

## 基于迭代非对称盲点网络的低剂量CT重建算法

郭广行, 阴桂梅, 刘晨旭, 段永红, 强彦, 王艳飞, 王涛

引用本文

郭广行, 阴桂梅, 刘晨旭, 段永红, 强彦, 王艳飞, 王涛. [基于迭代非对称盲点网络的低剂量CT重建算法](#)[J].

计算机科学, 2023, 50(12): 221-228.

GUO Guangxing, YIN Guimei, LIU Chenxu, DUAN Yonghong, QIANG Yan, WANG Yanfei, WANG Tao.

[Low-dose CT Reconstruction Algorithm Based on Iterative Asymmetric Blind Spot Network](#)[J].

Computer Science, 2023, 50(12): 221-228.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

**Similar articles recommended (Please use Firefox or IE to view the article)**

[使用RAP生成可传输的对抗网络流量](#)

Generate Transferable Adversarial Network Traffic Using Reversible Adversarial Padding

计算机科学, 2023, 50(12): 359-367. <https://doi.org/10.11896/jsjcx.221000155>

[面向工业图像异常检测的连续密集标准化流模型](#)

Continuous Dense Normalized Flow Model for Anomaly Detection in Industrial Images

计算机科学, 2023, 50(12): 212-220. <https://doi.org/10.11896/jsjcx.221000183>

[基于双空间共轭自编码器的多时相高光谱异常变化检测](#)

Multi-temporal Hyperspectral Anomaly Change Detection Based on Dual Space Conjugate Autoencoder

计算机科学, 2023, 50(12): 175-184. <https://doi.org/10.11896/jsjcx.221100092>

[基于特征融合与边界修正显著性目标检测](#)

Feature Fusion and Boundary Correction Network for Salient Object Detection

计算机科学, 2023, 50(12): 166-174. <https://doi.org/10.11896/jsjcx.221100203>

[基于Transformer特征融合的时间序列分类网络](#)

Transformer Feature Fusion Network for Time Series Classification

计算机科学, 2023, 50(12): 97-103. <https://doi.org/10.11896/jsjcx.221100112>

# 基于迭代非对称盲点网络的低剂量 CT 重建算法

郭广行<sup>1,2</sup> 阴桂梅<sup>3</sup> 刘晨旭<sup>3</sup> 段永红<sup>2</sup> 强彦<sup>4</sup> 王艳飞<sup>4</sup> 王涛<sup>4</sup>

1 太原师范学院地理科学学院 山西 晋中 030619

2 山西农业大学资源环境学院 山西 太谷 030801

3 太原师范学院计算机学院 山西 晋中 030619

4 太原理工大学信息与计算机学院 山西 晋中 030600

(614364849@qq.com)

**摘要** 针对通过机器学习方法进行低剂量 CT 重建的算法过度依赖成对图例的问题,提出了一种基于迭代非对称盲点网络的低剂量 CT 重建算法。首先,通过像素混洗下采样盲点网络对低剂量 CT 进行自监督训练,得到初步重建的 CT 图像;其次,建立迭代模型,迭代使用前一网络得到的结果图像作为本网络的低剂量输入进行训练,以得到最终网络模型;最后,采用非对称的方式,对像素混洗下采样的步幅进行调整,以尽可能地减少混叠伪影,得到最终的可用模型。理论分析和实验结果表明,与传统低剂量 CT 重建算法相比,基于迭代非对称盲点网络算法可以极大地减少低剂量 CT 重建算法对成对图例的依赖,且其生成结果在在图像质量、纹理特征和结构方面优于传统方法。

**关键词:** 低剂量;深度学习;重建;自监督

**中图分类号** TP399

## Low-dose CT Reconstruction Algorithm Based on Iterative Asymmetric Blind Spot Network

GUO Guangxing<sup>1,2</sup>, YIN Guimei<sup>3</sup>, LIU Chenxu<sup>3</sup>, DUAN Yonghong<sup>2</sup>, QIANG Yan<sup>4</sup>, WANG Yanfei<sup>4</sup> and WANG Tao<sup>4</sup>

1 School of Geography Science, Taiyuan Normal University, Jinzhong, Shanxi 030619, China

2 School of Resources and Environment, Shanxi Agricultural University, Taigu, Shanxi 030801, China

3 School of Computer, Taiyuan Normal University, Jinzhong, Shanxi 030619, China

4 College of Information and Computer, Taiyuan University of Technology, Jinzhong, Shanxi 030600, China

**Abstract** Aiming at the problem that the method of low-dose CT reconstruction by machine learning method relies too much on pairwise legends, a low-dose CT reconstruction algorithm based on iterative asymmetric blind spot network is proposed. Firstly, low-dose CT is self-supervised by pixel-mixed washing sampling blind spot network, and the preliminarily reconstructed CT images are obtained. Secondly, an iterative model is established, and the result image obtained by the previous network is used as the low-dose input of the network for training to obtain the final network model. Finally, the asymmetric method is used to adjust the stride of the sampling under pixel mixing to minimize aliasing artifacts and obtain the final usable model. Theoretical analysis and experimental results show that compared with the traditional low-dose CT reconstruction algorithm, the iterative asymmetric blind spot network algorithm can greatly reduce the dependence of the low-dose CT reconstruction algorithm on pairwise legends, and can generate images similar to or even better than the traditional method in terms of image quality, texture features and structure.

**Keywords** Low dose, Deep learning, Reconstruction, Self-supervision

计算机断层扫描(Computed Tomography, CT)是现代临床医学影像学诊断方法中一种常见的方法。CT 成像中辐射剂量的暴露,具有诱发癌症的潜在可能性。为了降低对人体的损害,临床医学正在尝试降低 X 射线 CT 的扫描剂量,但这也对重建算法提出了更高的要求。重建算法主要分为:投影域去噪算法、迭代重建算法和图像域去噪算法<sup>[1]</sup>。

投影域的算法<sup>[2]</sup>也被称为预处理的过程,其将投影数据作为输入进行推理。Hsieh<sup>[3]</sup>利用自适应平衡算法去噪,并且使用平均滤波投影数据以减少重建图像中产生的条形伪影。重建算法的优势在于速度快;但更为明显的缺点是在投影域降噪的过程中,数据的不一致容易在重建图像中引入新的噪声或伪影。近十年来,迭代重建算法<sup>[4]</sup>在 LDCT 领域得到了

到稿日期:2023-03-01 返修日期:2023-04-09

基金项目:国家自然科学基金面上项目(61872261)

This work was supported by the National Natural Science Foundation of China(61872261).

