

引用格式: 冉瑞生, 张思文, 李进, 等. 基于特征融合的低剂量 CT 图像降噪方法[J]. 微电子学与计算机, 2024, 41(5): 11-21.
RAN R S, ZHANG S W, LI J, et al. A noise reduction method of low-dose CT images based on feature fusion[J]. Microelectronics & Computer, 2024, 41(5): 11-21.
DOI: 10.19304/J.ISSN1000-7180.2023.0279

基于特征融合的低剂量 CT 图像降噪方法

冉瑞生¹, 张思文¹, 李进¹, 房斌²

(1 重庆师范大学 计算机与信息科学学院, 重庆 401331;

2 重庆大学 计算机学院, 重庆 400044)

摘要: 近年来低剂量 CT (Low Dose CT, LDCT) 被广泛应用于临床诊断中, 但 LDCT 会产生不规则的噪声。已有的降噪方法往往缺乏对全局特征信息的考虑, 以及不注重边缘特征信息和重建图像的视觉效果。为此, 提出了一种基于特征融合的低剂量 CT 图像降噪方法。首先, 利用 Transformer 优异的全局感受野提取图像的全局特征信息, 并利用卷积神经网络 (Convolutional Neural Network, CNN) 良好的局部特征提取能力提取图像的局部特征信息。在 Transformer 模块中加入维度变换思想, 以更好地抑制噪声; 在 CNN 模块中使用稠密连接的方式将浅层网络的特征信息复用于深层网络中, 以此保存更多的特征信息。其次, 为了获取更加丰富的图像细节特征, 使用了改进的索伯边缘增强算子来加强模型对边缘特征信息的提取能力。最后, 将 Transformer 模块和 CNN 模块获取的特征信息进行融合并输出重建图像。此外, 为了使降噪重建后的图像有更好的质量和视觉效果, 设计了一个多尺度复合损失函数。实验表明: 在 AAPM-Mayo 数据集的降噪实验中, 与当前主流的 LDCT 图像降噪方法相比, 本文方法取得了更好的降噪效果。

关键词: 图像降噪; 低剂量 CT; 特征融合; Transformer; CNN; 边缘增强; 损失函数

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2024)05-0011-11

A noise reduction method of low-dose CT images based on feature fusion

RAN Ruisheng¹, ZHANG Siwen¹, LI Jin¹, FANG Bin²

(1 College of Computer and Information Science, Chongqing Normal University, Chongqing 401331, China;

2 College of Computer Science, Chongqing University, Chongqing 400044, China)

Abstract: Low-Dose CT (LDCT) is widely used in clinical diagnosis, but it also generates some irregular noise. The existing noise reduction methods often lack consideration of global feature information, as well as do not focus on edge feature information and the visual effect of the reconstructed image. Then, a feature fusion-based noise reduction method for low-dose CT images is proposed. Firstly, the excellent global receptive field of Transformer is used to extract the global feature information of the image, and the good local feature extraction ability of Convolutional Neural Network (CNN) is used to extract the local feature information of the image. The dimensional transformation idea is added to the Transformer module to better suppress the noise; a dense connection in CNN is used to reuse the feature information from the shallow network to the deep network, so as to achieve pre-fusion of features and preserve more feature information. An improved Sobel edge enhancement operator is used to enhance the model's ability to extract edge feature information. Finally, the feature information acquired by the Transformer module and the CNN module is fused and the reconstructed image is output. In addition, a multiscale composite loss function is designed to make the reconstructed image with better quality and

收稿日期: 2023-04-04; 修回日期: 2023-05-03

基金项目: 重庆市技术创新与应用发展专项 (面上) (cstc2020jscx-msxmX0190); 重庆市教委科学技术研究计划 (KJZD-K202100505, KJQN202100515)

<http://www.journalmc.com>

visual effect after noise reduction. The experiments show that in the noise reduction experiments of AAPM-Mayo dataset, compared with the current mainstream LDCT image noise reduction methods, the method in this paper achieves a better noise reduction effect.

Key words: image denoising; low-dose CT; feature fusion; Transformer; CNN; edge enhancement; loss function

1 引言

在 20 世纪 90 年代, Naidich 等^[1]首次提出低剂量 CT(Low-Dose CT, LDCT)。该方法采用低剂量的 X 射线对检测部位进行诊断, 然而在降低辐射量的同时也导致成像后有明显的噪声问题^[2]。为了在低剂量的条件下也能获得与标准剂量的 CT 图像(NDCT^[1])相近的成像效果, 人们进行了大量研究, 研究方法主要包括: 投影域方法、统计迭代方法和后处理方法^[3]。

2006 年深度学习的概念被提出, 处理方法获得了更加广泛的应用。基于深度学习的 LDCT 图像降噪方法在面对复杂的图像噪声时, 表现出了更好的处理效果^[4]。这些方案大多基于卷积神经网络(Convolutional Neural Network, CNN)^[5]和生成对抗网络(Generative Adversarial Network, GAN)^[6]。如 Chen 等^[4]利用自编码器和残差思想提出了编码器-解码器结构的卷积神经网络 REDCNN(Residual Encoder Decoder Convolutional Neural Network)。Yang 等^[7]利用 Wasserstein 距离原理构建了生成对抗网络, 其方法极大地改善了生成对抗网络训练不稳定的问题, 使生成器和判别器在博弈中能更好地进行迭代。Liang 等^[8]提出的 EDCNN(Edge enhancement-based Densely Connected Neural Network)利用边缘增强和多层堆叠的全卷积网络结构实现了端到端的降噪模型结构, 并取得了不错的噪声抑制效果。Geng 等^[9]提出了一种多模式生成器的生成对抗网络 CNCL(Content Noise Complementary Learning), 该方法使用 3 种经典的去噪模型作为生成器的预测器, 以此来分析多种噪声类型, 通过这种多模型分析策略使模型具有较强的泛化能力。Choi 等^[10]提出一种自监督降噪方法, 在训练过程中采用了分阶段训练策略, 即离线训练和在线微调方法。不同以往的大多数监督训练方法, 该方法采用自监督的方法, 使得训练过程不再依赖低剂量 CT 和正常剂量 CT 的图像数据集。

尽管当前的去噪方法已经能够取得不错的效果, 但是仍面临着一些挑战。首先, 现有方法主要集中在

在图像局部特征信息的提取上, 而未充分考虑全局特征信息。其次, 这些方法在提取图像局部和细节信息时, 往往忽略了图像边缘特征信息。此外, 这些方法也未能充分注重重建图像的视觉效果, 从而导致图像降噪后的重建效果欠佳, 如放大后失真、图像细节丢失以及整体视觉效果模糊等问题。

