E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS

中图法分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1006-8961(2025)11-3694-13

论文引用格式:Zhang L, Chen S and Hua Y Y. 2025. Multi-scale attention encoder-decoder network for low-dose dental CBCT denoising. Journal of Image and Graphics, 30(11):3694-3706(张立, 陈胜, 华旖筠. 2025. 面向低剂量口腔CBCT去噪的多尺度注意力编解码网络. 中国图象图形学报, 30(11):3694-3706)[DOI:10.11834/jig.250055]

面向低剂量口腔CBCT去噪的多尺度注意力编解码网络

张立,陈胜*,华旖筠

上海理工大学光电信息与计算机工程学院,上海　200093

摘　要:**目的**　低剂量计算机断层扫描(low-dose computed tomography, LDCT)技术因辐射剂量低备受关注,但其图像噪声和伪影问题严重影响诊断准确性。尽管肺部LDCT去噪技术已取得显著进展,针对低剂量口腔锥形束计算机断层扫描(cone beam computed tomography, CBCT)图像的去噪研究仍较少。口腔CBCT图像因高密度牙齿组织与低密度软组织之间的宽动态范围差异,以及根管等细微结构的低对比度特性,导致目前LDCT去噪方法在口腔CBCT应用中易出现过度平滑和细节丢失现象。针对上述挑战,提出一种基于多尺度特征融合与方向注意力机制的口腔低剂量CBCT图像去噪网络模型。**方法**　模型采用编码器—解码器架构实现端到端的噪声学习与去除。通过多尺度特征融合模块(multi-scale feature fusion module, MFFM)提取口腔内不同尺度特征信息,并结合方向注意力特征细化模块(directional attention feature refinement module, DAFRM)动态增强对牙釉质—牙本质界面及牙髓区域的特征表达。为进一步优化网络去噪性能,设计包含像素损失、平滑损失以及结构相似性损失的联合损失函数,通过权重分配实现噪声抑制与细节保留之间的平衡。**结果**　在口腔低剂量CBCT数据集上与7种常见方法进行对比。实验结果表明,相较于其他LDCT去噪方法,本文模型在各项评价指标上均取得显著提升。峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似性(structural similarity index measurement, SSIM)指标值分别达到38.40 dB和0.952 4,相比RED-CNN(residual encoder-decoder convolutional neural network)和WGAN(Wasserstein generative adversarial network), PSNR分别提高约3.54 dB和14 dB;SSIM分别提升3%和18%。**结论**　所提方法能够有效地减少图像噪声和伪影,且视觉效果更清晰。

关键词:低剂量计算机断层扫描(LDCT);锥形束计算机断层扫描(CBCT);图像去噪;编码器;解码器;注意力机制

Multi-scale attention encoder-decoder network for low-dose dental CBCT denoising

Zhang Li, Chen Sheng*, Hua Yiyun

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai　200093, China

Abstract: **Objective** Low-dose computed tomography (LDCT) technology has attracted considerable attention in medical imaging because of its capability to substantially reduce radiation exposure while preserving clinical diagnostic value. In dental medicine, cone-beam CT (CBCT), as a pivotal 3D imaging technology, encounters a critical challenge: optimizing

收稿日期:2025-02-21;**修回日期**:2025-04-08;**预印本日期**:2025-04-15

*通信作者:陈胜　chnshn@163.com

基金项目:国家自然科学基金项目(81101116)

