

基于双流 Transformer 结构的多能计算机断层扫描成像投影数据去噪方法

欧阳顺馨¹, 史再峰^{1,3*}, 孔凡宁¹, 张丽丽¹, 曹清洁²

¹天津大学微电子学院, 天津 300072;

²天津师范大学数学科学学院, 天津 300387;

³天津市成像与感知微电子技术重点实验室, 天津 300072

摘要 多能计算机断层扫描(Computed tomography, CT)技术可以更加精确地分辨出人体组织对不同能量 X 射线光子的吸收情况, 是医学成像领域的重要发展方向。为了解决因量子噪声等非理想效应加重导致重建图像质量急剧退化的问题, 提出了一种基于移位窗口多头自注意力机制的双流 Transformer 网络结构。该结构利用移位窗口 Transformer 和局部增强窗口 Transformer 分别提取投影数据的全局和局部特征, 充分利用投影数据的非局部自相似性以保留投影数据的内部结构; 然后通过残差卷积融合提取的特征; 最后使用带有非局部全变分的混合损失函数来监督网络模型的训练, 提升该网络模型对投影数据内部细节的敏感程度。实验结果表明, 所提方法处理后的重建图像峰值信噪比(PSNR)值、结构相似性(SSIM)值和特征相似度(FSIM)值分别达到 37.7301 dB、0.9944 和 0.9961。与目前先进的多能 CT 去噪方法相比, 所提方法在去除低剂量多能 CT 投影数据噪声的同时, 可保留更多的细节特征, 有利于后续的精确诊断。

关键词 图像处理; 计算机断层扫描成像; 低剂量; 卷积神经网络; Transformer 网络

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP231426

Projection Domain Denoising Method for Multi-Energy Computed Tomography via Dual-Stream Transformer

Ouyang Shunxin¹, Shi Zaifeng^{1,3*}, Kong Fanning¹, Zhang Lili¹, Cao Qingjie²

¹School of Microelectronics, Tianjin University, Tianjin 300072, China;

²College of Mathematical Science, Tianjin Normal University, Tianjin 300387, China;

³Tianjin Key Laboratory of Imaging and Sensing Microelectronics Technology, Tianjin 300072, China

Abstract The multi-energy computed tomography (CT) technique can resolve the absorption rates of various energy X-ray photons in human tissues, representing a significant advancement in medical imaging. By addressing the challenge of swift degradation in reconstructed image quality, primarily due to non-ideal effects such as quantum noise, a dual-stream Transformer network structure is introduced. This structure utilises the shifted-window multi-head self-attention denoising approach for projection data. The shifted windows Transformer extracts the global features of the projection data, while the locally-enhanced window Transformer focuses on local features. This dual approach capitalizes on the non-local self-similarity of the projection data to maintain its inherent structure, subsequently merged by residual convolution. For model training oversight, a hybrid loss function incorporating non-local total variation is employed, which enhances the network model's sensitivity to the inner details of the projected data. Experimental results demonstrate that our method's processed projection data achieve a peak signal to noise ratio (PSNR) of 37.7301 dB, structure similarity index measurement (SSIM) of 0.9944, and feature similarity index measurement (FSIM) of 0.9961. Relative to leading denoising techniques, the proposed method excels in noise reduction while preserving more inner features, crucial for subsequent accurate diagnostics.

Key words image processing; computed tomography; low dose; convolutional neural network; Transformer network

收稿日期: 2023-06-01; 修回日期: 2023-07-27; 录用日期: 2023-08-08; 网络首发日期: 2023-08-21

基金项目: 国家自然科学基金(62071326)

通信作者: *shizaifeng@tju.edu.cn

1 引言

多能计算机断层扫描(Computed tomography, CT)技术可以分辨不同能量X射线光子在人体中的衰减情况,具有能谱分辨率和辐射剂量效率较高的特点,目前已成为医学成像领域的主要发展方向之一。随着辐射剂量的降低,量子噪声、光子饥饿等非理想效应对成像质量的影响越来越严重。因此,如何去除低剂量多能CT成像结果的噪声从而提升重建图像的信噪比具有重要的研究意义。

目前低剂量CT图像去噪算法主要分为基于投影数据的预处理方法^[1-2]、迭代重建方法^[3-4]和基于重建图像的后处理方法^[5-6]等。其中,迭代重建方法虽然能直接实现从投影数据到CT图像的高质量重建,但该方法强烈依赖先验信息且会消耗大量的计算资源。而基于重建图像的后处理方法通常借助字典学习、非局部滤波等方法处理已重建的低剂量CT图像,但在保留细节结构和特征方面依然存在不足。与这两类方法相比,基于投影数据的处理方法可以直接利用投影数据的噪声来构建噪声分布模型,这使得该方法具有更高的计算效率。

近年来,基于数据驱动的深度学习技术发展迅速并广泛应用到图像处理领域中,同时在物质分解^[7]、CT重建^[8]等医学影像处理领域也取得了很多研究成果。在CT图像去噪方面,Trung等^[9]提出了基于空洞残差卷积神经网络(DRN)的低剂量CT图像去噪方法。该方法结合下采样和空洞卷积,在不增加卷积层数的情况下即可实现噪声抑制,有效提升重建图像质量。Jiao等^[10]提出了一种基于卷积神经网络(CNN)的CT重建框架,即智能反向投影网络(iBP-Net)。该框架将多个CNN子模块融合成端到端网络,在投影层输出中恢复重建图像,在降低网络训练难度的同时取得了良好的重建结果。基于CNN的CT重建框架常通过卷积块的叠加来提取最深层次的视觉特征,但未能挖掘出不同层次特征信息之间的内在联系。而自注意力机制结构通过将图像非局部自相似性和CNN相结合可以有效解决上述问题,因此以自注意力机制为基础的Transformer结构被广泛应用于图像分割^[11]、图像识别^[12]等领域。Li等^[13]提出了一种基于Transformer的端到端低剂量CT去噪网络并取得了优异的性能,证明了Transformer结构在医学影像去噪方面的潜力。Yang等^[14]基于低剂量CT正弦图的统计特征提出了采用Transformer结构的投影域去噪网络,可以在去除正弦图噪声的同时保留其内部结构,显著提高图像质量,该方法证明基于Transformer结构的网络模型有助于保留投影数据的内部结构,在提升重建图像质量方面更具优势。然

而,Transformer结构在计算图像自注意力权重时需要进行大量的乘加运算,导致这类网络模型有非常大的算力需求,不利于推广和实际应用。

为了解决上述问题,本文提出了一种双流Transformer增强型投影域去噪网络(Dual-stream transformer enhanced projection denoising network, DTE-PDN)。其中,双流Transformer结构基于滑窗运算多头自注意力机制来实现,可以充分利用多能CT投影数据自相似性并提升模型的运算效率,两种Transformer模块可增强网络模型的局部和全局表征能力;残差卷积结构用于改善模型梯度,并减少由网络层数增加引起的浅层信息丢失。最终所提模型能实现对投影数据噪声的精确去除,有效提高重建图像信噪比。

