



计算机科学

COMPUTER SCIENCE

基于梯度引导的低剂量CT超分辨率重建算法

徐颖, 张道强, 葛荣骏

引用本文

徐颖, 张道强, 葛荣骏. 基于梯度引导的低剂量CT超分辨率重建算法[J]. 计算机科学, 2024, 51(8): 143-151.

XU Ying, ZHANG Daoqiang, GE Rongjun. [Super-resolution Reconstruction for Low-dose CT Based on Guidance of Gradient](#) [J]. Computer Science, 2024, 51(8): 143-151.

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)

[基于高深约束与边缘融合的单目3D目标检测](#)

Monocular 3D Object Detection Based on Height-Depth Constraint and Edge Fusion

计算机科学, 2024, 51(8): 192-199. <https://doi.org/10.11896/jsjcx.230500071>

[基于颜色流模型的非配对医学图像颜色迁移方法](#)

Color Transfer Method for Unpaired Medical Images Based on Color Flow Model

计算机科学, 2024, 51(8): 176-182. <https://doi.org/10.11896/jsjcx.230700088>

[轻量级深度神经网络模型适配边缘智能研究综述](#)

Lightweight Deep Neural Network Models for Edge Intelligence:A Survey

计算机科学, 2024, 51(7): 257-271. <https://doi.org/10.11896/jsjcx.240100045>

[无人机辅助边缘计算安全通信能力最大化方案](#)

Scheme for Maximizing Secure Communication Capacity in UAV-assisted Edge Computing Networks

计算机科学, 2024, 51(6A): 230800032-7. <https://doi.org/10.11896/jsjcx.230800032>

[边缘计算下差分隐私的应用研究综述](#)

Survey of Application of Differential Privacy in Edge Computing

计算机科学, 2024, 51(6A): 230700089-9. <https://doi.org/10.11896/jsjcx.230700089>

基于梯度引导的低剂量 CT 超分辨率重建算法

徐颖¹ 张道强¹ 葛荣骏²

1 南京航空航天大学计算机科学与技术学院 南京 211106

2 东南大学仪器科学与工程学院 南京 210096

(2942828142@nuaa.edu.cn)

摘要 低剂量 CT(Low-dose CT, LDCT)扫描在临床实践中起着关键作用,其能有效降低放射科医生和患者的患癌风险。然而,低剂量射线的使用会给生成的 CT 图像引入明显的噪声,这一问题凸显了 LDCT 降噪重建的必要性。图像重建领域中的另一个重要任务是超分辨率(Super-resolution, SR)重建,其目标是在减少计算开销的同时实现高分辨率的 CT 成像。高分辨率 CT 图像能够更准确地捕捉复杂的解剖细节。尽管这些任务在各自领域取得了显著进展,但目前仍缺乏能够有效利用这两个任务之间固有相关性并同时处理它们的有效方法。文中将边缘信息作为两个任务之间的纽带,并利用梯度提取强相关特征。这使得 LDCT 降噪重建过程能够辅助超分辨率重建过程,并最终生成具有清晰边缘的结果图像。文中提出的降噪和超分辨率重建网络(NRSR-Net)包括 3 个组成部分:1)边缘增强框架,该框架利用梯度信息引导和提取相关特征,从而充分利用两个任务之间的相关性,使降噪任务能够辅助超分辨率任务实现更好的性能;2)梯度门控融合模块(Gradient Guided Fusion Block, GGFB),该模块增强高度相关的边缘特征并抑制无关特征,从而实现边缘区域的有效重建;3)梯度损失,该损失函数为模型引入更加丰富的梯度特征,并使网络重点还原边缘区域。一系列的实验表明,NRSR-Net 在定量评估中取得了令人满意的峰值信噪比(PSNR)、结构相似性(SSIM)和学习感知图像块相似度(LPIPS),并获得了高质量的可视化结果。这些优势表明 NRSR-Net 在临床 CT 成像中具有巨大潜力。

关键词:低剂量 CT;超分辨率;梯度引导;多任务;边缘

中图分类号 TP391

Super-resolution Reconstruction for Low-dose CT Based on Guidance of Gradient

XU Ying¹, ZHANG Daoqiang¹ and GE Rongjun²

1 College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

2 College of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

Abstract Low-dose CT(LDCT) scan plays a pivotal role in clinical practice, effectively decreasing cancer risks for radiologists and patients. However, the utilization of low-dose radiation introduces notable noise into the resulting CT images, highlighting the necessity of low-dose CT reconstruction. Another important task in the field of image reconstruction is super-resolution(SR), with the aim of achieving high-resolution CT imaging while minimizing computational expenses. High resolution CT images afford the capacity to capture intricate anatomical details in greater fidelity. Although significant progress has been made in their respective domains, there is still a lack of effective methodologies that can effectively harness the inherent correlation between these tasks and handle them concurrently. We employ edge information as a link between the two tasks, and utilize gradients to extract shared features from both tasks. This allows the LDCT reconstruction process to assist the SR reconstruction process and generate resulting images with sharp edges. Our work consists of three components: 1) Edge-enhanced framework. The framework fully exploits the correlation between the two tasks by extracting relevant features using gradient information, enabling the denoising(DN)task to assist the SR task in achieving superior performance. 2) Gradient guided fusion block(GGFB), which enhances the highly correlated edge features while suppressing irrelevant features, thereby enabling effective reconstruction in edge regions. 3) Gradient loss, which introduces richer gradient features into the model and guides the network to prioritize the reconstruction of edge regions. Extensive experiment demonstrates that our noise reduction and super resolution reconstruction network(NRSR-Net) achieves promising PSNR, SSIM, and LPIPS in quantitative evaluations, as well as gains high-quality readable

到稿日期:2023-07-21 返修日期:2023-09-28

基金项目:江苏省自然科学基金(BK20210291);国家自然科学基金(62101249);中国博士后科学基金(2022M721611, 2021TQ0149)

This work was supported by the Natural Science Foundation of Jiangsu Province, China(BK20210291), National Natural Science Foundation of China(62101249) and China Postdoctoral Science Foundation(2022M721611, 2021TQ0149).

通信作者:葛荣骏(rongjun_ge@seu.edu.cn)

visualizations. All of these advantages demonstrate the great potential of NRSR-Net in clinical CT imaging.

