

动态可控残差卷积神经网络的低剂量 CT 图像处理

夏振宇¹, 刘 进^{1,2*}, 亢艳芹^{1,2}, 孙 宇¹, 强 俊¹, 刘 涛¹

1. 安徽工程大学 计算机与信息学院, 安徽 芜湖 241000

2. 东南大学 计算机网络和信息集成教育部重点实验室, 南京 210096

摘 要:针对计算机断层扫描(Computed Tomography, CT)中因采用低剂量扫描方式,导致图像噪声伪影干扰,尤其是不同部位噪声和伪影强度存在较大差异这一问题,提出了一种基于动态可控残差的卷积神经网络(DC-ResNet)算法。DC-ResNet 的主要思想是在常规残差网络连接中添加一个图像质量指导的控制变量,以允许获取残差的加权和,从而实现残差特征的动态可控。所设计的 DC-ResNet 网络是一种包括两个子网络的组合型结构,一个是作为主干网络的基础子网络,在该子网络中使用全局动态残差块和局部动态残差块来实现低剂量 CT 图像质量的提高;另一个是作为辅助网络的条件子网络,用来生成基础子网络中不同动态可控残差块的权值,辅助基础子网络的学习。通过 Mayo 与 UIH 数据实验验证,其视觉结果表明:处理后的不同部位 CT 图像噪声伪影均能够得到较好的抑制,并能有效地保留结构细节及组织纹理;量化结果表明:处理后的 CT 图像峰值信噪比(Peak-Signal to Noise Ratio, PSNR)和结构相似性(Structure Similarity, SSIM)均优于对比方法。

关键词:低剂量 CT; 动态可控残差; 条件子网络; 基础子网络

中图分类号:OTN911.73;TP391 文献标识码:A doi:10.16055/j.issn.1672-058X.2023.0002.010

Low Dose CT Image Processing Based on Dynamic Controllable Residual Convolution Neural Network

XIA Zhenyu¹, LIU Jin^{1,2*}, KANG Yanqin^{1,2}, SUN Yu¹, QIANG Jun¹, LIU Tao¹

1. School of Computer and Information, Anhui Polytechnic University, Anhui Wuhu 241000, China

2. Key Laboratory of Computer Network and Information Integration of Ministry of Education, Southeast University, Nanjing 210096, China

Abstract: Aiming at the problem of image noise and artifact interference caused by the low-dose scanning method in Computed Tomography (CT), especially the large difference in noise and artifact intensity in different parts, a convolutional neural network algorithm based on dynamic controllable residual (DC-ResNet) was proposed. The main idea of DC-ResNet is to add an image quality-guided control variable to the regular residual network connection to allow a weighted sum of residuals to be obtained, thereby achieving dynamic controllability of residual features. The designed DC-ResNet network was a combined structure including two sub-networks: one was the basic sub-network as the backbone network, in which the global dynamic residual block and the local dynamic residual block were used to achieve the improvement of low-dose CT image quality; the other was a conditional sub-network, as an auxiliary network, which was used to generate the weights of different dynamically controllable residual blocks in the base subnet and assist the learning of the base subnet. Through the experimental verification of Mayo and UIH data, the visual results showed that the noise and artifacts of the processed CT images in different parts were well suppressed, and the structural details and tissue texture were effectively preserved. The quantification results showed that the peak-signal to noise ratio (PSNR) and structural similarity (SSIM) of the processed CT images were better than those of the comparison methods.

Keywords: low-dose CT; dynamic controllable residuals; conditional sub-network; basic sub-network

收稿日期:2022-03-05 修回日期:2022-05-18 文章编号:1672-058X(2023)02-0064-09

作者简介:夏振宇(1997—),男,安徽合肥人,硕士研究生,从事计算机视觉图像处理方面的研究。

引用格式:夏振宇,刘进,亢艳芹,等. 动态可控残差卷积神经网络的低剂量 CT 图像处理[J]. 重庆工商大学学报(自然科学版), 2023,40(2):64—72.

XIA Zhenyu, LIU Jin, KANG Yanqin, et al. Low dose CT image processing based on dynamic controllable residual convolution neural network[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2023,40(2):64—72.

1 引言

计算机断层扫描(CT)是当前临床应用中的主要医学成像方法,且应用场景多。检查者进行CT扫描会受到X射线辐射,会对身体健康带来一定的风险,为此,CT扫描中的X射线伤害一直以来受到广泛的关注^[1]。当前,降低辐射剂量最常用的方法是减少X射线管工作电流和照射时间。一般情况下,X射线通量越弱,重建的CT图像中噪声伪影强度越大,且不同人体部位出现的伪影和噪声强度不同,条形伪影及斑点状噪声与人体组织结构有一定的关联,影响诊断效果^[2]。为解决这一固有物理问题,人们提出了许多方法来改善低剂量CT(Low dose CT, LDCT)的图像质量。这些算法大致可分为3类:投影域去噪^[3-4]、迭代重建^[5-6]和图像后处理^[7-8]。投影域去噪算法是直接对扫描获得的投影数据进行去噪,随后利用传统滤波反投影算法进行重建^[9],但此类方法处理后容易出现明显的图像分辨率丢失现象。迭代重建类算法,主要思想是结合噪声的统计特性及重建图像的先验信息,实现图像的迭代估计及噪声的伪影抑制^[10]。该类方法能充分利用投影域噪声分布的统计规律及先验模型,但由于CT扫描过程中包含的噪声统计特性复杂,重建图像多样,超参数多,计算量大,导致该类方法很难在临床实际中得到广泛应用。

随着图像处理技术的发展,LDCT的后处理方法逐渐成为热点,该类方法不依赖于原始数据,现有的技术大多可直接应用于LDCT图像处理。如Kang等^[11]使用块匹配3D(Block Method of 3-Dimension, BM3D)算法处理LDCT图像,Ma等^[12]把非局部均值(Non-Local Means, NLM)算法应用于CT图像去噪,Chen等^[13]受稀疏表示理论的启发,采用了一种快速字典学习进行LDCT图像去噪,这些传统基于模型的后处理方法能在一定程度上减少噪声和伪影,但这类方法难以在图像解剖细节、组织纹理特征保持与噪声伪影抑制之间取得平衡,在不同密度噪声与伪影的处理过程中,往往会出现过度平滑或者噪声伪影残留等现象。

伴随大数据集和大样本库的建立,深度学习技术在低剂量CT图像处理中也得到了全面应用。2017年,Chen等^[14]首次将卷积神经网络应用于LDCT图像去噪,初步实验结果表明该网络与传统方法相比拥有一定的优势;He等^[15]提出的ResNet通过残差学习解决了深度网络造成梯度爆炸和梯度消失的退化问题,让研究人员可以训练出更深的网络;随后,Zhang等^[16]提出一种DnCNN网络,采用卷积式端到端的残差学习,从函数回归角度将噪声图像中的噪声分离,其残差结构大幅提升了网络的去噪性能。此后,一系列基于卷

