

## 基于深层残差编解码的显微 CT 图像去噪算法

符慧娟<sup>1</sup>, 席晓琦<sup>1</sup>, 韩玉<sup>1</sup>, 李磊<sup>1</sup>, 王鑫光<sup>2</sup>, 闫宾<sup>1\*</sup><sup>1</sup>信息工程大学信息工程学院, 河南 郑州 450001;<sup>2</sup>河南省文物考古研究院, 河南 郑州 450001

**摘要** 针对实验室显微 CT 扫描过程会产生噪声, 导致重建后 CT 图像质量下降的问题, 提出一种深层多残差编解码卷积去噪网络。以原始的残差编解码网络为基础, 首先通过增加网络的卷积层数, 引入多残差映射, 实现对实验室级显微 CT 图像中噪声分布特性的有效学习; 其次设计了专用的混合损失函数, 增强网络对图像细节信息的保留能力。实验结果表明, 所提方法对 CT 图像中的噪声具有显著的抑制效果, 同时能够最大程度地保留图像的结构信息和特征信息。

**关键词** 图像处理; 图像去噪; 显微 CT; 无损检测; 残差学习; 混合损失函数

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP221785

## Micro-CT Image Denoising Algorithm Based on Deep Residual Encoding-Decoding

Fu Huijuan<sup>1</sup>, Xi Xiaoqi<sup>1</sup>, Han Yu<sup>1</sup>, Li Lei<sup>1</sup>, Wang Xinguang<sup>2</sup>, Yan Bin<sup>1\*</sup><sup>1</sup>College of Information System Engineering, Information Engineering University, Zhengzhou 450001, Henan, China;<sup>2</sup>Henan Provincial Institute of Cultural Heritage and Archaeology, Zhengzhou 450001, Henan, China

**Abstract** Aiming at the problem that noise will be generated during micro-CT scanning in the laboratory, resulting in the decline of CT image quality after reconstruction, this paper proposes a deep multi-residual encoding-decoding convolutional denoising network. This method is based on the original residual encoding-decoding convolutional network. First, we increased the number of convolutional layers and introduced the multiple residuals to realize effective learning of noise distribution characteristics in lab-level micro-CT images. Second, a special mix-loss function was designed to strengthen the network's ability to retain image details. Experimental results show that the proposed method has a significant effect on noise suppression and can greatly preserve the structure and feature information of CT images.

**Key words** image processing; image denoising; micro-CT; non-destructive testing; residual learning; mix-loss function

## 1 引言

X 射线三维成像能够无损表征物体内部信息, 已成为人类探索微观世界的重要手段<sup>[1]</sup>。随着光子计数探测技术和百纳米尺寸焦斑 X 射线源的出现, 实验室级显微分辨率 CT 技术得到快速发展<sup>[2]</sup>。由于超微焦斑的实验室级显微 CT 射线源的光子数比普通锥束 CT 低 1 到 2 个数量级, 在相同曝光时间下, 实验室级显微 CT 探测器接收到的光子数较少, 导致重建出的 CT 图像包含大量噪声, 造成图像空间分辨率和对比度下降, 细节信息损失。通常可以通过增加曝光时间

来提高图像的空间分辨率, 但这会导致时间分辨率下降<sup>[3]</sup>。因此, 不可避免地要在空间分辨率与效率之间进行权衡, 保证效率的同时提高实验室级显微 CT 的图像质量具有重要意义<sup>[4]</sup>。

针对 CT 图像去噪问题, 研究人员提出了许多去除噪声的方法。在传统去噪方法中, Dabov 等<sup>[5]</sup>提出的改进三维块匹配滤波算法最为经典。近年来, 深度学习在图像处理中得到了广泛应用, 对图像去模糊<sup>[6-8]</sup>和去噪<sup>[9-11]</sup>等取得了不错的成果。2017 年, Chen 等<sup>[12]</sup>做了大量的研究工作, 将卷积神经网络(CNN)应用于低剂量 CT 图像去噪。同年, Chen 等<sup>[13]</sup>提出了一种具

收稿日期: 2022-06-06; 修回日期: 2022-07-29; 录用日期: 2022-09-13; 网络首发日期: 2022-09-23

基金项目: 国家重点研发项目(2020YFC1522002)