Keywords Low-dose CT, Super-resolution, Gradient guidance, Multi-task, Edge

1 引言

低剂量 CT(LDCT)降噪和 CT 超分辨率的准确重建是医学图像处理领域中两个重要的热点研究方向。利用 X 射线、 γ 射线或其他射线类型的能量,CT 已成为一种强大的诊断手段,用于可视化人体内部复杂的结构。这些高质量和高保真度的 CT 图像提供了全面的信息,并在临床诊断中得到广泛应用。然而,获取清晰、高分辨率的 CT 图像会产生大量的计算开销,而降低成像分辨率会导致图像包含的细节减少;同时必要的辐射剂量增加了患癌风险,而减少射线的剂量会给生成的 CT 图像引入噪声。为了克服这些挑战,研究人员在 LDCT 重建和 CT 超分辨率重建领域开展了广泛的研究,并取得了显著进展。LDCT 重建旨在消除 CT 图像中的噪声,而 SR 重建旨在根据低分辨率(LR)图像恢复真实的高分辨率(HR)图像。

深度学习具有强大的特征表达能力,已被广泛应用于医学图像处理中^[1-4],在上述两个领域也取得了显著进展。在 LDCT 重建领域,现有的方法主要可以分为 3 类:基于单弦域、基于单图像域和基于图像和弦双域。基于单弦域,Ghani 等^[5]在弦域使用深度卷积网络来生成高质量的弦图,最终生成高质量的 CT 图像。Ma 等^[6]采用一种注意力残差密集卷积神经网络对 LDCT 弦图去噪。基于单图像域,Chen 等^[7]将自编码器、反卷积网络和残差连接结合在一起,为 LDCT 降噪构建了残差编码器-解码器卷积神经网络(Residual Encoder-Decoder Convolutional Neural Network, RED-CNN)。基于图像和弦双域,Zhang 等^[8]采用生成对抗网络(Generative Adversarial Networks, GAN),分别在图像域和正弦域训练生成器,并计算对应的生成对抗损失对生成器同时进行优化;Ge 等^[9]使用三重交叉注意力机制实现了双域特征的融合,进一步结合双域图像各自的优势。

在 CT 超分辨率重建领域,You 等^[10]使用 GAN 作为构建模块,并通过 Wasserstein 距离在循环一致性方面加强了非线性的从噪声 LR 输入图像到去噪和去模糊 HR 输出的端到端映射。为了降低内存开销并利用体素间距信息,Peng 等^[11]首先在 CT 体积的矢状和冠状方向进行插值,然后在轴向方向融合之前的插值结果。Zhang 等^[12]从输入的全局空间信息中提取注意力并通过图卷积网络获取全局描述符,增强了网络关注图像中信息更丰富的区域和结构的能力。Li 等^[13]利用小波变换的高频特性,提出了一个双通道模型,将小波变换与新的交叉注意力变换器相结合,从而使生成的图像保留更多的高频细节。

在发现降噪任务可以辅助超分辨率任务实现对 LDCT 图像的高质量超分辨率重建之后^[14],研究人员开始探索 LDCT 降噪和超分辨率重建的结合。Yim 等^[15]线性组合了降噪自编码器和超分辨率卷积神经网络,并针对降噪和超分辨率分别单独训练一个网络。然而,上述方法存在以下局限性:1)对图像先降噪再应用超分辨率算法的级联模型存在

一个缺点,即在去噪步骤中,随着噪声的去除,图像的高频内容往往也会损失一部分。2)忽略了降噪和超分辨率任务之间的潜在关联性。Hou 等^[16]通过一个双通道神经网络解决了这些问题。该网络包含一个共享的编码器和两个并行独立的解码器,每个解码器分别执行降噪和超分辨率任务。通过同时训练两个解码器分支,实现了图像超分辨率和降噪的联合学习目标。

尽管双通道神经网络^[16]在 LDCT 图像重建方面取得了成功,但仍有改进的空间。我们认识到,对高频区域特别是边缘区域的复原程度是评判重建结果好坏的重要标准,且高度复原是降噪和超分辨率重建任务的共同目标。因此,本文提出将梯度特征作为引导机制,引导降噪特征和超分辨率特征进行有效融合,使降噪过程更显著地辅助超分辨率过程,生成清晰、高分辨率的 CT 图像。最终,本文提出了同时对 LDCT 进行降噪和超分辨率的降噪和超分辨率重建网络(Noise Reduction and Super Resolution Reconstruction Network, NRSR-Net),该网络充分利用梯度特征从降噪和超分辨率过程中提取共享的高频信息。本文的主要贡献为以下 3 点:

1)提出了边缘增强框架。该框架在梯度特征的引导下提取降噪和超分辨率过程中的相关特征,使得降噪任务对超分辨率任务起到积极影响,从而增强 LDCT 的超分辨率重建性能。具体而言,该框架将梯度信息整合到网络架构和损失函数中。

2)提出了梯度引导融合模块(Gradient Guided Fusion Block,GGFB)。该模块将梯度作为门控机制来引导降噪和超分辨率重建特征的融合,从而增强重要的边缘特征并抑制不重要的无关特征,使融合结果包含更多的边缘信息。

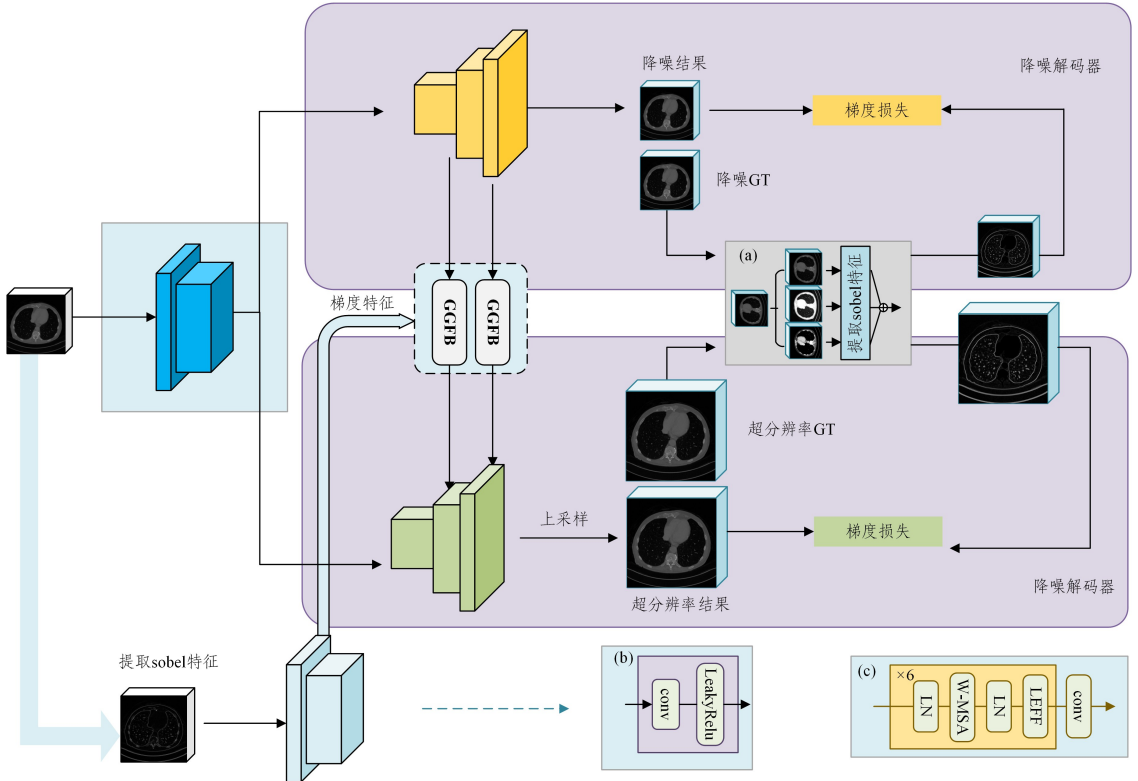
3)提出了梯度损失。为了计算梯度损失,我们从不同的 HU 窗口下的 CT 图像中提取梯度特征,并将其作为预测结果与标签之间 L1 损失的权重。梯度损失引入了更丰富的梯度信息,并引导网络重点重建边缘区域。