积神经网络结构的改进工作被相继提出,并逐步应用于LDCT图像处理中。如Chen等^[17]提出残差编解码网络(RED-CNN),该网络采用基于对称跳跃链接的深度卷积编解码框架,可极大程度实现LDCT图像噪声伪影的去除;此外,生成对抗网络(Generative Adversarial Network, GAN)也开始应用于LDCT图像处理,该方法缓解了LDCT图像处理后失真的问题^[18];此外,Yang^[19]在GAN网络的基础上设计了一种WGAN网络,并用于LDCT图像处理,实验结果表明该网络更好地保留了图像细节。上述方法虽然可以有效在LDCT图像的处理上取得较好的效果,但在针对不同部位出现不同强度伪影和噪声的情况时,需要针对不同噪声伪影水平数据进单独训练模型,缺乏灵活性,难以直接应用于退化过程复杂的真实LDCT图像处理中。

为解决上述问题,在常规ResNet基础上添加基于图像质量的控制权值,构建成动态残差块,以处理不同部位不同退化程度的LDCT图像。采用改进的动态可控残差网络(DC-ResNet),该网络包括两部分:基础子网络用于LDCT图像处理,采用权重动态可控制的残差连接;条件子网络通过图像退化程度参数来估计可控权重。利用基础子网络和条件子网络的组合优势,构建端到端的LDCT图像处理模型,可有效缓解图像处理过程中易出现细节丢失和过度平滑的问题,进一步提高处理后图像的质量。

2 基本原理

假设 $P=R^{m \times n}$ 为对应的LDCT图像, $Q=R^{m \times n}$ 为对应理想高质量CT图像,其中 m 和 n 分别代表CT图像的高度和宽度。如式(1)所示,LDCT图像退化过程可以表示为

$$P=T(Q) \quad (1)$$

其中, $T(\cdot)$ 表示低剂量扫描时,CT图像退化的非线性函数。在LDCT图像处理过程中,LDCT图像的噪声和伪影分布较为复杂,同时,不同人体部位的低剂量CT图像呈现的伪影和噪声水平具有较大差异。如图1所示:肩部和肺部出现比较密集的条形伪影,但腹部分布不同水平的斑状噪声。针对这一现象,传统的基于模型的方法难以做到噪声伪影抑制及细节保留的统一。与传统方法相比,基于数据驱动的深度学习方法能更好地对多种类型的数据及模式进行统计特性组合和建模。为此,基于卷积神经网络LDCT图像处理方法,就是通过训练,获取图像 $P \rightarrow Q$ 之间映射关系。

针对上述问题,本文通过输入层、输出层及动态可控残差模块的组合,构建一种用于LDCT图像处理的动态可控残差网络DC-ResNet,以实现对于LDCT图像中不同退化强度的伪影和噪声的抑制。

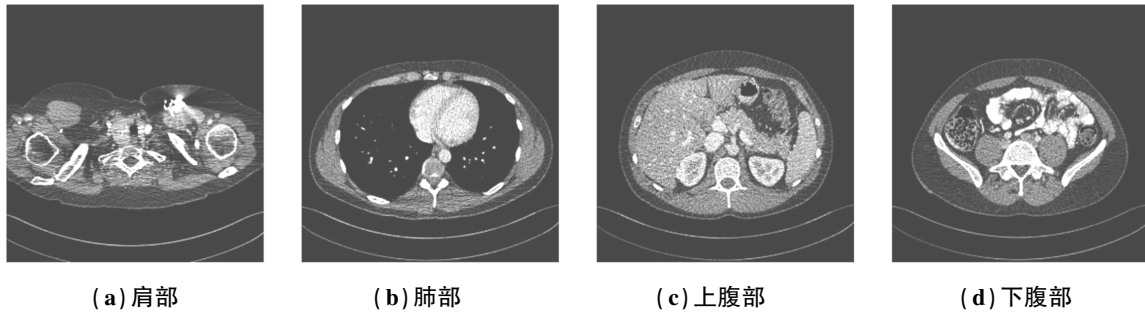


图 1 同一剂量下不同部位的 LDCT 图像

Fig. 1 LDCT images of different body parts for the same dose

2.1 动态可控残差

ResNet 可以在一定程度上减缓网络训练时的梯度退化现象,该网络不再使用多个直接堆叠的卷积层来直接实现特征映射,而是采用拟合特征残差的形式。如式(2)所示,单个残差块可表示为

$$Y = F(X, W_i) + X \quad (2)$$

其中, X 和 Y 分别为残差块的输入和输出, $F(X, W_i)$ 是残差块需学习的映射函数。通过引入跳跃连接,实现其恒等映射,在图像处理及分析中具有较大优势图 2(a)。但传统的残差块恒等映射形式单一,难以平衡不同特征的权重,或者说不同特征权重一致,不适用于退化复杂的 LDCT 图像处理。为此,本文将在常规残差连接上添加控制权值,建立一种动态可控的残差模块,可适应不同强度和形式的特征映射图 2(b)。

在保留传统残差特点的基础上,动态可控的残差模块通过引入可控制剩余连接的权值 α , α 的范围为 0 到 1, 实现残差映射 $F(X, W_i)$ 加权求和,模块输出 Y 的大小也变为 X 到 $F(X, W_i) + X$ 之间。如式(3)所示,动态残差块表示为

$$Y = F(X, W_i) \times \alpha + X \quad (3)$$

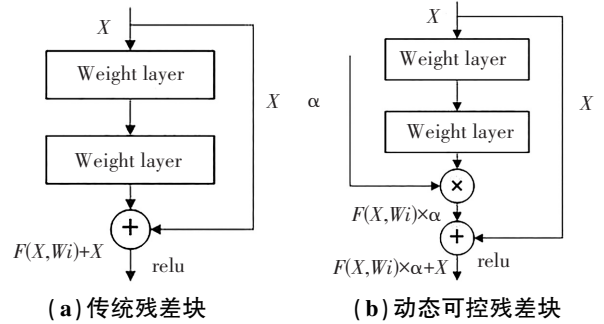


图 2 残差块结构图

Fig. 2 Architectures of residual blocks

2.2 DC-ResNet 网络结构

针对传统神经网络方法不能很好去除同一剂量下不同水平伪影和噪声这一问题,本文在上述动态可控残差块思想指导下,设计了一种组合型的动态可控残差网络 (DC-ResNet), 网络结构如图 3 所示。DC-ResNet 采用基础子网络和条件子网络两部分的联合训练方法,实现了 LDCT 图像质量的提高,可以降低附加参数。基础子网络为主干网络,用于 CT 图像处理;条件子网络为辅助网络,用来生成基础子网络中不同动态可控残差块的权值。

