

基于全卷积神经网络的低剂量CT去噪算法

洪启帆¹, 玄祖兴², 李雅馨¹

1. 北京联合大学智慧城市学院, 北京 100101; 2. 北京联合大学数理与交叉科学研究院, 北京 100101

【摘要】针对低剂量医学CT图像因减少辐射剂量而引入大量噪声,导致图像质量显著下降的问题,提出一种基于残差注意力机制和自适应特征融合的低剂量CT图像去噪算法,该算法使用全卷积神经网络来完成低剂量CT图像去噪。在网络框架中引入残差注意力机制和选择性内核特征融合模块,以过滤噪声信息,提取有效特征并自适应地融合图像特征,避免重建过程中的细节损失,提高图像质量,使去噪后的图像在感知上更接近原始图像。定性和定量实验表明,与现有的算法相比,在真实的临床数据集上,所提出的算法能够有效地抑制噪声,并恢复低剂量CT图像中更详细的纹理。与低剂量CT图像相比,所提出的算法将峰值信噪比提高14.94%,结构相似度提高4.68%,均方根误差降低40.11%,可以满足医学成像的诊断要求。

【关键词】低剂量CT;全卷积神经网络;噪声;注意力机制;特征融合

【中图分类号】R318;TP391.4

【文献标志码】A

【文章编号】1005-202X(2023)06-0695-06

Fully convolutional neural network based algorithm for low-dose CT image denoising

HONG Qifan¹, XUAN Zuxing², LI Yaxin¹

1. Smart City College, Beijing Union University, Beijing 100101, China; 2. Institute of Fundamental and Interdisciplinary Sciences, Beijing Union University, Beijing 100101, China

Abstract: A low-dose CT image denoising algorithm based on a residual attention mechanism and adaptive feature fusion is proposed to address the problem of the decline of image quality caused by a large amount of noise introduced by low-dose CT images due to reduced radiation dose. The algorithm uses a fully convolutional neural network to accomplish low-dose CT image denoising. A residual attention mechanism and a selective kernel feature fusion module are introduced into the network framework to remove noise, extract effective features and adaptively fuse image features to avoid detail loss during reconstruction, thereby improving image quality and making the denoised images perceptually closer to the original images. The qualitative and quantitative experiments show that the proposed algorithm can effectively suppress noise and recover more detailed textures in low-dose CT images as compared with existing algorithms on real clinical datasets. Compared with low-dose CT images, the proposed algorithm increases the peak signal-to-noise ratio by 14.94%, improves the structural similarity by 4.68%, and reduces the root mean square error by 40.11%, meeting the diagnostic requirements.

Keywords: low-dose computed tomography; full convolutional neural network; noise; attention mechanism; feature fusion

前言

由于低剂量计算机断层扫描 (Low-Dose

Computed Tomography, LDCT) 图像具有减少 X 射线辐射的潜力,因此 LDCT 问题在社会上得到了重点关注。然而,与正常剂量计算机断层扫描 (Normal-Dose Computed Tomography, NDCT) 图像相比,LDCT 图像在临床应用中存在严重的噪声和伪影,严重影响医生对病情的诊断^[1]。因此,研究抑制 LDCT 图像中的噪声并尽可能多地保留原始图像细节特征的算法具有重大的研究意义^[2]。传统的 LDCT 图像去噪算法大致可以分为两大类^[3]: 图像域和投影域。投影域去噪算法^[4]是在 CT 图像重建之前对投影相关数据进行去噪,然后将去噪后的信息进行反投影运算完成图像重建,该算法借助噪声信息在投影域分

【收稿日期】2022-12-23

【基金项目】北京市优秀人才培养资助青年拔尖个人项目 (2018000026833ZK57);北京联合大学人才强校优选计划 (BPHR2020EZ01);北京联合大学学科团队一体化建设计划项目 (ZB10202001)