Transformer 最早被用于自然语言处理(NLP)领域^[11], 凭借其优异的全局感受野取得了巨大成功。随后 Dosovitskiy 等^[12]提出了 ViT, 将 Transformer 引入到计算机视觉(CV)领域并取得了优异成绩。CNN 具有良好的局部特征提取能力, 但缺乏全局感受野, 而 Transformer 的注意力机制可以弥补这一缺陷^[13]。因此, 本文模型使用 Transformer 结合 CNN 作为主干网络。在训练过程中, 首先将低剂量 CT 图像分别输入到 Transformer 模块和 CNN 模块中。在 Transformer 模块中通过分块和维度变换来抑制噪声, 并在残差网络(ResNet)^[14]的辅助下实现对全局特征的学习和对全局噪声的抑制。在 CNN 模块中通过稠密连接(DenseNet)^[15]将边缘增强模块提取的边缘特征信息和 CNN 模块进行前期的特征融合, CNN 模块实现对局部特征的学习和对局部噪声的抑制。最后通过一个后期特征融合 CNN 模块来拟合双分支的输出结果。在损失函数方面, 本文设计了一种新的复合损失函数, 来增强图像质量和视觉效果。实验表明, 在不同剂量的 LDCT 图像实验中, 本文方法均有稳定的噪声抑制效果和更好的成像效果。

2 相关工作

2.1 Transformer Encoder

在计算机视觉领域中, Transformer Encoder 凭借其良好的动态全局性被广泛应用于各种视觉任务中。图 1 展示了 Transformer Encoder 的内部结构, 其中 Transformer 的多头自注意力机制(Multi-Head Self-Attention)^[11]使其拥有更大的感受野, 而前馈神经网络(Feed Forward Network, FFN)则可以提升特征的非线性表达能力。本文将对原始的 Transformer Encoder 进行相应的改进, 使其更加适用于低剂量 CT 图像的降噪任务。

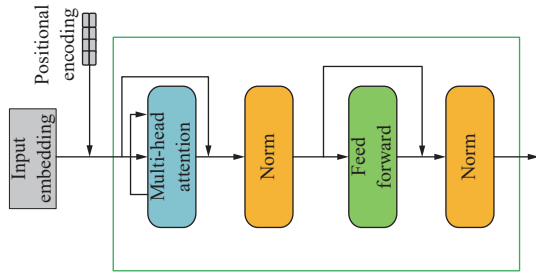


图1 Transformer Encoder 层

Fig. 1 Transformer Encoder layer

2.2 索伯算子

图像边缘特征作为图像的一部分, 通常携带了重要特征信息。索伯算子^[16]可作为一种 CT 图像边缘特征提取方法。如图 2 所示, 索伯算子由两个 3×3 的卷积核组成, 根据图像垂直方向和水平方向的差异, 分为纵向卷积核 g_x 和横向卷积核 g_y 。在进行特征提取时, 利用两个卷积核与图像矩阵进行平面卷积运算, 得到每个像素点的纵向和横向梯度值, 当梯度值较大时, 该点即为图像的边缘界点。

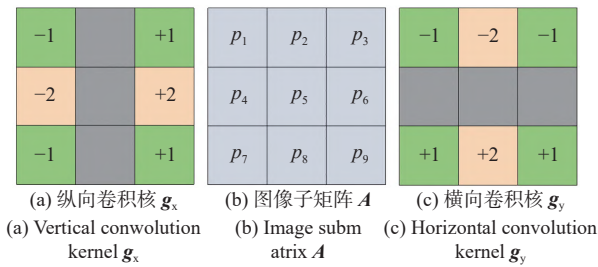


图2 索伯卷积核和图像子矩阵

Fig. 2 Sobel convolution kernel and image submatrix

在图 2(b) 中 A 为整张图像的一个 3×3 的子矩阵, 其中 $p_1 \sim p_9$ 表示相邻的 9 个像素点。 G_x 为图 2(b) 中 p_5 在垂直方向的梯度值, G_x 的表达式为

$$G_x = g_x \cdot A \quad (1)$$

$$G_x = (p_3 - p_1) + 2(p_6 - p_4) + (p_9 - p_7) \quad (2)$$

同理可得 p_5 在水平方向的梯度值 G_y 。得到 G_x 和 G_y 后, 可得总梯度 G , 其表达式为

$$G = \sqrt{G_x^2 + G_y^2} \quad (3)$$

目标点的总梯度较大时, 该点为界点。

2.3 损失函数

2.3.1 均方差损失函数

均方差损失 (Mean Square Error, MSE) 函数^[17]反映了输出降噪图像和原图像在每一个像素点之间的方差。方差越小真实值与预测值越接近, 其公式如下:

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

式中: y_i 表示真实值; \hat{y}_i 表示预测值。

2.3.2 平均绝对误差 L1 损失

平均绝对误差损失 (Mean Absolute Error, MAE) 函数^[18]也称为 L1 损失, 是一种相对简单的回归模型损失函数, 其结果为真实值和预测值之间的绝对差值之和, 表达式为

$$L_{L1} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

2.3.3 多尺度相似度损失函数

多尺度相似度损失函数 (Multiscale Structure Similarity Loss, MS-SSIM Loss) 是基于结构相似度 (SSIM) 的损失函数^[19]。SSIM 会计算真实图像和预测图像之间的亮度 (l)、对比度 (c) 和结构 (s) 的差值。MS-SSIM 则是在 SSIM 的基础上, 为了降低抖动剧烈的高频噪声, 使用低通滤波进行迭代, 并将过滤后的图像进行下采样, 将原始图像尺度定为尺度 1, 此后经过 $M-1$ 次的迭代后形成尺度 M 。MS-SSIM 使用多尺度代替单一尺度, 来适应不同尺度下结构相似度的变化, 在不同尺度下的亮度 (l) 不会发生改变。MS-SSIM 的公式如下:

$$F_{MS-SSIM} = [l(y, \hat{y})]^{\alpha_M} \cdot \prod_{j=1}^M [c_j(y, \hat{y})]^{\beta_j} \cdot [s_j(y, \hat{y})]^{\gamma_j} \quad (6)$$