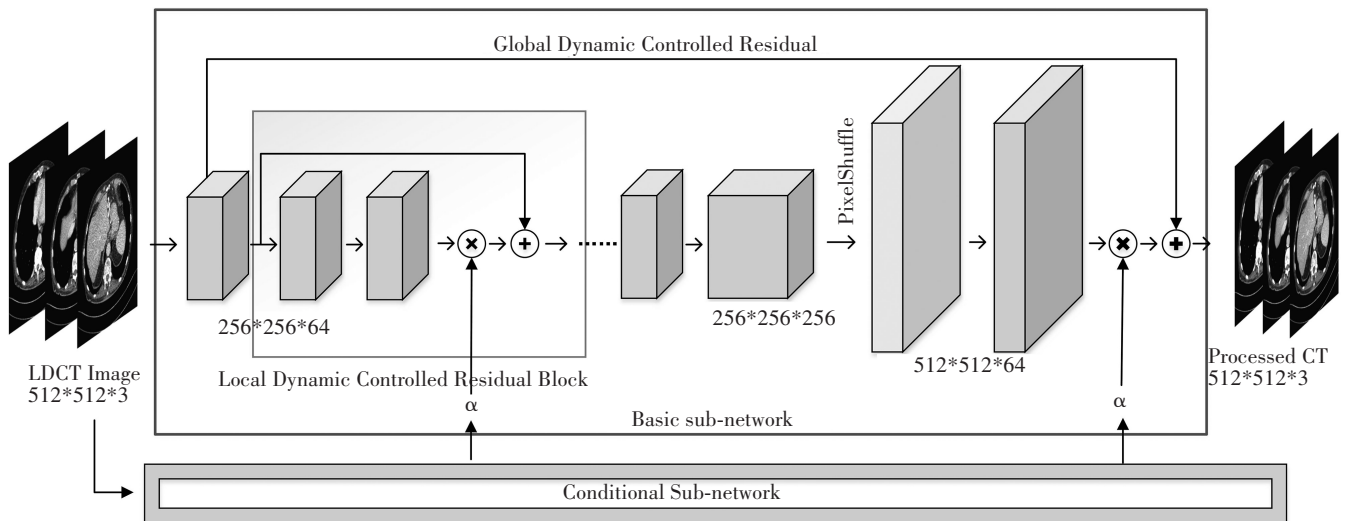


图 3 DC-ResNet 网络结构图

Fig. 3 DC-ResNet network architecture

动态残差块的基础子网络将使用两种不同形式的动态可控残差。

全局动态可控残差:当改变 α 值时,输出将表现为不同程度的处理效果。当 $\alpha = 1$ 时,网络输出处理后的图像,可实现单位映射;当 $\alpha = 0$ 时,禁用残差分支。此外,基础子网络主要考虑学习低剂量和高质量图像之间的高频残差信息,而大部分区域的残差特征接近于零,可避免训练数据的全部信息,降低模型的复杂度和学习难度。

局部动态可控残差:在每个网络模块上增加局部动态可控残差连接,可实现特定强度及类型的噪声和伪影特征提取。通过动态控制这些网络模块的剩余连接,实现不同噪声和伪影的抑制。此外,这里的局部动态可控残差用来缓解网络深度增加带来的模型退化问题,同时可自由增加网络模块数量以适应不同任务。

基础子网络:基础子网络主要采用 EDSR (Enhanced Deep Residual Networks for Single Image Super-Resolution) 形式的结构^[20],卷积层均采用大小为 3×3 的小卷积核,特征图均采用边界 0 填充的形式,以确保前后数据维数相同。网络输入大小为 $512 \times 512 \times 3$ 的 CT 图像序列,通过图像层间连贯性增加数据的冗余。图 3 可以看出:先使用全局动态可控残差块学习 LDCT 图像和标签 CT 图像间的高频残差,再经过上采样层进行特征提取得到 $512 \times 512 \times 64$ 的特征图,接下来送入 32 个局部动态可控残差块进行学习,将其再进行一次维度上的上采样之后产生 $512 \times 512 \times 256$ 的特征图,采取步长为 2 的 Pixel-Shuffle 方法生成 $512 \times 512 \times 64$ 的高分辨率特征图,最后经过全局动态可控残差块得到网络处理后的 CT 图像。

条件子网络:条件子网络由全连接网络组成,即在基础子网络中的每一个局部动态可控残差块和全局动态可控残差块部位增加一个全连接层,获得动态可控权值,其数量与残差块数量相同。因同一剂量下不同部位 CT 图像的噪声和伪影分布级强度不同,同时为提高基础子网络的特征提取能力,通过输入低剂量 CT 图像的噪声和伪影的水平估计值(该水平值组成条件向量信息输入条件子网络中),来学习权值向量 α (α 的值在 0~1 之间,且维度大小与残差块保持一致),辅助基础子网络的学习。

3 结果讨论与分析

3.1 实验设计

为了验证本文所设计的 DC-ResNet 网络的有效性,实验选用 BM3D、NLM 和 RED-CNN 3 种方法对比,其中 BM3D 和 NLM 是非深度学习的传统去噪方法,RED-CNN 是一种基于传统残差架构的神经网络,具有

优异的 LDCT 图像处理性能。实验将采用 Intel (R) Core(TM) i5, 3.0GHz, 32 GB 内存, NVIDIA GTX TITAN X 显卡,为加速模型收敛,所有训练及测试均是在显卡上完成的。实验采用自适应矩估计优化器(Adam)来训练网络,其中 $\beta_1 = 0.9$, $\beta_2 = 0.999$; BatchSize 大小为 16, 学习率是 1×10^{-4} ; 训练中采用像素级 L_1 损失函数 L , 如式(4)所示:

$$L(x_i, y_i; \Theta) = \sum_{i=1}^n |x_i - y_i| \quad (4)$$

其中, x_i 和 y_i 分别是网络第 i 个的常规剂量 CT 图像数据(标签)与对应的输出估计的 CT 图像, n 为样本数量,它是把真实目标值 x_i 与估计值 y_i 的绝对差值的总和(L)最小化, Θ 为网络参数。 L_1 损失表示目标值与估计值的绝对差值和,可更好地处理数据中的异常值,能够使得网络训练的结果拥有较好的鲁棒性,提升网络训练的收敛性。

实验结果将从主观视觉效果和客观量化指标两个方面来进行评价分析。客观量化采用 R_{PSNR} 和 R_{SSIM} 两个指标来评估处理后 CT 图像的质量, R_{PSNR} 和 R_{SSIM} 数值越大,图像量化质量越高。 R_{PSNR} 与 R_{SSIM} 值的计算如式(5)、式(6)和式(7)所示:

$$R_{\text{PSNR}} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_x^2}{R_{\text{MSE}}} \right) \quad (5)$$

$$R_{\text{SSIM}} = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (6)$$

$$R_{\text{MSE}} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [x(i, j) - y(i, j)]^2 \quad (7)$$