Supported by: National Natural Science Foundation of China (81101116)

radiation dose control not only impacts image quality but also directly affects patient safety. However, dose reduction inevitably intensifies quantum noise and artifact generation, which creates a dual challenge in oral imaging due to the complex anatomical characteristics of regions. Although denoising techniques for low-dose pulmonary CT imaging have advanced considerably, research on noise suppression in low-dose oral CBCT remains limited. The wide dynamic range between high-density structures such as teeth and alveolar bone and low-density soft tissues including gingiva and mucosa, along with the inherently low contrast of delicate anatomical features such as root canals, often leads to over-smoothing and loss of critical details in current deep learning-based low-dose CT denoising algorithms. These limitations hinder clinical diagnostic accuracy, which underscores the need for tailored noise reduction strategies specific to the unique imaging challenges of oral CBCT. Therefore, this study aims to explore an efficient denoising method for low-dose oral CBCT images. This method achieves effective suppression of image noise and artifacts while maximally preserving detailed information of dental edges, pulp chambers, and soft tissues, under the premise of ensuring radiation safety. **Method** A multi-scale attention-based encoder-decoder network model is proposed to address the challenges of excessive noise, prominent artifacts, and blurred details in low-dose dental CBCT images. The model adopts an end-to-end encoder-decoder architecture, which utilizes multi-layer convolution, downsampling, and upsampling operations to achieve direct mapping from noisy images to high-quality outputs. During the encoding phase, the network extracts low-, medium-, and high-level feature information through multi-layer convolution and pooling operations, which progressively captures global and local semantic information. In the decoding phase, upsampling and residual connections are employed to restore image details while preserving high-resolution information. The core innovations of the model lie in the integration of the multi-scale feature fusion module (MFFM) and the directional attention feature refinement module (DAFRM). The MFFM utilizes parallel convolutional kernels of varying sizes to extract multi-scale features, followed by self-attention mechanisms to dynamically weight and fuse these features. Thus, it comprehensively integrates information from tooth contours, bone structures, and soft tissue textures. The DAFRM focuses on enhancing regions with distinct spatial directional characteristics, such as the enamel-dentinal junction and pulp chambers. By leveraging adaptive pooling and attention-weighted strategies along horizontal and vertical orientations, this module strengthens the representation of critical anatomical details. To further optimize denoising performance, a hybrid loss function is designed, which combines pixel-wise mean squared error, smoothness loss, and structural similarity loss. Through multi-objective joint optimization, the network achieves noise suppression while maintaining pixel-level accuracy, structural continuity, and realistic restoration of edge details. This approach ensures a balance between radiation safety and diagnostic fidelity, which provides a robust framework for high-quality image reconstruction in low-dose dental CBCT applications. **Result** Experiments were conducted on a low-dose dental CBCT dataset to systematically validate the effectiveness of the proposed denoising algorithm, and its performance was compared with those of seven state-of-the-art denoising methods. Experimental results demonstrate significant improvements across all evaluation metrics. The proposed model achieves a peak signal-to-noise ratio of 38.40 dB, which outperforms RED-CNN and WGAN by 3.54 dB and 14 dB, respectively. For the structural similarity index, the model attains a value of 0.952 4, which represents improvements of 3% and 18% over the aforementioned baselines. Visual comparisons further confirm that the proposed method effectively preserves fine structures such as tooth edges and root canals while eliminating noise, avoiding common over-smoothing artifacts, and yielding denoised images with superior visual clarity. To further verify the robustness of the algorithm, additional tests were performed on CBCT datasets under varying dose levels. The results indicate that the proposed model consistently delivers high-quality denoising performance across higher and lower doses, which demonstrates strong adaptability to variations in noise distribution and intensity. This performance highlights the robust generalization capability and reliability of the model in clinical scenarios with heterogeneous imaging parameters. **Conclusion** The proposed multi-scale attention encoder-decoder network model for low-dose oral CBCT denoising effectively reduces noise and artifacts induced by low-dose imaging while preserving critical structural details. This model provides high-quality imaging foundations for subsequent clinical applications such as endodontic diagnosis and orthodontic treatment planning.