通信作者:强彦(614364849@qq.com)

广泛的关注。Sidky 等<sup>[5]</sup>使用全变分(Total Variation, TV)对图像进行处理。这些算法能有效提高重建质量,但算法时间复杂度较高。LDCT 的另一个选择是重建图像的后处理,其不依赖于原始数据,目的是消除直接应用于 LDCT 图像的伪影和噪声。Li 等<sup>[6]</sup>借鉴稀疏表示理论,引入非局部均值,用于重构图像邻域内的特征相似度。Aharon 等<sup>[7]</sup>将字典学习应用到 LDCT 去噪中,获得了更好的图像质量。但是该方法并没有考虑到 CT 数据噪声在统计学视角下与自然图像等是完全不同的,这会导致在去噪过程中丢失大量的图片纹理细节,严重影响医学判断。

20 世纪末以来,深度学习方法受到人们关注并快速发展,带着神经网络在图像处理领域的广泛应用。Chen 等<sup>[8]</sup>采用最简单的卷积神经网络(Convolutional Neural Networks, CNN)对低剂量 CT 图像直接去噪,虽然用到的技术较为简单,但是与传统方法相比,该方法已经取得了极好的效果。Wu 等<sup>[9]</sup>提出了一种堆叠式的学习模型用于 LDCT 图像去噪,级联网络在不同的噪声水平下具有多种特性,比单一网络具有更高的图像质量。Chen 等<sup>[10]</sup>提出了一种基于残差解码器(RED-CNN)结构的 UNet,其结果在视觉和量化指标上都取得了极大的进步,领先于当时最先进的技术,但这一模型也存在着复杂度过高和计算时间过长的问题。Yang 等<sup>[11]</sup>提出了基于生成对抗网络的 CT 图像的去噪方法;Wolterink<sup>[12]</sup>训练卷积神经网络作为生成器,将低剂量 CT 图像转换为常规剂量 CT 图像,并训练 CNN 检测器(Adversary CNN Detector)检测输出图像是否与常规剂量 CT 图像足够接近。但是,训练过程不稳定是 GAN 网络的通病,存在较大的不确定性;并且由于成对的低剂量 CT 数据难以获取,许多深度模型存在过拟合等问题,通过监督方法进行去噪无疑是费时费力的工作。

基于迭代重建算法的方法可以较好地保留图像细节,完成辅助诊断任务。但是,以传统的基于模型的迭代重建(Model-based Iterative Reconstruction, MBIR)方法为例,由于它们依赖于重复的前向和反向投影,其对于每张低剂量 CT 图像都需要消耗大量的时间才能得到目标结果。将 MBIR 与深度学习相结合,Gupta 等<sup>[13]</sup>和 Wu 等<sup>[14]</sup>首先提出使用 CNN 和自动编码器在 MBIR 框架中对正则化项进行建模。Adler 等<sup>[15]</sup>将 MBIR 的优化过程展开到  $N$  级网络,以平衡重建质量和重建速度。尽管相比传统的 MBIR 方法有所改进,但它们仍然存在迭代过程的计算成本过高的问题。作为替代方案,DLR 通常被表述为图像后处理。Jin 等<sup>[16]</sup>和 Chen 等<sup>[17]</sup>提出使用 UNet<sup>[18]</sup>和残差 UNet 对稀疏视图 CT 中的噪声/伪影进行后处理。Wang 等<sup>[19]</sup>和 Yang 等<sup>[20]</sup>提出将密集块和小波分解结合到 UNet 中,以便为重建提供更稳健的特征学习;还提出了直接正弦图反转和正弦图补全策略。这些方法虽然将迭代重建的思想引入深度学习模型中,但仍然无法摆脱模型对成对 CT 数据的依赖。

自然图像领域的新近研究表明,深度模型可以将噪声图像同时作为输入和标签使用,以进行自监督训练,并且取得了不错的效果。本研究将这一思想迁移至低剂量 CT 去噪领域,并结合本领域特点,实现自监督的低剂量 CT 重建方法。

综上所述,本文的主要工作如下:

- 1)改进自监督的经典方法盲点网络对低剂量 CT 进行重建,摆脱 CT 重建对成对数据的依赖性;
- 2)在自监督领域首次引入数据一致性方法,以保证 CT 图像的细节数据不丢失;
- 3)将迭代重建的传统 CT 重建思路首次引入深度学习自监督方法中,使得深度模型更容易学习到 CT 图像的高频特征。

## 1 模型相关结构

### 1.1 低剂量 CT 去噪模型

CT 是一种常见的医学成像模型,在高剂量下可以产生更清晰的图像,但人体受辐射的风险较高。但是,降低辐射剂量会降低信噪比,产生强噪声和伪影,降低 CT 图像的质量。传统方法由于仅针对具体的统计模型,对 LDCT 的重建性能并不理想<sup>[21]</sup>。深度学习方法由于具有端到端的特性,并不会受到这一问题的影响,训练样本的输入分布和标签分布才是决定深度学习的因素。

假设  $x \in R^{H \times W}$  表示低剂量的 CT 图像,  $y \in R^{H \times W}$  表示相应的正常剂量图像,则可以得到以下关系:

$$x = \sigma(y) \quad (1)$$

其中,  $\sigma: R^{m \times n} \rightarrow R^{m \times n}$  表示正常剂量 CT 图像引入量子噪声到低剂量 CT 图像的转换。降噪问题可以转换为寻找目标函数  $f$  的任务。

$$f = \arg \min \| f(x - y) - y \|_2^2 \quad (2)$$

其中,  $f$  被视为  $\sigma^{-1}$  的最佳近似。

### 1.2 盲点网络

BSN<sup>[22]</sup>是传统 CNN 的变体,该模型通过遮盖感受野中的中心像素来预测相应的输出像素。已有研究<sup>[22-24]</sup>表明,BSN 可以学习以自我监督的方式对噪声图像  $I_N \in R^{H \times W}$  进行去噪。要训练 BSN,必须满足以下两个假设:噪声是空间上的,即像素级的、独立的和零均值的。在这样的假设下,众所周知<sup>[23,25]</sup>,最小化自监督损失  $L_{\text{self}}$  等价于传统的监督学习,如下所示:

$$\begin{aligned} L_{\text{self}} &= E_{I_N} \| B(I_N) - I_N \|_2^2 \\ &= E_{I_N, I_C} \| B(I_N) - I_N \|_2^2 + c = L_{\text{super}} + c \end{aligned} \quad (3)$$