通信作者: \*ybspace@hotmail.net

有编解码结构的卷积神经网络(RED-CNN),该方法虽去除了CT图像中的噪声与伪影,但预测的结果中会出现边缘模糊和部分细节丢失的问题。近几年,对抗性训练也得到广泛的应用,Chen等<sup>[14]</sup>针对显微CT降噪研究设计了一个C-GAN模型;Kang等<sup>[15]</sup>提出CycleGAN网络,采用无监督学习的方式产生更清晰的样本;朱斯琪等<sup>[16]</sup>将CycleGAN应用于低剂量CT图像去噪。但生成对抗网络(GAN)存在模型大、训练困难的问题,在较大运行内存的专用计算设备上才能训练出理想的训练模型,这些因素使得GAN应用困难。

针对原始残差编解码网络<sup>[13]</sup>去噪后造成图像细节模糊的问题,本文设计一个轻量型的CNN用于实验室级显微CT图像去噪问题,降低训练过程对硬件设备的要求,提高CNN去噪方法的实用性。本文提出了一种深层残差编解码网络(DRED-CNN),以原始残差编解码器为基础结构,设计一个由18个卷积块组成的残差编解码网络,在没有先验知识的情况下,实现带噪声图像与无噪声图像之间端到端的学习,成功地实现盲去噪的操作。此外,引入残差学习,且采用L1损失与

多尺度结构相似性(MS-SSIM)损失函数的混合损失函数,这更有利于在去除噪声的同时极大地保留图像结构信息。

## 2 基于改进的残差编解码网络的降噪方法

### 2.1 深度残差编解码网络的设计

Bengio和LeCun曾在文章中证明采用浅层结构解决复杂问题是不可行的,深层结构是当前研究的重点<sup>[17]</sup>。深层模型意味着更好的非线性表达能力,扩大了感受野,可以学习更加复杂的变换,从而更好地拟合特征输入。因此,本文设计了加深网络层数的多残差编解码网络,网络的总体架构如图1所示。网络结构中 $K$ 为卷积核尺寸, $C$ 为输出通道数。DRED-CNN是一个端到端CNN,包含1个输入层、1个输出层和18个中间隐藏层。其中隐藏层由9个卷积块和9个反卷积块组成,连续的卷积块与反卷积块可以看作网络编码与解码的过程。每一层卷积核大小均设置为 $3 \times 3$ ,卷积步长设置为1,填充像素为0。

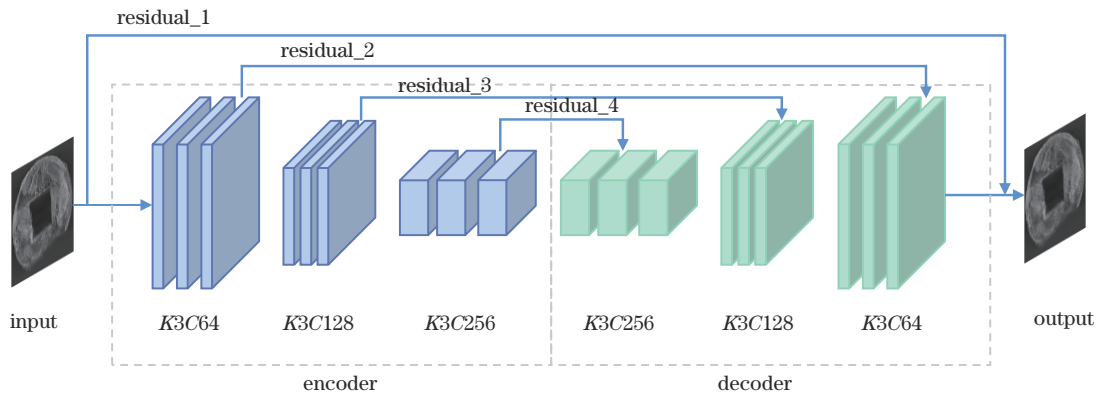


图1 DRED-CNN的结构

Fig. 1 DRED-CNN structure

网络前半部分,9个卷积块构成堆叠式编码器,编码器中卷积层的通道数分别设置为64、64、64、128、128、128、256、256和256,逐层提取图像特征数量,图像的噪声和伪影从低到高被逐级抑制,以保留提取图像块中的基本信息。由于卷积层之后的池化层可能会损失结构细节<sup>[18]</sup>,因此在DRED-CNN的编码器中依然没有使用池化层。虽然DRED-CNN去除了池化操作,但一系列卷积会减少输入信号的结构细节。所以在网络后半部分,将9个反卷积块集成到模型中,构成堆叠式解码器,用于结构细节的恢复,可以将其视为提取特征后的图像重建操作。为了保证网络的输入和输出准确匹配,卷积层和反卷积层使用大小相同的卷积核。卷积块和反卷积块在网络中是对称的,因此解码器中反卷积的输出通道数分别设置为256、256、256、128、128、128、64、64和64。网络架构中每层的卷积块如图2所示,每个卷积块都是由一个卷积层和ReLU