式中: α 、 β 、 γ 为重要性比重, 且均大于 0。因为 MS-SSIM 的损失值与效果呈正相关, 所以 MS-SSIM 损失公式如下:

$$L_{MS-SSIM} = 1 - F_{MS-SSIM} \quad (7)$$

3 基于特征融合的 LDCT 图像降噪方法

3.1 模型结构

本文的整体模型结构如图 3 所示, 模型主干由 Transformer 模块、边缘增强模块、CNN 模块 (前期和后期的特征融合模块) 组成。在前期和后期的特征融合模块中, 除了最后一层只有一个卷积核为 3×3 的 CNN 组成外, 其余每层均由两个 CNN 组成, 其卷积核分别为 3×3 和 1×1 。其中 3×3 的卷积核用于图像特征学习, 1×1 的卷积核则用于融合前一层以及稠密连接的特征信息。具体地, 将 LDCT 图像作为整个网络的输入, 分别输入到改进的 Transformer 模块、边缘增强模块和 CNN 模块中。LDCT 首先通过一个预处理模块做分块处理, 然后进入 Transformer 模块, 利用改进的 Transformer 来获取

<http://www.journalmc.com>

图像的全局信息, 并实现初步的全局降噪; 在另一个分支中 LDCT 进入边缘增强模块, 获取到更加准

确且丰富的边缘信息, 随后将获取的边缘信息和 LDCT 图像一起输入到 CNN 首层。

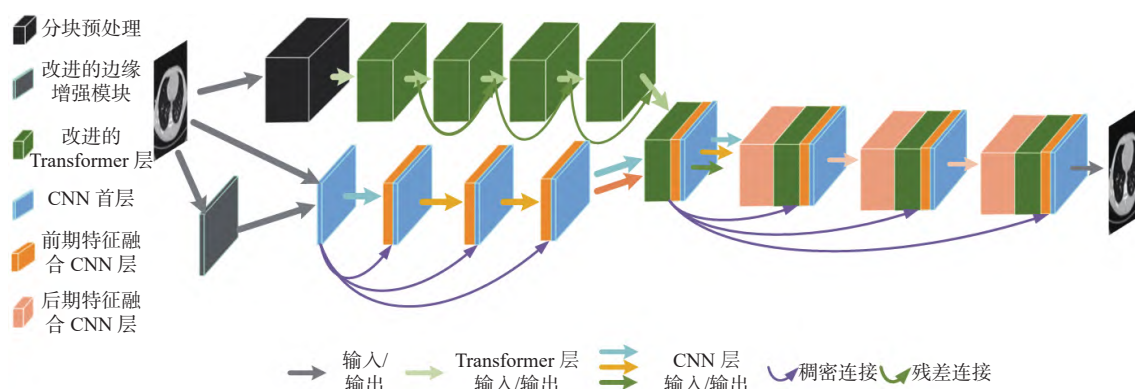


图3 本文整体模型结构图

Fig. 3 Overall architecture of the model in this paper

在 CNN 模块中, 利用两个特征融合模块进行数据增强, 特征融合 CNN 层会将稠密连接的特征信息, 在通道维度上进行融合叠加, 这样既丰富了图像的特征信息也保证了原始信息不会在降噪的过程中丢失细节, 同时只将第一层 CNN 进行稠密连接的方法也降低了计算复杂度。在图 3 中, CNN 首层之后的 3 层表示衔接前一层输出的特征融合 CNN 层, 首层通过稠密连接的方式融入特征融合 CNN 层中。经过稠密连接的特征融合 CNN 层, 增强了模型获取局部细节信息和噪声分布信息的能力。在经过前期特征融合 CNN 层的噪声消减后, 再将前期特征融合 CNN 模块获取的边缘信息和局部信息, 以及 Transformer 模块获取的全局信息一起输入到后期特征融合 CNN 模块中。在后期特征融合 CNN 模块中同样也使用了稠密连接的方式, 将第一层和其后的后期特征融合 CNN 层进行融合, 再经

过多层后期特征融合 CNN 后, 低剂量 CT 图像的噪声被进一步消减处理并最终输出重建图像。

3.2 改进的 Transformer 模块

为了更好地抑制 LDCT 图像的噪声, 本文改进了 Transformer 的编码器(Encoder)部分。如图 4 所示, 改进的 Transformer 由一个降维的 Encoder 层和一个升维的 Encoder 层拼接而成, 多个 Transformer 层连接形成 Transformer 模块。借鉴降噪自编码器的思想, 在每个 Transformer 层内部用一个可升降维的 DL(Dimension transformation Layer)层代替原 Encoder 结构的 FNN(Feed Forward Network)层, 激活函数选用含有正则化思想的 GeLU 函数。另外, 在每个 Encoder 层中, 使用了前后两个多头注意力机制计算模块, 并在两个模块前后以及每层 Transformer 的前后进行残差连接。

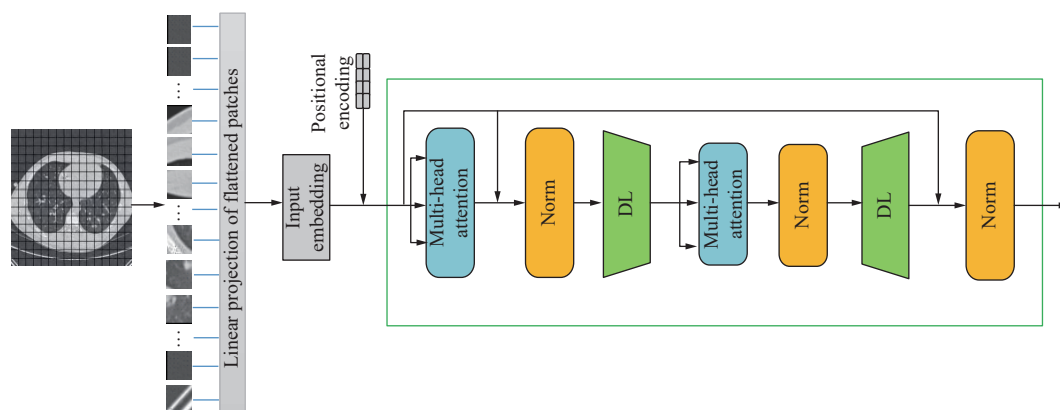


图4 改进的 Transformer 层

Fig. 4 Improved Transformer layer

模块流程如图 4 所示, 首先将图像分成同样大

小的图像块, 然后将图像块通过线性映射的方式作

<http://www.journalmc.com>

扁平化处理(Linear Projection of Flattened Patches)。然后进行一个可训练的线性映射, 再经过投影后将这个投影的输出(Embedded Patches)加上位置编码(Positional Encoding)输入到 Transformer 内部结构中。

