

低剂量 CT 图像全变分深度展开去噪网络

吴 涵^{1,2}, 张鹏程^{1,2}, 桂志国^{1,2}, 刘 祎^{1,2}

(1. 中北大学 信息与通信工程学院, 太原 030051;

2. 中北大学 生物医学成像与影像大数据山西重点实验室, 太原 030051)

摘要: 对低剂量 CT 图像去噪进行了研究, 分析了神经网络去噪在伪影抑制中计算性能低、泛化性不足的问题; 采用各向异性全变分深度展开去噪网络, 新方法结合图像相邻体素的边缘特性, 引入各向异性 TV 正则项保留图像结构信息, 避免各向同性 TV 导致的边缘模糊, 并通过 Chambolle-Pock 算法求解数学模型, 适配深度展开到卷积神经网络; 此外, 结合像素注意力机制进行网络优化, 捕捉图像中的重要细节; 经实验测试, 基于 Mayo 2016 数据集, 该方法在图像去噪效果上优于传统方法及其他先进网络模型, 在 PSNR、SSIM 和 VIF 等指标上表现更优, 满足低剂量 CT 图像高质量重建的需求。

关键词: 图像去噪; 计算机断层扫描; 原始对偶算法; 模型驱动; 注意力机制

Deep Total Variation Denoising Network for Low-Dose CT Images

WU Han^{1,2}, ZHANG Pengcheng^{1,2}, GUI Zhiguo^{1,2}, LIU Yi^{1,2}

(1. College of Information and Communication Engineering, North University of China, Taiyuan 030051, China;

2. Shanxi Key Laboratory of Biomedical Imaging and Big Data, North University of China, Taiyuan 030051, China)

Abstract: Research on denoising low-dose CT images was conducted, the issues of low computational performance and insufficient generalization in neural network denoising for artifact suppression were analyzed. An anisotropic total variation deep unfolding denoising network was adopted, with the new method incorporating the edge characteristics of adjacent voxels, the anisotropic TV regularization term was introduced to preserve the structural information of images and avoid the edge blurring caused by isotropic TV. The Chambolle-Pock algorithm was employed to solve the mathematical model, suitable for deep unfolding into convolutional neural networks. Additionally, the pixel attention mechanism was integrated for the network optimization to capture the important information of images. Through experimental tests on the Mayo 2016 dataset, this method has advantages over traditional methods and other advanced network models in image denoising, showing a better performance in the indicators of PSNR, SSIM, and VIF. This method meets the requirements for high-quality reconstruction of low-dose CT images.

Keywords: image denoising; computed tomography scanning; primal-dual algorithm; model-driven; attention mechanism

0 引言

计算机断层扫描技术能够清晰地显示解剖结构, 传统的 CT 扫描在获得最佳图像质量时通常需要高剂量的辐射。然而, 高剂量辐射会增加患癌症的风险, 尤其是对于弱势群体, 其影响更为显著。为了降低辐射剂量, 一些 CT 扫描可能会采用低剂量的方案, 但这可能会导致图像中噪声的增加, 从而影响图像质量。因此, 研发降噪算法和技术变得尤为重要, 以有效地去除图像中的噪声, 同时保持图像的清晰度和准确性。这不仅可以帮助减少患者的辐射暴露风险, 还能提高医学影像的诊断质量, 为临床诊断和治疗提供更可靠的依据。

传统的低剂量 CT 图像去噪算法^[1]有块匹配的 3D 滤波算法^[2] (BM3D) 方法、非局部自相似稀疏模型^[3]、全变

分^[4] (TV) 模型和低秩^[5] 模型等等。尽管这些算法在一定程度上可以减少噪声, 但解决复杂的优化问题需要耗时的计算过程, 并且解决非凸问题通常需要自行调试参数以增强去噪性能。随着神经网络的发展, 卷积神经网络被证明具有良好的图像去噪的潜力。着重于学习如何从噪声图像中重建清晰图像的映射。这些方法^[6-7] 通过分层网络学习图像特征, 能够有效处理各种复杂的噪声。文献 [8] 提出了一种残差编码器-解码器卷积网络 (REDCNN, residual encoder-decoder CNN), 融合了自动编码器、反卷积网络和高效连接的优势, 旨在提升神经网络的去噪能力并促进深度特征提取。文献 [9] 介绍了一种边缘增强模块, 并提出了基于此模块的密集连接卷积网络 (EDCNN, edge enhancement based densely connected CNN)。该模块有效地将边缘信息融入去噪过程中, 为图像处理领域带来了全新的思路。

收稿日期: 2024-05-28; 修回日期: 2024-07-09。

基金项目: 山西省基础研究计划项目 (202303021211148, 202103021224204, 20210302124403); 山西省回国留学人员科研资助项目 (2021-111)。

作者简介: 吴 涵 (2000-), 男, 大学本科。

通讯作者: 桂志国 (1972-), 男, 博士, 教授。

引用格式: 吴 涵, 张鹏程, 桂志国, 等. 低剂量 CT 图像全变分深度展开去噪网络[J]. 计算机测量与控制, 2024, 32(12): 229-235.