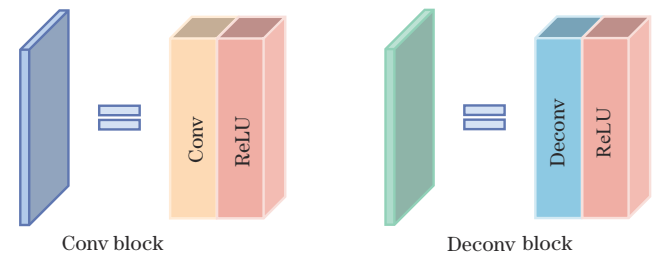


图2 卷积块与反卷积块组成

Fig. 2 Composition of convolution block and deconvolution block

激活函数组成。ReLU激活函数具有良好的非线性特性<sup>[19]</sup>,可以在保留更多浅层特征信息的基础上对深层次的特征更具判别力,进而提高网络的非线性拟合效果,使得模型的收敛速度维持在一个稳定状态。

卷积过程会损失部分图像细节信息,虽然反卷积层可以恢复一些细节,但当网络深度加深时,累积的损

失对图像恢复可能会很不理想。此外,当网络深度增加时,梯度消失会使网络训练参数无法更新。残差网络<sup>[20]</sup>解决了上述问题,在深度学习领域得到广泛应用。因此 DRED-CNN 在对称的卷积块和反卷积块之间引入残差学习,通过快捷连接来学习残差。残差的加入不会给网络增加额外的参数量和计算量,却可以提升模型的训练速度,提高训练效果。残差编解码网络学习的是  $x$  到  $x + F(x)$  之间的映射,只需要让  $F(x)$  近似等于 0,即要求解的映射等于上一层输出的特征映射  $x$ ,使得这一层残差块的网络状态仍是最佳的一个状态,如图 3 所示。

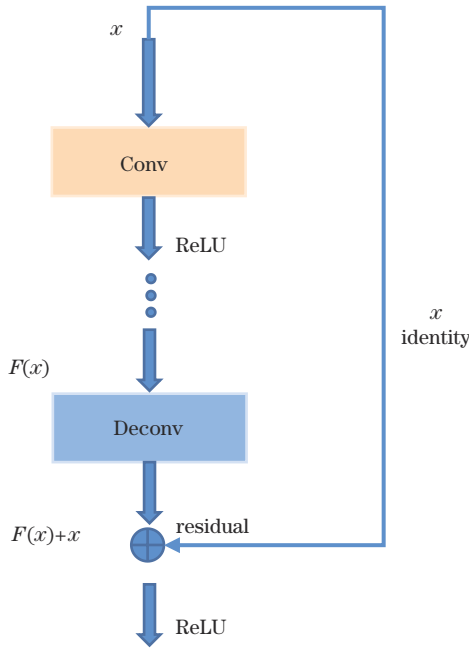


图 3 残差学习网络结构

Fig. 3 Residual learning network structure

### 2.2 混合损失函数的设计

在原始 RED-CNN 中使用均方误差 (MSE) 损失函数,但卷积神经网络处理涉及图像质量的任务时, MSE 损失函数与人类感知的图像质量之间相关性较差<sup>[21]</sup>,预测结果中会出现图像边缘模糊和细节丢失的现象。损失函数在网络性能中起着关键作用,为了解决原始 RED-CNN 去噪后的模糊现象,并且确保去噪后的 CT 图像纹理和结构细节得到保留,所提方法使用多尺度结构相似性 (MS-SSIM) 损失和 L1 损失构成的混合损失函数。MS-SSIM 损失容易造成亮度的改变,但它能保留图像的高频信息,即图像的边缘和细节;而 L1 损失函数有助于提高信噪比并较好地保持图像亮度。因此二者混合的损失函数可以强化网络对图像的保留能力,混合损失函数的计算公式为