2 模型

NRSR-Net 的整体结构如图 1 所示,该网络包含一个共享的编码器和分别用于处理降噪和超分辨率任务的两个独立的解码器。该整体架构的有效性已在之前的工作^[16]中得到证明。与单一超分辨率解码器的结构相比,它表现出了更优越的性能。编码器和解码器使用 Transformer 模块^[17]作为基本模块,以获得更大的感受野。为了降低全局自注意力引起的计算开销并充分利用局部上下文信息,我们采用了 Lewin Transformer 模块^[18]。该模块包含两个关键组成部分:1)非重叠窗口多头自注意力(Non-overlapping Window-based Multi-head Self-Attention, W-MSA),它只在局部窗口内进行自注意力计算,从而极大地减少了计算量;2)局部增强前馈网络(Locally-enhanced Feed-Forward Network, LeFF),它在传统的前馈网络(Feed-Forward Network, FFN)^[17]中加入 Depthwise 卷积,增强了网络获取局部相关性的能力。

在两个编码器之间,通过 GGFB 模块融合两个过程的特征,提取出具有强相关性的边缘特征。本文进一步将梯度

特征的使用拓展到损失函数中,将其作为权重以强调边缘区域的重建。



注:(a)从降噪以及超分辨率标签 CT 中提取不同 HU 窗的梯度特征;(b)用于处理 sobel 特征的基本模块;(c)编、解码器的基本模块。

图 1 NRSR-Net 的整体结构

Fig. 1 Overall architecture of NRSR-Net

2.1 边缘增强框架

边缘增强框架利用输入的 LDCT 的边缘信息作为降噪和超分辨率重建任务之间的纽带,使网络能充分利用两个任务之间的固有相关性。正如我们所知,边缘区域的复原程度是评判重建质量的重要标准,且高度复原是所有重建任务的共同目标。本文框架通过将 Sobel 梯度特征融入网络结构和最终的损失函数中的方式来利用边缘信息。在这些梯度特征的引导下,降噪和超分辨率的过程朝着恢复真实和清晰边缘的共同目标优化,降噪过程中产生的与超分辨率任务强相关的特征将对超分辨率过程起到正向促进作用。为了提取这些梯度特征,我们采用了一个四方向的 Sobel 算子。计算梯度特征的过程如下:

$$G_1 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * A \quad (1)$$

$$G_2 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * A \quad (2)$$

$$G_3 = \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} * A \quad (3)$$

$$G_4 = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} * A \quad (4)$$

$$G = |G_1| + |G_2| + |G_3| + |G_4| \quad (5)$$

其中 G_1, G_2, G_3, G_4 分别表示图像沿 y 轴、 x 轴、 45° 方向、 135° 方向的梯度。将4个方向的梯度相加得到最终的梯度,该梯度的视觉效果如图2所示。

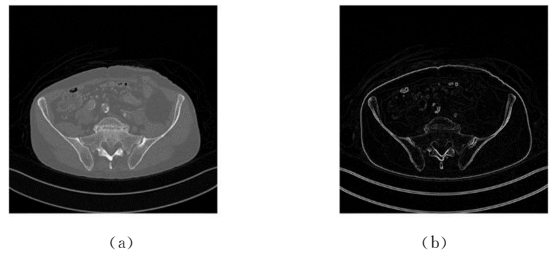


图 2 原 CT 图像与提取出的梯度特征

Fig. 2 Original CT image and extracted gradient features

2.2 梯度引导融合模块

GGFB 包含两个关键模块:空间门控前馈网络(Spatial Gated Feed-Forward Network, SGFN)和通道转置交叉注意力模块(Channel-wise Transposed Cross Attention, CTCA)。这两个模块分别在空间维度和通道维度上调整融合的特征,调整后的特征通过加法操作融合,从而使 GGFB 能够综合考虑这两个维度的信息。

2.2.1 通道转置交叉注意力模块

为了利用梯度特征增强边缘信息并抑制通道维度上的无关细节,我们在通道维度上采用了转置交叉注意力机制。如图3(b)所示,该模块具有两个关键特性:1)该模块将梯度特征

映射为注意力计算中的 query(Q)矩阵,并通过计算交叉注意力图谱得到梯度特征与主线管道特征之间的相关性。通过这种方式,梯度特征能够引导主线管道特征的生成,从而增强生成特征中通道维度包含的梯度信息。2)为实现通道维度的特征调整,我们采用了转置的注意力机制。与传统的注意力不同,特征在空间维度上被展平,而注意力图谱体现的是不同通道之间的相关性。在进行注意力计算之前,通过 3×3 的 Depthwise 卷积对特征进行处理,以保留局部上下文信息。

对于输入的梯度特征、降噪特征和超分辨率特征 X_{grad} , X_{dn} , $X_{sr} \in R^{\hat{C} \times \hat{H} \times \hat{W}}$, 首先将 X_{dn} 和 X_{sr} 拼接起来并使用一个 1×1 的 Pointwise 卷积对降噪和超分辨率特征进行初步融合,得到 X_{fuse} 。经过标准化后, X_{grad} 被映射 Q 矩阵, X_{fuse} 被映射为 K 和 V 矩阵。在映射过程中,我们使用了级联的 1×1 Pointwise 卷积和 3×3 Depthwise 卷积,其中 Pointwise 卷积用于聚合像素级的跨通道上下文信息,随后的 Depthwise 卷积用于编码通道级的空间上下文信息,避免空间维度上的展平操作导致的空间信息丢失。这一过程可用式(6)和式(7)描述:

$$X_{fuse} = W_{p1} \text{concat}(X_{dn}, X_{sr}) \quad (6)$$

$$Q = W_d^Q W_p^Q \text{norm}(X_{grad}) \quad (7)$$

$$K = W_d^K W_p^K \text{norm}(X_{fuse}) \quad (8)$$

$$V = W_d^V W_p^V \text{norm}(X_{fuse}) \quad (9)$$

其中, $W_p^{(Q)}$, $W_p^{(K)}$, $W_p^{(V)}$ 和 W_{p1} 表示 Pointwise 卷积的参数; $W_d^{(Q)}$, $W_d^{(K)}$, $W_d^{(V)}$ 表示 Depthwise 卷积的参数。接下来的处理包括对 Q, K, V 矩阵在空间维度上进行展平,以及对 Q 和 K 矩阵进行点积计算,以得到各个通道之间的相关性,这将生成一个维度为 $R^{\hat{C} \times \hat{C}}$ 的注意力图谱。随后,将注意力图谱与 V