2 多能CT投影数据噪声去除方法

2.1 DTE-PDN模型架构

与常规图像相比,多能CT各能量区间的投影数据具有相关性,因此在通过神经网络处理投影数据时,所提方法将各能量区间信息堆叠为一个高维张量作为网络模型的输入数据,依靠网络模型可以提取并融合各通道数据特征的优势,以更好地去除噪声。所提网络模型整体结构如图1所示,该网络模型主要由 m 个移位和局部增强Transformer窗口模块(Shifted-windows and local-enhanced-window transformer block, SLB)和 n 个残差卷积模块(Residual convolution block, RCB)构成。其中,每个RCB由2组 3×3 卷积层、批归一化层和Leaky ReLU激活层组成。SLB和RCB均采用残差结构,残差结构有助于避免网络结构加深导致的细节结构丢失问题,并有利于改善网络模型的梯度。

在每个SLB中,首先通过卷积层来提取输入数据特征,然后经拆分级联操作将数据特征拆分,并将特征信息分别输入至移位窗口Transformer单元(STU)和局部增强窗口Transformer单元(LTU)。该操作通过使用 1×1 卷积层将第 i ($1 < i \leq m$)级SLB的特征信息 \mathbf{X}_i 平均拆分为两个特征图组 \mathbf{X}_{i1} 和 \mathbf{X}_{i2} ,并分别经STU和LTU处理得到输出特征 $\mathbf{X}_{s,i}$ 和 $\mathbf{X}_{l,i}$,该过程为

$$(\mathbf{X}_{i1}, \mathbf{X}_{i2}) = \text{Split}[\text{Conv}(\mathbf{X}_i)], \quad (1)$$

$$\mathbf{X}_{s,i} = \text{Swin}(\mathbf{X}_{i1}), \quad \mathbf{X}_{l,i} = \text{Lewin}(\mathbf{X}_{i2}), \quad (2)$$

式中:Split(\cdot)为随机平均拆分操作;Conv(\cdot)为卷积操作;Swin(\cdot)为移位窗口Transformer运算;Lewin(\cdot)为局部增强窗口Transformer运算。将两个部分的输出特征 $\mathbf{X}_{s,i}$ 和 $\mathbf{X}_{l,i}$ 通过连接层合并,并通过卷积层进行特征融合后,与本级SLB输入数据加和后输出至下一模块。因此,第 i 级SLB的最终输出 \mathbf{Z}_i 可表示为

$$\mathbf{Z}_i = \text{Conv}[\text{Concat}(\mathbf{X}_{s,i}, \mathbf{X}_{l,i})] + \mathbf{X}_i, \quad (3)$$

式中:Concat(\cdot)为合并两个或多个特征图的函数。

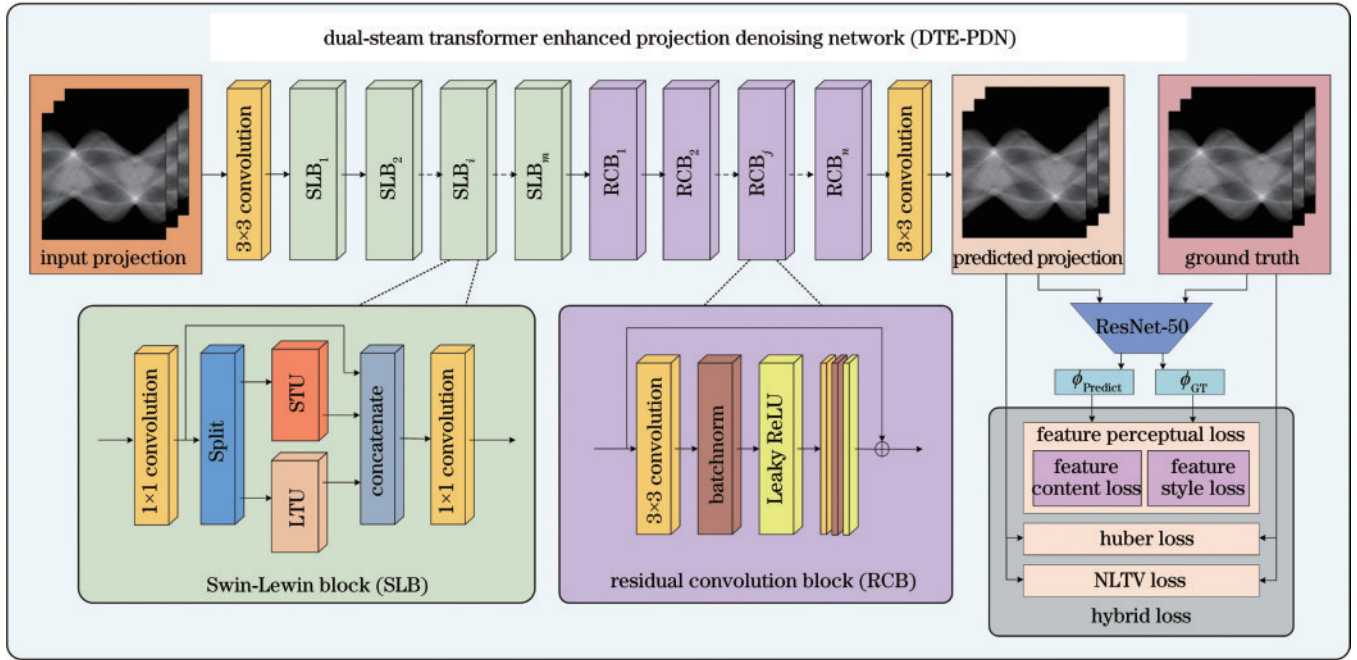


图 1 基于 DTE-PDN 的低剂量多能 CT 去噪方法结构图

Fig. 1 Structure diagram of the proposed DTE-PDN-based low-dose multi-energy CT denoising method

2.2 双流 Transformer 结构

相比于计算机视觉领域中的常规图像,投影数据具有特殊的内部结构,在去除噪声的同时尽可能保留这些内部结构,有助于提升重建图像质量^[14]。Transformer 结构利用图像的自相似性,在保留图像细节结构方面具有良好的效果,因此可用于保留多能 CT 投影域内部结构的细节信息。然而,Transformer 结构在计算量方面的劣势导致基于该结构的网络模型实用性受限。为解决上述问题,基于滑窗多头自注意力机制的移位窗口 Transformer 结构^[15]被提出,该结构可显著降低网络模型的运算量以提高数据处理效率,并利用自注意力机制原理来保证网络模型的非局部建模能力。因此,所提方法采用该结构来充分利用投影数据的非局部自相似性以更好保留其内部细节,具体结构如图 2 所示。