其中,  $I_C \in R^{H \times W}$  是噪声输入  $I_C$  的干净真实值,  $L_{\text{super}}$  是监督去噪损失函数,  $c$  是常数。

在低剂量 CT 重建领域中,相邻像素之间的噪声强度基于它们的相对距离显示出不可忽略的相关性,这显然不符合训练盲点网络的基本假设。为了消除这种相关性噪声的影响,Zhou 等<sup>[26]</sup>提出像素混洗下采样(Pixel-shuffle Downsampling, PD)的概念来打破低剂量 CT 图像噪声中的空间相关性。PD 通过使用固定步幅因子对噪声图像进行子采样来创建子图像,从而增加了噪声信号之间的实际距离。

PD<sub>s</sub> 可以被视为像素混洗<sup>[27]</sup>的逆运算,步幅因子为  $s$ 。由于现实世界的噪声信号与少数相邻像素相关,PD 过程中的二次采样可能会破坏它们之间的依赖关系。对此,将传统的去噪算法应用于下采样图像,其中 PD 的逆运算  $PD_s^{-1}$  用于重建全尺寸输出。为了保留图像纹理和细节,Zhou 等<sup>[26]</sup>

将步幅因子设置为 2, 即  $PD_2$ , 以获得最佳性能。

### 1.3 迭代数据细化

受到传统迭代重建思想的启发, 本研究提出了一种用于自监督深度学习模型的迭代方案, 称之为迭代数据细化。具体方案内容如下:

首先, 基于初始噪声数据集  $\{x_i + n_i, x_i\}$  训练模型  $F_0$ , 其中  $x_i$  表示原始低剂量 CT 图像,  $n_i$  表示添加的噪声。根据已有的自监督研究, 它能够在一定程度上对噪声图像  $\{x_i\}$  去噪。因此, 本研究通过使用结果训练得到的深度网络对原始的 CT 图像进行去噪, 得到一组新的精炼训练集:

$$\{F_0(x_i) + n_i, F_0(x_i)\} \quad (4)$$

新构建的数据集与原始噪声数据集共享相同数量的图像和场景内容。然而, 这个新数据集与噪声干净的数据集的数据偏差较小, 因为训练的去噪模型  $F_0$  减少了目标的  $L_2$  距离。

然后, 使用上述偏差较小的数据集进行构造, 可以在下一轮从头开始训练一个新的去噪模型  $F_1$ :

$$F_1 \leftarrow \{F_0(x_i) + n, F_0(x_i)\} \quad (5)$$

由于新数据集  $\{F_0(x_i) + n, F_0(x_i)\}$  和理想的噪声干净数据集  $\{y + n, y\}$  之间的数据偏差得到缓解, 根据第 3.1 节中的发现, 与在第一轮噪声较大的数据集上训练的  $F_0$  相比, 训练模型  $F_1$  可以更好地泛化到实际噪声图像  $\{x_0\}$ 。

基于改进的模型  $F_1$ , 上述过程可以通过细化训练目标和训练更好的模型来迭代交替执行, 如图 1 所示。从模型的角度来看, 一旦有了一个去噪模型  $F_m$ , 一个更好的模型  $F_{m+1}$  就可以使用构建的数据集进行训练:

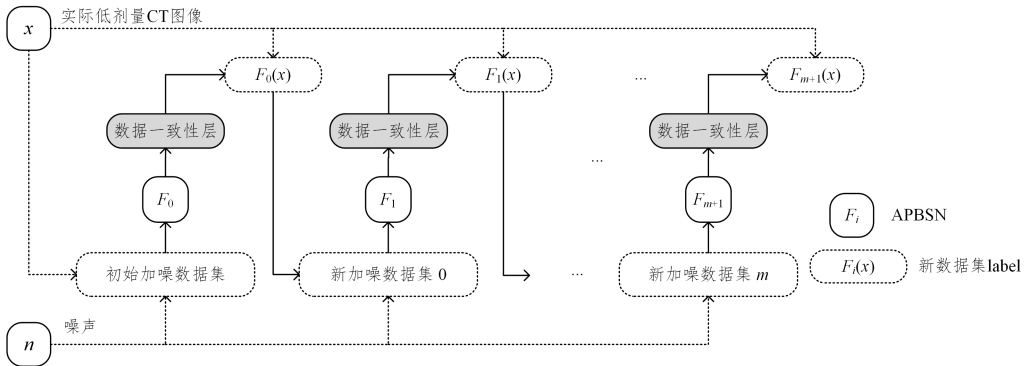


图 1 低剂量 CT 图像重建模型的总体框架

Fig. 1 Overall framework of low-dose CT image reconstruction model

该方法首先以原始低剂量 CT 图像加噪后得到的数据集  $\{x_i + n_i, x_i\}$  训练非对称盲点网络  $F_0$  (见 2.2 节), 并经过数据一致性层; 然后, 用该网络以  $x_i$  为输入进行预测, 得到  $\{F_0(x_i)\}$  图像; 接着, 以该图像加噪后得到的数据集  $\{F_0(x_i) + n_i, F_0(x_i)\}$  训练非对称盲点网络  $F_1$ , 重复此过程  $m$  次得到  $m+1$  个网络以及最后的输出结果  $\{F_{m+1}(x_i)\}$ 。

### 2.2 非对称 PD 盲点网络

本研究的目标是以自我监督的方式在低剂量 CT 图像上推广 BSN。为此, 采用 PD 并最小化以下损失  $L_{BSN}$  来训练 BSN。

$$\begin{aligned} L_{BSN} &= \| PD_s^{-1}(B(PD_s(I_N))) - I_N \|_1 \\ &= \| I_{BSN}^s - I_N \| \end{aligned} \quad (7)$$