相乘以实现原主线管道特征在通道维度上的加权。以上过程可用式(10)和式(11)表示:

$$X_{channel} = W_{p2} \text{Attention}(\hat{Q}, \hat{K}, \hat{V}) + X_{fuse} \quad (10)$$

$$\text{Attention}(\hat{Q}, \hat{K}, \hat{V}) = \text{Softmax}(\hat{Q} \cdot \hat{K} / \alpha) \cdot \hat{V} \quad (11)$$

其中, W_{p1} 是 Pointwise 卷积; $\hat{Q} \in R^{\hat{C} \times \hat{H} \times \hat{W}}$, $\hat{K} \in R^{\hat{H} \times \hat{W} \times \hat{C}}$, $\hat{V} \in R^{\hat{C} \times \hat{H} \times \hat{W}}$ 是被展平后的 Q, K, V ; α 是一个可学习的参数。

2.2.2 局部增强前馈网络

在 SGFN 中,梯度特征作为门控被用于在空间维度上调整主线管道特征,以增强边缘的重建效果,还原清晰的边缘。SGFN 是基于 FFN^[17] 模块改进而来的,如图 3(a)所示。相较于原始的 FFN 模块,SGFN 引入了一个 Depthwise 卷积层,以初步调整空间信息。经过 Gelu 激活函数处理后,梯度特征通过逐元素乘法作为门控应用于主线管道特征上,实现了对不同空间位置的特征加权。这种门控机制允许主线管道特征受到梯度特征的引导,更加关注具有边缘属性的区域,并抑制那些与边缘无关的属性。通过这种方式,SGFN 在空间维度上实现了对特征的细粒度调整,使得网络能够更好地捕捉到图像中的边缘信息,并提升重建结果的清晰度和准确性。该过程可用式(12)一式(15)表示:

$$X_{spatial} = W_{p3} \text{Gating}(\hat{X}_{fuse}, \hat{X}_{grad}) + X_{fuse} \quad (12)$$

$$\hat{X}_{grad} = W_d W_p \text{norm}(X_{grad}) \quad (13)$$

$$\hat{X}_{fuse} = W_d W_p \text{norm}(X_{fuse}) \quad (14)$$

$$\text{Gating}(\hat{X}_{fuse}, \hat{X}_{grad}) = \text{Gelu}(\hat{X}_{grad}) \odot (\hat{X}_{fuse}) \quad (15)$$

其中, W_p 和 W_{p3} 是 Pointwise 卷积的参数, W_p 将通道数扩展为 γ 倍, W_{p3} 将通道数压缩回原来的大小; W_d 表示 Depthwise 的参数; \odot 表示逐像素相乘。

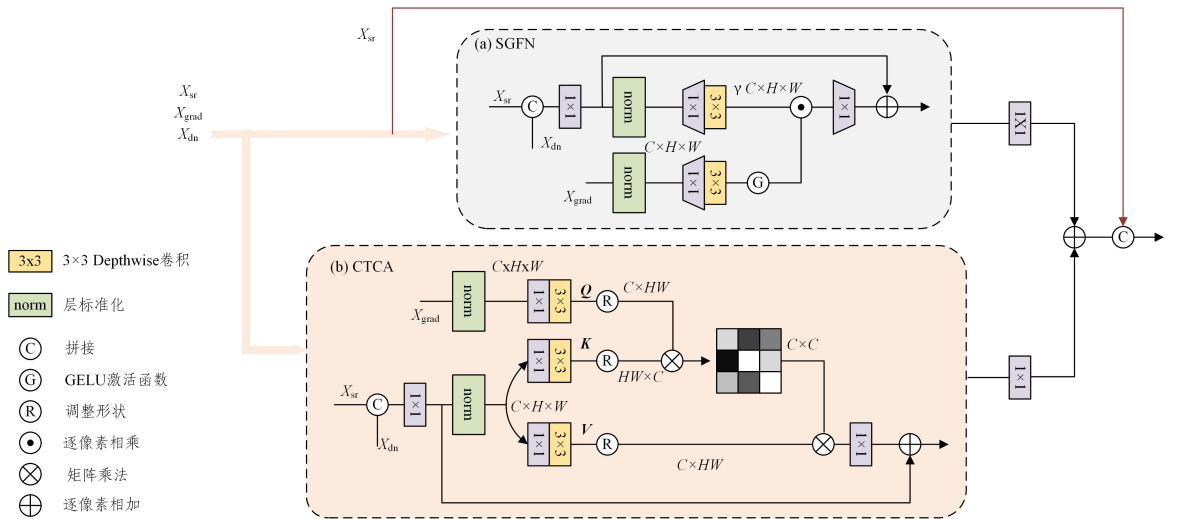


图3 GGFB模块结构

Fig. 3 Architecture of GGFB

2.3 梯度损失

为了提升梯度信息对训练过程的影响,本文将对梯度特征的使用拓展到损失函数中。我们从不同 HU 窗口的 CT 图像中提取出 sobel 特征,并将其作为损失函数的权重。这个改动是基于实际医学诊断中使用 CT 的实践经验,CT 成像的 HU 值取决于人体部位的密度。对于某一特定器官,医生在

使用 CT 图像时通常会将 HU 值调整到该器官对应的 HU 值范围内,以避免其他部位无关信息的干扰。通过调整 HU 的窗宽窗位,CT 图像能更清晰、准确地显示出目标部位的解剖结构,从而帮助医生更直观地了解患者的身体状况。梯度损失使网络侧重于边缘区域的重建,由于边缘区域常常包含重要的结构和特征信息,因此,梯度损失函数的使用能进一步

增强生成 CT 图像的清晰度和细节保持能力。

对于给定的 CT 图像,我们将其 HU 值调整到不同的范围,并提取相应的 sobel 特征来捕捉特定部位的梯度信息。这些梯度信息随后被作为权重应用于损失函数中,使边缘区域的复原程度对网络的优化有更大影响,从而迫使网络侧重于对边缘的重建。超分辨率任务的损失函数为:

$$L_{sr} = \alpha_1 |I_{SR} - I_{HR}|_1 + \alpha_2 L_{percep} + \alpha_3 L_{gen} \quad (16)$$