在 Transformer 层的内部结构中, 特征向量首先通过多头注意力机制计算捕获丰富的全局特征信息, 随后进行归一化处理, 得到的输出结果进入 DL 层进行降维。降维后再通过多头注意力机制计算和归一化处理, 最后再通过 DL 层还原至初始维度, 并进行归一化处理。公式如下:

$$z_0 = XE + E_{\text{pos}} \quad (8)$$

$$z_l = \text{MSA}(z_{l-1}) + z_{l-1} \quad (9)$$

$$z_l^* = \text{DL}(\text{LN}(z_l)) \quad (10)$$

$$z_l' = \text{DL}(\text{LN}(\text{MSA}(z_l^*) + z_l^*)) \quad (11)$$

$$z_{l+1} = \text{LN}(z_l' + z_0) \quad (12)$$

式中: X 为图像分块后的序列集合, $X = \{x^1, x^2, \dots, x^N\}$; E 为线性映射; E_{pos} 为位置编码; z_0 为输入序列; z_l 中 l 表示编码器的第 l 层; z_l^* 表示降维后的序列; z_l' 表示维度还原后的序列; MSA 为多头注意力机制; DL 为维度变换层; LN 为归一化层。

3.3 改进的索伯边缘增强模块

为了能够帮助模型更好地提取边缘特征信息, 本文改进了传统索伯算子的卷积核。相较于原始的索伯算子, 本文使用了横向、纵向、主对角线方向和次对角线共计 4 个方向的索伯卷积核, 并使用 5×5 的卷积核来扩大索伯算子的感受野, 如图 5 所示。

新的索伯算子总梯度公式为

$$G = \sqrt{G_x^2 + G_y^2 + G_{xy}^2 + G_{yx}^2} \quad (13)$$

式中: G_{xy} 和 G_{yx} 分别为主对角线和次对角线上的梯度值。

为了进一步增强边缘增强模块的泛化能力, 设计两个可训练的参数因子 a 和 b , 并根据每个位置的相对重要性, 设计相应的参数比重, 以此来增强模型对不同边缘信息的提取能力。为了保证训练有效性, 设置参数 a 的比重大于参数 b 的。

在训练过程中, 输入图像至边缘增强模块中, 图像经过与 5×5 自适应参数卷积核的一系列平面卷积运算后, 输出得到一组携带边缘信息的特征图, 如图 6 所示。图 6 中左侧部分为两幅不同的 CT 图像, 右侧部分为相应的边缘信息特征图。

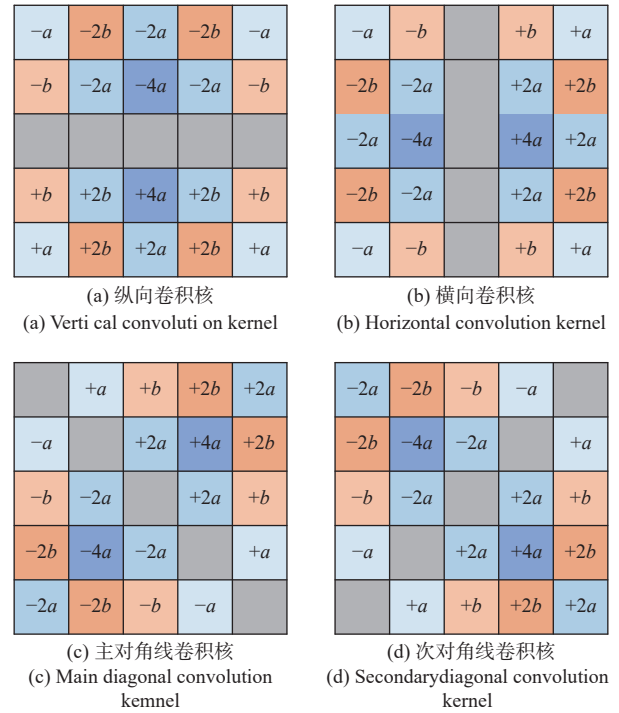


图 5 5×5 自适应参数卷积核

Fig. 5 5×5 Adaptive parameter convolution kernel

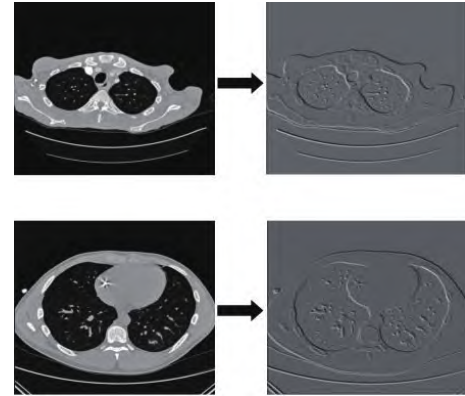


图 6 CT 图像边缘特征图

Fig. 6 CT image edge feature map

由图 6 可知, 提取的边缘特征信息较为清晰, 从而表明通过 4 个方向上的 5×5 索伯卷积核, 并通过两个可变参数的调节, 提高了模型提取边缘信息的准确性, 进而增强了模型的特征提取能力。

3.4 复合损失函数的设计

在已有的降噪模型中, 损失函数通常以基于误差敏感度的 MSE 函数为主, 但此类损失函数都不会涉及图像结构的相关性。然而图像结构中包含了图像的许多重要信息, 这些信息往往对视觉效果影响很大, 所以在一些客观指标较好的情况下, 重建图像仍会出现模糊、失真等问题。

鉴于上述问题, 本文首先使用 MS-SSIM 损失

<http://www.journalmc.com>

函数与 MSE 损失函数结合的方式, MS-SSIM 损失函数通过计算结构相关性损失的方式来改善图像的模糊感, 同时 MS-SSIM 作为 SSIM 的扩展版, 通过使用不同滤波算子的方式来扩大感受野, 并实现对图像细节等高频信息的保存。但在实验中发现该方法存在亮度和颜色的偏移问题。在文献 [20] 的启发下, 本文使用 L1 函数来固定图像的亮度和颜色, 并使用文献 [20] 中设定的高斯分布参数 G_σ 来优化 L1 损失梯度不变和参数比例的问题。最后本文结合了 MSE 损失、MS-SSIM 损失和 L1 损失, 提出一种新的复合损失函数, 简称为 MML(以上 3 种函数的首字母), MML 公式如下:

$$L_{\text{MML}} = m \cdot L_{\text{MS-SSIM}} + (n - m) \cdot G_\sigma \cdot L_{\text{L1}} + (1 - n) L_{\text{MSE}} \quad (14)$$

式中: m 、 n 为 MML 中调节比重的两个参数; L_{MML} 、 $L_{\text{MS-SSIM}}$ 、 L_{L1} 和 L_{MSE} 分别为 MML、MS-SSIM 损失、L1 损失和 MSE 损失。