【作者简介】洪启帆,硕士,研究方向:图像处理、深度学习等, E-mail: 18810536635@163.com

【通信作者】玄祖兴,博士,教授,研究方向:深度学习、应用数学, E-mail: zuxingxuan@163.com

布的特点,能更加容易分离噪声且计算效率较高^[5],但是当图像中含有分布较为复杂的噪声时,就很难获得清晰的去噪图像^[6-8]。图像域去噪算法^[9]是直接对低剂量CT图像去噪。然而,这种传统的去噪算法难以精准区分图像中的噪声和纹理信息,常常会导致去噪后的图像出现过度模糊的现象。随着深度学习在医学图像处理任务中展现出巨大的潜力,卷积神经网络^[10](Convolutional Neural Network, CNN)和生成对抗网络^[11](Generative Adversarial Networks, GAN)已经成为LDCT图像去噪任务的主流算法框架。Chen等^[12]率先将卷积神经网络应用于LDCT图像的去噪,并且还通过实验证明该算法优于传统的去噪算法。为了实现更高的图像质量,Chen等^[13]随后在传统卷积神经网络^[14]的基础上,提出一种结合残差网络的编码解码框架(Residual Encoder-Decoder Convolutional Neural Network, REDCNN)用于快速重建对应于LDCT的NDCT图像,该网络结合编码和解码的网络结构以及跳跃连接,并取得了更显著的效果。为了解决因数据集小而导致的模型训练不足和拟合不充分的问题,Wolterink等^[15]设计一个基于GAN的LDCT图像降噪模型,其灵感来自于风格迁移CNN,但图像去噪效果不尽人意。为了加强GAN模型的降噪效果,Yang等^[16]提出一个带有感知损失函数的生成对抗网络框架,该算法可以有效抑制LDCT图像中的噪声,同时也能保留图像的轮廓信息。Wang等^[17]提出一种不需要像素配对CT图像的GAN算法框架,将循环一致生成对抗网络^[18]应用于低剂量CT图像去噪任务中,通过均方误差函数不断矫正LDCT与NDCT之间的差异,拟合目标图像的分布函数,从而达到噪声抑制的目的。

现有的深度学习算法虽然优于传统算法,可以有效减少LDCT图像中的噪声,但总的来说,仍然存在一些缺陷,如对上下文信息的使用不足,图像特征的融合简单化,网络模型的表现力有限,以及对纹理的细节处理模糊不清。因此,本文提出一种用于LDCT图像的新型去噪算法,该算法采用了全卷积神经网络的框架,并根据现有算法的问题进行了有针对性的改进,在框架中引入残差注意力机制模块(Residual Attention Module, RAM)和选择性内核特征融合^[19](Selective Kernel Feature Fusion, SKFF)模块,用于噪声信息的过滤和图像高维特征信息的提取,并能有效地结合模型中的上下文信息进行自适应融合,避免丢失细节信息。为了让图像的空间信息不丢失,使用跳跃连接的方式将原始图像的空间信息加入在图像重建过程中,在一定程度上提升了图像质量,去噪图像也获得更好的观感。

1 本文算法

本文提出一种基于残差注意力机制与自适应特征融合的去噪网络(Residual Attention-Denoising Network With Adaptive Feature Fusion, RAFDN),其输入为LDCT图像,输出为去噪图像。RAFDN模型的整体结构和其主要模块RAM、SKFF模块设计如下所示。

1.1 RAFDN结构

RAFDN整体结构如图1a所示,主要由1个3×3卷积、1个1×1卷积、5个去噪模块(Denoising Module, DM)组成。其主要流程首先是输入一组尺寸为512×512×1的LDCT图像(因为图像为黑白图像,所以通道数为1),通过一个3×3卷积将图像通道数变为32通道获得多通道特征图,然后进入去噪模块进行特征提取、噪声去除及特征融合。DM具体结构如图1b所示,其组成成分可分为两个部分,分别是特征提取模块和SKFF模块。图1c所示为特征提取过程,为了得到图像中最为显著的特征信息,使用最大池化将输入的多通道特征图进行下采样。接着在ARM中将这信息中的噪声信息去除,最后通过反卷积操作和ARM的筛选进行特征图的重建,为了防止空间结构信息的丢失,在此过程中加入了前几层的结构信息。在特征融合阶段,将两个特征提取模块的信息进行自适应融合使其具有更加丰富的表达能力。在重复5次相同的操作后,最终采用1×1的卷积对特征图进行通道层面的压缩,从而得到降噪图像,为了避免图像质量的下降,在去噪图像后添加跳跃连接。在模型学习过程中,平均绝对误差(Mean Absolute Error, MAE)作为优化RAFDN算法端到端图像去噪的损失函数,如下所示:

$$L_{MAE} = \frac{1}{m} \sum_{i=1}^m |x_i - \hat{x}_i| \quad (1)$$

其中, x 表示全剂量CT图像, \hat{x} 表示经过模型算法去噪后的CT图像, i 表示图像在组中的编号, m 表示一组图像的数量。

1.2 RAM结构

RAM结构主要可以分为两个部分:特征提取部分和计算加权特征图部分,具体结构如图2所示。在特征提取部分首先通过3个空洞卷积来增加模型的感受野以获得更加丰富的局部信息,然后将其输出的特征图进行通道层面的叠加,最后通过激活函数和7×7的卷积进行特征图的压缩,再添加跳跃连接从而完成特征的提取。加权特征图的计算则是先通过3个残差块来计算特征图之间的关联性,再通过1×1卷积和激活函数获得注意力权重矩阵,将其与输入的特征图相乘再相加,重复两次获得最终的加权特征图。

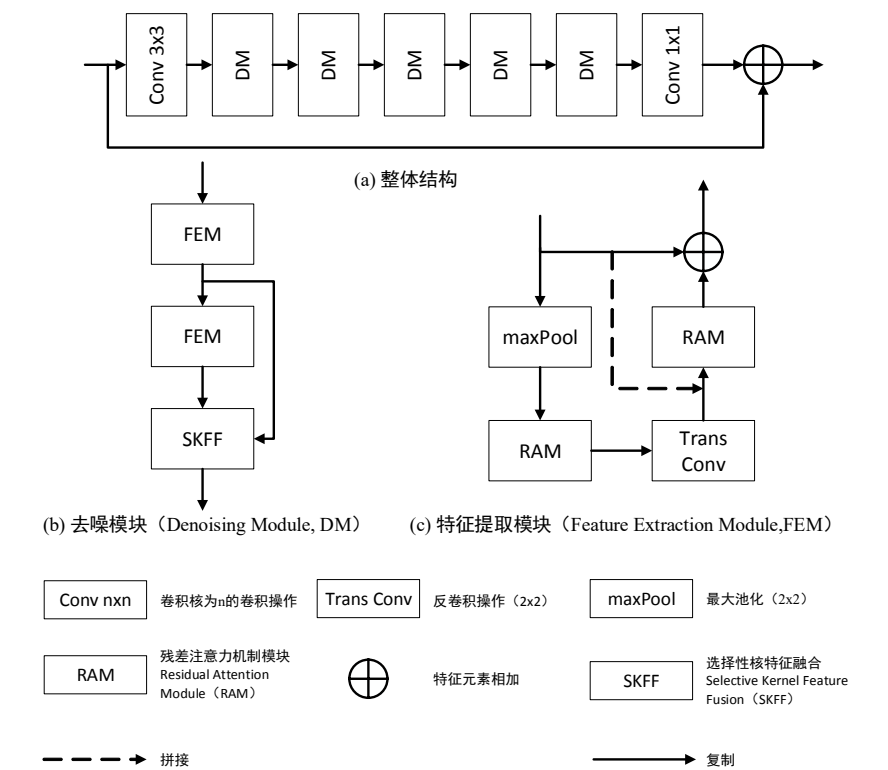


图1 整体模型框架

Figure 1 Framework of the overall model

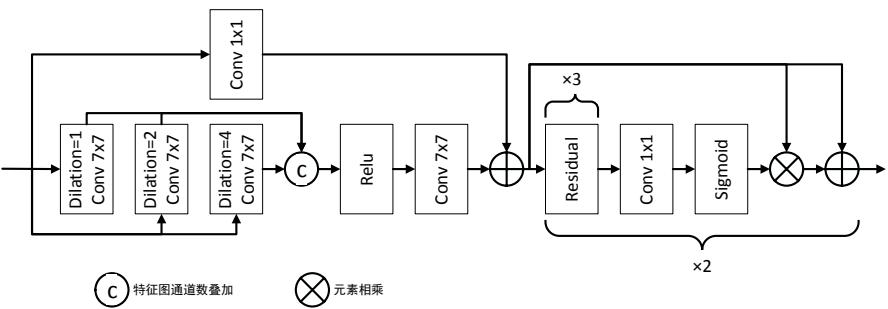


图2 RAM 结构图

Figure 2 Residual attention module structure

1.3 SKFF 结构

常用的特征融合方式包括特征图简单串联和特征图元素直接相加,但是这些方式使网络模型的表达能力受限,因此引入一种自适应特征融合结构SKFF来增加网络模型的非线性能力。SKFF的具体结构如图3所示,首先将输入的特征图进行元素相加,特征图再经过全局平均池化层后尺寸将变为1×1,以减少计算量,便于计算特征相关性,接着通过1×1卷积减少通道数,用激活函数实现非线性映射,然后再用1×1卷积进行通道还原,用Softmax函数计算关联矩阵,经过对应相乘再相加,最终获得自适应融合后的特征图。

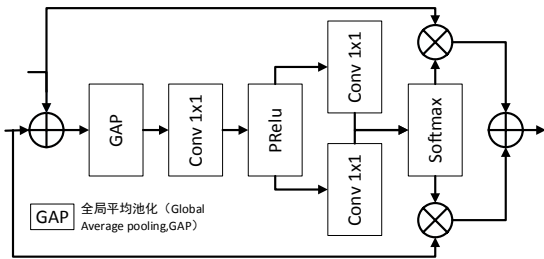


图3 SKFF 模块结构图

Figure 3 Selective kernel feature fusion module structure

2 实验分析与结果

2.1 实验数据集及实验环境设置