Key words: low-dose computed tomography (LDCT); cone-beam computed tomography (CBCT); image denoising; encoder; decoder; attention mechanism

0 引 言

随着计算机技术的发展，口腔数字化技术在口腔医学中的应用日益广泛。锥形束计算机断层扫描（cone beam computed tomography，CBCT）因其成像速度快、辐射剂量低等优势（赵阳 等，2024），已成为重要的诊断工具。然而，传统计算机断层扫描（computed tomography，CT）和CBCT在重复扫描时可能增加患者的辐射风险。在ALARA（as low as reasonably achievable）（Brenner和 Hall，2007）指导原则下，低剂量CT（low-dose computed tomography，LDCT）的研究受到国内外的广泛关注。目前降低辐射剂量最常见的方法就是减少X射线管的工作电流和缩短曝光时间以减少X射线的通量。然而，X射线通量的减弱通常会导致重建的CT图像噪声增加，从而降低图像的信噪比，进而导致医生诊断结果的高假阳性率。为解决这一固有的物理问题，国内外提出很多算法来改善LDCT图像的质量，这些算法按照图像重建前后大致可分为以下3类：重建前正弦域滤波、迭代重建和重建后的图像后处理。

正弦域滤波技术通过在图像重建前对原始数据或其对数变换进行操作，如滤波反投影（Pan 等，2009），来改善图像质量，但其缺点是会导致图像分辨率下降和边缘模糊。过去几十年中，迭代重建算法（iterative reconstruction，IR）（Willemink 和 Noël，2019）在低剂量CT图像重建中取得显著进展，通过结合数据统计特性、图像先验信息和成像系统参数，有效提高重建图像质量。然而，IR方法的高计算成本和参数的不公开很大程度上限制了其实际应用。

与正弦域滤波和IR方法不同，图像后处理技术直接作用于重建后的图像，不依赖原始数据。例如，非局部均值（Zhang 等，2014；Diwakar 等，2020）方法适用于CT图像去噪，K奇异值分解（Aharon 等，2006）能够减少CT图像中的伪影，三维块匹配（Feruglio 等，2010）算法用于多个CT成像任务中的图像恢复。通过这些图像后处理方法，使得LDCT图像的质量得到明显提高，但处理后的图像中仍然存在平滑度和结构残差问题。由于CT图像噪声的分布不均匀，导致目前这些问题难以得到解决。

近年来，深度学习在医学图像处理领域崭露头角，从低级的图像去噪、去模糊到高级的分割、检测任务，深度学习无处不在。它模拟人类的信息处理过程，通过分层网络框架从复杂的像素数据中有效学习高级特征。针对低剂量CT去噪问题，深度学习技术同样展现出强大的潜力。目前主流的低剂量CT图像去噪方法在深度学习领域大致可以分为以下3类：基于卷积神经网络（convolution neural network，CNN）（Kim 等，2019；You 等，2018）的去噪模型、基于生成对抗网络（generative adversarial network，GAN）（Huang 等，2022；Marcos 等，2021）的去噪模型和基于视觉Transformer（vision Transformer，VIT）的去噪模型。

基于CNN的低剂量CT去噪网络，通过学习噪声图像与清晰图像之间的映射关系，利用卷积提取图像特征并逐步优化网络来减少噪声。以U-Net（Ronneberger 等，2015）为例，其独特的编码器—解码器对称结构可以同时处理全局和局部信息，显著提升去噪图像质量。Chen 等人（2017）同样采用编码器—解码器网络结构，通过添加残差连接有效地提取图像特征，重构出高质量图像。Liang 等人（2020）则通过边缘增强模块，在去噪的同时，保留图像的边缘信息。Won 等人（2019）将输入特征图分为高频和低频两部分，有效地分离噪声和处理图像细节。Gu 等人（2024）通过引入分布回归层和预测性预训练框架，进一步优化去噪效果。