其中,  $I_{BSN}^s$  是  $PD_s$  和 BSN 的输出, 即  $PD_s$ -BSN。模型使用  $L_1$

$$F_{m+1} \leftarrow \{F_m(x_i) + n_i, F_m(x_i)\} \quad (6)$$

先前的模型  $F_m$  有助于缩小新的噪声较大的数据集  $\{F_m(x_i) + n_i, F_m(x_i)\}$  和理想的噪声干净的数据集  $\{y + n, y\}$  之间的差距。在下一轮训练中, 新创建的噪声数据集的改进有助于新训练的模型  $F_{m+1}$  的推进。另一方面, 从目标的角度, 中间去噪模型会产生一系列中间细化目标  $F_m(x_i)$  (对于  $m=0, 1, \dots$ )。此外, 与传统的迭代方法多次对一张噪声图像去噪并导致大量纹理丢失不同, 本方法在推理过程中改进了训练数据集并对噪声图像进行了一次去噪。

## 2 迭代非对称盲点网络

尽管 CT 技术极大地促进了医学领域的快速发展, 但是由于其本身的辐射剂量高, 因此患者需要承担未知的健康风险。基于模型的传统去噪方法需要昂贵的计算机设备, 简单的图像域去噪方法又受限于成对数据较少。

### 2.1 总体设计

为了解决这些问题, 提出了一种无监督盲点网络和迭代数据细化相结合的重建方法。该方法只需要将噪声图像作为输入, 即可在减少辐射剂量的同时获得高质量的图像。模型主要包括 3 个层次核心设计, 从外到内依次是: 1) 基于迭代重建思想的迭代数据细化整体架构; 2) 基于非对称像素混洗的无监督盲点网络, 被用于每个迭代中的核心模型; 3) 引入数据一致性概念, 在盲点网络输出结果中加入一致性保证项。整个方法的总体框架如图 1 所示。

范数来更好地泛化。首先将给定的噪声图像  $I_N$  分解为  $s^2$  个子图像。注意到  $PD_s(I_N)$  是这些子图像的平铺<sup>[26]</sup>  $I_{sub}^s \in R^{H/s \times W/s}$ 。将 BSN 应用于子图像并使用 PD 逆运算  $PD_s^{-1}$  重建输出  $I_{BSN}^s$ 。

然而, 将 PD-BSN 直接应用于低剂量 CT 图像非常困难。研究中观察到, PD-BSN 使用方程式中的自监督损失进行训练, 并不适用于低剂量 CT 图像。无论 PD 步幅因子  $s$  是多少,  $PD_2$ -BSN 和  $PD_5$ -BSN 都无法从给定的噪声输入中恢复清晰锐利的图像。

Lee 等<sup>[28]</sup> 提出的非对称像素混洗盲点网络 (AP-BSN) 成功将 PD-BSN 方法应用于自然图像领域。受这项研究的启发, 首先讨论 PD-BSN 关于步幅因子  $s$  的两个重要方面。

1)打破空间相关性。最初,PD 被提出来是用于减少现实世界图像中相邻噪声信号之间的空间相关性。Zhou 等<sup>[26]</sup>将步幅因子设置为 2。为了确定合适的步幅因子,本文针对两个数据集进行了一项噪声间相关性与距离间关系的实验,得到如图 2 所示结果。

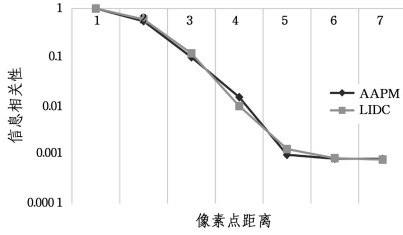


图 2 噪声相关性与距离的关系实验

Fig. 2 Experiment on the relationship between noise correlation and distance

显而易见,当噪声像素间距离大于或等于 5 时,两个像素点之间的信息相关性较低。当选用的步幅因子大于 5 时,

组成每个子图像的相邻像素点在原图中的间距保证了两者的信息相关性可以忽略。由此,本研究认为步幅因子应该设为 6,以最大限度地打破图像的空间相关性。

2)混叠伪影。随着步幅因子  $s$  变大,来自 PD 的子图像会遭受更强程度的混叠。从信号处理的角度来看,当原始信号没有得到适当的带宽限制时,下采样图像会出现混叠。由于 PD 过程在二次采样之前没有利用低通滤波器,在实验中发现应用大步长 PD(例如  $s=6$ )时,混叠会以一种噪声形式出现。

通过上述讨论,容易得出 PD<sub>s</sub> 的步幅因子存在对打破空间相关性和混叠伪影的权衡。Lee 等<sup>[28]</sup>提出了一种非对称像素混洗下采样方法,即对训练和推理过程分别使用步幅因子 5 和 2,并取得了良好的效果。本文参考使用该非对称技术,但对步幅因子的设定进行了独立的实验以使得该技术适用于低剂量 CT 图像,最终得到步幅因子 6 和 2 为最优方案。图 3 说明了 AP-BSN 的非对称训练和推理方案。

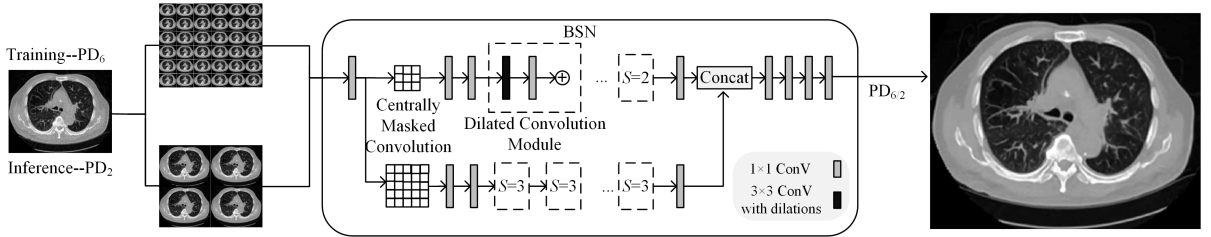


图 3 AP-BSN 的总体结构

Fig. 3 Overall structure of AP-BSN

### 2.3 数据一致性层

为了防止 CT 图像在网络推理过程中可能出现的关键信息丢失的问题,我们利用 CT 数据独有的正弦域数据来确保关键信息的传递。受 Wu 等的启发<sup>[29]</sup>,在 CT 图像经过 AP-BSN 推理之后加入数据一致性层:

$$\mathbf{x}_o \rightarrow \mathbf{x}_o - \lambda_k \cdot \text{FBP}(\text{FBP}(\mathbf{x}_o) - P(\mathbf{x}_i)) \quad (8)$$