该部分包含一个基于移位窗口的多头自注意力 (Shifted-window-based multi-head self-attention, SW-MSA) 单元、两个 LayerNorm 层和一个多层感知机 (Multi-layer perception, MLP) 层。输入特征向量经 LayerNorm 层送入 SW-MSA 单元,由 SW-MSA 单元处理后的特征向量和输入进行残差连接以减少信息丢失。SW-MSA 单元将输入特征分割成不同的窗口分区,并在计算过程中使窗口分区产生位移,然后在下一层中的移位窗口位置产生新的窗口分区,新窗口中的自注意力计算能跨越前一层中的窗口边界,实现和前一层窗口的连接。每个 SW-MSA 单元和 MLP 层之前的模块使用 LayerNorm 层进行归一化。假设 $M \times M$ 为非重叠窗口大小,与普通的多头自注意力模块相比,

SW-MSA 单元的复杂度由 $O(H^2W^2C)$ 减小到 $O(M^2HWC)$,其中, H 、 W 、 C 分别为特征图的长度、宽度和通道数。SW-MSA 单元采用多头自注意力机制结构,可在若干特征空间中并行提取多组特征,每个特征空间的自注意力计算过程为

$$\text{Attention}(\mathbf{Q}_\tau, \mathbf{K}_\tau, \mathbf{V}_\tau) = \text{Softmax}\left(\frac{\mathbf{Q}_\tau \cdot \mathbf{K}_\tau^T}{\sqrt{d}} + \mathbf{B}_\tau\right) \cdot \mathbf{V}_\tau, \quad (4)$$

式中: \mathbf{Q}_τ 、 \mathbf{K}_τ 和 \mathbf{V}_τ 分别为特征空间 τ 的查询向量、键向量和值向量; d 为查询向量 \mathbf{Q}_τ 和键向量 \mathbf{K}_τ 的维度; \mathbf{B}_τ 为相对位置偏差。

移位窗口 Transformer 结构虽然有助于保留投影数据的内部细节结构,但在提取局部特征依赖性方面效果不佳,而局部特征信息在图像恢复任务中具有重要作用,通过退化像素邻域可有效提升恢复效果。因此,所提方法在采用移位窗口 Transformer 结构提取全局特征的同时,还通过局部增强窗口 Transformer 结构^[16]来提升网络模型对局部特征的获取能力。该结构在移位窗口 Transformer 结构的基础上引入卷积算子,用基于深度卷积层构建的局部增强前馈网络 (Locally-enhanced feed-forward network, LeFF) 单元替代原先的 MLP 层。该单元首先利用全连接层将输入数据提升到更高维度,然后对高维特征数据进行深度卷积,最后使用全连接层将其恢复至初始维度并与输入数据相加。该单元所采用的升维、降维和卷积操作可以有效提升该部分网络模型的代表能力,并且在全连接层和深度卷积层之后均引入了批归一化层和 Leaky ReLU 激活层以优化训练过程。

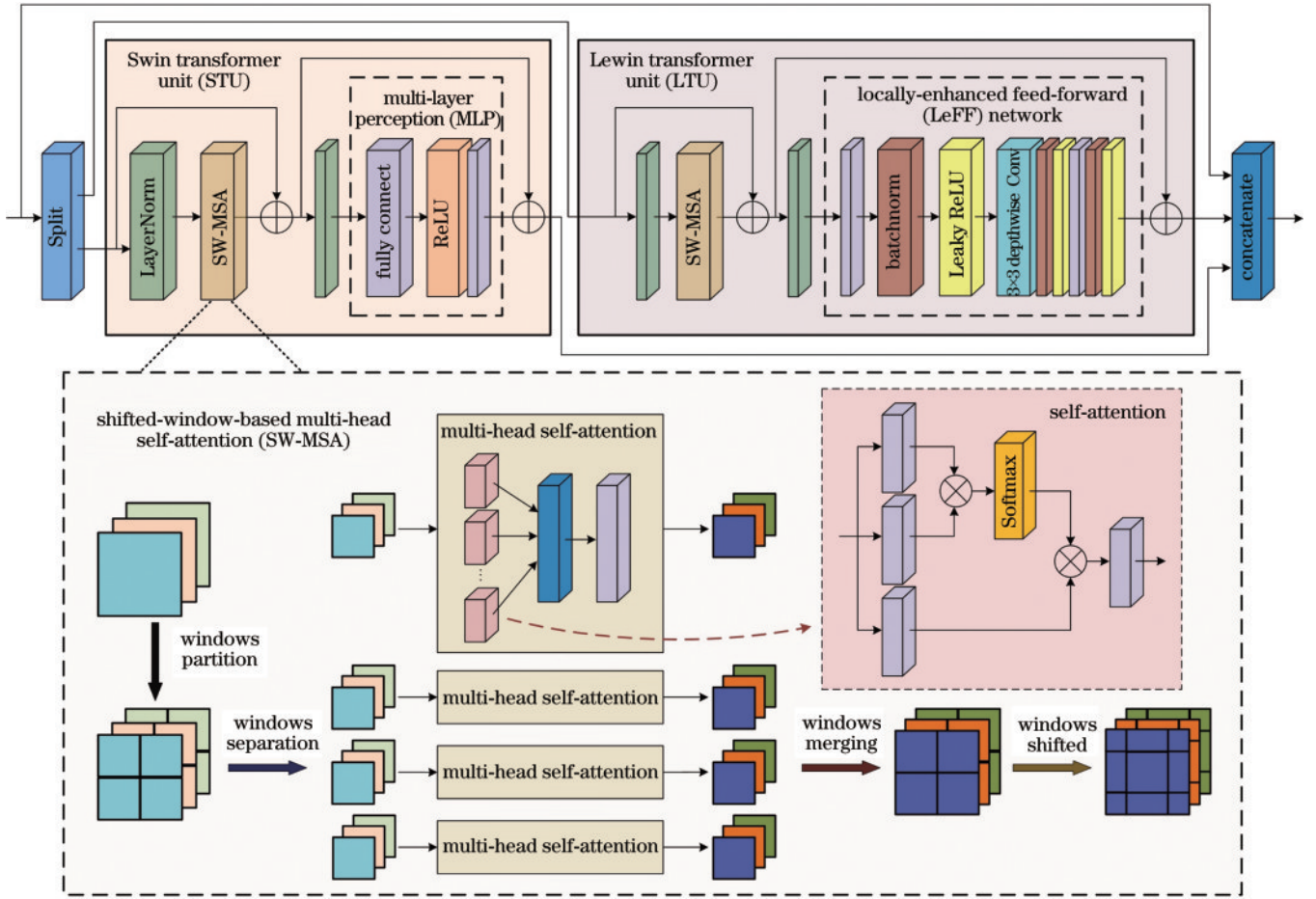


图 2 双流 Transformer 结构图
Fig. 2 Structure diagram of the dual-stream Transformer

2.3 混合损失函数

采用由平滑平均绝对误差 (Huber) 损失、特征感知损失和非局部全变分 (Non-local total variation, NLTV) 损失组成的混合损失函数。其中, Huber 损失是平均绝对误差 (MAE; L_1) 的一种可微变体, 相比于均方误差 (MSE; L_2), 在对异常点的处理上具有更好的鲁棒性, 其可以表示为

$$L_{\text{Huber}} = \frac{1}{HWC} \sqrt{\| \mathbf{P}_{\text{Predict}} - \mathbf{P}_{\text{GT}} \|^2 + \epsilon^2}, \quad (5)$$