其中, x 和 y 分别为常规剂量 CT 图像(标签)与模型处理后的 CT 图像, $m \times n$ 是图像的宽高, x_{max} 为图像 x 中最大像素值。其中 μ_x 和 μ_y 分别表示两个图像局部均值,用来估计亮度; σ_x^2 和 σ_y^2 分别表示两个图像的局部区域方差,用来估计对比度; σ_{xy} 为协方差,用来估计结构相似度; c_1 和 c_2 为常数。

3.2 数据选取

为评估所提网络模型的处理性能,实验使用两种不同的数据集进行验证。为保证实验的公平性,两种数据集中低剂量数据均为 0.25 倍常规剂量,图像大小保持一致,且进行相同的预处理。同时为了将相邻 CT 图像间存在的冗余信息进行充分利用,以增加局部特征,实验将所有的 CT 图像制作成 $512 \times 512 \times 3$ 的数据块输入基础子网络中。

通过由 Mayo Clinic 授权的“2016 年 NIH-AAPM-Mayo LDCT 图像处理比赛”的数据集(简称 Mayo 数据)进行实验,完成对 DC-ResNet 网络性能的评估。该数据集中包括与 10 名匿名患者匹配的 CT 图像(即每一

幅常规剂量 CT 图像都有与之对应的 LDCT 图像),其中 LDCT 图像为通过模拟 0.25 倍常规剂量情况下的噪声伪影而获得。选取 9 位病人共 5 326 张匹配的 CT 图像作为训练集,再采用区别性特征表示(Discriminative Feature Representation, DFR)的方法来标定每张 LDCT 图像的噪声及伪影强度,并归一化标定值^[21],最后将 L333 号病人的 310 张 LDCT 图像作为测试集。

为补充 Mayo 数据集患者中缺少其他部位的特征信息,及验证不同数据集下的结果,实验增加联影医疗(United Imaging Healthcare)公司授权的 CT 图像数据集(简称 UIH 数据),以进一步验证 DC-ResNet 网络的有效性。该数据集包含 16 位患者的 LDCT 图像及与对之应的常规剂量 CT 图像,实验选取含有肩部图像患者的 390 张 LDCT 图像作为测试集,剩下的 5 681 张匹配的 CT 图像作为训练集,进行与 Mayo 数据相同的预处理,并标定噪声伪影强度,以完成实验的训练与验证。

3.3 Mayo 数据结果

测试数据集中,选取肺部、腹部和胯部 3 个不同部

位的横断面 LDCT 图像处理后的结果进行比较,如图 4 所示,其中显示窗为 $[-150\text{HU}, 250\text{HU}]$,参考图为常规剂量 CT 图像。从视觉效果上来看,实验选取的 4 种方法均可实现不同程度的噪声伪影抑制。通过比较可以发现:传统 BM3D 处理和 NLM 处理后,CT 图像中仍然存在部分噪声、伪影残留和过平滑现象,如图 4(c)、图 4(d)的局部放大图所示;而采用深度学习的 RED-CNN 对 LDCT 图像的噪声和伪影抑制效果要优于 BM3D,但处理后的 CT 图像中仍出现了不同程度的噪声和伪影残留,如图 4(e)中腹部图像区域所示。DC-ResNet 网络考虑到不同部位 CT 图像的退化差异,处理效果均优于对比方法。从放大的局部区域看:如图 4(f)所示肺部区域所示,DC-ResNet 在该部位抑制条形伪影的效果最为出色;如图 4(f)中腹部图像区域,DC-ResNet 在保证降低噪声强度的同时,保留图像细节,肝静脉血管纹理清晰,图像的均匀度保存更加完整;如图 4(f)胯部图像区域所示,处理后的图像拥有清晰的轮廓和细节,对比度较高,显著优于对比方法。

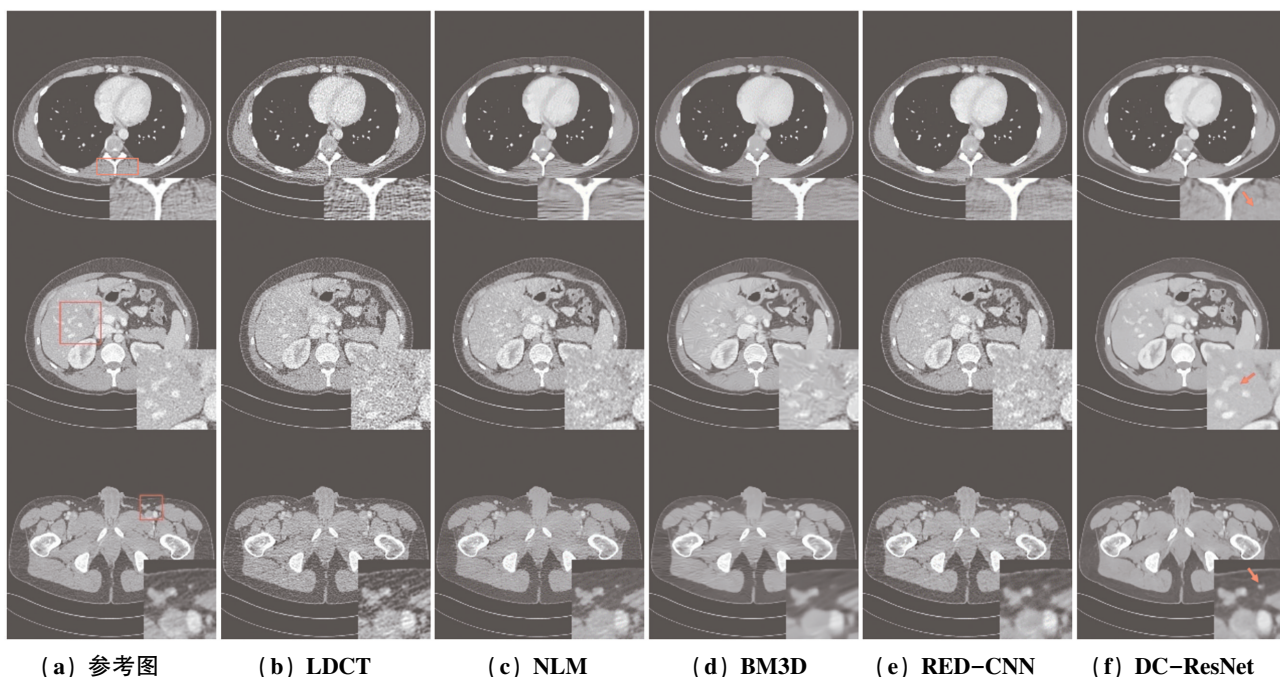


图 4 Mayo 肺部、腹部、胯部横断面 LDCT 图像处理结果

Fig. 4 Processing results of Mayo LDCT images of cross-sections of lung, abdomen, and crotch