$$Y_{\text{SSIM}}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (1)$$

$$Y_{\text{MS-SSIM}} = \prod_{j=1}^M Y_{\text{SSIM}}(x_j, z_j), \quad (2)$$

$$L_{\text{SL}} = 1 - Y_{\text{MS-SSIM}}, \quad (3)$$

$$L_1 = \frac{1}{HW} |y - x|, \quad (4)$$

$$L_{\text{mix}} = \alpha L_{\text{SL}} + (1 - \alpha) L_1, \quad (5)$$

式中:  $x$  是网络输出的估计图像;  $y$  是网络学习的无噪声图像;  $\mu_x$  与  $\sigma_x^2$  是输出图像的均值与方差;  $\mu_y$  与  $\sigma_y^2$  是标签图像的均值与方差;  $\sigma_{xy}$  表示标签图像与预测结果之间的方差;  $x_j$  和  $z_j$  为第  $j$  层的局部图像内容;  $M$  为尺度层的数量;  $H$  和  $W$  是图像的高度和宽度; MS-SSIM 损失的占比为  $\alpha$ , L1 损失的占比为  $1 - \alpha$ 。Zhao 等<sup>[22]</sup>的大量研究工作验证出  $\alpha$  设置为 0.84 时能获得较好的损失函数,因此在训练时将  $\alpha$  设置为 0.84。在保证网络架构不变的情况下,使用合适的损失函数,卷积神经网络可以自动学习端到端之间的噪声映射,输出结果的质量也会显著提高。

### 2.3 网络训练参数配置

网络在 AMAX 工作站的 PyTorch 框架下完成训练和测试, AMAX 工作站的 CPU 型号为 Intel Xeon Gold 5118, 可用内存为 64 GB。网络训练和测试使用了型号为 GeForce RTX 2080Ti 的计算显卡。在网络参数训练优化实验中, 批量大小均设置为 16, 共学习了此网络模型中的  $58.6 \times 10^4$  个参数。训练过程迭代轮数共有 40, 网络的训练需要大约 2.5 h。

## 3 实验与结果分析

### 3.1 实验数据集制作

由于仿真的噪声图像与真实扫描出的噪声图像在噪声分布特性上仍具有一定的差别, 因此实验采用由 ZEISS Xradia 510 Versa 3D X 射线显微 CT 扫描的真实青铜钱币图像作为网络训练的数据集, 成像系统如图 4 所示。在数据集的制作过程中, 采用了较为常见的  $512 \text{ pixel} \times 512 \text{ pixel}$  的重建图像尺寸, 扫描样本包含青铜残币与完整的青铜钱币。

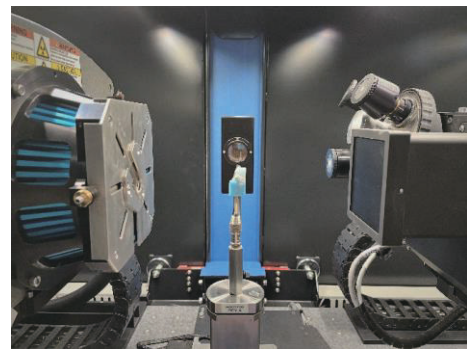


图 4 实验室级显微 CT 系统的组成

Fig. 4 Composition of lab-level micro-CT system



对青铜钱币的扫描参数的设置如表 1 所示,扫描过程中光源到旋转轴的距离(SOD)设置为 80.01 mm,光源到探测器的距离(SDD)设置为 200.01 mm,体素大小为 27.26  $\mu\text{m}$ ,管电压设置为 140 kV,射线源的功率设置为 10 W。仅更改扫描过程中的曝光时间,设置长曝光时间为 3 s,短曝光时间为 0.1 s,通过这种采集方式获取大量无噪声与有噪声的数据对。3 s 曝光时间扫描一组数据需要 5 h,0.1 s 曝光时间扫描一组数据需要 76 min。共采集到 2048 对长曝光时间和短曝光时间的 CT 图像,随机选择 1900 对重建图像用作网络训练集,其余 148 对重建图像用作网络测试集。

表 1 对青铜钱币的扫描参数  
Table 1 Scanning parameters for bronze coins

SOD / mm	SDD / mm	Voxel size / $\mu\text{m}$	Tube voltage /kV	Power / W	Time of exposure /s
80.01	200.01	27.26	140	10	3.0/0.1