其中,FBP 表示滤波反投影,用于将正弦数据转换为图像数据;P 为投影过程,用于将图像数据转化为正弦数据,表示 AP-BSN 推理之后得到的图像; $\mathbf{x}_i$  表示推理前的图像。通过数据一致性层,可以确保 CT 图像在经过网络处理后的正弦数据与之前的正弦数据一致,找回由于网络推理导致的步幅信息丢失的问题。大量研究表明,这一举措可以显著提高图像最终的质量。

## 3 实验相关设置

本章介绍了用于训练和评估网络的数据集,并讨论实验中数据预处理和超参数的选择。

### 3.1 数据集

实验主要使用 2016 年梅奥医学中心 (NIH-AAPM Mayo) 低剂量 CT 大挑战中发布的公开数据集<sup>1)</sup>。该数据集包含 10 个匿名患者低剂量腹部 CT 图像,每个病例都提供 3 种数据:投影数据、DICOM (Digital Imaging and Communi-

cations in Medicine) 数据和临床数据报告。同时,还在不同数据集上进行了实验,来验证模型的泛化能力。LIDC-IDRI 数据集中提供 DICOM 数据和 XML 文件报告。对于不同的数据集,采用插入泊松噪声的方式,以达到与 1/4 剂量水平相对应的噪声水平,模拟 LDCT 数据。对于每个病例,采用 3mm 和 1mm 的切片厚度。在训练中,从数据集中 6 例患者的扫描图像中随机抽取 2560 张 1/4 剂量 CT 图像作为输入;从 4 例患者的扫描图像中抽取 1382 对和 1345 对 CT 图像进行验证和测试实验。

### 3.2 网络和训练设置

APBSN 网络输入大小为  $256 \times 256$ ,为了在保证 CT 图像大小符合网络要求的同时扩大数据集,实验中将原本为  $512 \times 512$  大小的图像进行步长为 64 的切分,得到 40000 余张训练低剂量 CT 图像,并对这些图像做 radon 变换,得到对应的正弦数据。通过 Adam 优化算法进行 APBSN 网络训练,初始学习率从  $1 \times 10^{-4}$  开始线性下降至  $1 \times 10^{-6}$ ,每次迭代以最小化损失函数  $L_{\text{BSN}}$  (见式(7)) 为目标训练单个网络, batch size 设置为 16, epoch 设置为 300。

### 3.3 评价方法

在图像重建领域中,峰值信噪比 (Peak Signal Noise Ratio, PSNR) 通常用来评价压缩后的图像与原始图像的质量<sup>[30]</sup>。压缩后的图像会比原始图像质量差,根据去噪后图像

<sup>1)</sup> <http://www.aapm.org/GrandChallenge/165 LowDoseCT/#>

与原始参考图像之间的像素值对比结果  $x_0$  进行衡量。PSNR 越高,压缩后失真越小<sup>[31]</sup>,推理之后的图像中信息占比越大。均方误差(Mean Squared Error,MSE)代表 NDCT 与经过网络推理后的 LDCT 之间的均方误差。均方根误差(Root Mean Squared Error,RMSE)主要是根据标准图和推理结果图之间的离散程度来量化误差值,其取值范围为  $[0, +\infty)$ ,当预测值与真实值完全吻合时等于 0,即完美模型;误差越大,该值越大。RMSE 是重建领域的一项非常重要的指标。数学定义如下:

$$R_{\text{MSE}} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i,j) - K(i,j)\|^2 \quad (9)$$

$$R_{\text{RMSE}} = \sqrt{R_{\text{MSE}}} \quad (10)$$

$$\begin{aligned} R_{\text{PSNR}} &= 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{R_{\text{MSE}}} \right) \\ &= 20 \cdot \log_{10} \left( \frac{\text{MAX}_I}{\sqrt{R_{\text{MSE}}}} \right) R_{\text{RMSE}} = \sqrt{R_{\text{MSE}}} \end{aligned} \quad (11)$$

在相关研究中,评价指标 PSNR 使用得最多。PSNR 的取值范围通常为 20~40,取值越大,质量越好。但其值并不能很好地反映人眼的主观感受,人体的视觉感知不仅依靠图像本身,而且很容易受到临近区域的像素影响。结构相似性就可以作为另一种评价方式,来检测图像质量的好坏。

结构相似性(Structural Similarity,SSIM)是用于比较处理前和处理后的图像的相似性指标<sup>[32]</sup>,计算比较复杂,主要针对人眼观察效果进行量化。SSIM 通常取到 0~1,取值越大,质量越好。结构相似理论认为,过去的量化指标大多针对数据相似性,并不考虑人眼的角度。但是事实证明,即使均方误差较低,高 SSIM 的图像也更利于医学诊断。在进行结构信息计算时,不应受到照明的影响,同时为了保证计算的准确性,还需在计算结构信息之前归一化图像。可以采用以下的计算方式来进行结构信息的计算:

$$R_{\text{SSIM}}(\mathbf{X}, \mathbf{Y}) = l(\mathbf{X}, \mathbf{Y}) \times c(\mathbf{X}, \mathbf{Y}) \times s(\mathbf{X}, \mathbf{Y}) \quad (12)$$

$$l(\mathbf{X}, \mathbf{Y}) = \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_2} \quad (13)$$

$$c(\mathbf{X}, \mathbf{Y}) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (14)$$

$$s(\mathbf{X}, \mathbf{Y}) = \frac{2\sigma_x \sigma_y + C_3}{\sigma_x \sigma_y + C_3} \quad (15)$$