实验各选取一张冠状面和矢状面图像,比较 LDCT 图像数据在不同切面上的处理效果,其结果如图 5 所示。从图中可以看出:对比方法虽然在一定程度上抑制了噪声位移,但整体图像质量不高,部分细节丢失,同时在胯部位置中存在伪影残留,而本文方法显著降低了各部位噪声和伪影,同时也更好地保留了图像组织区域的边缘特征,处理后的图像具有较好的视觉效果。

实验采用差异图的形式以比较不同方法的处理效

果,结果如图 6 所示。差异图由处理后的图像与常规剂量 CT 图像相减得到,差异图反映处理后的图像与常规剂量 CT 图像之间的接近程度。从图 6 中第 2—4 列可以看出:处理后的图像残留伪影较多,与常规剂量 CT 图像差异较大,而本文方法的差值图(6(e))仅余下部分伪影且整体较均匀平滑,显著优于其他处理方法。综上,从视觉效果来看,本文方法的处理效果要优于对比方法。

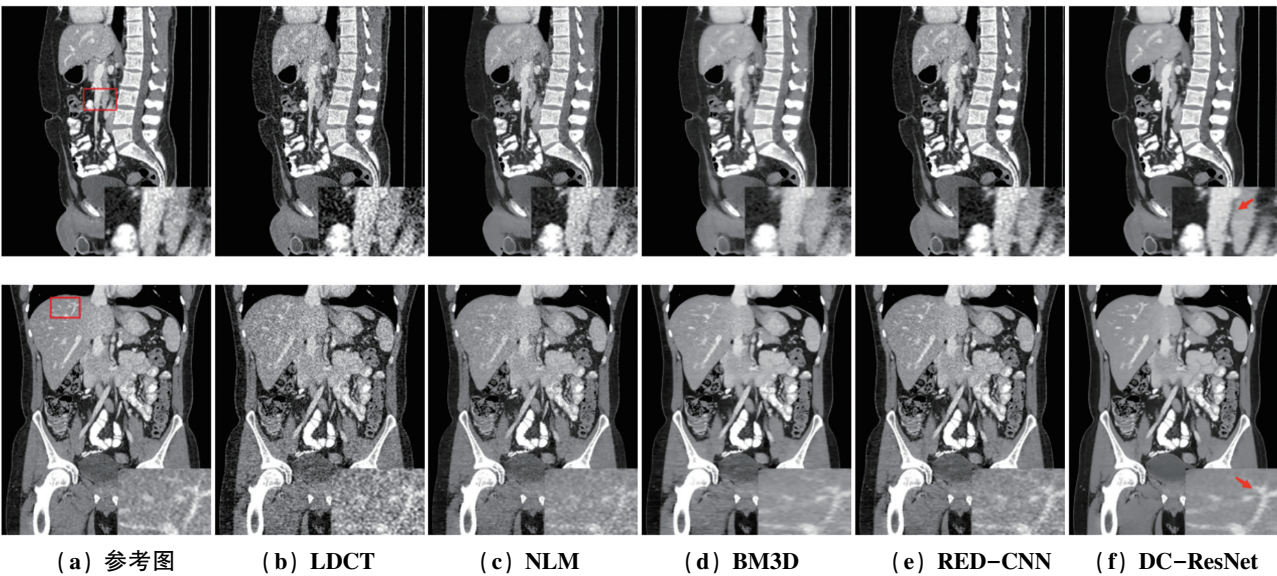


图 5 Mayo 矢状面和冠状面 LDCT 图像处理结果

Fig. 5 Processing results of Mayo LDCT images in sagittal and coronal planes

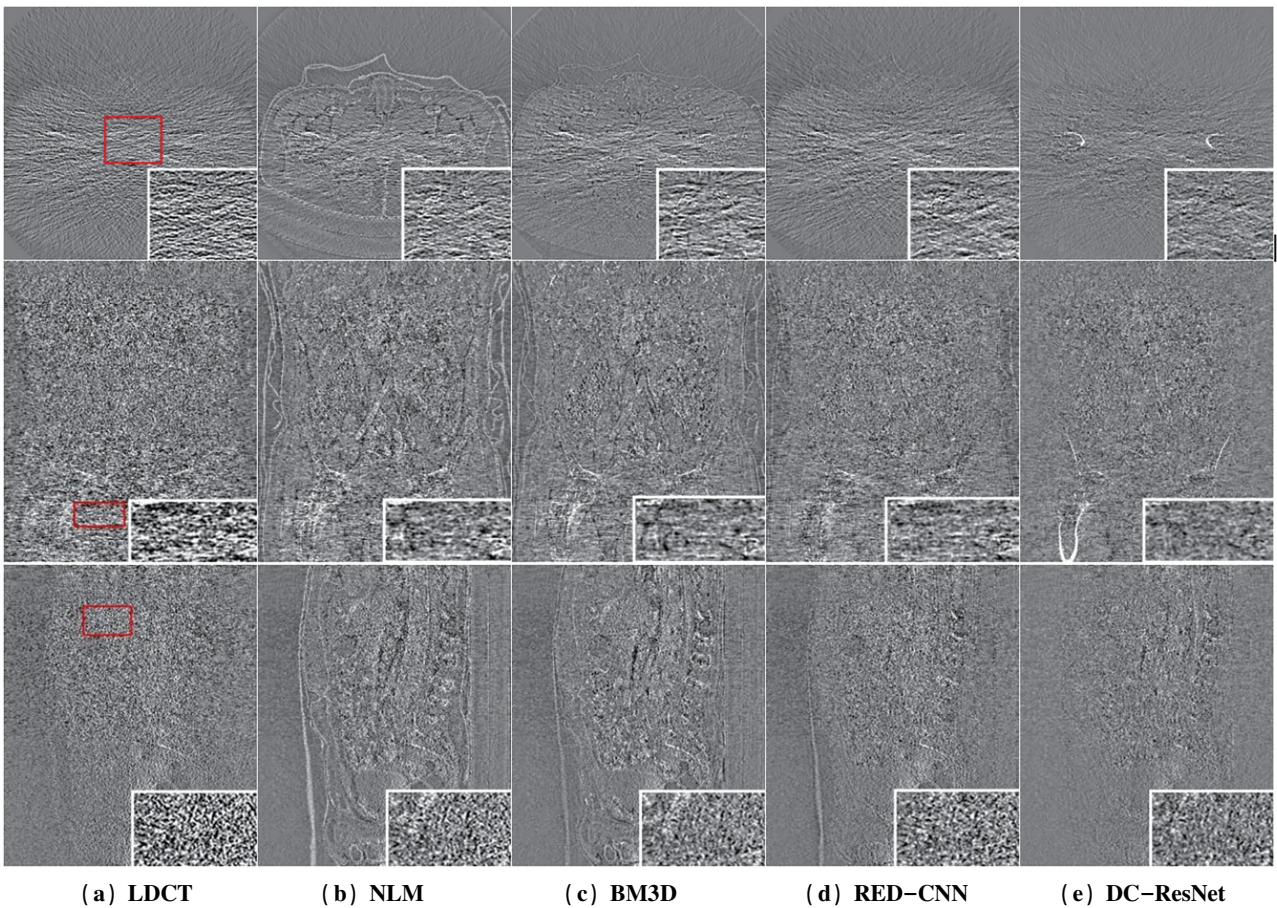


图 6 Mayo LDCT 图像处理结果的差值图

Fig. 6 The difference images relative to the reference image (RDCT FBP image) for Mayo LDCT images