基于GAN的低剂量CT去噪网络采用生成对抗的方式实现图像去噪。其核心架构主要由两个部分组成：一个生成器网络和一个判别器网络。生成器网络专注于生成去噪后的图像，而判别器网络则评估这些图像与正常剂量图像（normal-dose CT，NDCT）之间的差异。在训练过程中，生成器不断学习如何生成更加逼真的去噪图像，而判别器则不断提高其鉴别能力来准确区分生成图像与真实图像。然而，这种对抗训练机制在提升图像去噪质量的同时，也面临着训练过程不稳定的挑战。为解决这一问题，WGAN（Wasserstein GAN）（Yang 等，2018）通过引入Wassertein距离度量并结合感知损失函数，在有效缓解GAN训练过程不稳定的同时，进一步提升生成图像的视觉效果和细节保留能力。

Transformer最初用于自然语言处理（natural language processing，NLP）领域，其核心的自注意力机制允许模型在处理序列数据时捕捉长距离依赖关系。随着对Transformer架构理解的不断加深，其设计理

念和技术优势创新性地迁移至计算机视觉领域，VIT技术应运而生。不同于CNN以及GAN的局部特征提取，VIT将图像分割成多个块，通过自注意力层来捕捉块与块之间的关系，从而实现对图像全局信息的理解。在低剂量CT去噪中，VIT可以更全面地考虑图像中的噪声和结构信息，从而生成精确的去噪结果。Wang等人（2021）提出的不包含任何卷积层的纯Transformer架构网络，充分展示了VIT在低剂量CT去噪中的优越性能（Wang等，2023）。

尽管基于GAN和Transformer的低剂量CT去噪方法在性能上表现出色，但其训练过程通常依赖大规模数据集以捕捉数据分布规律和复杂噪声模式，从而确保模型的泛化能力并避免过拟合现象。相比之下，基于CNN的网络模型凭借其局部感受野和层次化特征抽象的优势，在医学图像处理领域积累丰富的实践经验并得到广泛认可。

针对现有低剂量CT去噪方法在口腔CBCT图像去噪中存在的过度平滑及细节丢失问题，本文提出一种多尺度注意力编解码网络。该网络在编解码器之间设计多尺度特征融合模块（multi-scale feature fusion module，MFFM），通过对不同尺度特征信息的捕捉与融合，MFFM能够确保特征的丰富性与完整性，从而有效避免图像关键细节丢失。同时，本文引入复合损失函数，强化去噪后图像的整体结构一致性。在此基础上，进一步设计了方向注意力特征细化模块（directional attention feature refinement module，DAFRM）。DAFRM通过对特征图沿不同空间方向进行分组聚合，并结合自适应注意力机制动态调整特征权重，实现对牙齿及牙髓区域的重点关注。

1 方 法

LDCT图像去噪的整个过程可以简化为以下问题，对于一个LDCT图像，可将其定义为$\boldsymbol{X} \in \mathbf{R}^{m \times n}$，定义$\boldsymbol{Y} \in \mathbf{R}^{m \times n}$是一个与之对应的NDCT图像，其中$m$，$n$为图像的宽和高，它们之间的关系可以表述为

$$\boldsymbol{X} = \sigma(\boldsymbol{Y}) \tag{1}$$

式中，$\sigma: \mathbf{R}^{m \times n} \rightarrow \mathbf{R}^{m \times n}$表示LDCT由NDCT图像噪声的映射，涉及量子噪声和其他因素的退化过程，因此去噪问题可以转化为寻求一个函数f，使得

$$f = \underset{f}{\operatorname{argmin}} \left\| f(\boldsymbol{X}) - \boldsymbol{Y} \right\|_2^2 \tag{2}$$

f可以视为σ的最优近似，从LDCT到NDCT之间的映射关系很难用数学公式来表示，导致噪声模型很难确定，而使用深度学习的方法不需要对噪声图像的统计分布有先验知识，可以直接学习映射关系。

与常规医学影像不同，口腔CBCT成像具有显著的解剖性特征差异，其影像中同时包含牙齿、牙槽骨等高密度组织以及牙龈、黏膜等低密度软组织，组织间密度差异显著且边界过渡复杂（韦雪琼和王远军，2021）。同时，牙髓等细微结构对清晰度要求也较高。现有低剂量CT深度学习方法在应用于口腔CBCT时，由于未能充分考虑上述解剖特性，往往在软组织过渡区域产生过度平滑伪影，并导致细微结构细节丢失。

