域突出了 FIND-Net 增强边缘清晰度与减少过度平滑的能力，保留了解剖边界。这确保了在高对比度和低对比度区域中优越的保留和伪影抑制。

图 4 显示了我们 29 张图像数据集中一张真实世界 CT 扫描的结果，该数据集没有无伪影的参考。尽管 FIND-Net 和 DICDNet 减少了伪影，DICDNet 和 OSCNet 的变体在非损坏区域引入了过度的平滑。蓝色区域的分析证实了 FIND-Net 的卓越保持能力，达到了最低的 MAE 和最高的 SSIM 和 PSNR。

为了评估对干净解剖区域的影响，我们分析了来自七幅图像的无伪影补丁。一个理想的 MAR 方法应该让这些补丁保持不变。OSCNet、OSCNet+ 和 DICDNet 表现相似 ($MAE \approx 0.007$, $SSIM \approx 0.94$, $PSNR \approx 40.4$)，而没有高斯滤波的 FIND-Net 更好地保留了结构 (MAE 0.006, $SSIM$ 0.96, $PSNR$ 41.81)。FIND-Net 表现优于所有模型，取得了最佳的分数 (MAE 0.005, $SSIM$ 0.97, $PSNR$ 43.23)，并展示了以最小的非预期修改实现稳健抑制的能力。

3 结论

本文介绍了 FIND-Net，这是一种基于深度学习的 MAR 框架，结合了 FFC 和可训练的高斯滤波来增强伪影抑制。通过结合空间和频域处理，FIND-Net 优于现有的 MAR 方法，相较于 DICDNet，FIND-Net 实现了统计上显著的 3.07 % MAE 降低、0.18 % SSIM 提升和 0.90 % PSNR 提升。综合实验确认了

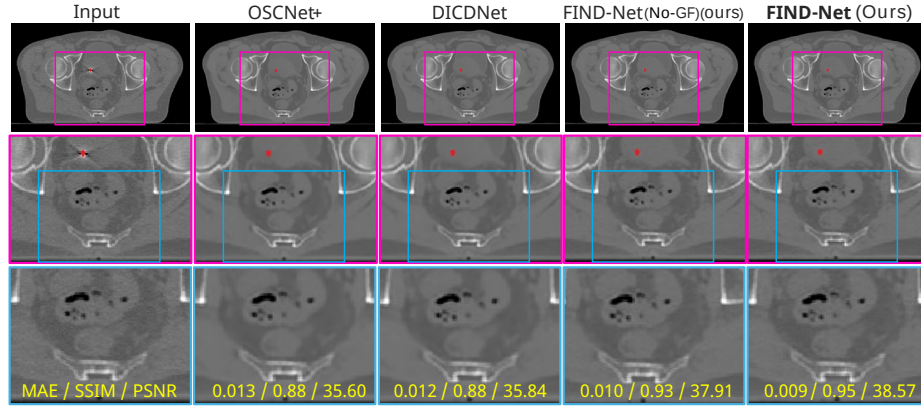


Fig. 4. 对真实世界 CT 扫描中的 MAR 方法进行定性比较（没有可用的真实值）。红色蒙版突出显示了分割后的金属。蓝色区域（输入中无伪影）用于评估结构保留。

FIND-Net 在处理不同复杂程度的伪影时具有稳健性，在抑制条纹伪影的同时保留解剖结构。尽管由于频域操作增加了推理时间（每张图像 1.07 秒），FIND-Net 降低了计算复杂度，并为 MAR 性能设立了新的基准。

Acknowledgments. The authors gratefully acknowledge the scientific support and HPC resources provided by the Erlangen National High Performance Computing Center of the Friedrich-Alexander-Universität Erlangen-Nürnberg. The hardware is funded by the German Research Foundation. This research has received funding from the scholarship program “Dhip campus-bavarian aim” .