文献 [10] 引入了内容-噪声互补学习 (CNCL, content-noise complementary learning) 策略, 这一策略的独特之处在于同时学习图像域噪声和内容, 以提高去噪效率。随着 Transformer 技术与计算机视觉任务的结合, 文献 [11] 利用无卷积 T2T 视觉原理推出了 Transformer 编码器-解码器扩展网络 (TED-Net, transformer encoder-decoder dilation network)。这为传统卷积方法提供了一种替代方案。但神经网络有一些局限, 包括不透明的网络设计、次优的系统优化、有限的网络泛化, 因此, 追求可解释、泛化性强的网络模型仍是一个突出且持续的挑战。最近, 有研究将迭代算法展开^[12-13]到神经网络中, 即将迭代优化算法的每一步都表示为神经网络的层次结构。网络中的一个 Layer 代表迭代算法的一次迭代。这样, 整个优化过程可以被看作是一个深度神经网络的前向传播过程。例如, 受到交替方向乘子算法启发的网络 (ADMM-Net)^[14]将稀疏变换算子通过卷积的形式表示; 由迭代收缩阈值算法启发的网络 (ISTA-Net+)^[15]利用简单的 CNN 卷积层模拟模型中的非线性变换。这些迭代算法展开到神经网络虽然提高了性能并增强了网络可解释性, 但拟合度较低, 网络不能很好地表示算法的优化细节。

为优化上述存在的一系列问题, 本文提出了一种基于 CP 算法的各向异性全变分的去噪模型, 并深度展开到卷积神经网络 (DTV-Net)。该方法考虑相邻图像体素之间的边缘特性, 并在不同方向上引入不同的权值, 自适应地从每个迭代图像中获得权值, 以避免各向同性 TV 导致的边缘模糊。与其他方法不同, 该方法展开到网络的拟合度高, 适应性强, 有更高的泛化能力和更好的网络解释性。通过实验结果可以表明, DTV-Net 较其他实验方法相比, 去噪效果更好。

1 去噪模型及优化算法

在低剂量 CT 图像去噪领域, 全变分模型^[16]和 Chambolle-Pock 算法作为重要的方法得到了广泛关注。全变分模型因其在保留图像边缘和细节方面的优越性能, 成为图像去噪的重要技术。而 Chambolle-Pock 算法作为一种高效的优化方法, 能够有效求解包括全变分模型在内的各种图像处理问题。因此, 在本章中, 将详细介绍这两种技术的基本原理, 作为后续研究的基础。

1.1 全变分模型

LDCT 图像去噪的目的是从带噪声的图像中提取有效信息, 抑制噪声, 从而改善图像质量。目前针对去噪已有很多方法, 如偏微分方程, 小波变换^[17]和全变分。其中全变分方法在去除噪声的基础上可以保留图像的边缘信息, 本研究引用的是 Rudin, Osher 和 Fatemi 提供的全变分模型, 即:

$$\min_{u \in Y} \left\{ \frac{1}{2} \|u - g\|^2 + \lambda \|(|\nabla u|)\|_1 \right\} \quad (1)$$

对于式 (1) 的优化问题, 目前已有很多求解算法, 例如交替方向乘子法^[18], 迭代收缩/阈值算法^[19], 和原始对偶

梯度下降算法^[20]。但针对 ROF 去噪模型, 本文选择使用参数更少、收敛性更强的 CP 算法^[21-22]。

1.2 Chambolle-Pock 算法

CP 算法适用于原始最小化^[23]的一般形式:

$$\min_{u \in X} [F(Kx) + G(x)] \quad (2)$$

对偶最大化的一般形式为:

$$\max_{u \in Y} [-F^*(y) - G^*(-K^T y)] \quad (3)$$

式 (2) 和式 (3) 中, x 和 y 分别为空间 X 和 Y 中的矢量, K 为 X 到 Y 的线性变换, 符号 “ T ” 为转置, $*$ 表示凸共轭。 G 和 F 均是不平滑的凸函数。凸函数 $R(z)$ 的共轭定义为式 (4):

$$R^*(z) = \max_z [(z, z') - R(z')] \quad (4)$$

式中, 符号 (\cdot, \cdot) 表示两个向量的内积。由此原始对偶问题可转为鞍点最优化问题。其形式如下:

$$\min_x \max_y \{ (Kx, y)_Y + G(x) - F^*(y) \} \quad (5)$$

然后对其求最大化 y 和最小化 x , CP 算法如下:

CP 算法框架 (N 步迭代)

Pseudocode for N steps of the Chambolle-Pock algorithm.

1) $L \leftarrow \|K\|_2; \tau \leftarrow 1/L; \sigma \leftarrow 1/L; \theta \leftarrow 1; n \leftarrow 0$

2) initialize x_0 and y_0 to zero value

3) $\bar{x}_0 \leftarrow x_0$

4) Repeat

5) $y_{n+1} \leftarrow \text{prox}_\sigma[F^*](y_n + \sigma Kx_n)$

6) $x_{n+1} \leftarrow \text{prox}_\tau[G](x_n - \tau K^T y_{n+1})$

7) $\bar{x}_{n+1} \leftarrow x_{n+1} + \theta(x_{n+1} - x_n)$

8) $n \leftarrow n + 1$

9) until $n \geq N_{\text{iter}}$

CP 算法中, L 为矩阵 K 的最大奇异值。详细求解步骤见文献 []。附录 A。 $\text{prox}_\sigma[F^*]$ 和 $\text{prox}_\tau[G]$ 是最邻近映射操作。 prox 最邻近映射的表达如式 (6) 所示:

$$\text{prox}_a[H](z) = \underset{z'}{\text{argmin}} \left[H(C) + \frac{\|z - z'\|_2^2}{2a} \right] \quad (6)$$

该式实质上是一个最优化问题^[25], 但由于 z 是一个变量, 因此其结果是关于 z 的函数。

2 本文方法

低剂量 CT 图像在减少辐射剂量的同时不可避免地引入噪声, 影响了图像的诊断质量。尽管现有去噪方法在一定程度上改善了图像质量, 但在平衡去噪效果与边缘保留方面仍有提升空间。此外, 这些方法在计算效率和泛化能力上也存在不足, 需要进一步优化。