### 3.2 青铜钱币的实验结果分析

利用主观和客观的评价指标来检验所提方法的实

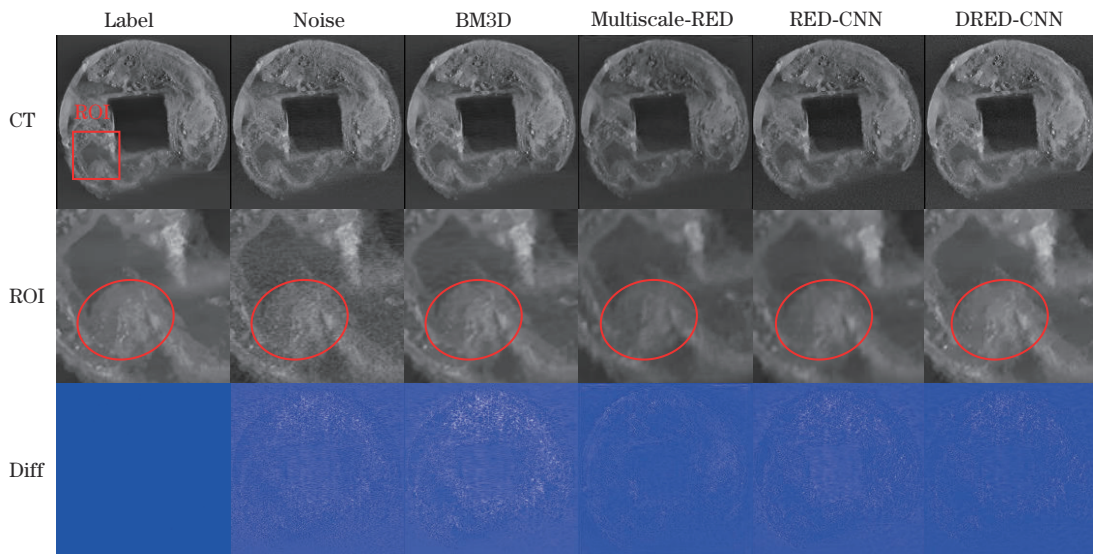


图 5 不同方法对青铜钱币的去噪结果

Fig. 5 Denoising results of different methods for bronze coins

图 6 为不同方法对青铜残币的去噪结果。仍可看出:BM3D 算法与多尺度残差编解码网络去噪结果的模糊现象严重,没有很好地保留真值图像中的细节信息;原始 RED-CNN 去噪后虽然图像清晰度升高,但输出结果较平滑,可以明显地观察到 ROI 中的结构信息丢失。通过比较去噪方法对青铜残币去噪后输出结果与长曝光结果的差值图,也可以看出 BM3D 与原始 RED-CNN 的差值图包含了大量的图像信息,而改进的深层残差编解码网络的差值图包含的图像信息最少。

客观评价中,对所有方法进行量化评估,以验证所提方法的优势。采用结构相似性(SSIM)和峰值信噪比(PSNR)来评价所提方法的性能。SSIM 通过图像的结构信息来衡量两幅图像的相似度,SSIM 值越大

实验效果,通过对比 BM3D、原始 RED-CNN 与所提方法的去噪结果来验证所提方法的有效性。

主观评价如图 5 所示,从左至右分别为青铜钱币的长曝光结果、短曝光结果、BM3D 去噪结果、多尺度残差编解码去噪结果、RED-CNN 去噪结果、所提方法去噪结果。为了更好地分析实验结果,选择了一个感兴趣区域(ROI)进行展示。从图 5 中看出:BM3D 算法的去噪结果中仍带有一些条形伪影,与多尺度残差编解码网络效果类似,边缘均过于平滑,输出结果模糊比较严重;原始 RED-CNN 去噪后虽保留了较多腐蚀程度信息,但仍然存在部分腐蚀信息丢失现象,不利于日后文物工作者对青铜器腐蚀病害的研究。为了更好地比较不同方法在保留细节信息能力方面的差异,还展示了去噪方法输出结果与长曝光结果的差值图,差值图像包含的信息越少,表明去噪算法对信息的保留能力越好。可以看出 BM3D 与原始 RED-CNN 的差值图中包含了较多的图像信息,而改进的深层残差编解码网络的差值图包含的图像信息较少。

表明两幅图像的差值越小;PSNR 通常用来衡量图像的失真程度,PSNR 值越大失真越小。PSNR 的公式为

$$R_{\text{PSNR}} = 10 \cdot \log_{10} \left[ \frac{N \cdot (2^n - 1)^2}{\sum_{i=1}^N (X_i - Y_i)^2} \right], \quad (6)$$