为更加直观及定量地比较 Mayo 数据中不同算法的处理效果,本文采用 R_{PSNR} 和 R_{SSIM} 两个指标来评估实验结果。表 1 为胸部、腹部及胯部 LDCT 图像处理结果的 R_{PSNR} 和 R_{SSIM} 值,可以发现,含有较多噪声和伪影的腹部和胯部 BM3D 和 RED-CNN 处理结果量化值差距较小,而且在所选部位以及所有图像中,DC-ResNet 的量化值均高于对比方法,具有明显的量化优势。

表 1 不同方法使用 Mayo 测试数据处理后的量化结果

Table 1 Quantitative results associated with different methods for the images in the Mayo testing dataset

		LDCT	NLM	BM3D	RED-CNN	DC-ResNet
胸部	R_{PSNR}	38.053 5	39.966 8	41.219 8	41.959 0	43.468 0
	R_{SSIM}	0.844 3	0.893 0	0.913 8	0.928 3	0.947 3
腹部	R_{PSNR}	35.900 4	39.619 1	40.448 7	40.783 0	42.013 4
	R_{SSIM}	0.758 5	0.861 0	0.880 8	0.892 2	0.914 0
胯	R_{PSNR}	37.939 7	41.382 6	42.161 7	42.327 6	44.111 3
	R_{SSIM}	0.829 8	0.912 2	0.919 1	0.927 6	0.947 7
整体	R_{PSNR}	38.080 2	40.946 5	42.200 6	42.331 9	43.729 5
	R_{SSIM}	0.830 5	0.908 4	0.919 1	0.927 5	0.944 1

3.4 UIH 数据结果

为评估 DC-ResNet 网络中 LDCT 图像处理方法在真实数据中的性能,实验进一步采用 UIH 数据来验证其有效性。实验选取 3 个不同部位的轴向图进行比较,结果如图 7 所示。由图 7(b)可看到:在低剂量扫描下,肩部图像相较于其他部位存在严重的条形伪影,处理难度增大。从处理结果来看,本文方法与优秀的

RED-CNN 相比,具有更优的视觉效果,图像边界细节信息也得到了很好的保留,没有出现组织细节的丢失和边缘特征的扩散,更有利于诊断。

所选取的冠状面和矢状面处理结果如图 8 所示。从图 8 可以看到:在 LDCT 冠状图中黑色静脉血管区域已经难以辨认,组织和细微结构也很难诊断,DC-ResNet 取得了较处理好处理结果。图 9 为各方法结果与常规剂量 CT 图像之间的差异图,对比 LDCT 图像与不同方法处理后的图像差异图,可以发现本文方法处理后,差异图强度较小,这表明处理后的图像更加接近常规剂量 CT 图像。

表 2 为 UIH 数据经过不同方法处理后的 R_{PSNR} 和 R_{SSIM} 值,可以看到,在胸部和腹部图中,DC-ResNet 处理后图像量化指标最高;而在胯部图中,本文方法的量化指标略低于 RED-CNN 处理后的量化指标,但在视觉效果上更加具有临床诊断意义。综上,本文所提的 DC-ResNet 网络具有较强的鲁棒性,对不同部位的 LDCT 图像进行处理,均能获得优于对比方法的 CT 图像。

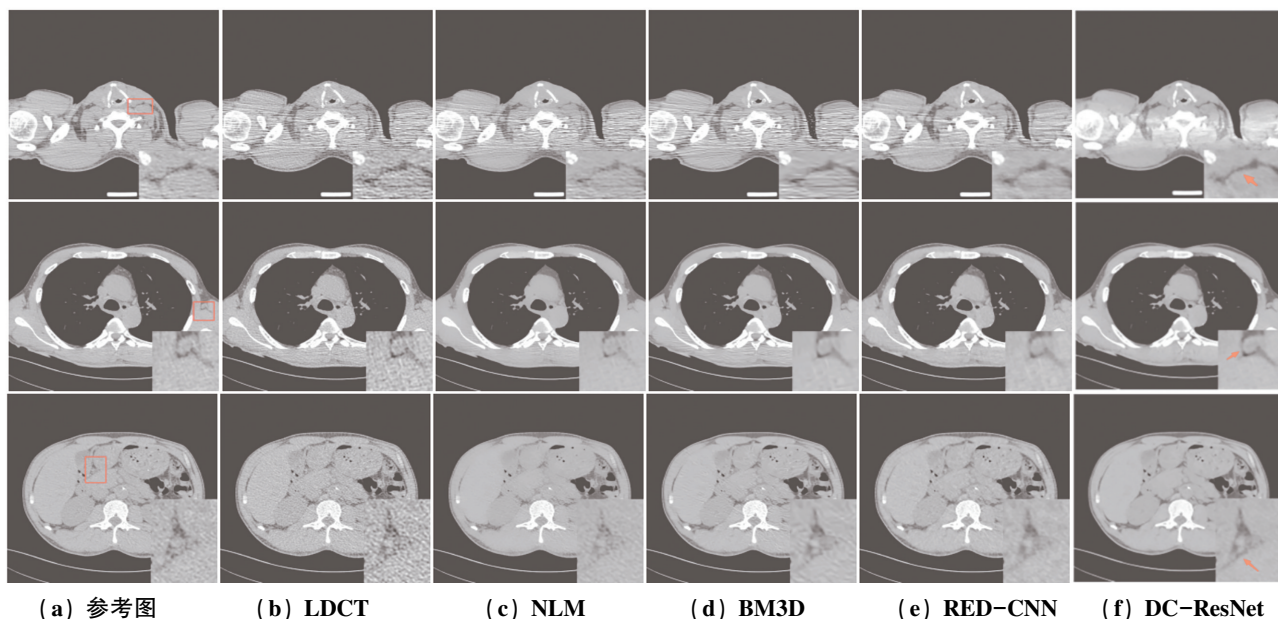


图 7 UIH 肩部、胸部、腹部横断面 LDCT 图像处理结果

Fig. 7 Processing results of UIH LDCT images of cross-sections of shoulder, chest, and abdomen