为了获取相对更优的参数值, 本节还将对 MML 函数的参数 m 和参数 n 进行探讨。实验选取 4 组不同的参数值进行比较, 为了保证实验的有效性, 参数设置时要保持 $n > m$, 且 m 的值与 n 的值差距不大, 这是因为 L1 损失在 MML 函数中不起主导作用, 如果 m 的值和 n 的值差距过大, 将会使 L1 的参数比重过大, 从而影响实验。参数分组实验结果如表 1 所示。

表 1 参数分组实验
Tab. 1 Experimental results in subgroups

Group	m	n	PSNR	SSIM
1	0.168	0.200	27.631 4	0.859 7
2	0.294	0.350	27.475 2	0.861 8
3	0.420	0.500	27.431 9	0.858 3
4	0.672	0.800	27.207 4	0.867 5

实验中选用 PSNR^[21] 和 SSIM 指标(见 4.2 节)作为评价指标, 表 1 参数分组中, 由于保持 L1 的参数比重, 所以参数 m 和参数 n 同增同减。观察表 1 分组实验结果可知, 当参数 m 和参数 n 的值逐渐变大时, PSNR 的值逐渐变小, 而 SSIM 的值则出现了变大后又变小的情况, 这使得在面对不同任务 and 不同环境时, 需要多次实验, 找到最优参数值。经过综合比较, 本文选用了分组 2 的参数值。

4 实验结果

4.1 实验设置

本文选用 NIH AAPM-Mayo Clinic 低剂量 CT <http://www.journalmc.com>

挑战赛数据集进行试验。Mayo 数据集中含有标准剂量的 CT 图像以及对应的低剂量 CT 图像, 图像大小为 512×512 。近几年, AAPM-Mayo 数据集不断扩大, 并于 2021 年发布了更新的数据集^[22]。本文选用 25% 剂量的腹部 CT 图像(Data A)和 10% 剂量的胸部 CT 图像(Data B)进行训练, 各个剂量的 CT 图像分别随机选用 20 个病人的 5 000 组 CT 图像作为训练集, 随机选用 10 个病人的 2 500 组 CT 图像作为测试集。整个实验在 Python 环境下进行, 使用的 GPU 型号为 NVIDIA A100 PCIe, 显存为 40G, 图像尺寸为 512×512 。在 Transformer 模块中将每张图像切分成尺度为 32×32 的小块, 在 DL 层将尺度为 32×32 的小块降维到尺度 16×16 。训练轮数为 100, 学习率为 0.001, Batch size 为 10, 优化器使用 Adam 算法。

4.2 评价指标

本文选取降噪模型中主流的 3 种评价指标来判断图像降噪和重建的效果。

(1) 峰值信噪比(PSNR)是针对噪声的客观评价指标, 该指标通过比较像素点的方式评判图像的降噪效果。峰值信噪比与图像的降噪效果呈正相关。

(2) 结构相似度(SSIM)是一种更接近人眼观察的评测标准。结构相似度从亮度、对比度和结构上对降噪效果进行评判。结构相似度与图像的重建效果呈正相关。

(3) 均方根误差(RMSE)主要是通过降噪图像和正常图像的标准均方误差来评判降噪图像的效果。均方根误差与图像的降噪效果呈负相关。

4.3 对比实验

本文与几种主流的降噪方法在 Data A 和 Data B 两个分组中进行了对比实验。对比的主流方法包括基于 CNN 的 REDCNN 方法、EDCNN 方法和基于 GAN 的 WGAN 方法、CNCL 方法以及自监督方法 CS2-TP 方法、CS2-FT 方法。表 2 给出了对比结果。从表 2 可知, 本文方法在 PSNR、SSIM、RMSE 中均获得了更佳的效果。其中基于 GAN 的两种方法总体降噪效果要低于基于 CNN 的两种方法。而 CS2-PT 方法和 CS2-FT 方法作为一种同源但不同阶段的自监督方法, PSNR 数据较好, 获得了不错的降噪效果, 但是 SSIM 要弱于部分降噪方法。

在图 7 中展示了本文方法与部分可复现方法的可视化降噪 CT 图像。为了更好地观察细节, 将第一行和第三行的红框内的部分放大后展示在第二行

和第四行中。图 7(g) 为本文改进索伯算子的边缘特征图, 仅用于参照观察, 不作为实验对比。

表 2 在 AAPM-Mayo 数据集上各个方法的比较和结果
Tab. 2 Comparison and results of each method on the AAPM-Mayo dataset

Method	PSNR/dB		SSIM		RMSE	
	Data A	Data B	Data A	Data B	Data A	Data B
LDCT	35.835 5	18.576 3	0.940 3	0.683 6	0.016 6	0.103 1
REDCNN ^[4]	41.761 2	26.667 1	0.981 9	0.829 7	0.008 1	0.046 9
WGAN ^[7]	38.867 5	23.746 7	0.959 6	0.788 2	0.011 2	0.069 7
EDCNN ^[8]	41.554 8	26.193 5	0.983 5	0.839 5	0.008 5	0.049 8
CNCL ^[9]	39.154 1	24.861 2	0.962 7	0.805 8	0.010 1	0.061 5
CS2-PT ^[10]	—	27.100 0	—	0.809 0	—	—
CS2-FT ^[10]	—	27.200 0	—	0.812 0	—	—
Ours	42.681 1	27.473 2	0.988 3	0.881 3	0.007 8	0.045 9

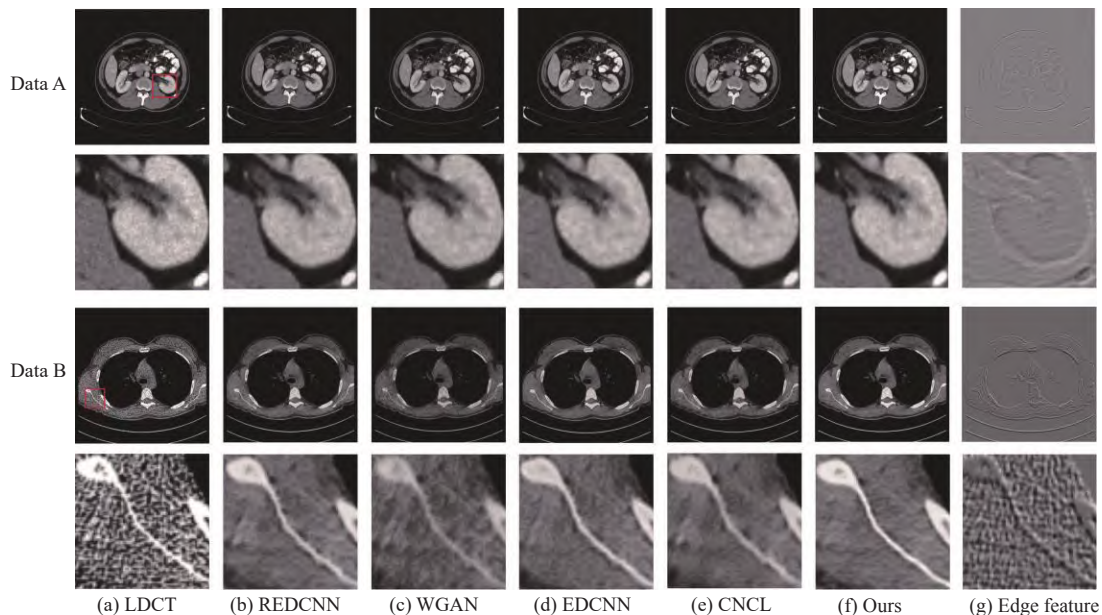


图 7 不同方法在 AAPM 数据集的降噪重建图像

Fig. 7 Different methods of noise reduction in AAPM dataset for image reconstruction