DTV-Net 的设计思路包括以下几个方面: 结合各向异性 TV 正则项以避免边缘模糊; 采用 Chambolle-Pock 算法提高优化效率; 将迭代优化过程深度展开至卷积神经网络中, 增强模型的可解释性和拟合度; 结合像素注意力机制, 捕捉图像中的重要细节。实验结果表明, DTV-Net 在图像去噪和边缘保留方面表现出色, 同时在计算效率和泛化能力上也有显著提升。

2.1 CP 算法求解 DTV 模型

为了更好地保留图像结构细节, 本文在 ROF 模型的基础上, 改进成各向异性全变分正则项, 如下:

$$\min_{u \in X} \left\{ \frac{1}{2} \|u - g\|^2 + \lambda_1 \|(|\nabla_x u|)|\|_1 + \lambda_2 \|(|\nabla_y u|)|\|_1 \right\} \quad (7)$$

其中: X 作为有限维向量空间, $\|\cdot\|$ 表示为 1 范数, ∇_x 和 ∇_y 分别代表水平和垂直方向上的梯度算子, λ_1 和 λ_2 为正则项参数。

如 1.2 节所述, 使用 CP 算法去除噪声的关键所在是 F 的凸共轭函数的求取。将式 (7) 根据 (2) 式改写, 即:

$$J(x, y, z) = G(x) + F_1(y) + F_2(z) \quad (8)$$

其中:

$$G(x) = \frac{1}{2} \|x - g\|^2, F(y, z) = F_1(y) + F_2(z) \quad (9)$$

$$F_1(y) = \lambda_1 \|(|y|)|\|_1, F_2(z) = \lambda_2 \|(|z|)|\|_1 \quad (10)$$

$$y = \nabla_x u, z = \nabla_y u, K = \begin{pmatrix} \nabla_x \\ \nabla_y \end{pmatrix} \quad (11)$$

式 (9) 中, $J(x, y, z)$ 是 3 个凸函数的和。根据式 (6) 可得:

$$G^*(x) = \frac{1}{2} \|x\|_2^2 + (x, g) \quad (12)$$

$$F_1^*(y) = \delta_{\text{Box}(\lambda_1)}(|y|), F_2^*(z) = \delta_{\text{Box}(\lambda_2)}(|z|) \quad (13)$$

$\delta_{\text{Box}(a)}(|x|)$ 是一个指示函数其定义如下:

$$\delta_{\text{Box}(a)}(|x|) \equiv \begin{cases} 0 & \|x\|_\infty \leq a \\ \infty & \|x\|_\infty > a \end{cases} \quad (14)$$

根据式 (6) 可得 $F_1^*(y)$ 和 $F_2^*(z)$ 以及 $G(x)$ 的 $prox$ 近端映射如下:

$$prox_\sigma[G^*](x) =$$

$$\underset{x'}{\operatorname{argmin}} \left\{ \frac{1}{2} \|x' - g\|_2^2 + (x', g) + \frac{\|x - x'\|_2^2}{2\sigma} \right\} = \frac{y - \sigma g}{1 + \sigma} \quad (15)$$

$$prox_\sigma[F_1^*](y) = \frac{\lambda_1 y}{\max \cdot (\lambda_1, I, |y|)} \quad (16)$$

$$prox_\sigma[F_2^*](z) = \frac{\lambda_2 z}{\max \cdot (\lambda_2, I, |z|)} \quad (17)$$

由公式 (15) ~ (17) 可得到 J^* 的近端映射:

$$prox_\sigma[J^*](x, y, z) =$$

$$\left(\frac{x - \sigma g}{1 + \sigma}, \frac{\lambda_1 y}{\max \cdot (\lambda_1, I, |y|)}, \frac{\lambda_2 z}{\max \cdot (\lambda_2, I, |z|)} \right) \quad (18)$$

式中, I 的元素值全为 1。式 (18) 可以代入 CP 算法, 得到 DTV-CP 算法实例。

DTV-CP 算法实例:

Pseudo-code of the DTV-CP algorithm for solving Eq. (8).

1) $L \leftarrow \|K\|_2; \tau \leftarrow 1/L; \sigma \leftarrow 1/L; \theta \leftarrow 1; n \leftarrow 0$

2) Initialize, u_0, p_0 and q_0 to zero value

3) $\bar{u}_0 \leftarrow u_0$

4) Repeat

5) $p_{n+1} \leftarrow \lambda_1 (p_n + \sigma \Delta \nabla_x \bar{u}_n) / \max \cdot (\lambda_1, |p_n + \sigma \nabla_x \bar{u}_n|)$

6) $q_{n+1} \leftarrow \lambda_2 (p_n + \sigma \nabla_y \bar{u}_n) / \max \cdot (\lambda_2, |q_n + \sigma \nabla_y \bar{u}_n|)$

$$7) u_{n+1} \leftarrow (u_n + \sigma(\bar{u}_n - g)) / (1 + \sigma)$$

$$8) \bar{u}_{n+1} \leftarrow (u_n - \tau \nabla_x^T p_{n+1} - \tau \nabla_y^T q_{n+1} - \tau u_{n+1})$$

$$9) n \leftarrow n + 1$$

$$10) \text{until } n \geq N_{\text{iter}}$$

τ 和 σ 的值一直都是 $1/L$, θ 值为 1, n 为 0, u_0, p_0, q_0 素均为 0。表 2 中第 5~8 行分别是对 p_0, q_0, u_0 和 \bar{u}_0 的更新, 最后的迭代结果为 \bar{u}_{n+1} 。