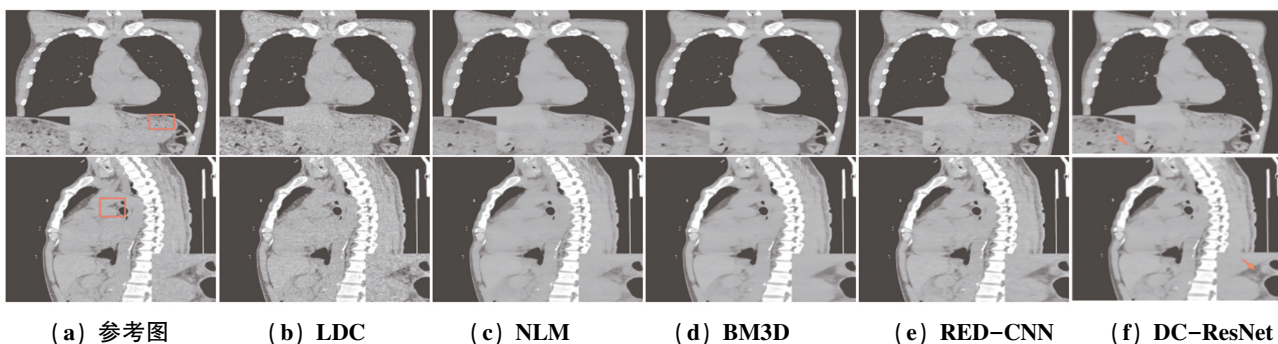


图 8 UIH 矢状面和冠状面 LDCT 图像处理结果

Fig. 8 Processing results of UIH LDCT images in sagittal and coronal planes

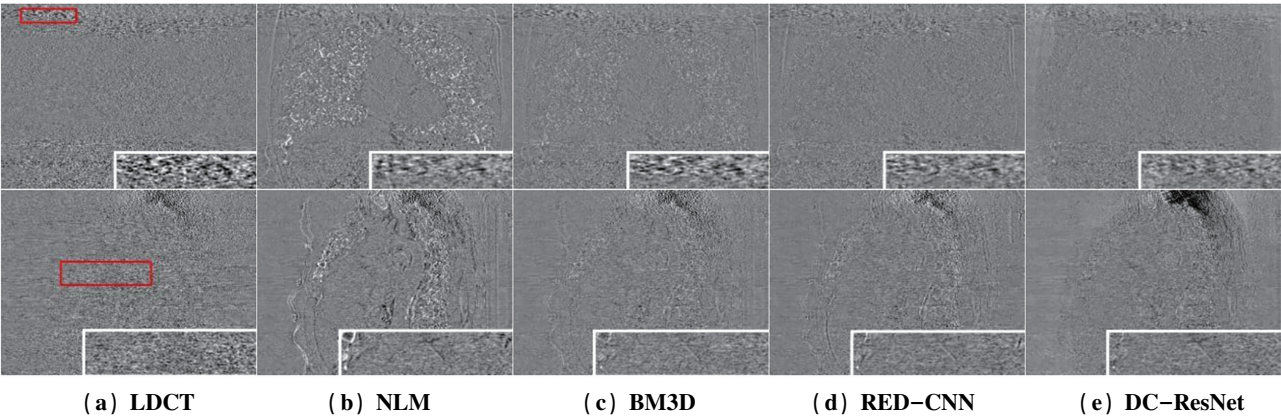


图 9 UIH LDCT 图像处理结果的差值图

Fig. 9 The difference images relative to the reference image (RDCT FBP) for UIH LDCT images

表 2 不同方法使用 UIH 测试数据处理后的量化结果

Table 2 Quantitative results associated with different methods for the images in the UIH testing dataset

		LDCT	NLM	BM3D	RED-CNN	DC-ResNet
胸部	R_{PSNR}	34.195 8	35.023 1	35.713 7	38.938 5	37.654 5
	R_{SSIM}	0.822 2	0.858 1	0.862 59	0.910 7	0.894 1
腹部	R_{PSNR}	42.048 8	43.833 6	44.552 5	45.257 8	45.855 2
	R_{SSIM}	0.930 2	0.955 3	0.961 8	0.969 33	0.970 9
胯部	R_{PSNR}	40.798 4	43.272 9	44.057 6	44.874 7	45.413 1
	R_{SSIM}	0.888 4	0.921 8	0.939 5	0.950 3	0.962 1
整体	R_{PSNR}	41.322 6	43.223 6	44.012 3	44.779 7	45.165 2
	R_{SSIM}	0.908 6	0.931 2	0.948 5	0.958 6	0.955 1

3.5 网络性能分析

从 DC-ResNet 网络结构图可以看出:该网络级联了多个局部动态可控残差块,局部动态可控残差块的数量对训练时间与预测效果均有一定的影响。为此,实验进一步测试了不同局部动态可控残差块数量对于网络整体性能的影响。实验以 P_{SNR} 和 R_{SSIM} 为指标,在局部动态可控残差块数量 N 为 4、8、16、32、64 的情况下进行验证,与此同时,统计各模型训练时间与参数量,实验结果如表 3 和图 10 所示。从图 10 可以看出:当局部动态可控残差块数量较少时,增加 N , P_{SNR} 和 R_{SSIM} 变化较大;当 N 大于 16 时, P_{SNR} 和 R_{SSIM} 变化较小。从表 3 中可以发现:增加 N ,网络参数逐渐增多,训练时间显著延长。因此综合考虑网络训练时间和处理性能,实验选择采用 32 个局部动态可控残差块用来实现 LDCT 图像的处理。

表 3 不同数量局部动态可控残差块网络模型训练参数与时间对比结果

Table 3 Comparison of training parameters and time associated with different numbers of local dynamic controllable residual block network models

	$N=4$	$N=8$	$N=16$	$N=32$	$N=64$
参数	483 596	779 011	1 369 856	2 551 555	4 914 947
时间/min	61	86	132	213	368

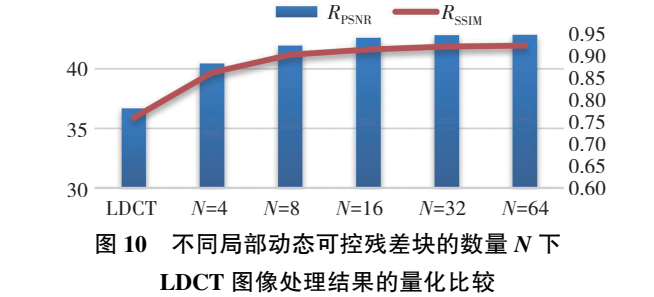


图 10 不同局部动态可控残差块的数量 N 下 LDCT 图像处理结果的量化比较

Fig. 10 Quantitative comparison of LDCT image processing results under the number of different locally dynamically controllable residual block N