式中: \mathbf{P}_{GT} 和 $\mathbf{P}_{\text{Predict}}$ 分别为理想标签和网络模型输出的数据; ϵ 为正则化常数。

Huber 损失能够在像素级别上度量目标图像和生成图像, 然而仅优化该损失可能会忽略图像纹理等高频细节, 使得图像过于平滑。特征感知损失利用具有预训练权重的网络模型作为特征提取器, 将需要训练的模型输出的数据与标签分别输入至特征提取模型中, 并提取部分隐藏层的输出特征, 描述二者差异并将其最小化来减少细节特征丢失。所提方法采用 ResNet-50^[17] 作为特征提取器, 通过计算两组数据特征图的内容损失 L_{Content} 和风格损失 L_{Style} 来描述差异, 具体表示为

$$L_{\text{Feature}} = L_{\text{Content}} + \lambda L_{\text{Style}} = \frac{1}{L} \sum_l \frac{1}{h_l w_l c_l} \times \left(\|\phi_{\text{Predict}, l} - \phi_{\text{GT}, l}\|_1 + \lambda \|G(\phi_{\text{Predict}, l}) - G(\phi_{\text{GT}, l})\|_1 \right), \quad (6)$$

式中: L 为特征提取模型中抽取特征图的总层数; h_l, w_l 和 c_l 分别为输出数据和标签在特征提取模型中第 l 层特征图 $\phi_{\text{Predict}, l}$ 和 $\phi_{\text{GT}, l}$ 的长度、宽度和通道数; $G(\cdot)$ 为计算所提取的特征图数据 Gram 矩阵的算子; λ 为两部分损失归一化后的权重参数; $\|\cdot\|_1$ 为 L_1 范数算子。

为进一步利用投影数据的非局部自相似性, 在去除噪声的同时保留更多的内部细节结构, 引入 NLTV 损失作为混合损失函数的正则化项。NLTV 是全变分 (Total variation, TV) 的一种变体, 该算子将非局部算子与 TV 框架有机结合, 可以显著改善因 TV 最小化导致的纹理特征丢失现象^[18]。最小化该部分损失有助于减少投影数据内部细节信息丢失, 进而提升重建图像质量。NLTV 损失可表示为

$$L_{\text{NLTV}} = \frac{1}{HWC} \left[Z_{\text{NLTV}}(\mathbf{P}_{\text{Predict}}) + \mu Z_{\text{NLTV}}(\mathbf{P}_{\text{Predict}} - \mathbf{P}_{\text{GT}}) \right], \quad (7)$$

式中: $Z_{\text{NLTV}}(\cdot)$ 为计算给定图像 I 的 NLTV 的算子; μ 为权重参数。所提方法还在长度、宽度和通道数维度

上对 NLTV 计算均值,以平衡该部分损失与其他部分损失的数量级。在计算给定图像 I 的 NLTV 时,假设所划分的图像块大小为 $p \times p$ 。对于图像中的目标图像块 $U_x (U_x \in \mathbb{R}^{p \times p})$,其中像素点位于图像位置 x 处。通过计算相似性权重来筛选 x 周边局部范围 $r \times r$ 内与 U_x 相似的图像块 $U_y (U_y \in \mathbb{R}^{p \times p})$, U_x 与 U_y 的相似性权重可表示为

$$w_{xy} = \exp\left(-\frac{g_\sigma \times |U_x - U_y|^2}{2h_0^2}\right), \quad (8)$$

式中: g_σ 是一个标准差为 σ 的高斯卷积核; h_0 为全局滤波参数。图像 I 的 NLTV 表示为

$$Z_{\text{NLTV}}(I) = \sum_{x \in \Omega} \sqrt{\sum_{y \in \Omega_x} w_{xy} \cdot (U_x - U_y)^2}, \quad (9)$$

式中: Ω 为给定图像的整体空间; Ω_x 为像素点 x 周围的图像域空间。因此,用于训练 DTE-PDN 的混合损失函数整体可表示为

$$L_{\text{Total}} = L_{\text{Huber}} + \lambda_1 L_{\text{Feature}} + \lambda_2 L_{\text{NLTV}}, \quad (10)$$

式中: λ_1 和 λ_2 分别为权重参数。

3 实验与结果分析

3.1 数据集与模型配置

所提方法通过大量数据集来训练、验证和测试所提出的网络模型,这些数据集基于杜克大学提出的 XCAT^[19] 人体模型生成。所提方法选取的人体模型分别来自于 12 名不同性别、年龄、身高、体重的患者,由 120 kVp 管电压驱动的 GE_Maxiray_125 球管进行扫描,采用探测器^[20] 仿真模型模拟 X 射线的探测过程并提供投影数据。探测器能量阈值分别设为 80 keV 和 100 keV,并忽略能量在 20 keV 以下的光子,三个能量区间分别为 20~80 keV、80~100 keV 和 100~120 keV,切片厚度为 1.5 mm,扫描范围为 360°。本实验通过减少光子通量的方式来模拟低剂量扫描条件,初始入射光子数设为 1×10^5 。将来自 10 组人体模型的 1200 组低剂量和常规剂量扫描结果用于网络模型的训练和验证,随机抽取其中 240 组扫描结果作为验证集,另外 2 组人体模型的 240 组扫描结果用于测试网络模型的性能。本实验采用补丁训练策略,将 1200 组训练集和验证集划分成尺寸为 64×64 的重叠补丁,最终共有 48000 组补丁用于网络模型的训练和验证。

所提出的网络模型基于 TensorFlow 框架实现,训练周期为 1200,最小训练批次为 32,并采用 Adam^[21] 算法来优化网络模型的训练,初始学习率设为 1×10^{-5} ,指数衰减参数分别设为 0.5 和 0.99,混合损失函数的权重参数 λ_1 和 λ_2 分别设置为 0.3 和 1×10^{-5} 。本实验采用 Intel i5-9400f CPU 搭载的 NVIDIA 公司 GeForce RTX 2080 GPU 来训练网络模型,其运行内存大小为

32 GB,显存大小为 8 GB,操作系统为 Windows 10, 64 bit, CUDA API 的版本为 9.0, Python 版本为 3.9.16, TensorFlow-GPU 版本为 2.6.0。基于仿真数据集的 DTE-PDN 模型训练过程损失收敛曲线如图 3 所示,可以看出,训练过程中损失函数收敛速度较快,说明网络模型具有较为平滑的梯度。