其中, I_{SR} 是超分辨率解码器的输出结果; I_{HR} 是超分辨率任务的标签图像; L_{percep} 是文献[19]中提出的感知损失,用于衡量在 VGG 特征空间内的距离; L_{gen} 表示对抗损失,用于复原接近原图像的分布。这些损失函数通过 $\alpha_1, \alpha_2, \alpha_3$ 这 3 个超参数来平衡。降噪任务的损失函数可以表示为:

$$L_{dn} = \beta_1 \sqrt{\|I_{DN} - I_{ND}\|^2 + \epsilon^2} \quad (17)$$

其中, ϵ 为经验常数,在本文方法中设为 10^{-3} , β_1 为超参数。梯度损失函数表示为式(18):

$$L_{GL} = \alpha_4 \sum_i G_{HR}^{(i)} |I_{SR} - I_{HR}|_1 + \beta_2 \sum_i G_{ND}^{(i)} |I_{DN} - I_{ND}|_1 \quad (18)$$

其中, I_{DN} 是降噪解码器的输出结果, I_{ND} 是对应的正常剂量图像, $G_{HR}^{(i)}$ 和 $G_{ND}^{(i)}$ 分别表示从 I_{ND} 和 I_{HR} 的不同 HU 窗 CT 图像中提取出的梯度, α_4 表示超分辨率任务中梯度损失的权重, β_2 表示降噪任务中梯度损失的权重。最终的损失函数可以表示为:

$$L = L_{sr} + L_{dn} + L_{GL} \quad (19)$$

3 实验和结果

3.1 数据集

本文的实验使用两个 CT 数据集来验证模型的效果。

“the 2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge”公开数据集:该数据集是一个腹部数据集,包括了来自 10 名患者的 CT 扫描图像,涵盖了 LDCT 图像和其对应的常规剂量扫描图像,共计 4553 对 CT 图像,每张图像的分辨率为 512×512 。在实验前,手动对数据集进行了划分,选择来自 8 名患者的 3417 对图像作为训练集,将来自一名患者的 610 对图像作为验证集,将最后一名患者的 525 对图像作为测试集。

模拟数据集:模拟数据集包括 40 个病人的胸部 CT 扫描图像,原始数据来源于公开挑战赛“ISBI 2019 SegTHOR”。由于原始的数据集中只有正常剂量的 CT 图像,因此我们模拟生成不同剂量的低剂量 CT。在插入低剂量噪声时,我们使用了光子计数模型^[20]。该模型模拟了 LDCT 扫描过程,噪声插入过程表示为:

$$P_B = P_A + \sqrt{\frac{1 - a \exp(P_A)}{a} \frac{N_{0A}}{N_{0A}}} \left(1 + \frac{1 + a N_e \exp(P_A)}{a} \frac{N_{0A}}{N_{0A}} \right) x \quad (20)$$

其中, P_A 是正常剂量的对数转换投影数据, N_{0A} 为光子的入射数, P_B 为模拟的低剂量投影数据, a 表示模拟扫描的剂量水平, x 是服从标准正态分布的随机变量, N_e 为电子噪声的噪声等效量子^[21]。 N_{0A} 设置为 10^5 , N_e 设置为 $10^{[20]}$ 。在实验中,我们模拟了 10% 和 20% 剂量的低剂量 CT,并手动对数据集进行了划分:选择来自 28 名患者的 5181 对 CT 扫描作为训练集、来自 4 名患者的 649 对 CT 扫描作为验证集,来自剩下 8 名患者的 1590 对 CT 图像作为测试集。

在预处理阶段,使用传统的 Bicubic 算法将原始 CT 图像降采样 4 倍,得到分辨率为 128×128 的图像。在实验中,将低分辨率 LDCT 图像作为模型的输入,低分辨率 NDCT 作为降噪任务的标签,原始分辨率的 NDCT 作为超分辨率任务的标签。将本文提出的模型与目前最先进的 4 种超分辨率方法 SRGAN^[22], ESRGAN^[19], SPSR^[23] 和 SwinIR^[24] 进行了比较,它们都是使用 GAN 进行训练的。

3.2 实验设置

本实验采用 PyTorch 框架来实现、训练和测试所提出的模型,实验平台的配置为 Nvidia GeForce RTX 3090。在训练过程中,优化器采用了 Adam^[25],学习率为 0.0002,每一批包含 16 个样本。将所有模型训练 150 个周期,每 50 个周期学习率减半。在将 LDCT 送入网络之前,对图片进行了归一化处理,并随机进行旋转、翻转和裁剪,将其划分为大小为 32×32 的图像块。 $\alpha_1, \alpha_2, \alpha_3, \beta_1$ 和 β_2 的值分别设置为 0.01, 1, 0.005, 0.005 和 0.01。为评估本文方法的重建性能,我们采用了 3 个广泛使用的指标:峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似性指数(Structural Similarity Index, SSIM)和可学习感知图像块相似性(Learned Perceptual Image Patch Similarity, LPIPS)^[26]。

PSNR 能够衡量两个图像像素级的距离,其计算式为:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (21)$$

$$PSNR = 10 * \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (22)$$

其中, $I(i, j)$ 是真实的图像, $K(i, j)$ 是预测结果, MAX_I 表示真实图像中的最大像素值, MSE 表示均方差。

SSIM 从亮度、对比度和结构 3 个方面来衡量图像间的相似性,其计算式为:

$$SSIM(I, K) = \frac{(2 * \mu_I * \mu_K + c_1) * (2 * \sigma_{IK} + c_2)}{(\mu_I^2 + \mu_K^2 + c_1) * (\sigma_I^2 + \sigma_K^2 + c_2)} \quad (23)$$

其中, μ_I 和 μ_K 表示 I 和 K 的均值, σ_I^2 和 σ_K^2 表示 I 和 K 的方差, σ_{IK} 是二者的协方差。 c_1 的值为 $(0.01 * MAX_I)^2$, c_2 的值为 $(0.03 * MAX_I)^2$ 。

为了评估预测结果与真实图像在视觉感官上的相似性,本文还使用 LPIPS 作为衡量指标。LPIPS 计算的是两张图像在 VGG 特征空间上的距离。

3.3 与现有算法的比较

3.3.1 定量比较

本节实验中,将本文提出的 NRSR-Net 与目前超分辨率重建领域几种最先进的算法进行了比较,包括 SRGAN^[22], ESRGAN^[19], SPSR^[23] 和 SwinIR^[24]。值得强调的是, SwinIR 在训练中采用了 GAN 损失。我们分别在 AAPM 公开数据集和不同剂量的模拟低剂量 CT 数据集上对以上方法进行了比较,并使用 PSNR, SSIM 和 LPIPS 这 3 个指标对重建效果进行多维度的评价。定量比较的结果如表 1 所列,本文模型在大多数情况下都优于其他方法,在 20% 剂量的模拟数据集上, NRSR-Net 在 LPIPS 这一指标上相比 SPSR 不占优势,但在 PSNR 和 SSIM 上大幅领先。在 3 个指标上的综合提升证明,相较于其他方法, NRSR-Net 的重建结果有更高的准确性和视觉感知相似性。