在图 7 中, REDCNN 方法在两组实验对比中均有着不错的降噪效果, 但是在图 7(b) 中可以看出 REDCNN 的降噪图像都过于平滑, 视觉效果不佳, 存在模糊和细节丢失的问题。相比 REDCNN 方法, EDCNN 方法在 Data A 中视觉效果方面表现更佳, 但是在更大噪声的 Data B 中, EDCNN 也出现了与 REDCNN 相同的问题, 观察图 7(b) 和图 7(d) 第三、四行可以看出 EDCNN 在 Data B 中留有更多的噪点。

WGAN 作为一种典型的基于 GAN 的降噪方法, 在 LDCT 的两组降噪实验中虽然都保留了不错的细节和结构信息, 但是有更多的噪点, 对噪声抑制效

果差, 图像效果在同比方法中均为最低。CNCL 方法相比于 WGAN 有更好的噪声抑制效果, 整体质量有所提升, 但仍有诸多细节上的噪点。

通过对图 7 中图像细节的观察, 在噪声相对低的 Data A 中, 降噪方面和图像质量, 各个方法均能取得不错的效果, 但是在噪声相对高的 Data B 中, 在降噪效果、重建图像质量以及视觉效果方面, 各个方法表现各异, 差距较大。与部分可视化降噪 CT 图像对比中, 本文方法的噪声抑制效果、重建图像质量和视觉效果均取得了更佳的表现。在 MML 复合损失函数的辅助下, 本文方法得到的重

<http://www.journalmc.com>

建图像在放大后的失真程度最小。观察图 7(g), 改进的索伯算子提取了图像诸多细节特征, 使本文方法保留了更多的细节以及边缘特征。低剂量 CT 作为一种辅助医生诊断的图像技术, 成像的视觉效果十分重要, 通过与各个方法的对比可见, 本文方法具有更好的视觉效果。

此外, 为了验证本文方法在实际应用中的资源占用情况, 本研究将本文方法与基于卷积神经网络的 EDCNN 方法以及基于生成对抗网络的 CNCL 方法进行了比较, 并进一步比较了它们的降噪效果。本文方法中的 Transformer 模块采用了 64×64 的分块方式, 实验数据选用了 Data B 数据集, 图像尺寸为 512×512 。在评价指标方面, 我们采用了峰值信噪比(PSNR)、运行时间(Running time)以及 GPU 显存占用量(GPU Occupancy)。其中, 运行时间指的是处理 10 张图像所需的平均时间。资源占用和降噪效果的对比结果如表 3 所示。

在降噪效果方面, 由表 3 中峰值信噪比数据可知, 本文方法具有明显的优势。在运行时间方面, 本文方法与 EDCNN 和 CNCL 相当。对于 GPU 占用情况, 本文方法与 EDCNN 和 CNCL 相当或略大一点。由于本文改进的 Transformer 模块仅使用

Transformer 的 4 层编码器, 同时进行了轻量化处理, 在当前的算力资源下, 本文方法是可行的。

表 3 资源占用和降噪效果对比

Tab. 3 Comparison of resource consumption and noise reduction effect

Evaluation Metrics	EDCNN	CNCL	Our Method
PSNR/dB	26.193 5	24.861 2	27.208 1
Running time/s	5.05	5.41	5.38
GPU/G	14	15	17

4.4 消融实验

4.4.1 Transformer 消融实验与分块实验

为了探究本文改进的 Transformer 模块对模型性能的提升作用, 本次实验只将模型中改进的 Transformer 模块替换为原始的 Transformer, 其他模块、参数和实验环境均保持不变, 训练数据选用 Data A 和 Data B。因为改进 Transformer 模块的主要目的是提升模型的降噪能力, 所以实验以 PSNR 作为评价指标。为了方便, 我们将原始的 Transformer 模块记为 T_1 , 本文改进的 Transformer 模块记为 T_2 , 实验结果如图 8 所示。

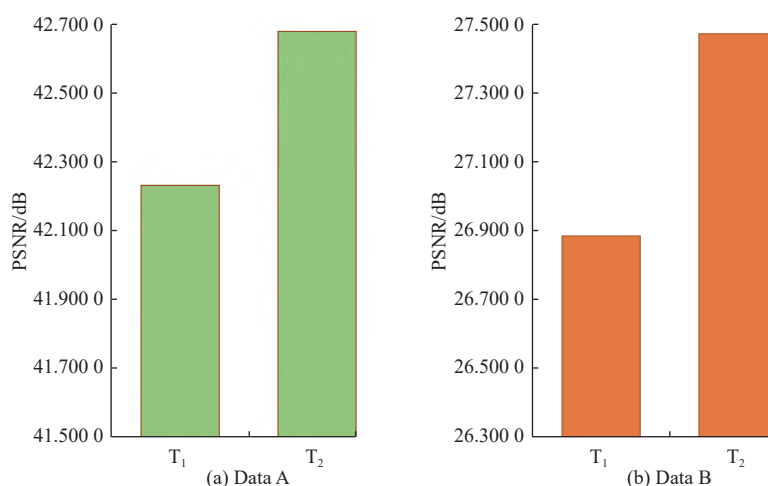


图 8 T_1 和 T_2 的 PSNR 柱状图

Fig. 8 PSNR histogram of T_1 and T_2

图 8 表明在 Data A 和 Data B 两组数据中, T_2 方法都明显高于 T_1 方法的 PSNR 值。由此可见, 本文改进的 Transformer 模块有助于模型提升降噪能力。

本小节除了验证了本文改进的 Transformer 模块对模型的影响外, 还探究了 Transformer 分块预处理阶段, 不同分块策略对 Transformer 和模型性能的影响。实验将图像块尺寸设置为 16×16 、 $32 \times$