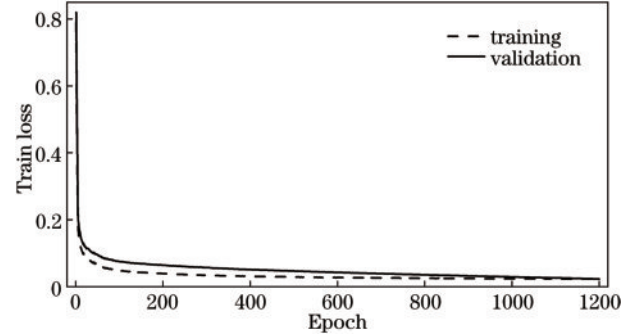


图 3 在仿真数据集上的损失收敛曲线

Fig. 3 Loss convergence curves on simulation dataset

3.2 对比实验结果分析

本实验选取具有代表性的 RICCS-SI^[22]、RED-CNN^[23]、DRN^[9]、TransCT^[24] 作为对比方法,并采用峰值信噪比 (Peak signal to noise ratio, PSNR)、结构相似性指数 (Structure similarity index measurement, SSIM) 和特征相似性指数 (Feature similarity index measurement, FSIM) 评估各方法处理结果。其中, PSNR 和 SSIM 常用于常规图像质量评价领域,可客观评估重建图像的噪声含量及其与参考图像之间的结构相似程度,FSIM 在医学影像评估领域已被证明是与医生主观评价结果相关程度较高的客观指标之一^[25],在实用潜力评估方面更具参考价值。与投影数据相比,重建图像在对病灶组织的观察与解析中能起到更加关键的作用,为更加直观地评价所提出的投影数据去噪方法,本实验使用滤波反投影 (FBP) 算法将各方法处理后的投影数据重建为多能 CT 图像,以更好地对比各方法对噪声的去除效果。

本实验分别选取腹部和胸部切片的投影数据来测试网络模型的去噪能力,各方法处理结果分别如图 4 和图 5 所示。其中,图 4(a) 和图 5(a) 为低剂量条件下各能量区间的重建图像,图 4(b) 和图 5(b) 为常规剂量条件下对应图像,该列图像可作为各方法处理结果的对比标准。为更加直观地对比各方法的处理效果,本实验在各重建图像中选取感兴趣区域 (Region of interest, ROI) 并进行放大,在图 4 和图 5 中分别标注为 ROI A 和 ROI B。与低剂量多能 CT 重建图像相比,各方法的处理均有效减少了重建图像中的噪声含量,而未经处理的低剂量多能 CT 重建图像中噪声含量较高,导致部分人体组织的边缘和细节结构难以观察。RICCS-SI 处理后得到的重建图像如图 4(c) 和图 5(c) 所示,该方法虽然去除了部分噪声,但在重建图像中的

残留噪声仍然影响着对病灶组织的观察。RED-CNN 方法的处理结果如图 4(d)和图 5(d)所示,可以看出在该方法处理的各部位重建图像中仍残留部分噪声,且边缘结构不够清晰。DRN 方法的处理结果如图 4(e)和图 5(e)所示,与 RED-CNN 方法相比,该方法得到的重建图像质量有较大的提升,内部噪声较少,但图中部分微小组织结构模糊,存在细节特征丢失的现象。图 4(f)和图 5(f)显示了 TransCT 方法的处理结果,可以看出采用非滑窗化 Transformer 结构的 TransCT 方

法有效利用了图像的非局部自相似性,可在去除噪声的同时保留更多细节特征。所提方法的处理结果如图 4(g)和图 5(g)所示,细节特征清晰,且噪声含量显著降低,取得了与 TransCT 方法相近的处理效果。网络模型中的 STU 和 LTU 采用滑窗运算多头自注意力结构可有效利用投影数据非局部自相似性,并具有比非滑窗化 Transformer 结构更少的运算量,所提方法所提出的网络模型因此具有更快的处理速度,而且通用性和实用潜力也更强。