表 1 在不同剂量、不同数据集下与其他方法的定量比较

Table 1 Quantitative comparison with the existing methods on different datasets at various doses

Dataset	Metric	Bicubic	SRGAN	ESRGAN	SPSR	SwinIR	Ours
AAPM	PSNR	31.8540±1.4160	34.0190±1.3620	37.2510±1.6690	<u>37.2630±1.6750</u>	37.0500±1.5220	37.8370±1.610
	SSIM	0.8647±0.0345	0.8846±0.0391	0.9146±0.0317	<u>0.9198±0.0317</u>	0.9160±0.0297	0.9244±0.0270
	LPIPS	0.2773±0.0387	0.1348±0.0340	0.1118±0.0261	0.1067±0.0239	<u>0.1026±0.0227</u>	0.1015±0.0229
20% dose	PSNR	33.2020±2.4930	36.5270±2.4620	36.8433±2.3770	<u>37.2970±2.4690</u>	37.1430±2.5860	37.4240±2.4960
	SSIM	0.9020±0.0280	0.9219±0.0343	0.9182±0.0266	0.9286±0.0248	<u>0.9342±0.0292</u>	0.9371±0.0252
	LPIPS	0.2045±0.0596	0.0846±0.0367	0.0793±0.0350	<u>0.0721±0.0301</u>	0.0734±0.0316	0.0683±0.0270
10% dose	PSNR	33.1020±2.4860	36.3290±2.3970	36.7820±2.4300	36.6200±2.4240	<u>36.9190±2.5340</u>	37.1790±2.4970
	SSIM	0.8841±0.0325	0.9198±0.0329	<u>0.9292±0.0294</u>	0.8518±0.0394	0.9249±0.0335	0.9395±0.0261
	LPIPS	0.2223±0.0595	0.0886±0.0383	0.0840±0.0363	0.0744±0.0292	0.0771±0.0323	<u>0.0748±0.0293</u>

注:加粗表示最好的结果,下划线表示次好的结果。

3.3.2 定性比较

为了获取更直观的比较效果,本节实验对不同方法生成的图像的解剖细节进行对比,如图 4 和图 5 所示。图 4 给出了 AAPM 数据集上不同方法的重建结果,其中图 4(a)展示了血管部位的重建效果。在图像重建任务中,由于血管具有形状尺寸多变、与周围部位对比不明显等特性,对血管部位的有效还原一直是一项具有挑战性的任务。在图 4(a)中可以

看到,NRSR-Net 比较清晰地恢复了血管轮廓;在图 4(b)中,对于医学图像中常见的微小结构,只有 NRSR-Net 成功重建出清晰的边缘。图 5 给出了各个模型在 10% 和 20% 的模拟数据集上的重建结果,可以看到,在不同的剂量下,对于不同的器官,NRSR-Net 相比其他方法有更好的重建效果。解剖细节的对比证明了 NRSR-Net 在重建边缘区域上的优势,这归功于将梯度特征融入中间特征的生成过程和损失函数这一做法。

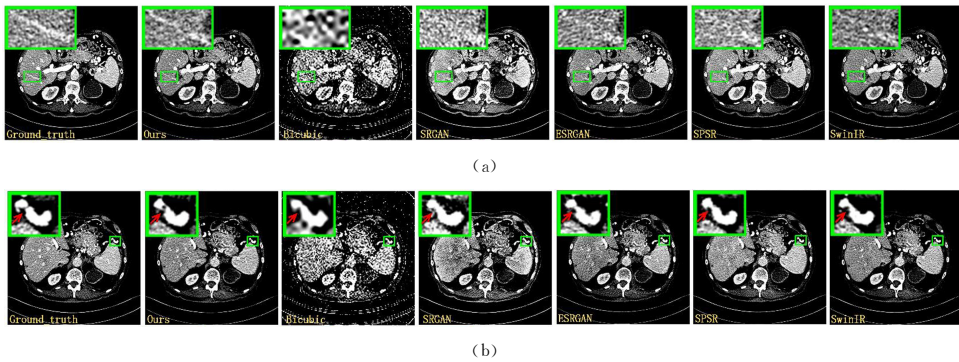


图 4 AAPM 数据集上不同方法重建结果的视觉效果

Fig. 4 Visual effects of reconstruction results by different methods on AAPM dataset

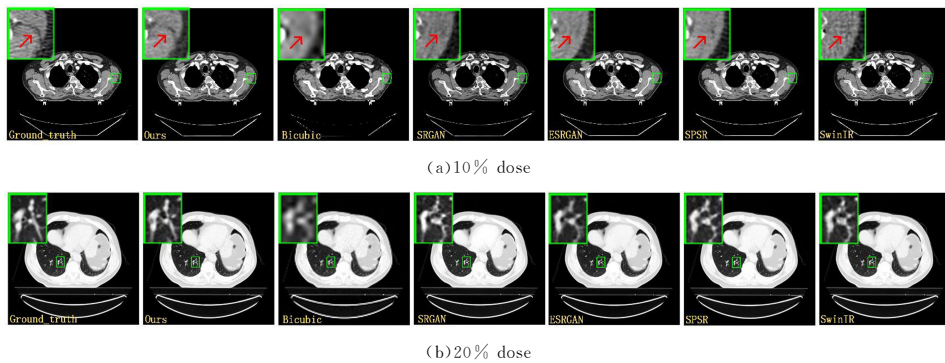


图 5 不同方法在 10% 和 20% 剂量的模拟数据集上的结果

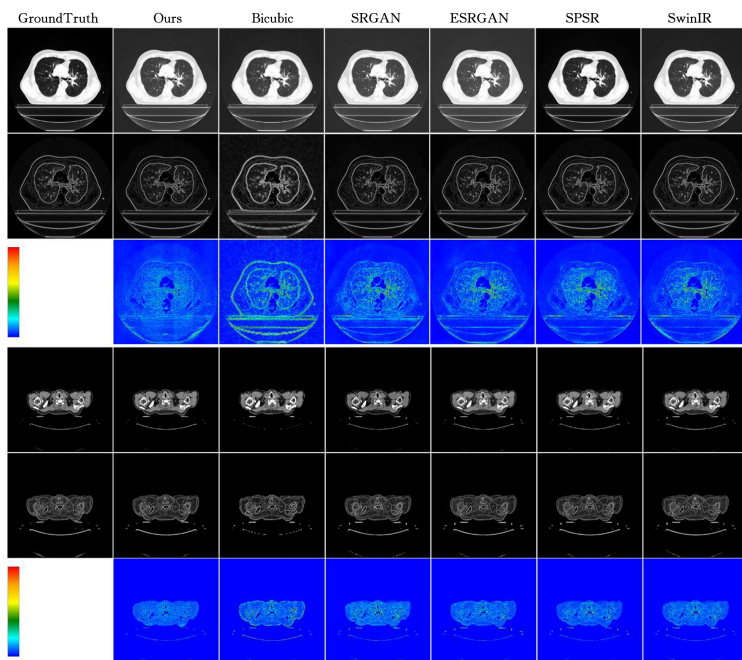
Fig. 5 Results of different methods on simulated datasets at 10% and 20% doses