2.2 算法展开

由 DTV-CP 算法可得, 将其迭代过程深度展开到卷积神经网络中, 构建 DTV-Net, 通过不同的卷积层模拟 CP 算法中 p_n, q_n, u_n 的迭代公式 (5) ~ (8) 进行运算, 如式 (19) ~ (21):

$$p_{n+1} = SC_PA\{\Psi_{n+1}[\lambda_1(p_n + \sigma \nabla_x u_n)]\} \quad (19)$$

$$q_{n+1} = SC_PA\{\Psi_{n+1}[\lambda_2(q_n + \sigma \nabla_y u_n)]\} \quad (20)$$

$$\bar{u}_{n+1} = u_n + \varphi_{n+1} \left[\frac{\tau \nabla_x^T P_{n+1} + \tau \nabla_y^T Q_{n+1} + \tau g}{(1 + \tau)} \right] \quad (21)$$

其中: $SC_PA(\cdot)$ 为像素注意力模块, $\Psi_{n+1}(\cdot)$ 和 $\varphi_{n+1}(\cdot)$ 表示不同结构的卷积层。将迭代公式展开到神经网络模型的过程描述如下。首先, 利用卷积层模块分别计算输入特征图 u_n 的梯度 $\nabla_x u_n$ 和 $\nabla_y u_n$, 具体而言, ∇_x 卷积层的卷积核为 $[-1 \ 1]$, 用于提取水平方向上的梯度;

而 ∇_y 卷积层的卷积核为 $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, 用于提取垂直方向上的梯度。

这些梯度信息将在后续步骤中用于更新变量 p_n 和 q_n 。然后, 通过卷积层模块 Ψ_{n+1} 对 $\lambda_1(p_n + \sigma \nabla_x u_n)$ 和 $\lambda_2(q_n + \sigma \nabla_y u_n)$ 进行处理, 这一步骤的目的是对输入特征进行卷积操作, 提取重要的特征信息, 并为下一步的像素注意力机制做好准备。接下来, 将卷积层输出的特征图输入到像素注意力模块 SC_PA 中, 进行加权处理。像素注意力机制通过对特征图进行加权和, 提升特定区域的特征表达能力, 从而有效替代原公式中的 $\max(\lambda_1, |p_n + \sigma \nabla_x u_n|)$ 和 $\max(\lambda_2, |q_n + \sigma \nabla_y u_n|)$ 操作。

像素注意力机制的合理性在于它能够自适应地调整不同位置上的特征权重, 从而实现与 \max 操作类似的效果。传统的 \max 操作选择最大值, 而像素注意力机制通过学习权重分布来强调重要特征, 具有更高的灵活性和表示能力。随后, 使用卷积层 ∇_x^T 和 ∇_y^T 对更新后的 p_{n+1} 和 q_{n+1} 进行转置梯度计算, 得到 $\nabla_x^T P_{n+1}$ 和 $\nabla_y^T Q_{n+1}$ 。最后, 将 $\tau \nabla_x^T P_{n+1} + \tau \nabla_y^T Q_{n+1} + \tau g$ 输入到卷积层模块 φ_{n+1} 中, 进行特征提取和线性组合操作, 最终通过参数优化得到新的特征图 \bar{u}_{n+1} , 网络训练时自动优化 λ_i, σ, τ 这些参数, 确保了更新和优化。经过 N 次迭代后, 利用 1×1 卷积融合图像数据, 得到最终输出图像。通过上述过程, 原始的迭代公式被成功展开为神经网络模型中的具体操作, 使每一步都对应相应的卷积层和像素注意力机制处理, 从而确保了模型的有效实现和优化。最后用 $\theta_i^1, \theta_i^2, \theta_i^p$ 表示 λ_i, σ, τ 和 θ 这些参数, 最终网络展开结构如式 (22):

$$\begin{cases} p_{n+1} = SC_PA\{\Psi_{n+1}[(p_n, \nabla_x u_n), \theta_i^1]\} \\ q_{n+1} = SC_PA\{\Psi_{n+1}[(q_n, \nabla_y u_n), \theta_i^2]\} \\ \bar{u}_{n+1} = u_n + \varphi_{n+1}[(\nabla_x^T P_{n+1}, \nabla_y^T Q_{n+1}, g), \theta_i^p] \end{cases} \quad (22)$$

将传统迭代算法展开为神经网络模型优点如下：神经网络具有强大的参数优化能力，能够自动调整算法中的参数，使得模型更好地适应不同的数据集和任务。其次，通过卷积层和像素注意力机制的引入，提升了模型的灵活性和表达能力，能够更有效地处理复杂的特征关系和局部信息。此外，神经网络模型的并行计算能力使得整个迭代过程可以高效地实现，加速了算法的收敛速度和处理效率。

2.3 网络架构

图 1 展示了本文的 DTV-Net 网络架构。一次迭代由 3 个模块 (Block) 构成：PBlock 和 QBlock 分别用于更新对偶变量 P_n 和 Q_n ，而 UBlock 则更新原始变量 u_n ，具体更新步骤如公式 (22) 所定义。原始模块由图 1 中虚线框内的 PBlock 和 QBlock 组成，其输入为 g ， P_n ， Q_n 和 u_n ，输出新的变量 u_{n+1} ，卷积层通道数分别为 49、48 和 24，激活函数采用 ReLU，以便捕捉输入特征中的非线性信息并促进梯度流动。对偶模块的架构与原始模块类似，在图 1 中表示为 UBlock，但其激活函数选择为 Tanh。选择 Tanh 函数的原因在于它能够维持输入和输出之间的非线性对应关系，同时抑制梯度消失问题。在初始化阶段，对偶数据直接初始化为零，这样可以简化计算并提供一个稳定的起点；而原始数据则通过线性映射来计算初始化值，以确保输入数据具有合理的初始状态，从而促进网络的有效学习和快速收敛。通过这种设计，网络能够充分利用卷积层和像素注意力机制的优势，将原始的迭代公式转化为具体的神经网络操作，确保模型的有效实现和优化，从而实现对低剂量图像的高质量重建。