其中,  $\mathbf{X}$  和  $\mathbf{Y}$  分别表示处理后的 LDCT 和 NDCT,  $l(\mathbf{X}, \mathbf{Y})$  表示亮度比值,  $c(\mathbf{X}, \mathbf{Y})$  表示结构差值,  $s(\mathbf{X}, \mathbf{Y})$  代表对比度;  $\mu_x$  和  $\mu_y$  表示图像像素点的平均数,  $\sigma_x$  和  $\sigma_y$  表示图像的方差,  $\sigma_x^2$  和  $\sigma_y^2$  表示图像的协方差;  $C_1$ ,  $C_2$  和  $C_3$  是为了增强结果的稳定性,其值为接近零的非零常数,用来防止计算过程中出现分母为零的情况。SSIM 值越接近 1, NDCT 和处理后的 LDCT 在结构上的相似性越强。

## 4 实验结果

### 4.1 参数选择

#### 4.1.1 步幅因子

APBSN 网络需要对训练和推理过程分别确定步幅因子,以最大程度地让训练和推理过程中的混叠噪声达到某种平衡,使得训练效果达到最佳。为此,对不同的步幅因子进行

实验,结果如图 4 所示。

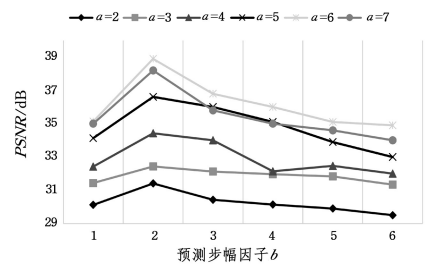


图 4 步幅因子  $a$  和  $b$  的定量实验

Fig. 4 Quantitative experiment of stride factor  $a$  and  $b$

通过实验结果可以观察到,当训练步幅因子  $a=6$ ,推理步幅因子  $b=2$  时,PSNR 达到最大值,因此本研究选用步幅因子 6 和 2。

#### 4.1.2 迭代次数

本文提出的方法受到迭代重建思想的启发,通过循环迭代的方法令多个相同架构的网络相互补充学习,以更好地学习到 CT 图像中的噪声信息并将其去除。本文设计实验来观察迭代次数与训练效果之间的关系以确定迭代次数,如图 5 所示。

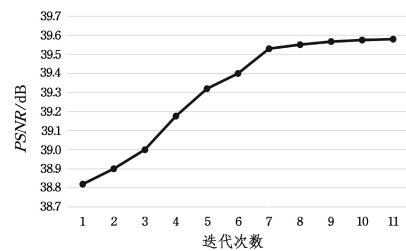


图 5 迭代次数的定量实验

Fig. 5 Quantitative experiment of the number of iterations

观察实验结果可以得出,当迭代次数等于 7 时,最终的 PSNR 开始收敛。此时增加迭代次数虽然可以提高 PSNR 值,但是需要更多的训练和推理时间。因此在综合考虑图像效果与时间因素的情况下,本研究选用迭代次数为 7 进行后续研究。

#### 4.1.3 扩张卷积模块

在盲点网络中,会多次使用扩张卷积块(如图 5 所示)。为了确定扩张卷积块的个数对最终成像质量的影响,本文进行了对比实验,结果如图 6 所示。

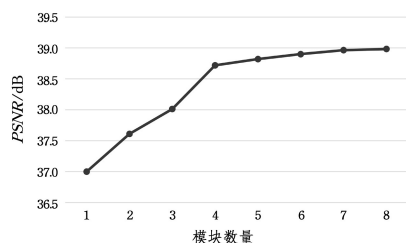


图 6 扩张卷积模块数量的定量实验

Fig. 6 Quantitative experiment on the number of expanded convolutional modules

容易观察到,当扩张卷积块数量大于或等于 4 之后,PSNR 值开始收敛。在考虑网络复杂度的因素之后,本研究

设置模块数量为 5。

#### 4.2 定量分析

为了验证本文模型在重建图像中的性能,本小节针对不同重建方法在相同数据集上的具体表现作出了分析。重建方法包括 CNN-VGG<sup>[33]</sup>, WGAN<sup>[11]</sup>, WGAN-MSE<sup>[12]</sup>, BM3D<sup>[34]</sup>, Eformer<sup>[35]</sup>, ADAGAN<sup>[36]</sup>和 LDCT。

通过计算不同模型的 PSNR, SSIM 和 RMSE, 来对模型的整体进行评价, 结果如表 1、表 2 所列。

表 1 不同重建方法在 NIH-AAPM Mayo 数据集上的定量实验结果对比

Table 1 Comparison of quantitative experimental results of different reconstruction methods on NIH-AAPM Mayo dataset

方法	PSNR	SSIM	RMSE
LDCT	34.3094	0.8276	0.0193
WGAN <sup>[11]</sup>	36.1047	0.8522	0.0108
WGAN-MSE <sup>[12]</sup>	38.1983	0.8917	0.0092
CNN-VGG <sup>[33]</sup>	38.3061	0.8720	0.0097
BM3D <sup>[34]</sup>	38.9903	0.8295	0.0112
Eformer <sup>[35]</sup>	39.0972	0.9376	0.0081
ADAGAN <sup>[36]</sup>	39.2876	<b>0.9382</b>	0.0072
本文方法	<b>39.5423</b>	0.9373	<b>0.0056</b>

表 2 不同重建方法在 LIDC-IDRI 数据集上的定量实验结果对比

Table 2 Comparison of quantitative experimental results of different reconstruction methods on LIDC-IDRI dataset

方法	PSNR	SSIM	RMSE
LDCT	34.5781	0.8381	0.0187
WGAN <sup>[11]</sup>	36.1786	0.8698	0.0113
WGAN-MSE <sup>[12]</sup>	38.2134	0.8968	0.0091
CNN-VGG <sup>[33]</sup>	38.3179	0.8781	0.0093
BM3D <sup>[34]</sup>	38.9892	0.8273	0.0115
Eformer <sup>[35]</sup>	39.0564	0.9325	0.0089
ADAGAN <sup>[36]</sup>	39.2653	<b>0.9341</b>	0.0077
本文方法	<b>39.5750</b>	0.9315	<b>0.0060</b>