为了更直观地描述重建图像的误差分布情况以及 NRSR-Net 对边缘区域的重建效果,我们计算了重建图像和真实图像的绝对误差,以及重建图像的 sobel 梯度特征,结果如图 6 所示。结合绝对误差图和梯度分布图可以发现,在边缘区域 NRSR-Net 的绝对误差更小,这表明本文模型能更好地重建边缘。

为了进一步对实验结果进行定性分析,本节实验还比较了不同结果图像的灰度变化曲线。如图 7 所示,灰度变化

曲线描述了沿着绿色线的像素的灰度值。为了描述这一变化,实验计算了曲线的均值±标准差。从图像中可以看到,Bicubic 和 SRGAN 由于存在较为明显的伪影,其标准差与真实值存在较大偏差。

将 NRSR-Net 与其他几种方法的结果进行对比可以看到,在标准差与真实值相差不大的情况下,NRSR-Net 的均值最接近真实值,这表明 NRSR-Net 生成的结果更符合真实分布。



注:每组图片的第二排为重建图像的 sobel 特征,第三排为与真实图像之间的绝对误差。

图 6 不同方法重建结果的梯度分布图和绝对误差图

Fig. 6 Gradient distribution maps and absolute error maps of reconstruction results by different methods

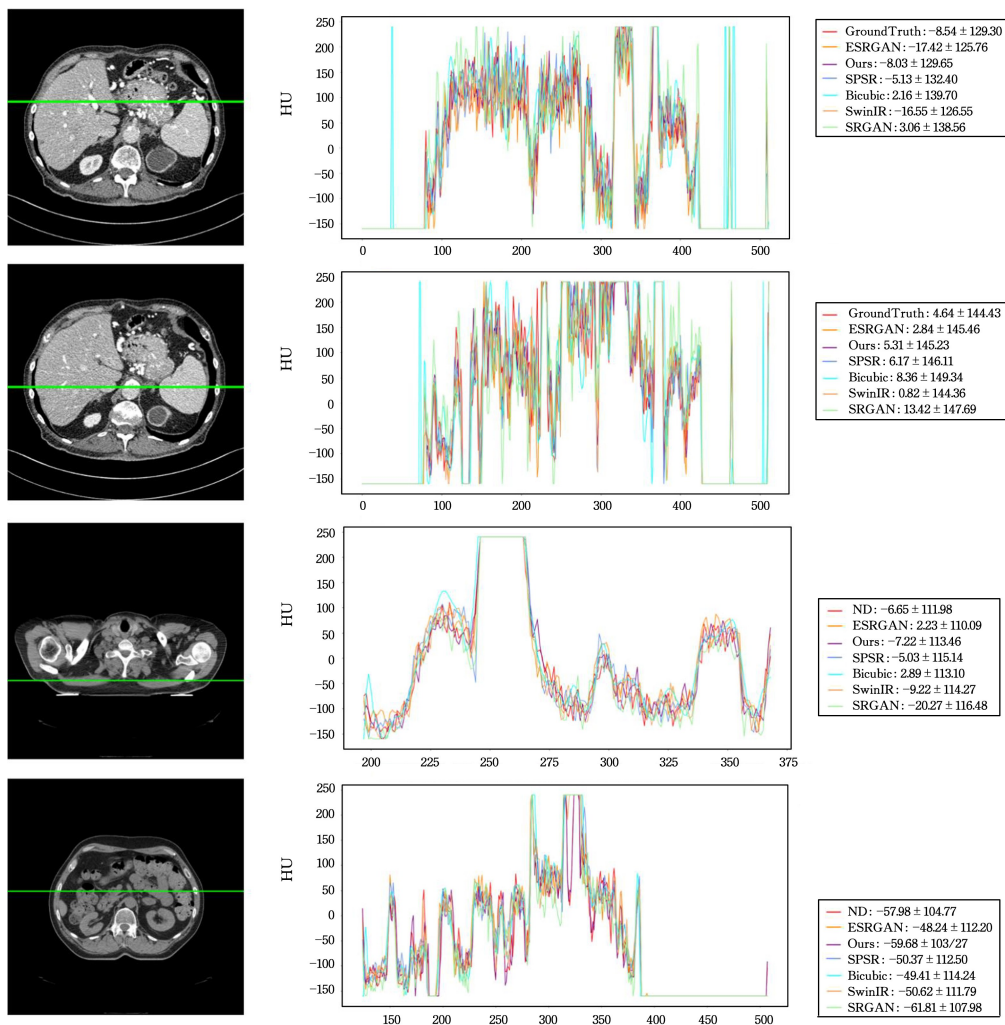


图 7 不同重建结果沿绿线的灰度值变化曲线(电子版为彩图)

Fig. 7 Grayscale variation curves along the green line of different SR reconstruction results

3.4 消融实验

为了探究每个模块各自的贡献,本文进行了一系列消融实验。此外,为了探究梯度特征对 GGFB 乃至整个网络性能的影响,本文设计了一个额外的 GGFB 变体模块,该模块结构与原 GGFB 大致相同,但输入特征不包括梯度特征。消融实验涵盖 5 种不同的网络变体:1) Baseline, 基准模型,该模型不包括 GGFB 和梯度损失;2) w/o GGFB, 不包括 GGFB 模块;3) w/o GL, 不包括梯度损失;4) GGFB(w/o grad), 包含 GGFB 变体模块以及梯度损失;5) Full Version, 本文提出的最终网络, 包含 GGFB 和梯度损失。

表 2 不同网络变体的消融实验

Table 2 Ablation study on different variant network

Model	GGFB	GGFB(w/o grad)	GL	PSNR	SSIM	LPIPS
Baseline	×	×	×	37.4940±1.5610	0.9215±0.0244	0.1066±0.0228
w/o GGFB	×	×	✓	37.5320±1.6560	0.9161±0.0286	0.1073±0.0241
w/o GL	✓	×	×	37.5270±1.6600	0.9172±0.0274	0.1035±0.0227
GGFB(w/o grad)	×	✓	✓	37.6930±1.6710	0.9216±0.0305	0.1029±0.0233
Full Version	✓	×	✓	37.8370±1.6060	0.9244±0.0270	0.1015±0.0229