1.1 网络概述

本文提出一种用于口腔低剂量CBCT图像去噪的多尺度注意力编解码网络，整体网络结构如图1所示。该网络采用端到端的可训练架构，由编码器、解码器、MFFM和DAFRM 4部分组成。解码器部分通过多层卷积与池化操作提取图像的多层次特征，逐步降低特征图的空间分辨率并增强语义信息。解码器则通过逐步上采样恢复图像的空间分辨率，并利用残差连接机制（He等，2016）将解码器特征与编码器对应层的特征进行融合，以保留更多细节信息。在编码器和解码器之间的MFFM，用于充分整合不同尺度特征信息，增强网络对多尺度细节和全局信息的建模能力，从而提升特征表达的丰富性。解码器之后的DAFRM通过结合注意力机制对解码器生成特征进行加权调整，着重强化牙齿边缘几何特征及解剖结构的空间方向性表达，确保关键区域细节得以保留，同时抑制非关键区域的噪声干扰。

1.2 多尺度特征融合模块MFFM

口腔内的组织密度和形态差异显著，如高密度的硬组织（如牙齿、牙槽骨）与低密度的软组织（如牙龈、黏膜），这种显著差异性要求网络能够针对不同组织特性进行多尺度的特征提取与融合。对于硬组织，大尺度特征有助于定位其空间位置并勾勒整体形态，而小尺度特征则能够刻画其精细边缘及内部结构；对于软组织，适中的尺度特征则更适合捕捉其柔和边界及内部纹理变化，从而在去噪过程中更准确地还原其真实形态。

本文在编解码器之间采用MFFM来捕捉和融合

不同尺度的特征信息。其主要由自注意力模块及多尺度卷积两部分组成。其中，自注意力模块用于捕获输入特征图的全局依赖关系，通过1×1的卷积操作计算输入特征图的$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$来生成注意力图，并基于注意力图对特征图进行重建。最后通过残差连接将输入特征图与重建特征图融合，保留原始信息并增强特征表达能力。自注意力模块可表示为

$$\boldsymbol{Y}=\boldsymbol{X}+\gamma\cdot\mathrm{softmax}\left(\frac{\boldsymbol{Q}\cdot\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d}}\right)\cdot\boldsymbol{V} \tag{3}$$

式中，$\boldsymbol{Y}\in\mathbf{R}^{C\times H\times W}$为输出特征图，$\boldsymbol{X}\in\mathbf{R}^{C\times H\times W}$为输入特征图，$C$、$H$和$W$分别为通道数以及特征图的高和宽，$\gamma$为用于调节自注意力输出对原始输入的影响程度的缩放因子，$\boldsymbol{Q}$、$\boldsymbol{K}$和$\boldsymbol{V}$分别通过1×1卷积生成，d为特征维度。

多尺度卷积操作通过不同尺寸的卷积核提取输入特征图的不同尺度特征，并将这些特征进行融合，最后通过批归一化和ReLU（rectified linear unit）激活函数来生成输出特征图，多尺度卷积特征融合操作可描述为

$$\boldsymbol{F}=ReLU(BN(C_{1\times1}\boldsymbol{Y}+C_{3\times3}\boldsymbol{Y}+C_{5\times5}\boldsymbol{Y})) \tag{4}$$

式中，$\boldsymbol{F}$为输出特征图，$ReLU$为线性整流单元，$\boldsymbol{Y}$为输入特征图，BN为批量归一化操作（batch normalization，BN），$C_{1\times1}$、$C_{3\times3}$和$C_{5\times5}$分别为卷积核大小为1×1、3×3和5×5的卷积操作。