在最终的结果中, WGAN 在 PSNR 和 SSIM 评价指标中表现最差, 原因是 WGAN 没有进行 MSE 和 VGG 正则化。WGAN-MSE 的改进版本虽对这一现象有所改进, 但仍不及较为先进的方法, 其主要原因在于 GAN 网络存在的不确定性和对 CT 图像纹理细节的关注不足。本文方法的 PSNR 与 RMSE 在所有模型中都取得了最优值, 但在 SSIM 上的表现却略微逊色于 Eformer 和 ADAGAN。这一现象是由于 Eformer 和 ADAGAN 都在监督方法的优势之下引入了细节结构的注意力通道, 使得模型可以更好地学习到图像结构方面的信息。值得注意的是, 本文提出的方法在 RMSE 上有非常大的优势, 这是因为本方法加入了数据一致性层。

#### 4.3 消融实验分析

为了验证本文采用迭代重建思想的迭代训练架构和数据一致性层在重建图像中的性能, 本文将模型拆分成 APBSN、APBSN+数据一致性层(DC)、APBSN+Iteration 和 IABSN 这 4 个版本, 并对它们在相同的数据集上的具体表现作出了分析, 如表 3 所列。

表 3 不同版本在 NIH-AAPM Mayo 数据集上的实验结果对比  
Table 3 Comparison of experimental results of different versions on NIH-AAPM Mayo dataset

模块	PSNR	SSIM	RMSE
APBSN	38.8194	0.9006	0.0113
APBSN+DC	39.0112	0.9284	0.0062
APBSN+Iteration	39.4143	0.9301	0.0084
IABSN(本文)	<b>39.5423</b>	<b>0.9373</b>	<b>0.0056</b>

由实验结果可以看出, 单纯使用 APBSN 进行训练, 3 项指标都不太理想; 在加入数据一致性层(DC)后, RMSE 和 SSIM 有明显提升, 说明数据一致性层在训练中可以较好地保留低剂量图像的纹理特征, 并且降低了正常剂量图像间的像素差异; 而在加入迭代思想后, PSNR 有着显著的提高, 表明本方法通过迭代可以更好地捕捉低剂量 CT 图像中的噪声信息。

为了展示本文提出的方法所拥有的去噪性能, 选择一张测试图像进行测试, 结果如图 7 所示。所有网络都具有一定的去噪能力, 然而 CNN-MSE 对图片进行了模糊, 并产生了条形伪影。对于图 8 放大后的感兴趣区域, 可以很容易观察到放大后的 ROI。WGAN 可以避免过度平滑和小条纹阴影。同时, 单独使用 WGAN 或 GAN 会产生更强的噪声, 在低剂量下产生了明显的带状伪影, 从而导致出现一些白色结构图像。CNN-VGG 图像在视觉上更接近于 NDCT 图像, 这是因为 CNN-VGG 中使用的 VGG 损失是在一个特征空间中计算的, 这个特征空间之前是在一个非常大的自然图像数据集上训练的。利用 VGG 损失, 将嵌入 VGG 网络的人类感知知识转移到 CT 图像质量评价中, 它的性能本身是不可接受的, 因为它只将数据分布从输入图像映射到给定标准图像, 而不能保证图像内容的完整性。而 Eformer 和 ADAGAN 虽然较好地保留了图像边缘细节, 但是对 CT 图像整体的噪声去除略有不足。相比之下, 本文提出的方法可以得到效果相近甚至更好的 CT 图像; 并且由于使用自监督模式, 本方法对数据集的要求显著降低, 便于收集更庞大的数据。

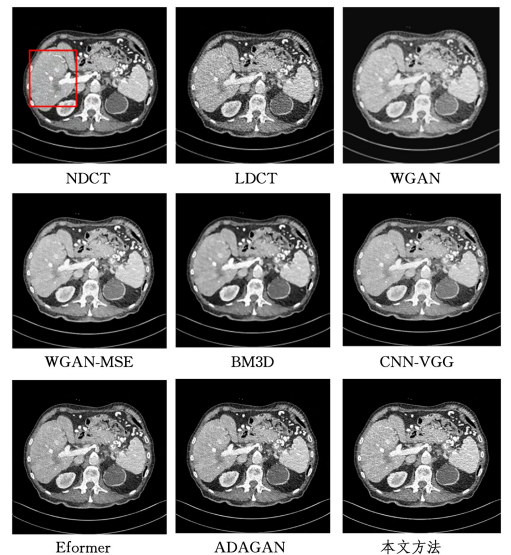


图 7 图像重建算法的实验结果

Fig. 7 Experimental results of image reconstruction algorithm

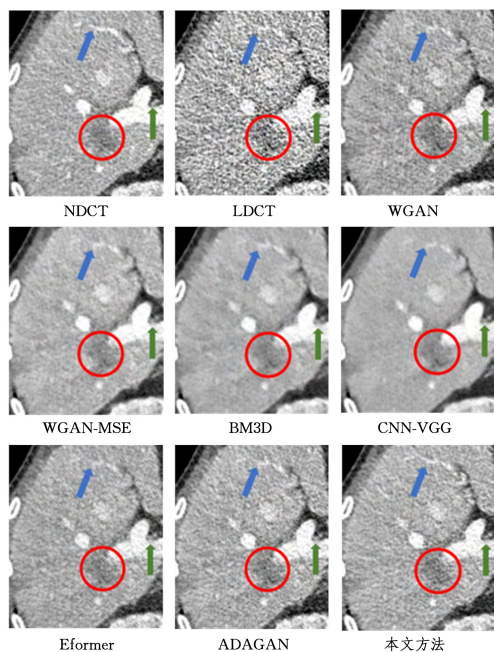


图 8 图像重建算法区域放大的实验结果

Fig. 8 Experimental results of region enlargement of image reconstruction algorithm

**结束语** 本文研究了 CT 图像的重建,在图像空间中图像进行去噪,提出了一种新的基于 BSN 的自监督 CT 图像重建方法。总的来说,该模型在定性可视化和定量验证方面优于传统的 CT 重建方法和其他最先进的方(如 Eformer 和 ADAGAN),并且有助于提高图像质量和统计属性。除了采用上述 BSN 体系结构,本文还增加了迭代重建的思想,使得本方法通过迭代可以更好地捕捉低剂量 CT 图像中的噪声信息,获得更好的定性和定量性能。此外,本文使用的数据一致性思想是 CT 重建领域所特有的性质,使得 CT 图像的重建质量比自然图像更有保证。