32 和 64×64 在 Data A 上进行对比实验。评价指标选用 PSNR、运行时间(Running time)和占用的 GPU 显存(GPU Occupancy), 其中运行时间指处理 10 张图像的时间。实验结果如表 4 所示。

结果表明, 在以上 3 种分块方法中, 图像块尺度为 16×16 时, PSNR 指数最高, 去噪效果最好, 但相应的运行时间长且内存占用大。所以在实际训

练中,如果运行时间和占用内存可接受,可以适当减小图像块尺度来增强模型性能。

表4 不同图像块尺度的模型性能对比

Tab. 4 Model performance comparison for different

image patch size

Performance	Scale		
	16×16	32×32	64×64
PSNR/dB	43.106 7	42.681 1	41.916 8
Running time/s	14.16	8.75	5.38
GPU/G	26	20	17

4.4.2 边缘增强模块消融实验

为了验证本文提出的模型性能,本节实验通过定量分析的方式对索伯边缘增强模块进行消融实验,为了更好地验证改进的索伯边缘增强模块的效果,本节分别在不使用索伯边缘增强(null)、使用传统3×3卷积核和使用本文改进的5×5自适应参数卷积核的情况下进行训练;训练数据选用Data A和Data B,训练过程中保持参数和实验环境一致,实验以PSNR作为评价指标,训练过程的PSNR曲线如图9和图10所示。

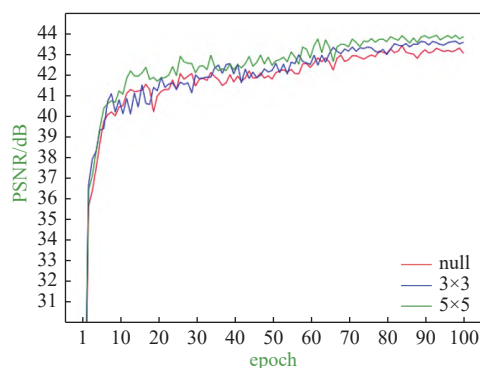


图9 Data A的PSNR曲线图

Fig. 9 PSNR curve of Data A

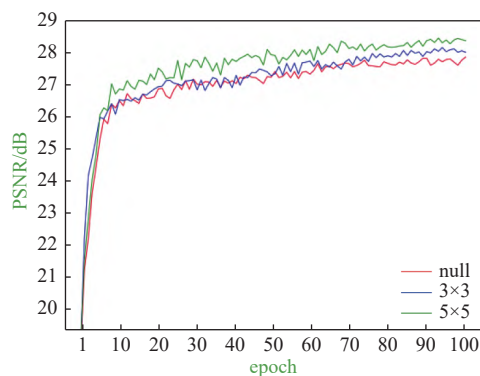
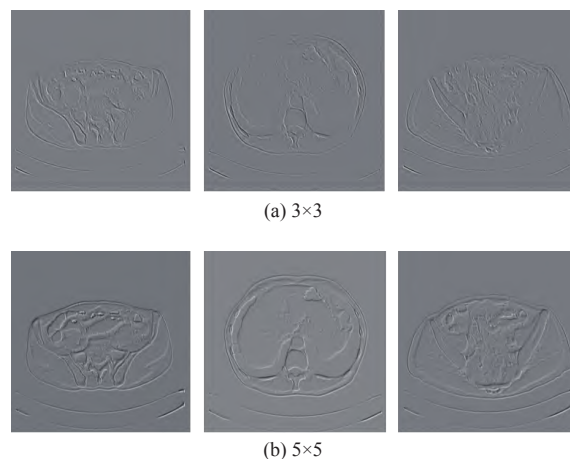


图10 Data B的PSNR曲线图

Fig. 10 PSNR curve of Data B

可以看到,在Data A和Data B中5×5自适应参数卷积核的边缘增强方法,要强于不使用边缘增强和使用传统3×3卷积核的边缘增强方法,并且在大部分情况下5×5曲线都高于null曲线和3×3曲线。

图11是边缘增强可视化图像,图11(a)为传统3×3卷积核的边缘增强方法,图11(b)为5×5自适应参数卷积核的边缘增强方法。从可视化图形中可知,5×5自适应参数卷积核的边缘增强方法具有更多的细节和更清晰的视觉效果。低剂量CT图像的边缘特征图中提取了CT图像大量的细节特征信息,而细节信息对于CT图像尤为重要。综上,本文设计的5×5自适应参数卷积核的边缘增强方法,能够使重建图像的边缘特征信息更加准确,对模型性能有更好的提升效果。



(b) 5×5

图11 边缘增强特征图

Fig. 11 Edge enhancement feature map

4.4.3 损失函数消融实验

为了证明本文提出的MML损失函数相较于其他损失函数具有优越性,本节将基于本文方法分别使用MSE、MS-SSIM+L1(同MML相同,L1不起主导作用,仅用于固定图像的亮度和颜色)和MML作为损失函数在Data B上进行消融实验。

实验结果如图12所示,其中PSNR、SSIM为图像降噪效果和图像视觉效果的评价指标,其值与效果呈正相关,RMSE为均方根误差,其值与效果呈负相关。由图12可知,将损失函数设置为MML时,评价指标PSNR、SSIM和RMSE表现均为最好。

5 结束语

综上,本文提出一种基于特征融合的LDCT图像降噪方法,将Transformer提取的全局特征、CNN获取的局部特征以及边缘增强模块获取的边缘

<http://www.journalmc.com>

特征进行特征融合,并在新的复合损失函数的帮助下重建 CT 图像。实验结果表明,与当前主流和最新方法相比,本文方法在 PSNR、SSIM 和 RMSE 这 3 个评价指标上取得了最佳的效果。通过 Transformer 模块和边缘增强模块的消融分析实验可得,

本文改进的 Transformer 模块和自适应参数卷积核的索伯算子对模型有良好的提升效果。此外对本文设计的多尺度复合损失函数进行消融实验和总结分析,结果表明该损失函数对重建图像的降噪效果和视觉效果都有较好的提升效果。