实验使用2016年NIH-AAPM-Mayo Clinic LDCT Grand Challenge^[20]公开发布的数据集进行模型训练和测试。该数据集包含2378张厚度为3.0 mm、分辨率为512×512的低剂量(1/4剂量)和正常剂量(全剂量)CT图像,来自10位匿名患者。在实验中,选择患者L506的数据共211张图像进行评估,而其余9名患者共2167张图像用于模型训练。训练时还采用了数据增强的方法,通过随机旋转(90°、180°或270°)和翻转(上/下、左/右)原始图像产生更多的训练图像。

在本次实验中,服务器的中央处理器和显卡分别为Intel 酷睿 i5 11400F @ 2.60 GHz 和 NVIDIA GeForce GTX 3060 12 G。所有模型都是基于深度学习框架Pytorch1.7所实现的,并在集成开发环境Pycharm2020中运行。同时,采用Adam^[21]优化器来优化模型拟合目标分布的过程^[22],其中指数衰减率 $\beta_1 = 0.9$ 、 $\beta_2 = 0.999$ 、学习率 $\alpha = 10^{-4}$ 。训练时将图像尺寸从512×512切割为64×64,批量尺寸大小为4,训练轮数epoch=80,一共迭代43 360次,每3 000次迭代,学习率衰减为原来一半。

2.2 实验结果

在本次实验中,峰值信噪比(Peak Signal-to-Noise

Ratio, PSNR)、结构相似度^[23](Structural Similarity, SSIM)和均方根误差(Root Mean Square Error, RMSE)被作为评价指标,更高的PSNR和SSIM表明图像质量更好,而RMSE越低则图像质量越好。

2.2.1 消融实验 本节将详细讨论各个算法模块的作用,并进行相应的消融实验。实验过程是在实验环境及相关参数不变的基础上,向基础模块依次加入算法模块来测试其对LDCT去噪效果的影响。为了避免实验的随机性,所有实验均重复10次,使用PSNR、SSIM、RMSE来衡量图像质量。为了验证提出模块的有效性,测试的基础模型设置是将提出的RAM用3×3的卷积进行替换,特征融合不使用SKFF模块而使用直接特征相加的方式,测试集的平均结果详细记录在表1中。从评价指标的角度来看,依次加入SKFF模块、RAM,图像质量逐渐提升,相比于基础模型,本文算法重建的去噪图像在PSNR、SSIM分别增长2.85%、1.02%,RMSE降低10.00%。其中RAM提升明显,可见前期的特征提取对最后图像重建影响较大。图4从观感上进行展示,从图中可看出,1/4剂量CT图像存在大量噪声,随着逐步加入算法模块,去噪后的图像细节特征从模糊变得逐渐清晰,总体上评价指标也处于上升的趋势。