## 参 考 文 献

- [1] YIN X,ZHAO Q,LIU J,et al. Domain Progressive 3D Residual Convolution Network to Improve Low-Dose CT Imaging[J]. IEEE Transactions on Medical Imaging, 2019, 38(12): 2903-2913.
- [2] HOYEON L,JONGHA L,HYEONGSEOK K,et al. Deep-neural-network based sinogram synthesis for sparse-view CT image reconstruction[J]. arXiv:1803.00694,2018.
- [3] HSIEH J. Adaptive streak artifact reduction in computed tomograph resulting from excessive X-ray photon noise[J]. Medical Physics, 1998, 25(11): 2139-2147.
- [4] THIBAULT J B, SAUER K D, BOUMAN C A, et al. A three-dimensional statistical approach to improved image quality for multislice helical CT[J]. Medical Physics, 2007, 34(11): 4526-4544.
- [5] SIDKY E Y, PAN X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization[J]. Physics in Medicine & Biology, 2008, 53(17): 4777.
- [6] LI Z, YU L, TRZASKO J D, et al. Adaptive non-local means filtering based on local noise level for CT denoising[C]// Proce-

dings Volume 8313, Medical Imaging 2012: Physics of Medical Imaging, 2012:83131H.

- [7] AHARON M, ELAD M, BRUCKSTEIN A.  $\ell_1$ -SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation[J]. IEEE Transactions on Signal Processing, 2006, 54: 4311-4322.
- [8] HU C, YI Z, ZHANG W, et al. Low-dose CT denoising with convolutional neural network[J]. arXiv:1610.00321v1, 2016.
- [9] WU D, KIM K, FAKHRI G, et al. A cascaded convolutional neural network for X-ray low-dose CT image denoising[J]. arXiv:1705.04267, 2017.
- [10] CHEN H, YI Z, KALRA M K, et al. Low-Dose CT with a Residual Encoder-Decoder Convolutional Neural Network (RED-CNN) [J]. IEEE Transactions on Medical Imaging, 2017, 36(99): 2524-2535.
- [11] 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 transactions on medical imaging, 2018, 37(6): 1348-1357.
- [12] 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.
- [13] GUPTA H, JIN K H, NGUYEN H Q, et al. CNN-based projected gradient descent for consistent CT image reconstruction[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1440-1453.
- [14] WU D, KIM K, EL FAKHRI G, et al. Iterative low-dose CT reconstruction with priors trained by artificial neural network[J]. IEEE Transactions on Medical Imaging, 2017, 36(12): 2479-2486.
- [15] ADLER J, ÖKTEM O. Learned primal-dual reconstruction[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1322-1332.
- [16] JIN K H, MCCANN M T, FROUSTEY E, et al. Deep Convolutional Neural Network for Inverse Problems in Imaging[J]. IEEE Transactions on Image Processing, 2016(99): 4509-4522.
- [17] 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.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[J]. arXiv: 1505.04597, 2015.
- [19] 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 Transactions on Medical Imaging, 2018, 37(6): 1348-1357.
- [20] WANG Y, CHAO L, SHAN W, et al. Improving the Quality of Sparse-view Cone-Beam Computed Tomography via Reconstruction-Friendly Interpolation Network[C]// Computer Vision - ACCV, 2022: 86-100.
- [21] YE D H, BUZZARD G T, RUBY M, et al. Deep Back Projection For Sparse-View Ct Reconstruction[C]// 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP).

- Anaheim, CA, USA, 2018.
- [22] KRULL A, BUCHHOLZ T O, JUG F. Noise2Void - Learning Denoising From Single Noisy Images [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.
- [23] BATSON J, ROYER L. Noise2Self: Blind Denoising by Self-Supervision[J]. arXiv:1901.11365, 2019.
- [24] KRULL A, VIČAR T, PRAKASH M, et al. Probabilistic noise2void; Unsupervised content-aware denoising[J]. Frontiers in Computer Science, 2020, 2; 5.
- [25] XIE Y, WANG Z, JI S. Noise2Same: Optimizing A Self-Supervised Bound for Image Denoising[J]. arXiv:2010.11971, 2020.
- [26] ZHOU Y, JIAO J, HUANG H, et al. When AWGN-Based Denoiser Meets Real Noises[C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2020; 13074-13081.
- [27] SHI W, CABALLERO J, HUSZÁR F, et al. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network[C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.
- [28] LEE W, SON S, LEE K M. Ap-bsn: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022; 17725-17734.
- [29] WU W, HU D, NIU C, et al. DRONE: Dual-domain residual-based optimization network for sparse-view CT reconstruction [J]. IEEE Transactions on Medical Imaging, 2021, 40 (11): 3002-3014.
- [30] HORÉ A, ZIOU D. Image quality metrics; PSNR vs. SSIM [C] // 20th International Conference on Pattern Recognition (ICPR 2010). IEEE Computer Society, 2010.
- [31] BOVIK H. Image information and visual quality [J]. IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society, 2006, 15(2): 430.
- [32] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment; from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [33] YOU C, YANG L, ZHANG Y, et al. Low-Dose CT via Deep CNN with Skip Connection and Network in Network[J]. arXiv: 1811.10564, 2018.
- [34] DABOV K, FOI A, KATKOVNIK V, et al. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering [C] // IEEE Transactions on Image Processing. IEEE, 2007; 2080-2095.
- [35] LUTHRA A, SULAKHE H, MITTAL T, et al. Eformer: Edge enhancement based transformer for medical image denoising[J]. arXiv:2109.08044, 2021.
- [36] ZHANG Z, HAN H, SHANGGUAN X, et al. Artifact and Detail Attention Generative Adversarial Networks for Low-Dose CT Denoising[J]. IEEE Transactions on Medical Imaging, 2021, 40(12): 3901-3918.



**GUO Guangxing**, born in 1978, Ph.D, associate professor. His main research interests include big data image processing and so on.



**QIANG Yan**, born in 1969, Ph.D, is a member of China Computer Federation. His main research interests include intelligent information processing, image recognition and machine learning.

(责任编辑:柯颖)