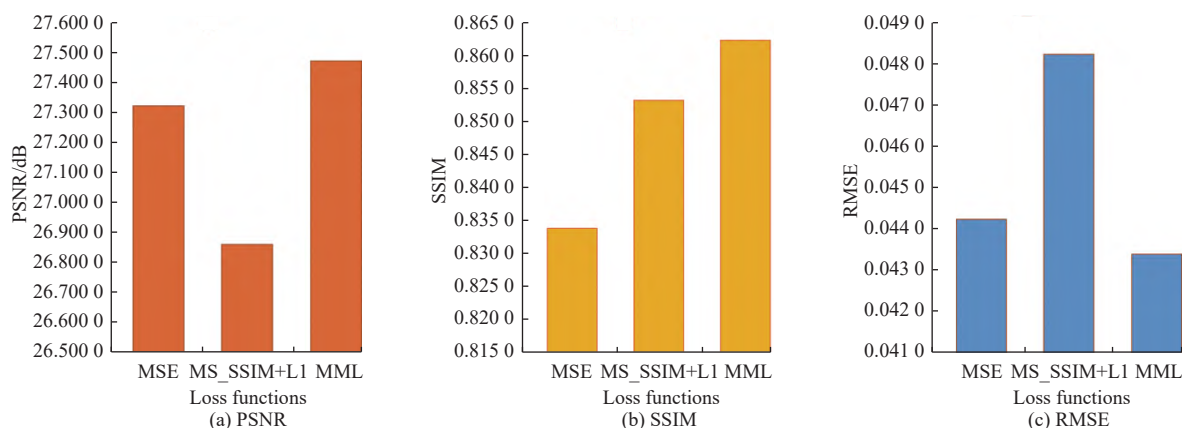


图 12 不同损失函数的客观比较

Fig. 12 Objective comparison of different loss functions

结合上述实验结果可知,本文提出的方法在 LDCT 降噪任务中取得了优异的表现。当然在此基础上,将模型的性能进一步提升仍是需要解决的问题,如 Transformer 网络在降维方面借鉴了数据降维相关方法的思想,从而增强了降噪能力。然而,对于 Transformer 网络本身以及降噪思路的改进,仍有进一步分析和改进的空间。今后我们将更加深入地研究 LDCT 图像的降噪方法,改进现有模型并探索模型在医学图像领域的更多应用场景。

参考文献:

- [1] NAIDICH D P, MARSHALL C H, GRIBBIN C, et al. Low-dose CT of the lungs: preliminary observations[J]. Radiology, 1990, 175(3): 729-731. DOI: 10.1148/radiology.175.3.2343122.
- [2] YANG L T, LI Z N, GE R J, et al. Low-dose CT denoising via sinogram inner-structure transformer[J]. IEEE Transactions on Medical Imaging, 2023, 42(4): 910-921. DOI: 10.1109/TMI.2022.3219856.
- [3] DECKER J A, BETTE S, LUBINA N, et al. Low-dose CT of the abdomen: Initial experience on a novel photon-counting detector CT and comparison with energy-integrating detector CT[J]. European Journal of Radiology, 2022, 148: 110181. DOI: 10.1016/j.ejrad.2022.110181.
- [4] 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. DOI: 10.1109/TMI.2017.2715284. <http://www.journalmc.com>
- [5] 岳晴,尹健宇,王生生.基于改进 CNN 的低剂量 CT 图像肺结节自动检测[J]. 计算机科学, 2022, 49(S1): 54-59. DOI: 10.11896/jsjcx.210400211.
YUE Q, YIN J Y, WANG S S. Automatic detection of pulmonary nodules in low-dose CT images based on improved CNN[J]. Computer Science, 2022, 49(S1): 54-59. DOI: 10.11896/jsjcx.210400211.
- [6] ZHANG H B, GUO W P, ZHANG S Q, et al. Unsupervised deep anomaly detection for medical images using an improved adversarial autoencoder[J]. Journal of Digital Imaging, 2022, 35(2): 153-161. DOI: 10.1007/s10278-021-00558-8.
- [7] 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. DOI: 10.1109/TMI.2018.2827462.
- [8] LIANG T F, JIN Y, LI Y D, 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 (IC-SP). Piscataway: IEEE, 2020: 193-198. DOI: 10.1109/IC-SP48669.2020.9320928.
- [9] GENG M F, MENG X X, YU J Y, et al. Content-noise complementary learning for medical image denoising[J]. IEEE Transactions on Medical Imaging, 2022, 41(2): 407-419. DOI: 10.1109/TMI.2021.3113365.
- [10] CHOI K, LIM J S, KIM S. Self-supervised inter-and intra-slice correlation learning for low-dose CT image restoration without ground truth[J]. Expert Systems with Ap-

- plications, 2022, 209: 118072. DOI: 10.1016/j.eswa.2022.118072.
- [11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017.
- [12] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv, 2020: 2010.11929. DOI: 10.48550/arXiv.2010.11929.
- [13] 欧阳勇, 万豆, 高榕, 等. 基于自注意力的 TCN-Transformer 的电网单相故障检测方法[J]. 微电子学与计算机, 2022, 39(9): 89-97. DOI: 10.19304/j.issn1000-7180.2021.1331.
- OUYANG Y, WAN D, GAO R, et al. Research on single-phase fault line selection in distribution network based on TCN+Transformer self-attention[J]. Microelectronics & Computer, 2022, 39(9): 89-97. DOI: 10.19304/j.issn1000-7180.2021.1331.
- [14] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778. DOI: 10.1109/CVPR.2016.90.
- [15] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 2261-2269. DOI: 10.1109/CVPR.2017.243.
- [16] FAN P, ZHOU R G, HU W W, et al. Quantum image edge extraction based on classical Sobel operator for NEQR[J]. Quantum Information Processing, 2019, 18(1): 24. DOI: 10.1007/s11128-018-2131-3.
- [17] NICOLSON A, PALIWAL K K. Deep learning for minimum mean-square error approaches to speech enhancement[J]. Speech Communication, 2019, 111: 44-55. DOI: 10.1016/j.specom.2019.06.002.
- [18] CHICCO D, WARRENS M J, JURMAN G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation[J]. Peerj Computer Science, 2021, 7: e623. DOI: 10.7717/peerj-cs.623.
- [19] PATEL Y, APPALARAJU S, MANMATHA R. Human perceptual evaluations for image compression[J]. arXiv preprint arXiv, 2019: 1908.04187. DOI: 10.48550/arXiv.1908.04187.
- [20] 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. DOI: 10.1109/TCI.2016.2644865.
- [21] SURIYAN K, RAMAINGAM N, RAJAGOPAL S, et al. Performance analysis of peak signal-to-noise ratio and multipath source routing using different denoising method[J]. Bulletin of Electrical Engineering and Informatics, 2022, 11(1): 286-292. DOI: 10.11591/eei.v11i1.3332.
- [22] MOEN T R, CHEN B Y, HOLMES III D R, et al. Low - dose CT image and projection dataset[J]. Medical Physics, 2021, 48(2): 902-911. DOI: 10.1002/mp.14594.

作者简介:

冉瑞生(通信作者) 博士, 教授, rshran@cqnu.edu.cn

张思文 硕士研究生, 2020210516071@stu.cqnu.edu.cn