式中: $X$  表示网络的输出估计图; $Y$  表示真值图像。表 2 计算了不同去噪算法输出结果的 PSNR 和 SSIM 数值,对无噪声 CT 图像、有噪声 CT 图像、传统 BM3D 去噪结果、Multiscale-RED 去噪结果、RED-CNN 去噪结果及所提方法去噪后的 CT 图像进行定量指标计算,对指标求均值后的结果如表 2 所示。

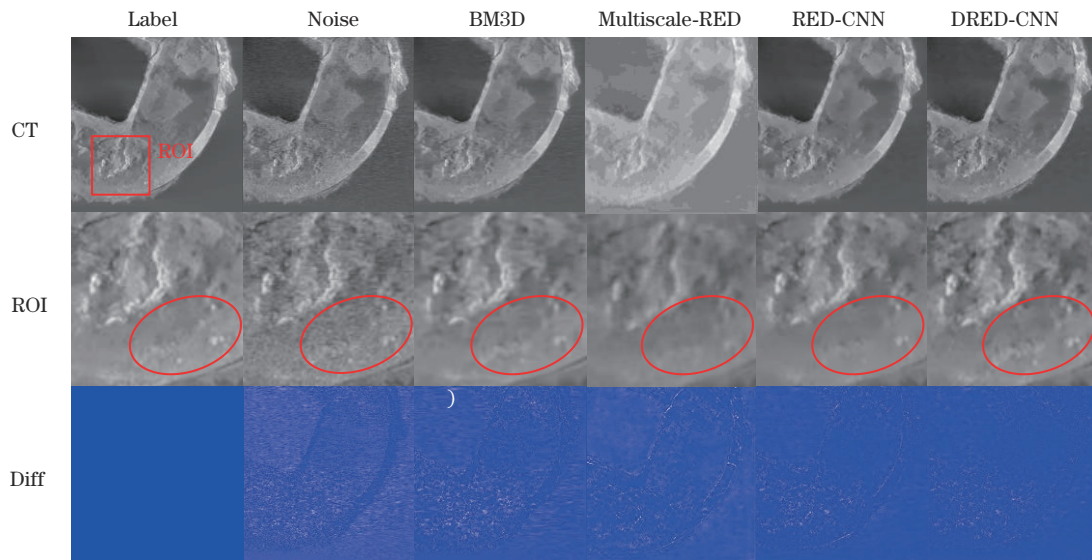


图 6 不同方法对青铜残币的去噪结果

Fig. 6 Denoising results of different methods for bronze residual coins

表 2 不同方法对图像的重建质量评价

Table 2 Reconstruction quality evaluation of different methods for images

Method	Average PSNR	Average SSIM
LDCT	33.8335	0.9163
BM3D	34.1415	0.9508
Multiscale-RED	35.0793	0.9424
RED-CNN	35.8979	0.9644
DRED-CNN	38.8658	0.9745

由表 2 可以看出,对比传统去噪方法 BM3D 与原始的 RED-CNN,所提 DRED-CNN 的客观指标均表现出最优的性能,具有最高的 PSNR 和 SSIM 值。从视觉主观分析与定量指标计算结果可以得出,所提 DRED-CNN 能够在去除噪声的同时,很好地保留青铜钱币的腐蚀锈迹,有效改善了 CT 图像质量。

## 4 结 论

针对实验室级显微 CT 扫描过程会产生噪声,导致重建后 CT 图像质量下降的问题,提出了一个深层多残差编解码卷积神经网络。首先,为了对噪声在图像域数据中的分布特性进行充分学习,加深卷积层数并引入残差学习;其次,使用 L1 损失与 MS-SSIM 损失的混合损失函数,强化网络对图像细节信息的保留能力。此外,为了保留更多的图像特征,所提 DRED-CNN 没有池化层。实验结果表明:DRED-CNN 方法对重建图像的去噪效果显著,主观视觉感知与客观定量指标均得到提升;在噪声抑制、结构保护方面都得到了良好的效果,极大地保留了图像的细节信息,避免了原始方法去噪后的模糊现象;PSNR 和 SSIM 客观指标均优于其他方法,DRED-CNN 的 PSNR 比原方法提高了 8.3%。未来将加入切片间信息,构建三维降噪