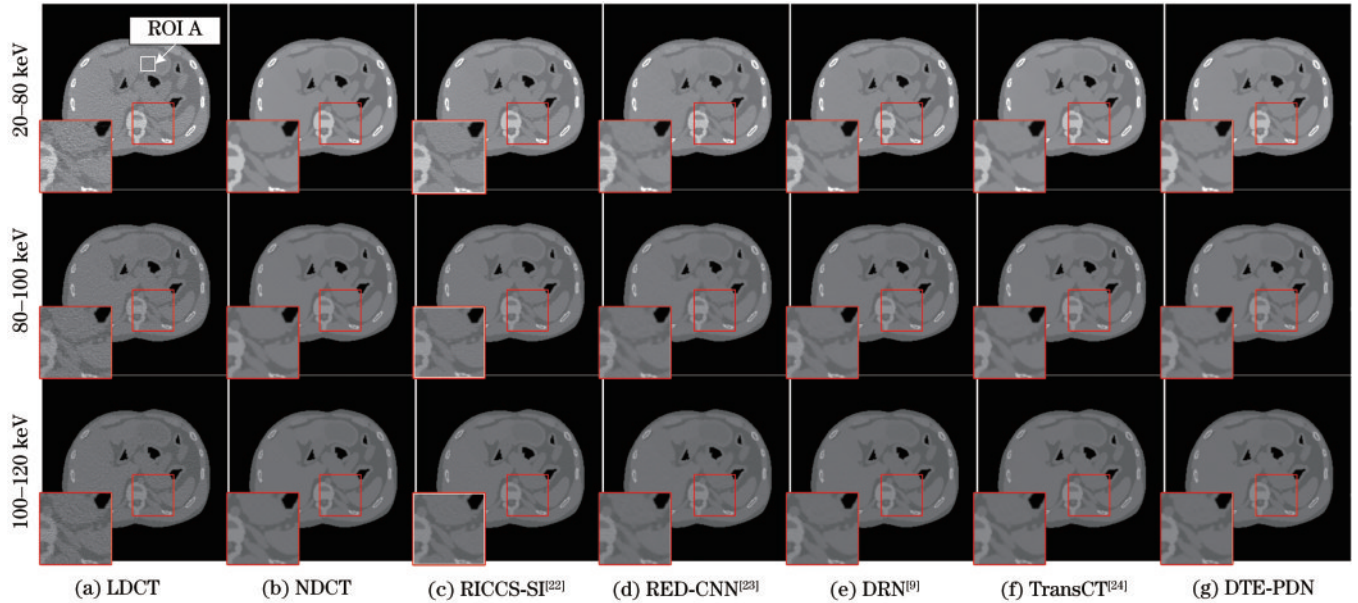


图 4 腹部切片不同算法的去噪效果对比

Fig. 4 Comparison of denoising effects of abdomen slice from different methods

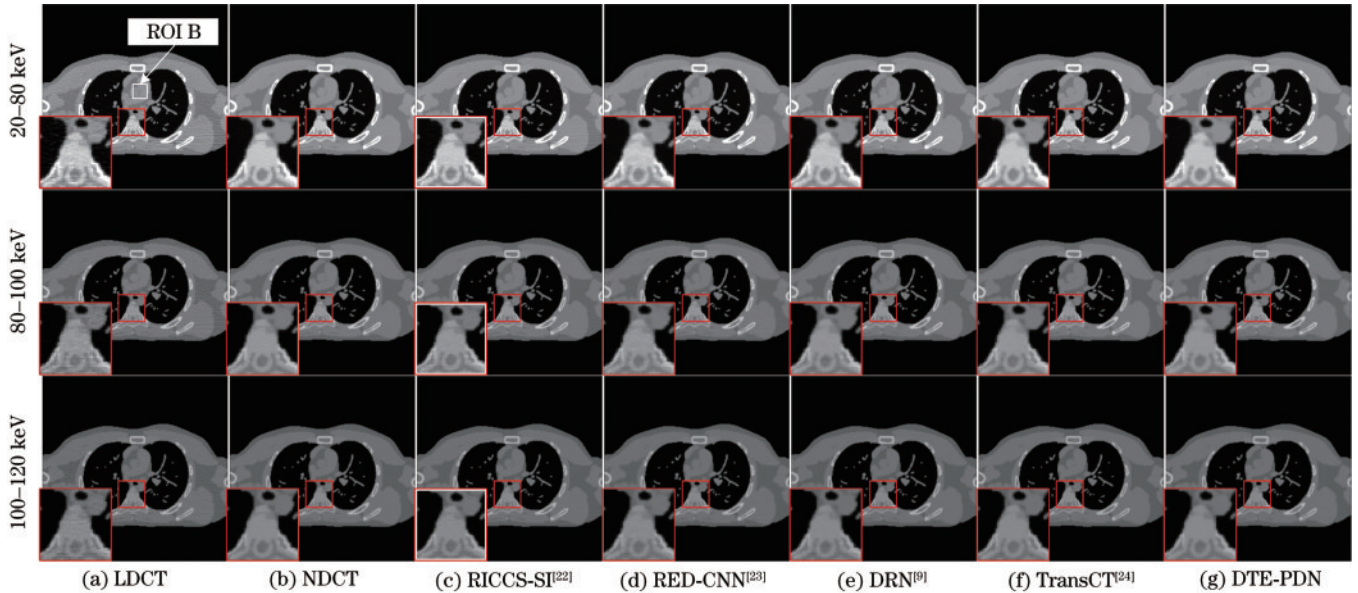


图 5 胸部切片不同算法的去噪效果对比

Fig. 5 Comparison of denoising effects of chest slice from different methods

本实验分别计算了测试集中腹部和胸部切片经各方法处理后重建图像的 PSNR、SSIM 和 FSIM 平均值,计算结果如表 1 和表 2 所示。可以看出,所提

出的 DTE-PDN 方法在 PSNR、SSIM 和 FSIM 上均高于其他对比方法,说明所提方法在去除多能 CT 图像噪声方面性能更优。相较于 RICCS-SI 方法,经所提

表 1 不同算法腹部切片测试结果客观指标

Table 1 Evaluation index of abdomen slice test results from different methods

Energy /keV	Index	RICCS-SI ^[22]	RED-CNN ^[23]	DRN ^[9]	TransCT ^[24]	DTE-PDN
20-80	PSNR /dB	33.9413	36.4773	37.0067	37.5022	37.8992
	SSIM	0.9692	0.9852	0.9893	0.9950	0.9958
	FSIM	0.9214	0.9515	0.9742	0.9963	0.9972
80-100	PSNR /dB	34.3815	37.7927	38.0993	38.9681	39.2680
	SSIM	0.9703	0.9883	0.9891	0.9975	0.9964
	FSIM	0.9257	0.9543	0.9760	0.9969	0.9976
100-120	PSNR /dB	35.0164	38.7362	39.0957	39.5153	39.6702
	SSIM	0.9755	0.9901	0.9933	0.9975	0.9977
	FSIM	0.9248	0.9556	0.9770	0.9971	0.9971

表 2 不同算法胸部切片测试结果客观指标

Table 2 Evaluation index of chest slice test results from different methods

Energy /keV	Index	RICCS-SI ^[22]	RED-CNN ^[23]	DRN ^[9]	TransCT ^[24]	DTE-PDN
20-80	PSNR /dB	34.8861	36.9783	37.3461	37.5323	37.7301
	SSIM	0.9720	0.9873	0.9895	0.9927	0.9944
	FSIM	0.9311	0.9550	0.9762	0.9977	0.9976
80-100	PSNR /dB	36.4913	38.2874	38.6820	39.6757	39.7969
	SSIM	0.9767	0.9892	0.9912	0.9956	0.9961
	FSIM	0.9368	0.9568	0.9774	0.9977	0.9981
100-120	PSNR /dB	37.3751	39.2031	39.6868	39.8529	39.8615
	SSIM	0.9793	0.9906	0.9931	0.9981	0.9984
	FSIM	0.9325	0.9566	0.9771	0.9983	0.9984

方法处理结果的 PSNR 值提升了 2.4864 dB 以上, SSIM 值提升了 0.0191 以上, FSIM 值提升了 0.0613 以上。相较于 DRN 方法, 经所提方法处理的腹部切片各能量区间投影信息的重建图像 PSNR 值提升了 0.1747 dB 以上, SSIM 值提升了 0.0044 以上, FSIM 值提升了 0.0201 以上, 说明 DRN 方法虽然可以通过扩大卷积核感受野来聚合更大范围的特征信息, 但对于图像中的细节特征和结构的保留能力依然不足。表 2 中胸部切片测试结果的各评价指标也证明了上述结论, 表明所提方法利用非局部自相似性可以有效保留更多投影数据内部的细节结构, 使重建图像具有更加清晰的细节特征。所提方法与 TransCT 方法测试结果的各评价指标非常接近, 表明所提方法采用的滑窗运算多头自注意力结构在降低

网络模型计算量的同时取得与非滑窗化 Transformer 结构相近的去噪效果。

为了进一步分析各方法测试结果中的噪声含量, 本实验分别选取腹部和胸部切片中较为平滑的部分作为 ROI 并计算该区域内的均值与标准差, 所选取的 ROI 在图 4 和图 5 中分别标注为 ROI A 和 ROI B, 结果如表 3 所示。由表 3 中数据可知, 所提出的 DTE-PDN 模型的测试结果标准差最低, 说明经所提方法处理的投影数据所重建的 CT 图像相比其他方法具有更低的噪声水平。采用非滑窗化 Transformer 结构的 TransCT 方法与所提方法所得结果方差相近, 说明所采用的滑窗运算多头自注意力结构可达到与非滑窗化 Transformer 结构相近的去噪效果, 进一步证明了所提方法在多能 CT 投影数据噪声去除方面的性能。

表 3 不同算法测试结果 ROI 的均值与标准差对比

Table 3 Comparison of the mean and standard deviation among ROIs in test results from different methods

ROI	Energy /keV	RICCS-SI ^[22]	RED-CNN ^[23]	DRN ^[9]	TransCT ^[24]	DTE-PDN
ROI A	20-80	0.2020±0.0042	0.2021±0.0030	0.2022±0.0027	0.2024±0.0019	0.2028±0.0017
	80-100	0.1771±0.0034	0.1769±0.0023	0.1770±0.0020	0.1773±0.0015	0.1770±0.0015
	100-120	0.1672±0.0030	0.1665±0.0021	0.1680±0.0018	0.1678±0.0014	0.1677±0.0013
ROI B	20-80	0.2116±0.0043	0.2116±0.0031	0.2115±0.0032	0.2116±0.0028	0.2117±0.0028
	80-100	0.1829±0.0036	0.1826±0.0027	0.1827±0.0026	0.1830±0.0023	0.1829±0.0022
	100-120	0.1728±0.0033	0.1728±0.0026	0.1731±0.0024	0.1732±0.0021	0.1730±0.0020