DC-ResNet 网络中使用了两种形式的动态可控残差块,即全局动态可控残差块和局部动态可控残差块。为验证其有效性,实验采用一组消融测试来评估这两种动态可控残差块对实际处理效果的影响。在不改变主要网络结构情况下,通过固定权值 α ,实现“全局动态可控残差块”、“局部动态可控残差块”两个特殊情况的网络,并在 Mayo 数据集上进行训练与测试。表 4 为相关实验结果。可以看出:全局动态可控残差块和局部动态可控残差块对最终处理效果起到了积极作用, R_{SSIM} 和 R_{PSNR} 值均得到了提高,高于无该模块的两种特殊情况网络结果。这也表明动态可控残差思想的有效性。

表 4 动态可控残差块对网络模型性能影响

Table 4 Influence of dynamic controllable residual blocks on network model performance

动态可控残差块		R_{PSNR}	R_{SSIM}
全局	局部		
	✓	43.555	0.938
✓		42.661	0.930
✓	✓	43.741	0.939

4 结 论

传统神经网络的 LDCT 图像处理方法难以适用多种水平的伪影和噪声情况,为此,本文研究采用一种基

于动态可控的残差网络 DC-ResNet, 以灵活处理 LDCT 图像中复杂多变的退化。DC-ResNet 网络包括条件子网络与基础子网络两部分, 基础子网络用于 LDCT 图像处理, 采用权值动态可控制的残差连接; 条件子网络通过图像退化程度参数来估计可控权值。利用基础子网络和条件子网络的组合优势, 构建了端到端的 LDCT 图像处理模型。在 Mayo 和 UIH 两个数据集上的对比实验结果表明: DC-ResNet 网络在不同部位 LDCT 图像处理中均能获得少噪声少伪影的图像, 且 R_{PSNR} 和 R_{SSIM} 指标均优于 RED-CNN 的结果。未来工作可进一步考虑改进网络结构, 扩大性能并应于与投影域数据预处理或其他医学影像图像退化环境中, 提高其应用面。

参考文献(References):

- [1] 张权. 低剂量 X 线 CT 重建若干问题研究[D]. 南京: 东南大学, 2015.
ZHANG Quan. A study on some problems in image reconstruction for low-dose CT system [D]. Nanjin: Southeast University, 2015.
- [2] 贾丽娜. 低剂量 X-CT 图像质量改善的后处理方法研究[D]. 太原: 中北大学, 2019.
JIA Li-na. A study on post-processing method for improving image quality of low-dose X-ray CT images [D]. Taiyuan: North University of China, 2019.
- [3] PARK H S, LEE S M, KIM H P, et al. CT sinogram-consistency learning for metal-induced beam hardening correction[J]. Medical Physics, 2018, 45 (12): 5376—5384.
- [4] 马晨欣, 胡君杰, 闫斌. CT 扇形束滤波反投影图像重建算法优化[J]. 激光与光电子学进展, 2012, 49(9): 86—91.
MA Chen-xin, HU Jun-jie, YAN Bin. Optimization of fan-beam CT filtered backprojection reconstruction algorithm[J]. Laser & Optoelectronics Progress, 2012, 49(9): 86—91.
- [5] 亢艳芹, 刘进, 王勇, 等. 联合卷积稀疏编码与梯度 L0 范数的低剂量 CT 三维重建[J]. 光学学报, 2021, 41 (9): 104—115.
KANG Yan-qin, LIU Jin, WANG Yong, et al. Low dose CT 3D reconstruction with convolution sparse coding and gradient L0 norm[J]. Journal of Optics, 2021, 41(9): 104—115.
- [6] 秦华锋, 刘霞. 基于稀疏自编码的手指静脉图像分割[J]. 重庆工商大学学报(自然科学版), 2019, 36(4): 1—8.
QIN Hua-feng, LIU Xia. Finger vein image segmentation based on sparse self coding[J]. Journal of Chongqing Technology and Business University(Natural Sciences Edition), 2019, 36 (4): 1—8.
- [7] 章云港, 易本顺, 吴晨玥, 等. 基于卷积神经网络的低剂量 CT 图像去噪方法[J]. 光学学报, 2018, 38(4): 115—121.
ZHANG Yun-gang, YI Ben-shun, WU Chen-yue, et al. Low-dose CT image denoising method based on convolutional neural network[J]. Acta Optica Sinica, 2018, 38(4): 115—121.
- [8] SHAN H, PADOLE A, HOMAYOUNIEH F, et al. Competitive performance of a modularized deep neural network compared to commercial algorithms for low dose CT image reconstruction[J]. Nature Machine Intelligence, 2019, 1(6): 269—276.
- [9] TAO X, ZHANG H, WANG Y B, et al. VVBP-tensor in the FBP algorithm: its properties and application in low-dose CT reconstruction[J]. IEEE Transactions on Medical Imaging, 2020, 39(3): 764—776.
- [10] 刘进, 亢艳芹, 顾云波, 等. 稀疏张量约束的低剂量 CT 图像重建[J]. 光学学报, 2019, 39(8): 167—176.
LIU Jin, KANG Yan-qin, GU Yun-bo, et al. Low dose computed tomography image reconstruction based on sparse tensor constraint [J]. Acta Optica Sinica, 2019, 39 (8): 167—176.
- [11] KANG D, SLOMKA P, NAKAZATO R, et al. Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm [C]//Medical Imaging 2013. Image Processing, Lake Buena Vista (Orlando Area). Florida, USA, 2013.
- [12] MA J H, HUANG J, FENG Q J, et al. Low-dose computed tomography image restoration using previous normal-dose scan [J]. Medical Physics, 2011, 38(10): 5713—5731.
- [13] CHEN Y, YING X D, SHI L Y, et al. Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing[J]. Physics in Medicine and Biology, 2013, 58(16): 5803—5820.
- [14] CHEN H, ZHANG Y, ZHANG W H, et al. Low-dose CT via convolutional neural network[J]. Biomedical Optics Express, 2017, 8(2): 679—694.
- [15] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770—778.
- [16] ZHANG K, ZUO W M, CHEN Y J, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising[J]. IEEE Transactions on Image Processing. Las Vegas, NV, USA, 2016, 26(7): 3142—3155.
- [17] CHEN H, YI ZH, KALRA M K, et al. Low-dose CT with a residual encoder-decoder convolutional neural network [J]. IEEE Transactions on Medical Imaging, 2017, 36 (99): 2524—2535.
- [18] 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 Si-qi, WANG Yu, CAI Yu-fang. Low-dose CT denoising algorithm based on improved cycle GAN [J]. Acta Optica Sinica, 2020, 40(22): 70—78.
- [19] 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.
- [20] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, HI, USA, 2017: 136—144.
- [21] GU Y B, HUI T, LU T L, et al. Discriminative feature representation for noisy image quality assessment[J]. Multimedia Tools and Applications, 2020, 79(1): 7783—7809.