网络,更高地提升显微 CT 重建后的成像质量。

## 参 考 文 献

- [1] Hsieh J. Computed tomography: principles, design, artifacts, and recent advances[M]. Bellingham: SPIE, 2009.
- [2] 张朝宗, 郭志平, 张朋. 工业 CT 技术和原理[M]. 北京: 科学出版社, 2009: 2-13.  
Zhang C Z, Guo Z P, Zhang P. Industrial technology and principle of CT[M]. Beijing: Science Press, 2009: 2-13.
- [3] Yao W, Chen L, Wu H, et al. Micro-CT image denoising with an asymmetric perceptual convolutional network[J]. Physics in Medicine and Biology, 2021, 66(13): 135018.
- [4] 闫镔, 李磊. CT 图像重建算法[M]. 北京: 科学出版社, 2014: 1-22.  
Yan B, Li L. CT image reconstruction algorithm[M]. Beijing: Science Press, 2014: 1-22.
- [5] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. IEEE Transactions on Image Processing, 2007, 16(8): 2080-2095.
- [6] Zhu L L, Han Y, Xi X Q, et al. Completion of metal-damaged traces based on deep learning in sinogram domain for metal artifacts reduction in CT images[J]. Sensors, 2021, 21(24): 8164.
- [7] Zhu L L, Han Y, Li L, et al. Metal artifact reduction for X-ray computed tomography using U-net in image domain [J]. IEEE Access, 2019, 7: 98743-98754.
- [8] Liu M N, Han Y, Xi X Q, et al. Multiscale dense U-net: a fast correction method for thermal drift artifacts in laboratory NanoCT scans of semi-conductor chips[J]. Entropy, 2022, 24(7): 967.
- [9] Wolterink J M, Leiner T, Viergever M A, et al. Generative adversarial networks for noise reduction in low-dose CT[J]. IEEE Transactions on Medical Imaging,

- 2017, 36(12): 2536-2545.
- [10] Yang Q S, Yan P K, Zhang Y B, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1348-1357.
- [11] 亢艳芹, 刘进, 王勇, 等. 联合卷积稀疏编码与梯度  $L_0$  范数的低剂量 CT 三维重建[J]. 光学学报, 2021, 41(9): 0911005.  
Kang Y Q, Liu J, Wang Y, et al. Low-dose CT 3D reconstruction using convolutional sparse coding and gradient  $L_0$ -norm[J]. Acta Optica Sinica, 2021, 41(9): 0911005.
- [12] Chen H, Zhang Y, Zhang W H, et al. Low-dose CT denoising with convolutional neural network[C]//2017 IEEE 14th International Symposium on Biomedical Imaging, April 18-21, 2017, Melbourne, VIC, Australia. New York: IEEE Press, 2017: 143-146.
- [13] 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.
- [14] Chen L J, Liang Z, Lian M S, et al. A C-GAN denoising algorithm in projection domain for micro-CT[J]. Molecular & Cellular Biomechanics, 2020, 17(2): 85-92.
- [15] Kang E, Koo H J, Yang D H, et al. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography[J]. Medical Physics, 2019, 46(2): 550-562.
- [16] 朱斯琪, 王珏, 蔡玉芳. 基于改进型循环一致性生成对抗网络的低剂量 CT 去噪算法[J]. 光学学报, 2020, 40(22): 2210002.  
Zhu S Q, Wang J, Cai Y F. Low-dose CT denoising algorithm based on improved cycle GAN[J]. Acta Optica Sinica, 2020, 40(22): 2210002.
- [17] Bengio Y, LeCun Y. Scaling learning algorithms towards AI[J]. Large-Scale Kernel Machines, 2007, 34(5): 1-41.
- [18] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [19] Agarap A F. Deep learning using rectified linear units (ReLU) [EB/OL]. (2018-03-22) [2022-03-05]. <https://arxiv.org/abs/1803.08375>.
- [20] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [21] Zhang L, Zhang L, Mou X Q, et al. A comprehensive evaluation of full reference image quality assessment algorithms[C]//2012 19th IEEE International Conference on Image Processing, September 30-October 3, 2012, Orlando, FL, USA. New York: IEEE Press, 2012: 1477-1480.
- [22] Zhao H, Gallo O, Frosio I, et al. Loss functions for image restoration with neural networks[J]. IEEE Transactions on Computational Imaging, 2017, 3(1): 47-57.