3.3 消融实验结果分析

本实验对所提出的网络模型中各模块与结构的性能展开消融研究,其结果如表 4 所示。由表 4 中数据可知,STU 和 LTU 均有助于提升模型的去噪性能,与 STU 相比,LTU 对局部特征的建模能力更强,对去除投影数据噪声与保留其内部细节结构和特征具有更好的效果。基于 STU 和 LTU 的双流 Transformer 结构能够使模型的去噪能力进一步提升,PSNR 值相比单一 Transformer 结构方案提升了 0.0763 dB 以上,SSIM 值提升了 0.0021 以上,FSIM 值提升了 0.0045 以上。DTE-PDN 采用的残差结构有助于改善网络模型的训练梯度,使模型具有更好的性能,表 4 中的结果也证明了上述观点。与非残差结构的方案相比,残差结构的加入使测试数据的 PSNR 值提升了 0.0719 dB 以上,SSIM 值提升了 0.0032 以上,FSIM 值提升了 0.0019 以上。由此可知,所提方法将双流 Transformer 结构与残差结构相结合具有最好的去噪效果。

为了进一步研究混合损失函数各部分对模型的作用效果,采用各项损失的不同组合来训练 DTE-PDN 模型,并在 20~80 keV 的腹部切片中随机选取投影路径作为 ROI,如图 6(a)标注线所示。图 6(b)显示了由不同损失函数组合训练得到的测试结果在该投影路径上线性衰减系数的变化。由图 6(b)中曲线可以看出,仅由 Huber 损失训练的测试结果过于平滑,在边缘部位与标签偏差较大。加入特征感知损失后,测试结果中较为平滑的组织部位数值变化较小,说明特征感知损失的加入有益于提升模型对细节特征信息的敏感性,但在减小重建图像与标签的差距上效果并不明显。对于 Huber 损失与非局部全变分损失组合训练的模型,图像边缘部位的准确程度显著提升,说明该损失组合提高了模型的边缘特征学习能力,避免了因重建图像过度平滑所导致的边缘不清晰等问题。由 Huber 损失、特征感知损失和非局部全变分损失构成的混合损失函数来训练模型可以获得最接近标签图像的处理效果,说明损失函数中各部分优势和性能达到了均衡,证明了这种组合为最佳损失函数方案。

表 4 DTE-PDN 模型中不同模块组合的腹部和胸部切片测试结果客观指标

Table 4 Evaluation index of abdomen and chest slices from DTE-PDN with different combination of blocks

Slice	STU	LTU	Residual structure	20~80 keV			80~100 keV			100~120 keV		
				PSNR / dB	SSIM	FSIM	PSNR / dB	SSIM	FSIM	PSNR / dB	SSIM	FSIM
Abdomen	✓			37.4853	0.9871	0.9796	38.8544	0.9847	0.9803	39.1781	0.9855	0.9814
		✓		37.4965	0.9872	0.9908	38.9568	0.9879	0.9856	39.2517	0.9867	0.9823
	✓	✓		37.6030	0.9922	0.9953	39.1961	0.9931	0.9939	39.5180	0.9919	0.9948
	✓	✓	✓	37.8992	0.9964	0.9972	39.2680	0.9964	0.9972	39.6702	0.9977	0.9971
Chest	✓			37.3112	0.9861	0.9788	39.0217	0.9876	0.9790	39.1171	0.9873	0.9819
		✓		37.4083	0.9873	0.9854	39.1445	0.9908	0.9835	39.2718	0.9895	0.9831
	✓	✓		37.4846	0.9907	0.9952	39.6001	0.9929	0.9953	39.6792	0.9923	0.9954
	✓	✓	✓	37.7301	0.9944	0.9976	39.7969	0.9961	0.9981	39.8615	0.9984	0.9984

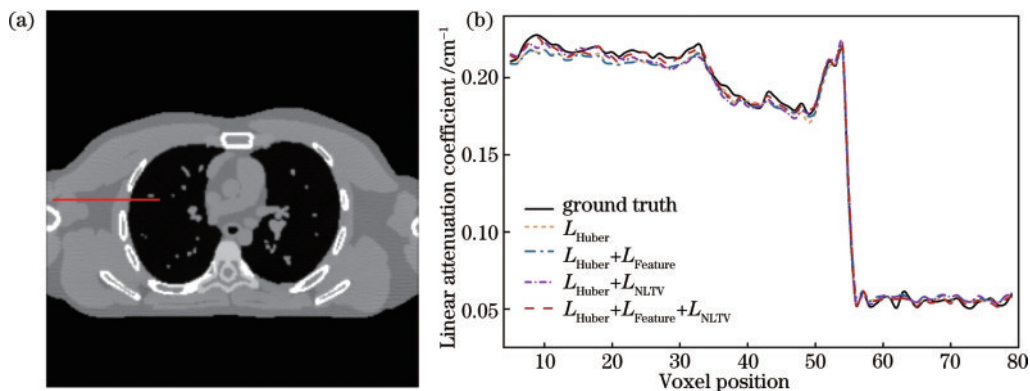


图 6 不同损失函数组合的消融实验结果。(a)选取的投影路径;(b)线性衰减系数曲线

Fig. 6 Ablation experiment results of different loss function combinations. (a) Selected projection path; (b) linear attenuation coefficient curves

为进一步验证混合损失函数各部分权重参数对网络模型训练效果的影响,设置了不同的 λ_1 和 λ_2 数值,

并分别使用不同权重参数组合的混合损失函数对网络模型进行训练。以 80~100 keV 的胸部切片为例,图 7

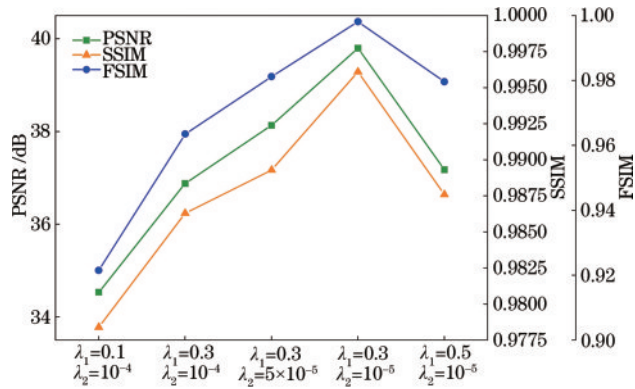


图7 不同权重参数组合的混合损失函数训练DTE-PDN处理胸部切片指标曲线

Fig. 7 Curves of the evaluation index of the chest slice processed by DTE-PDN trained by the hybrid loss with different weight combinations