表1 各个算法模块对图像去噪效果的影响
Table 1 Effect of each algorithm module on image denoising

| 评价指标 | 基础模型 | 基础模型+SKFF | 基础模型+SKFF+RAM |
|---------|------------------|------------------|------------------|
| PSNR/dB | 32.688 3±1.973 0 | 32.887 8±1.850 1 | 33.618 4±1.881 7 |
| SSIM | 0.907 6±0.029 8 | 0.908 0±0.030 8 | 0.916 9±0.029 8 |
| RMSE | 9.476 1±1.996 7 | 9.247 8±2.009 5 | 8.528 7±1.938 2 |

2.2.2 与其他方法相比 为了进一步证明所提出算法的优越性,将其与目前先进的算法进行对比,这些算法包括REDCNN^[13]、EDCNN^[24]、CTformer^[25],低剂量和正常剂量CT图像作为参考。表2记录了这些算法的去噪效果,其各种指标结果为测试集全部图像相应指标的平均值。首先对实验进行定量的分析,根据实验结果数据,与LDCT相比,本文算法PSNR提高14.94%,SSIM提高4.68%,RMSE降低40.11%。同时对比于其他算法,本文所提出的算法实现了最先进的结果,PSNR、SSIM、RMSE分别达到(33.618 4±1.881 7) dB、0.916 9±0.029 8、8.528 7±1.938 2,各项指标均表现得更为出色,重建出的去噪图像与正常剂量CT图像结构性差异更小。图5从视觉的角度定性地展示了不同算法去噪后图像的差别,通过放大的部分可以看出1/4剂量的CT图像存在大量斑点型噪声严重影响图像质量,EDCNN

和CTformer都能有效地去除图像中的噪声,但在噪声去除过程中也过度模糊了一些边缘轮廓信息,REDCNN对噪声的去噪效果较差,不能很好地处理图像中的噪声。图5f为本文去噪算法的放大图,相比于其他的算法,能更好地保留轮廓细节特征,并具有更高对比度和观感。

3 结 论

本研究针对LDCT图像去噪任务,提出深度学习模型RAFDN,该模型基于全卷积神经网络框架融入RAM和SKFF模块,相较于目前主流的LDCT图像去噪算法有较高的PSNR与SSIM,同时具有更小的RMSE和更好的观感。未来的研究将继续根据LDCT图像的特点,对模型结构进行针对性的优化,进一步提升LDCT图像的质量,帮助医生更准确地诊断疾病。