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Gjesteb, L., De Man, B., Jin, Y., Paganetti, H., Verburg, J., Giantsoudi, D., Wang, G.: Metal artifact reduction in CT: where are we after four decades? IEEE Access 4, 5826–5849 (2016). <https://doi.org/10.1109/ACCESS.2016.2608621>
2. Wang, H., Xie, Q., Zeng, D., Ma, J., Meng, D., Zheng, Y.: OSCNet: Orientation-Shared Convolutional Network for CT Metal Artifact Learning. IEEE Trans. Med. Imaging 43 (1), 489–502 (2024). <https://doi.org/10.1109/TMI.2023.3310987>
3. Anhaus, J., Hofmann, C.: New Approaches for Metal Artifact Reduction in Computed Tomography. In: Iniewski, K., Gadey, H. (eds.) Emerging Radiation Detection: Technology and Applications, pp. 131–146. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-63897-8_8
4. Kalender, W.A., Hebel, R., Ebersberger, J.: Reduction of CT artifacts caused by metallic implants. Radiology 164 (2), 576–577 (1987). <https://doi.org/10.1148/radiology.164.2.3602406>

5. Lyu, Y., Lin, W.A., Liao, H., Lu, J., Zhou, S.K.: Encoding Metal Mask Projection for Metal Artifact Reduction in Computed Tomography. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12262, pp. 147–157. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59713-9_15
6. Yu, L., Zhang, Z., Li, X., Xing, L.: Deep sinogram completion with image prior for metal artifact reduction in CT images. *IEEE Trans. Med. Imaging* 40 (1), 228–238 (2021). <https://doi.org/10.1109/TMI.2020.3025064>
7. 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 Trans. Med. Imaging* 41 (4), 869–880 (2022). <https://doi.org/10.1109/TMI.2021.3127074>
8. Richter, M.L., Schöning, J., Wiedenroth, A., Krumnack, U.: Should you go deeper? Optimizing convolutional neural network architectures without training. In: 20th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 964–971. IEEE, Pasadena, CA, USA (2021). <https://doi.org/10.1109/ICMLA52953.2021.00159>
9. Chi, L., Jiang, B., Mu, Y.: Fast Fourier Convolution. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M.F., Lin, H. (eds.) *Advances in Neural Information Processing Systems* 33 (NeurIPS 2020), pp. 4479–4488. Curran Associates, Inc. (2020).
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778 (2016). <https://doi.org/10.48550/arXiv.1512.03385>
11. Fu, J., Xie, Q., Meng, D., Xu, Z.: Rotation equivariant proximal operator for deep unfolding methods in image restoration. *IEEE Trans. Pattern Anal. Mach. Intell.* 46 (10), 6577–6593 (2024). <https://doi.org/10.1109/TPAMI.2024.3383532>
12. AAPM CT Metal Artifact Reduction (CT-MAR) Grand Challenge. <https://www.aapm.org/GrandChallenge/CT-MAR/>, last accessed 2025/02/01.
13. AAPM CT Metal Artifact Reduction (CT-MAR) Grand Challenge Benchmark Tool, https://github.com/xcist/example/tree/main/AAPM_datachallenge, last accessed 2025/02/01
14. Wu, M., FitzGerald, P., Zhang, J., Segars, W.P., Yu, J., Xu, Y., De Man, B.: XCIST - an open access x-ray/CT simulation toolkit. *Phys Med Biol.* 67 (19) (2022). <https://doi.org/10.1088/1361-6560/ac8c85>
15. Yan, K., Wang, X., Lu, L., Summers, R.M.: DeepLesion: Automated Mining of Large-Scale Lesion Annotations and Universal Lesion Detection with Deep Learning. *J Med Imaging.* 5 (3), 036501 (2018). <https://doi.org/10.1117/1.JMI.5.3.036501>
16. Goren, N., Dowrick, T., Avery, J., Holder, D.: UCLH Stroke EIT Dataset - Radiology Data (CT). *Zenodo* (2017). <https://doi.org/10.5281/zenodo.838704>