整个MFFM模块的输出为$\boldsymbol{Z}=ReLU(C_{3\times3}\boldsymbol{F})$，$\boldsymbol{Z}\in\mathbf{R}^{C\times H\times W}$。

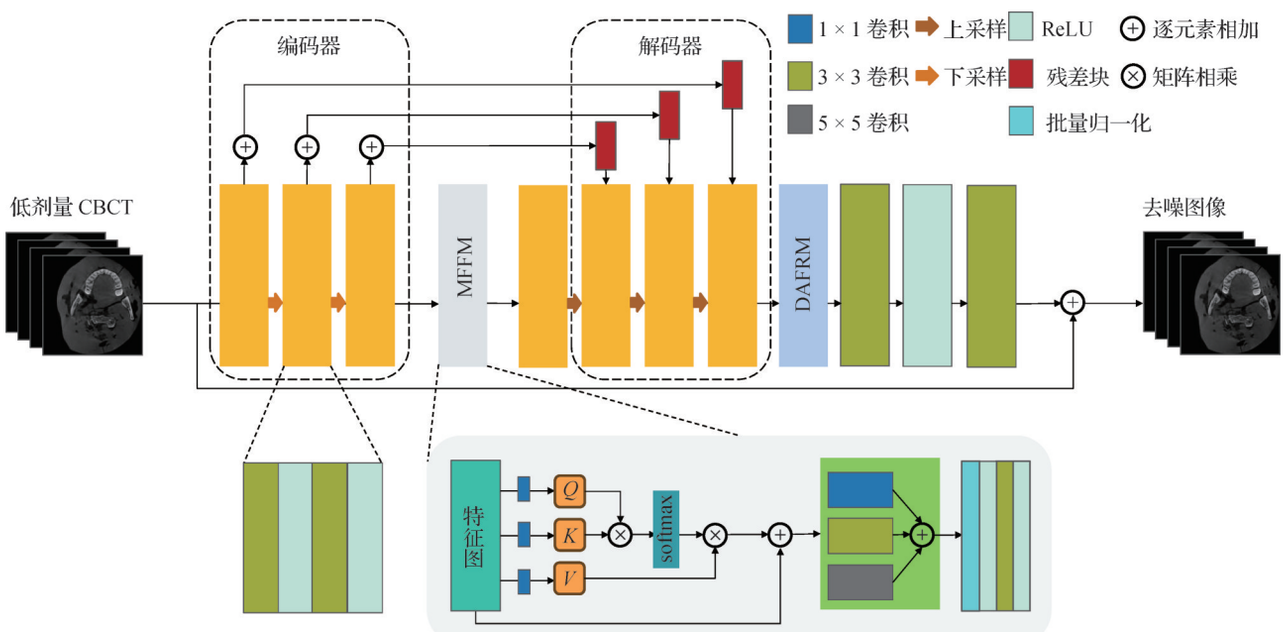


图1　网络架构

Fig. 1　Structure of network

1.3　方向注意力特征细化模块DAFRM

由于图像信息在逐层解码过程中会导致特征图中重要细节丢失以及结构特征模糊化，尤其是在口腔这一复杂的解剖结构中，解码器需要特别关注其明显的空间方向性特征，如牙齿的长轴方向和牙槽骨的弧度方向等。基于这一特性，本文在解码器部分设计一个DAFRM，以针对性地聚合和调整特征，其结构如图2所示。DAFRM借鉴坐标注意力的思想，通过捕捉输入特征图的空间方向性信息，生成方向感知特征图，并通过对分组特征图提取局部特征，共同构建动态权重图。通过调整特征图中的权重分布，使重建的特征图在不同方向上都能得到充分表达，从而更好地保留图像中的重要细节信息和结构特征。整个过程可以分为特征融合以及注意力权重构建两个阶段。