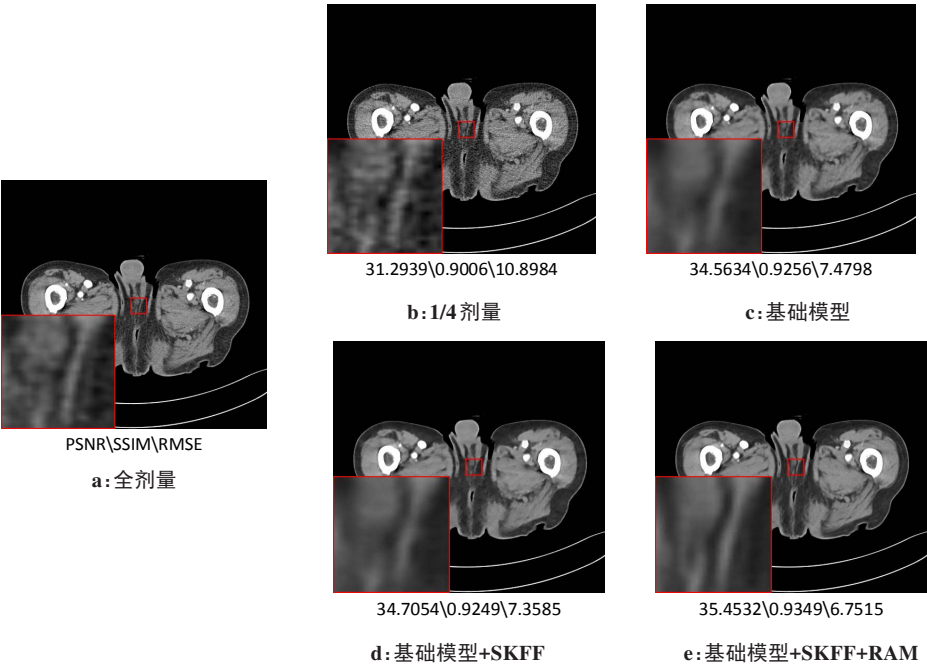


图4 消融实验去噪图像的视觉比较

Figure 4 Visual comparison of denoised images from ablation experiments

表2 与先进算法比较

Table 2 Comparison with advanced algorithms

| 方法 | PSNR/dB | SSIM | RMSE |
|----------|------------------|-----------------|------------------|
| LDCT | 29.248 9±2.110 0 | 0.875 9±0.038 6 | 14.241 6±3.952 3 |
| REDCNN | 32.329 0±1.833 0 | 0.905 3±0.028 8 | 9.863 0±2.084 1 |
| EDCNN | 32.308 0±2.139 9 | 0.907 3±0.031 8 | 9.964 6±2.543 9 |
| CTformer | 33.080 7±1.868 9 | 0.911 9±0.030 4 | 9.071 1±2.054 9 |
| 本文算法 | 33.618 4±1.881 7 | 0.916 9±0.029 8 | 8.528 7±1.938 2 |