在前两个模块中使用了像素注意力机制，如图 2 所示，其中 SC 表示自校准卷积 (Self-Calibrated Convolution)，PA 表示像素注意力。该模块通过堆叠 SC-PA 块，实现对输入特征的非线性映射，从而增强模型的表达能力。自校准卷积技术利用输入特征图的内部信息动态调整卷积核的

权重，以自适应方式改进特征表达，有效捕捉和强化数据中的复杂特征和模式。像素注意力机制则通过加权处理特征图中的不同像素，提升对关键像素区域的关注度，增强了模型在重要特征提取和任务感知方面的精度和效率。SC-PA 模块具备可堆叠的优势，多个 SC-PA 块可以串联使用，形成深层次的特征抽取和处理网络。这种层级的特征处理机制使得模型能够逐步优化特征表示，提升整体的数据解释能力和任务执行效果，使得网络能够逐步优化特征表示，更好地满足对 LDCT 图像去噪的需求。整个网络框架的输入为低质量图像 g ，输出为重建后的高质量图像 u_{out} ， P_n 、 Q_n 和 u_n 是迭代过程中的变量，用于逐步逼近最终输出。

3 实验设计

3.1 实验数据

使用了 Mayo Clinic 授权的 Mayo 2016 LDCT 数据集^[29]。数据集均采用了全剂量的 CT 图像和四分之一剂量 CT 图像，图片尺寸均为 512×512 。在网络训练时，根据针对十名病人的 CT 图片，共采用了 9 名患者的 5 000 幅低剂量的 CT 图片作为训练集，以及一名病人的 500 幅图片作为测试集。在训练前，采用了随机裁剪。通过随机裁剪的方式，使得模型能够学习到不同部位的特征，从而更全面地理解图像内容。这也有助于减轻过拟合的风险，使模型更加稳健可靠。

3.2 网络训练

采用了 Adam 优化算法，该算法在处理大规模数据和参数更新时表现出色。Adam 结合了动量和自适应学习率，可以加速收敛过程，减少训练时间。其学习率设定为 10^{-4} ，权重衰减参数 β_1 和 β_2 分别设置为 0.9 和 0.999。批量大小设置为 16，这是在考虑显存容量和训练效率的基础上确定的。较小的批量大小可以使模型更快地进行参数更新，但需要更多的训练迭代次数。采用了最小均方误差 (MSE) 作为损失函数，定义如下：

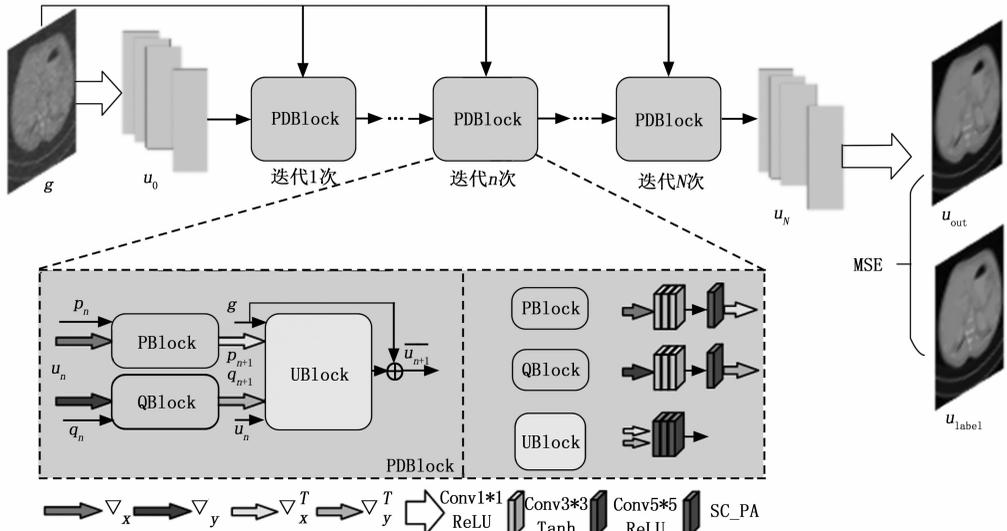


图 1 DTV-net 网络架构图

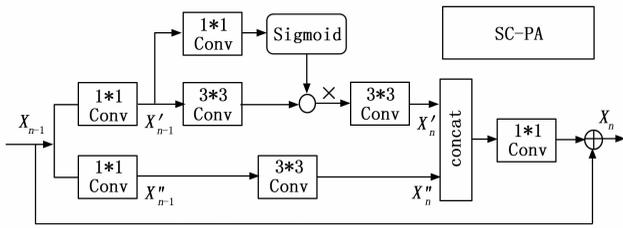


图 2 SC-PA 模块结构图

$$L_{MSE} = \|u_{out} - u_{label}\|_2^2 \quad (23)$$

其中: u_{out} 是网络测试结果, u_{label} 则是标签图像。

3.3 评估指标

本研究比较了 5 种去噪方法进行评价, 包括 EDCNN、REDCNN、BM3D、CNCL 以及 TEDNet 算法。采用 3 个客观评价指标, 即峰值信噪比 (PSNR)^[27]、结构相似性 (SSIM)^[27] 和视觉信息保真度 (VIF)^[28] 来定量分析 LDCT 图像的质量。VIF 计算公式引用文献 [28]。PSNR 和 SSIM 计算方式如式 (24) ~ (25):

$$PSNR(u_{label}, u_{out}) = 10 \times \lg \left[\frac{(2^n - 1)^2}{\|u_{label} - u_{out}\|_2^2} \right] \quad (24)$$

$$SSIM(u_{label}, u_{out}) = \frac{2\mu_{u_{label}}\mu_{u_{out}} + C_1}{\mu_{u_{label}}^2 + \mu_{u_{out}}^2 + C_1} \times \frac{2\sigma_{u_{label}u_{out}} + C_2}{\sigma_{u_{label}}^2 + \sigma_{u_{out}}^2 + C_2} \quad (25)$$