注:GGFB(w/o grad)是 GGFB 模块不使用梯度特征作为输入的变体,该变体模块用于验证使用梯度特征的有效性;GL 表示梯度损失。加粗的数字表示最好的结果。

结束语 本文提出了 NRSR-Net, 该网络利用降噪和超分辨率任务之间的强相关性,对 LDCT 同时进行降噪和超分辨率重建。本文的贡献可以总结为 3 点:1) 提出了边缘增强框架,利用边缘信息提取了降噪和超分辨率过程中高度相关的特征。通过共同优化这些相关特征,使降噪和超分辨率任务相互增强,从而提升了 LDCT 的降噪和超分辨率重建的性能。2) 提出了 GGFB, 它在两个任务的中间特征融合中起到引导作用,采用门控机制对特征进行微调,确保融合结果中包含丰富的边缘信息。3) 提出了梯度损失,对不同 HU 窗的 CT 图像计算梯度,获取不同部位结构的边缘信息,并将其作为最终损失函数中的权重。梯度损失给网络引入更全面的边缘信息,使网络优先重建边缘区域。本文在公开的腹部 CT 数据集以及不同剂量的模拟 CT 上进行了一系列实验,并使用图像重建领域常用的指标(包括 PSNR, SSIM 和 LPIPS)评估最终的重建结果。定量指标结果表明了 NRSR-Net 的有效性。本文还比较了视觉结果,发现 NRSR-Net 能够重建相对清晰的轮廓和锐利的边缘,即使对于血管等具有挑战性的复杂部位和微小结构的边缘区域也是如此。这些优势表明 NRSR-Net 在临床 CT 成像中具有巨大的潜力。

参考文献

[1] SUN F Q, CUI Z Q, ZOU P, et al. Brain Tumor Segmentation Algorithm Based on Multi-scale Features[J]. Computer Science, 2022, 49(S1): 12-16.

[2] YAO Z W, YANG F, HUANG J, et al. Improved CycleGANs for Intravascular Ultrasound Image Enhancement[J]. Computer Science, 2019, 46(5): 221-227.

[3] LIU P P, ZHANG W H, LU Z T, et al. Prediction of Malignant and Benign Gastrointestinal Stromal Tumors Based on Radiomics Feature[J]. Computer Science, 2019, 46(1): 285-290.

[4] WANG L F, WANG R F, LIN S Z, et al. Multimodal Medical Image Fusion Based on Dual Residual Hyper Densely Networks

网络变体的实验结果如表 2 所列,可以看出, w/o GL 相较于 Baseline 在 PSNR 和 LPIPS 上的领先证明了 GGFB 有效地整合了降噪和超分辨率特征,保留了重要的重建信息; Full Version 相对于 GGFB(w/o grad)的全面领先验证了梯度特征在引导融合过程和突出共享边缘特征方面的有效性;虽然 w/o GGFB 相比 Baseline 不占优势,但 Full Version 相比 w/o GL 的提升证实了梯度损失的有效性。本文对这一现象的解释是,尽管单独的梯度损失可能影响有限,但它可以配合 GGFB 增强边缘信息的提取,从而进一步增强模型的性能。

[J]. Computer Science, 2021, 48(2): 160-166.

[5] GHANI M U, KARL W C. CNN based sinogram denoising for low-dose CT[C]// Mathematics in Imaging. Optica Publishing Group, 2018.

[6] MA Y J, REN Y, FENG P, et al. Sinogram denoising via attention residual dense convolutional neural network for low-dose computed tomography[J]. Nuclear Science and Techniques, 2021, 32: 1-14.

[7] 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.

[8] ZHANG Y, HU D, ZHAO Q, et al. CLEAR: comprehensive learning enabled adversarial reconstruction for subtle structure enhanced low-dose CT imaging[J]. IEEE Transactions on Medical Imaging, 2021, 40(11): 3089-3101.

[9] GE R, HE Y, XIA C, et al. DDPNet: A Novel Dual-Domain Parallel Network for Low-Dose CT Reconstruction[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2022: 748-757.

[10] YOU C, LI G, ZHANG Y, et al. CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE)[J]. IEEE Transactions on Medical Imaging, 2019, 39(1): 188-203.

[11] PENG C, LIN W A, LIAO H, et al. Saint: spatially aware interpolation network for medical slice synthesis[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 7750-7759.

[12] ZHANG Y, LI K, LI K, et al. MR image super-resolution with squeeze and excitation reasoning attention network[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13425-13434.

[13] LI G, LYU J, WANG C, et al. WavTrans: Synergizing wavelet and cross-attention transformer for multi-contrast MRI super-

- resolution[C] // International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham:Springer Nature Switzerland,2022;463-473.
- [14] SINGH A,PORIKLI F,AHUJA N. Super-resolving noisy images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014;2846-2853.
- [15] YIM D,KIM B,LEE S. Simultaneous denoising and spatial resolution enhancement using convolutional neural network-based linear model in diagnostic CT images[C] // Medical Imaging 2020;Physics of Medical Imaging. SPIE,2020;1031-1036.
- [16] HOU H,JIN Q,ZHANG G, et al. CT image quality enhancement via a dual-channel neural network with jointing denoising and super-resolution[J]. Neurocomputing,2022,492;343-352.
- [17] VASWANI A,SHAZEER N,PARMAR N,et al. Attention is all you need[C]//Advances in Neural NIPS'17;Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017;6000-6010.
- [18] WANG Z,CUN X,BAO J, et al. Uformer: A general u-shaped transformer for image restoration [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022;17683-17693.
- [19] WANG X,YU K,WU S, et al. Esrgan; Enhanced super-resolution generative adversarial networks[C] // Proceedings of the European Conference on Computer Vision (ECCV) Workshops. 2018.
- [20] YU L,SHIUNG M,JONDAL D, et al. Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols[J]. Journal of computer Assisted Tomography,2012,36(4):477-487.
- [21] YANG L,LI Z,GE R, et al. Low-dose ct denoising via sinogram inner-structure transformer[J]. IEEE Transactions on Medical Imaging,2022,42(4):910-921.
- [22] LEDIG C,THEIS L,HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017;4681-4690.
- [23] MA C,RAO Y,CHENG Y, et al. Structure-preserving super resolution with gradient guidance [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020;7769-7778.
- [24] LIANG J,CAO J,SUN G, et al. Swinir; Image restoration using swin transformer[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021;1833-1844.
- [25] KINGMA D P,BA J. Adam; A method for stochastic optimization[J]. arXiv:1412. 6980,2014.
- [26] ZHANG R,ISOLA P,EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018;586-595.



XU Ying, born in 1997, postgraduate. His main research interests include medical image reconstruction and deep learning.



GE Rongjun, born in 1992, Ph.D, associate professor, is a member of CCF (No. 13248M). His main research interests include deep learning, intelligent reconstruction and analysis of medical image, medical information analysis and image processing.

(责任编辑:杨雪敏)