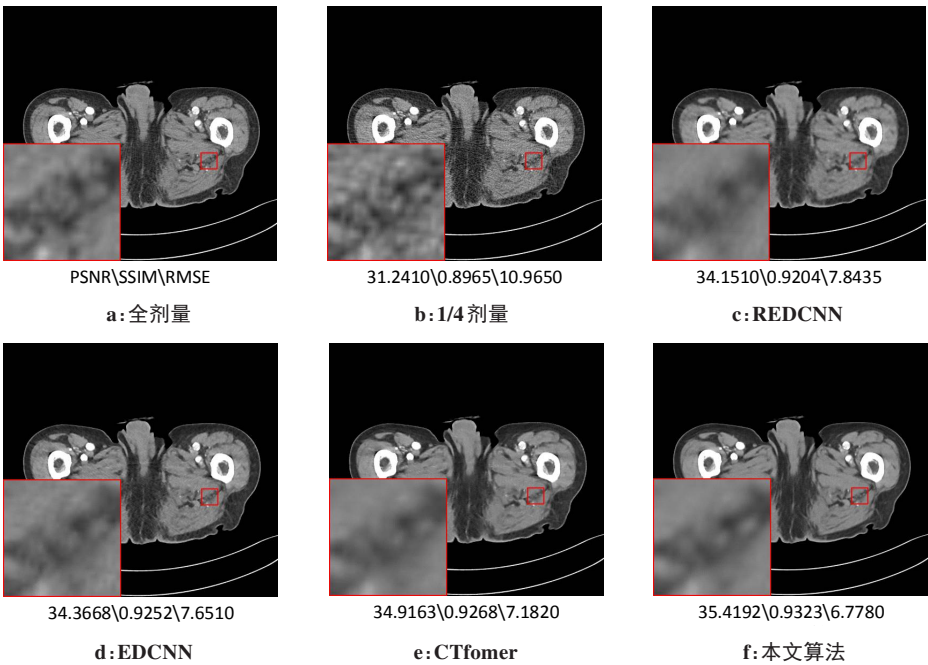


图5 与先进算法的视觉比较

Figure 5 Visual comparison with advanced algorithms

【参考文献】

- [1] Brenner DJ, Hall EJ. Computed tomography-an increasing source of radiation exposure[J]. N Engl J Med, 2007, 357(22): 2277-2284.
- [2] 欧阳婉卿, 张剑, 彭辉, 等. 基于改进生成对抗网络的低剂量CT去噪算法[J]. 光电子·激光, 2022, 33(2): 171-180.
Ouyang WQ, Zhang J, Peng H, et al. A low-dose CT denoising algorithm based on improved generative adversarial networks[J]. Journal of Optoelectronics-Laser, 2022, 33(2): 171-180.
- [3] 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): 70-78.
Zhu SQ, Wang J, Cai YF. Low-dose CT denoising algorithm based on improved recurrent consistent generative adversarial network[J]. Acta Optica Sinica, 2020, 40(22): 70-78.
- [4] Wang J, Li T, Lu H, et al. Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography[J]. IEEE Trans Med Imaging, 2006, 25(10): 1272-1283.
- [5] Whiting BR. Signal statistics in X-ray computed tomography[C]//Medical Imaging 2002: Physics of Medical Imaging. SPIE, 2002: 53-60.
- [6] Hsieh J. Adaptive streak artifact reduction in computed tomography resulting from excessive X-ray photon noise[J]. Med Phys, 1998, 25(11): 2139-2147.
- [7] Lu H, Hsiao T, Li X, et al. Noise properties of low-dose CT projections and noise treatment by scale transformations[C]//2001 IEEE Nuclear Science Symposium Conference Record (Cat. No. 01CH37310). IEEE, 2001, 3: 1662-1666.
- [8] Liu Y, Ma J, Fan Y, et al. Adaptive-weighted total variation minimization for sparse data toward low-dose X-ray computed tomography image reconstruction[J]. Phys Med Biol, 2012, 57(23): 7923.
- [9] Chen Y, Yin X, Shi L, et al. Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing[J]. Phys Med Biol, 2013, 58(16): 5803.
- [10] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proc IEEE, 1998, 86(11): 2278-2324.
- [11] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Commun ACM, 2020, 63(11): 139-144.
- [12] Chen H, Zhang Y, Zhang W, et al. Low-dose CT denoising with convolutional neural network [C]//2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE, 2017: 143-146.
- [13] Chen H, Zhang Y, Kalra MK, et al. Low-dose CT with a residual encoder-decoder convolutional neural network [J]. IEEE Trans Med Imaging, 2017, 36(12): 2524-2535.
- [14] Dong C, Loy CC, He K, et al. Image super-resolution using deep convolutional networks [J]. IEEE Trans Pattern Anal Mach Intell, 2015, 38(2): 295-307.
- [15] Wolterink JM, Leiner T, Viergever MA, et al. Generative adversarial networks for noise reduction in low-dose CT [J]. IEEE Trans Med Imaging, 2017, 36(12): 2536-2545.
- [16] Yang Q, Yan P, Zhang Y, et al. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss[J]. IEEE Trans Med Imaging, 2018, 37(6): 1348-1357.
- [17] Wang T, Lei Y, Tian Z, et al. Deep learning-based image quality improvement for low-dose computed tomography simulation in radiation therapy[J]. J Med Imaging, 2019, 6(4): 043504.
- [18] Zhu JY, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2223-2232.
- [19] Han H, Chung SW, Kang HG. MIRNet: learning multiple identities representations in overlapped speech[J]. arXiv preprint arXiv: 2008.01698, 2020.
- [20] McCollough CH, Bartley AC, Carter RE, et al. Low-dose CT for the detection and classification of metastatic liver lesions: results of the 2016 low dose CT grand challenge[J]. Med Phys, 2017, 44(10): e339-e352.
- [21] Singarimbun RN, Nababan EB, Sitompul OS. Adaptive moment estimation to minimize square error in backpropagation algorithm[C]//2019 International Conference of Computer Science and Information Technology (ICoSNiKOM). IEEE, 2019: 1-7.
- [22] 谢丰雪, 杨帆, 冯维, 等. 基于深度学习的低剂量CT图像去噪[J]. 中国医学物理学杂志, 2022, 39(5): 547-550.
Xie FX, Yang F, Feng W, et al. Low-dose CT image denoising based on deep learning[J]. Chinese Journal of Medical Physics, 2022, 39(5): 547-550.
- [23] Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Trans Image Process, 2004, 13(4): 600-612.
- [24] Liang T, Jin Y, Li Y, 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). IEEE, 2020: 193-198.
- [25] Wang D, Fan F, Wu Z, et al. CTformer: convolution-free Token2Token dilated vision transformer for low-dose CT denoising [J]. arXiv preprint arXiv: 2202.13517, 2022.

(编辑:陈丽霞)