其中: $\mu_{u_{label}}$, $\mu_{u_{out}}$ 分别为标签图像和网络测试图像的均值, $\sigma_{u_{label}}^2$ 和 $\sigma_{u_{out}}^2$ 是两者的方差, $\sigma_{u_{label}u_{out}}$ 代表协方差, C_1 和 C_2 为常数。

4 实验结果与分析

为了全面评估所提出的去噪方法的性能, 采用了定性和定量两种分析方式。定性分析主要通过视觉比较来评价去噪后的图像质量和细节保留情况, 而定量分析则使用 PSNR、SSIM 和 VIF 等客观指标来量化去噪效果。通过这两种分析方式, 能够综合评估各去噪方法的优缺点, 并验证所提出方法在 LDCT 图像去噪中的优势。

4.1 定性分析

图 3 展示了各种去噪方法对腹部切片图像 (包括 LDCT 和 NDCT 图像) 的视觉效果。LDCT 图像中的感兴趣区域 (ROI) 通过红框标出, 以便比较去噪效果和细节保留能力。研究结果显示, BM3D 方法被证明可以有效减少噪声, 但其可能在此过程中会模糊图像细节, 突显了在噪声去除和细节保留之间的权衡。相比之下, 基于深度学习的方法在去噪效果上超越了 BM3D, 但在保留图像细节方面表现不尽相同。以 EDCNN 方法为例, 虽然成功去除了部分散斑噪声, 但也不可避免地掩盖了一些图像细节, 这可能会对后续的医学诊断造成影响。另一方面, REDCNN 方法在去除噪声的同时平滑了图像, 导致了细节的模糊化, 这也是需要考虑的一个问题。因此, 在选择适当的去噪方法时, 需要综合考虑去噪效果和图像细节保留的平衡, 以确保最终得到的图像既清晰又准确。而本文提出的 DTV-Net 方法在细节保留方面优于 EDCNN, 并在图像平滑处理方面优于

REDCNN, 尤其是在虚线圆圈处, DTV-Net 能够最明显地显示细节差异, 这进一步证实了其在细节保留和噪声抑制方面的有效性。

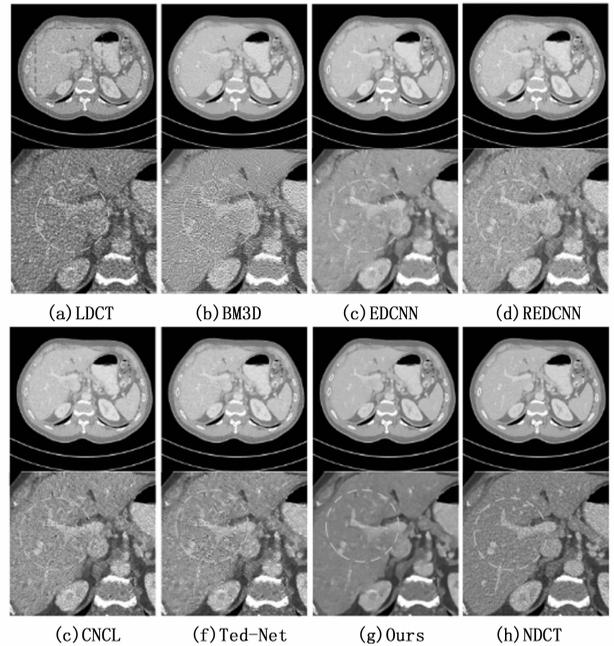


图 3 各方法对比实验去噪效果比较图 (腹部切片)

实验证实了迭代网络方法在消除条带伪影方面的有效性, 并且这一效果在针对胸部实验中也得到了验证。在图 4 中, 不同的降噪方法展现了不同的效果, 尤其是在图像的兴趣区域 (ROI) 中。BM3D 方法在图像平滑处理中引入了新的横向条纹伪影, 这可能会干扰后续图像分析和诊断。而 REDCNN、EDCNN 和 CNCL 方法虽然在降噪过程中取得了一定效果, 但也不可避免地牺牲了一些图像细节, 这对于需要高精度图像的医学诊断应用而言, 可能是不可接受的。然而, 本文提出的 DTV-Net 方法在实验中表现出

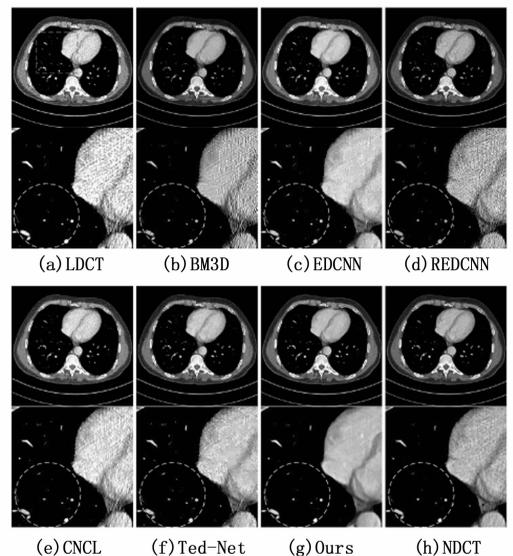


图 4 各方法对比实验去噪效果比较图 (胸部切片)