展示了经不同权重参数方案训练后 DTE-PDN 模型的处理效果。可以看出,所提方法所采用的权重参数方案具有最佳的训练效果,进一步证明了混合损失函数在提升网络模型去噪性能方面的优势。

4 结 论

辐射剂量的降低使多能 CT 扫描系统成像质量下降,不利于后续诊断工作中对病灶组织的准确观察。针对上述问题,提出了一种基于双流 Transformer 的投影数据噪声去除方法来提升重建图像质量。该方法通过同时提取并融合多能 CT 投影数据的局部和全局特征,有效提升了重建图像的 PSNR。该方法充分利用投影数据的稀疏性和自相似性,采用 NLTV 损失并结合 Huber 损失和特征感知损失来训练所提出的网络模型,有效提高了网络模型对投影数据内部细节结构的敏感程度,进一步缩小了重建图像与标签数据之间的结构特征差异。实验结果表明,经所提方法处理的投影数据重建出的 CT 图像细节结构清晰、噪声含量较低,说明该方法在显著去除投影噪声的同时最大限度地保留了内部细节结构,保证了在低剂量条件下多能 CT 图像的诊断价值。

参 考 文 献

[1] Zeng G L. Photon starvation artifact reduction by shift-variant processing[J]. IEEE Access, 2022, 10: 13633-13649.
 [2] 牛善洲, 张梦真, 邱洋, 等. 基于全广义变分约束加权最小二乘的低剂量计算机断层重建方法[J]. 激光与光电子学进展, 2023, 60(4): 0411001.
 Niu S Z, Zhang M Z, Qiu Y, et al. Total generalized variation constrained weighted least-squares for low-dose computed tomography reconstruction[J]. Laser & Optoelectronics Progress, 2023, 60(4): 0411001.
 [3] Wu D F, Kim K, Li Q Z. Low-dose CT reconstruction with Noise2Noise network and testing-time fine-tuning [J]. Medical Physics, 2021, 48(12): 7657-7672.

[4] He Z N, Zhang Y K, Guan Y, et al. Iterative reconstruction for low-dose CT using deep gradient priors of generative model[J]. IEEE Transactions on Radiation and Plasma Medical Sciences, 2022, 6(7): 741-754.
 [5] Zhang W, Chen X Z, Wang H Y, et al. Low-dose CT image reconstruction based on wavelet-TGV and non-convex penalty function[J]. Journal of Nonlinear and Convex Analysis, 2021, 22(10): 2219-2230.
 [6] 龙超, 金恒, 黎玲, 等. 基于特征融合的非局部均值 CT 图像降噪[J]. 光学学报, 2022, 42(11): 1134024.
 Long C, Jin H, Li L, et al. Non-local mean CT image denoising based on feature fusion[J]. Acta Optica Sinica, 2022, 42(11): 1134024.
 [7] Xie B Q, Zhu Y M, Niu P, et al. Super-energy-resolution material decomposition for spectral photon-counting CT using pixel-wise learning[J]. IEEE Access, 2021, 9: 168485-168495.
 [8] 刘昊鑫, 赵源萌, 张存林, 等. 基于改进 U-net 的牙齿锥形束 CT 图像重建研究[J]. 中国激光, 2022, 49(24): 2407207.
 Liu H X, Zhao Y M, Zhang C L, et al. Study on tooth cone beam CT image reconstruction based on improved U-net network[J]. Chinese Journal of Lasers, 2022, 49(24): 2407207.
 [9] Trung N T, Trinh D H, Trung N L, et al. Low-dose CT image denoising using deep convolutional neural networks with extended receptive fields[J]. Signal, Image and Video Processing, 2022, 16(7): 1963-1971.
 [10] Jiao F Y, Gui Z G, Li K P, et al. A dual-domain CNN-based network for CT reconstruction[J]. IEEE Access, 2021, 9: 71091-71103.
 [11] Di S H, Zhao Y Q, Liao M, et al. TD-net: a hybrid end-to-end network for automatic liver tumor segmentation from CT images[J]. IEEE Journal of Biomedical and Health Informatics, 2023, 27(3): 1163-1172.
 [12] Wu H J, Xu C, Liu H Z. SMART: semantic-aware masked attention relational transformer for multi-label image recognition[J]. IEEE Signal Processing Letters, 2022, 29: 2158-2162.
 [13] Li H R, Yang X M, Yang S H, et al. Transformer with double enhancement for low-dose CT denoising[J]. IEEE Journal of Biomedical and Health Informatics, 2023, 27(10): 4660-4671.
 [14] Yang L T, Li Z N, Ge R J, et al. Low-dose CT denoising via sinogram inner-structure transformer[J]. IEEE Transactions on Medical Imaging, 2023, 42(4): 910-921.
 [15] Liu Z, Lin Y T, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2022: 9992-10002.
 [16] Wang Z D, Cun X D, Bao J M, et al. Uformer: a general U-shaped transformer for image restoration[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-24, 2022, New Orleans, LA, USA. New York: IEEE Press, 2022: 17662-17672.
 [17] Liang T F, Jin Y, Li Y D, et al. EDCNN: edge

- enhancement-based densely connected network with compound loss for low-dose CT denoising[C]//2020 15th IEEE International Conference on Signal Processing (ICSP), December 6-9, 2020, Beijing, China. New York: IEEE Press, 2021: 193-198.
- [18] Feng J, Yu H J, Wang S Y, et al. Image-domain based material decomposition by multi-constraint optimization for spectral CT[J]. IEEE Access, 2020, 8: 155450-155458.
- [19] Segars W P, Sturgeon G, Mendonca S, et al. 4D XCAT phantom for multimodality imaging research[J]. Medical Physics, 2010, 37(9): 4902-4915.
- [20] Shi Z F, Yang H Y, Cong W X, et al. An edge-on charge-transfer design for energy-resolved X-ray detection [J]. Physics in Medicine and Biology, 2016, 61(11): 4183-4200.
- [21] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2014-12-22) [2023-03-05]. <https://arxiv.org/abs/1412.6980>.
- [22] Ren J R, Liang N N, Yu X H, et al. Projection domain processing for low-dose CT reconstruction based on subspace identification[J]. Journal of X-Ray Science and Technology, 2023, 31(1): 63-84.
- [23] Chen H, Zhang Y, Kalra M K, et al. Low-dose CT with a residual encoder-decoder convolutional neural network [J]. IEEE Transactions on Medical Imaging, 2017, 36 (12): 2524-2535.
- [24] Zhang Z C, Yu L Q, Liang X K, et al. TransCT: dual-path transformer for low dose computed tomography [M]//de Bruijne M, Cattin P C, Cotin S, et al. Medical image computing and computer-assisted intervention- MICCAI 2021. Lecture notes in computer science. Cham: Springer, 2021, 12906: 55-64.
- [25] Mason A, Rioux J, Clarke S E, et al. Comparison of objective image quality metrics to expert radiologists' scoring of diagnostic quality of MR images[J]. IEEE Transactions on Medical Imaging, 2020, 39(4): 1064-1072.