色。它不仅成功消除了大部分图像噪声，还有效地消除了条纹伪影，尤其是在保留了 NDCT 图像中感兴趣区域内的边缘信息的同时。从而证实了 DTV-Net 方法在医学图像处理中的潜在应用和优势，为进一步的研究和应用提供了有力支持。

为了全面验证 DTV-Net 的去噪效果及其在不同部位 CT 图像去噪的稳定性，还选取了盆腔区域 CT 图像作为实验对象，结果如图 5 所示。在盆腔区域的感兴趣区 (ROI) 图像中，可以明显看出，BM3D 方法在低对比度区域容易引起模糊，导致整体去噪效果不佳。相比之下，EDCNN 方法虽然能够平滑噪声，但过于平滑，可能会牺牲掉图像中的重要细节。图 5 (g) 展示了采用 DTV-Net 方法的去噪结果。与 REDCNN 相比，该方法不仅有效地去除了噪声，还能使黄圈内的结构信息更加清晰，保留了更多的细节和边缘信息。这表明 DTV-Net 在处理复杂结构和低对比度区域时表现出色，能够在去噪的同时保持图像的高质量 and 细节完整性。通过对比不同方法在盆腔区域 CT 图像中的表现，进一步验证了 DTV-Net 在实际应用中的优越性和稳定性。

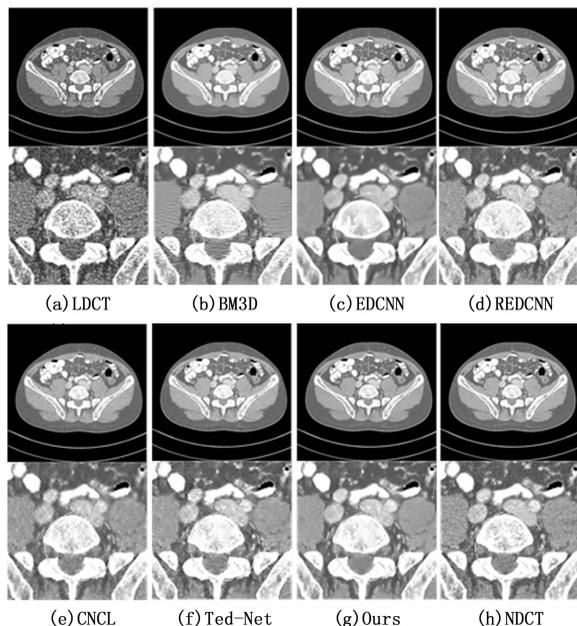


图 5 各方法对比实验去噪效果比较图 (盆腔切片)

4.2 定量分析

表 1 列出了使用 560 幅图像进行测试的评估指标结果，包括每种方法的平均值及其标准偏差。通过对这些数据及视觉效果进行综合分析，可以明显看出，在所比较的各种图像去噪方法中，DTV-Net 显示出了最显著的性能优势。具体来说，DTV-Net 在去除噪声的同时有效地保留了图像的边缘信息，这一特性与 NDCT 方法的表现非常相近。

5 结束语

为了解决神经网络的不透明性和各向同性 TV 导致的边缘模糊问题，本研究提出了 DTV-net，一种通过各向异性 TV 正则项增强的卷积神经网络。该网络利用 CP 算法，

表 1 各实验的图像质量指标 (均值±标准差)

方法	PSNR/dB	SSIM	VIF
LDCT	27.956±1.297	0.591±0.064	0.331±0.049
BM3D	32.739±1.038	0.755±0.051	0.587±0.048
EDCNN	33.722±1.097	0.805±0.036	0.517±0.053
REDCNN	34.228±1.155	0.811±0.038	0.586±0.052
CNCL	33.748±1.069	0.798±0.035	0.579±0.042
TED-NET	33.950±1.141	0.808±0.034	0.566±0.051
本文方法	34.536±1.152	0.830±0.031	0.643±0.042

通过卷积层和像素注意模块自动学习迭代信息，从而减轻神经网络的黑盒特性，并最大限度地减少边缘模糊。为了验证所提算法的性能和泛化能力，使用 Mayo 2016 低剂量 CT 数据集进行实验，并与 BM3D、REDCNN、EDCNN、CNCL 和 TED-Net 等算法进行了对比。结果显示，所提方法在各项图像评估指标上均优于其他方法，并通过视觉分析证实了其优越性。未来的研究探讨将该方法扩展到其他模型，并进一步验证算法的收敛性和网络优化，以取得更优异的结果。

参考文献:

- [1] TIAN C, FEI L, ZHENG W, et al. Deep learning on image denoising: an overview [J]. *Neural Networks*, 2020, 131: 251-275.
- [2] FERUGLIO P F, VINEGONI C, GROS J, et al. Block matching 3D random noise filtering for absorption optical projection tomography [J]. *Physics in Medicine & Biology*, 2010, 55 (18): 5401.
- [3] MEI Y, FAN Y, ZHOU Y. Image super-resolution with non-local sparse attention [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 3517-3526.
- [4] CAI W, JIANG J, OUYANG S. Hyperspectral image denoising using adaptive weight graph total variation regularization and low-rank matrix recovery [J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 19: 1-5.
- [5] WANG Z, QIAN C, GUO D, et al. One-dimensional deep low-rank and sparse network for accelerated MRI [J]. *IEEE Transactions on Medical Imaging*, 2022, 42 (1): 79-90.
- [6] ZHANG K, GAO X, TAO D, et al. Single image super-resolution with non-local means and steering Kernel regression [J]. *IEEE Transactions on Image Processing*, 2012, 21 (11): 4544-4556.
- [7] LIU P, ZHANG H, ZHANG K, et al. Multi-level wavelet-CNN for image restoration [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Piscataway: IEEE, 2018: 773-782.
- [8] 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.

- [9] LIANG T, JIN Y, LI Y, et al. Edcnn: Edge enhancement-based densely connected network with compound loss for low-dose ct denoising [C] //Proceedings of the 15th IEEE International Conference on Signal Processing. Piscataway: IEEE, 2020, 1: 193 - 198.
- [10] GENG M, MENG X, YU J, et al. Content-noise complementary learning for medical image denoising [J]. IEEE Transactions on Medical Imaging, 2021, 41 (2): 407 - 419.
- [11] WANG D, WU Z, YU H. TED-net: convolution-free T2T vision transformer-based encoder-decoder dilation network for low-dose CT denoising [C] //Proceedings of the 2021 International Workshop on Machine Learning in Medical Imaging. Springer, Cham, 2021: 416 - 425.
- [12] MONGA V, LI Y, ELDAR Y C. Algorithm unrolling: interpretable, efficient deep learning for signal and image processing [J]. IEEE Signal Processing Magazine, 2021, 38 (2): 18 - 44.
- [13] XIA W, SHAN H, WANG G, et al. Synergizing physics/model-based and data-driven methods for low-dose CT [J]. Arxiv Preprint Arxiv: 2203.15725, 2022.
- [14] YANG Y, SUN J, LI H, et al. ADMM-CSNet: A deep learning approach for image compressive sensing [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 42 (3): 521 - 538.
- [15] YOU D, XIE J, ZHANG J. ISTA-NET++: Flexible deep unfolding network for compressive sensing [C] //Proceedings of the 2021 IEEE International Conference on Multimedia and Expo. Piscataway: IEEE, 2021: 1 - 6.
- [16] LEE C O, PARK J. Fast nonoverlapping block Jacobi method for the dual Rudin-Osher-Fatemi model [J]. SIAM Journal on Imaging Sciences, 2019, 12 (4): 2009 - 2034.
- [17] TIAN C, ZHENG M, ZUO W, et al. Multi-stage image denoising with the wavelet transform [J]. Pattern Recognition, 2023, 134: 109050.
- [18] 王 心, 朱浩华, 刘光灿. 卷积鲁棒主成分分析 [J]. 计算机应用, 2021, 41 (5): 1314 - 1318.
- [19] LI J, ZHU Q, WU Y, et al. Image reconstruction based on deep iterative shrinkage network [C] //Proceedings of the 6th International Conference on Image, Vision and Computing. Piscataway: IEEE, 2021: 259 - 263.
- [20] CHAMBOLLE A, POCK T. A first-order primal-dual algorithm for convex problems with applications to imaging [J]. Journal of Mathematical Imaging and Vision, 2011, 40: 120 - 145.
- [21] PELT D M, BATENBURG K J. Improving filtered backprojection reconstruction by data-dependent filtering [J]. IEEE Transactions on Image Processing, 2014, 23 (11): 4750 - 4762.
- [22] 席雅睿, 乔志伟, 温 静, 等. 基于 Chambolle-Pock 算法框架的高阶 TV 图像重建算法 [J]. 计算机应用, 2020, 40 (6): 1793 - 1798.
- [23] LIU Y, MA J H, FAN Y, et al. Adaptive-weighted total variation minimization for sparse data toward low-dose x-ray computed tomography image reconstruction. [J]. Physics In Medicine and Biology, 2012, 57 (23): 7923 - 7956.
- [24] XU L, LU CW, XU Y, et al. Image smoothing via gradient minimization [J]. ACM Transactions on Graphics, 2011, 30 (6): 1 - 12.
- [25] MEINHARDT T, MOLLER M, HAZIRBAS C, et al. Learning proximal operators: using denoising networks for regularizing inverse imaging problems [C] //Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 1781 - 1790.
- [26] ZHAO H, KONG X, HE J, et al. Efficient image super-resolution using pixel attention [C] //Computer Vision-ECCV 2020 Workshops; Springer International Publishing, 2020: 56 - 72.
- [27] HORE A, ZIOU D. Image quality metrics: PSNR vs. SSIM [C] //Proceedings of the 20th International Conference on Pattern Recognition. IEEE, 2010: 2366 - 2369.
- [28] SHEIKH H R, BOVIK A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15 (2): 430 - 444.
- [29] MCCOLLOUGH C H, BARTLEY A C, CARTER R E, et al. Low-dose CT for the detection and classification of metastatic liver lesions: results of the 2016 low dose CT grand challenge [J]. Medical Physics, 2017, 44 (10): 339 - 352.
- [15] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications [EB/OL]. (2020 - 02 - 19) [2024 - 07 - 19]. <https://arxiv.org/abs/1704.04861>.
- [16] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 4510 - 4520.
- [17] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE, 2019: 1314 - 1324.
- [18] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 6848 - 6856.
- [19] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design [C] //Proceedings of the European conference on computer vision (ECCV). Munich, Germany: Springer, 2018: 116 - 131.
- [20] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [EB/OL]. (2016 - 04 - 30) [2024 - 07 - 19]. <https://arxiv.org/abs/1511.07122>.

(上接第 228 页)