在特征融合阶段，沿特征空间水平方向和垂直方向分别对每个通道进行编码，增强对方向特征的提取与有效利用。首先，对解码器部分输出的特征图$\boldsymbol{X}\in\mathbf{R}^{C\times H\times W}$进行分组，得到分组特征图$\boldsymbol{X}_g=\mathbf{R}^{C_1\times H\times W}$，$C_1=C/g$，$g$为通道分组因子。对分组特征图$\boldsymbol{X}_g$沿宽度维度进行自适应平均池化操作，生成高度为$h$的第$c$个通道的输出特征可以表示为$\boldsymbol{F}_c(h)=$

$\frac{1}{W}\sum_{w=0}^{W-1}x_c(h,w)$；相应地，沿垂直方向进行自适应平均池化操作，生成宽度为w的第c个通道的输出特征可以表示为$\boldsymbol{F}_c(w)=\frac{1}{H}\sum_{h=0}^{H-1}x_c(h,w)$。其中，$x_c(h,w)$表示输入特征图第$c$个通道在位置$(h,w)$的值。将水平方向特征$\boldsymbol{F}_h\in\mathbf{R}^{C_1\times H\times 1}$和垂直方向特征$\boldsymbol{F}_w\in\mathbf{R}^{C_1\times 1\times W}$在空间维度上拼接，生成融合特征$\boldsymbol{T}\in\mathbf{R}^{C_1\times(H+W)\times 1}$。

在注意力权重构建阶段，将融合特征$\boldsymbol{T}$沿空间维度拆分为水平方向特征$\boldsymbol{T}_h\in\mathbf{R}^{C_1\times H\times 1}$和垂直方向特征$\boldsymbol{T}_w\in\mathbf{R}^{C_1\times 1\times W}$，结合sigmoid激活函数生成注意力权重图$\boldsymbol{y}_h$和$\boldsymbol{y}_w$。将$\boldsymbol{y}_h$和$\boldsymbol{y}_w$分别沿宽度和高度维度扩展到与输入特征图$\boldsymbol{X}_g$形状一致，从而生成融合水平和垂直信息的注意力图。再将输入特征图与生成的注意力图进行逐元素相乘，得到输出特征图$\boldsymbol{Z}_g$。

当仅依赖水平与垂直方向的注意力权重图开展卷积操作时，其关注焦点往往侧重于通道间的关联，在空间维度上存在较大的局限性。这种局限性导致其难以有效捕捉到图像中诸如口腔牙齿这类具有复杂局部结构的细节信息。为了能够充分结合水平、垂直方向特征以及局部复杂空间特征，在注意力构建阶段对输入分组特征图$\boldsymbol{X}_g$以及$\boldsymbol{Z}_g$共同构建注意力权重，注意力权重计算为

$$\boldsymbol{W}=\varphi\left(\delta(\boldsymbol{Z}_g)\times[\boldsymbol{X}_g^3,\boldsymbol{X}_g^5]+\delta(\boldsymbol{X}_g^3)\times\boldsymbol{Z}_g+\delta(\boldsymbol{X}_g^5)\times\boldsymbol{Z}_g\right)\tag{5}$$

式中，$\boldsymbol{W}\in\mathbf{R}^{1\times H\times W}$，$\varphi(\cdot)$为动态权重函数，通过对3组特征图进行$1\times 1$卷积实现，$\delta(\cdot)$为非线性变换函数，$\boldsymbol{X}_g^3$，$\boldsymbol{X}_g^5$分别为对分组特征图$\boldsymbol{Z}_g$进行$3\times 3$卷积以及$5\times 5$卷积操作得到的特征图，$[\cdot,\cdot]$表示空间维度上的拼接操作。

整个DAFRM的输出计算为

$$Y_c(i,j)=X_c(i,j)\otimes W_c(i,j)\tag{6}$$

式中，$X_c(i,j)$表示输入分组特征图$\boldsymbol{X}_g$第c个通道在(i,j)位置的值，$\otimes$表示逐元素相乘，$W_c(i,j)$表示经过sigmoid操作后第c个通道在(i,j)位置生成的注意力权重。

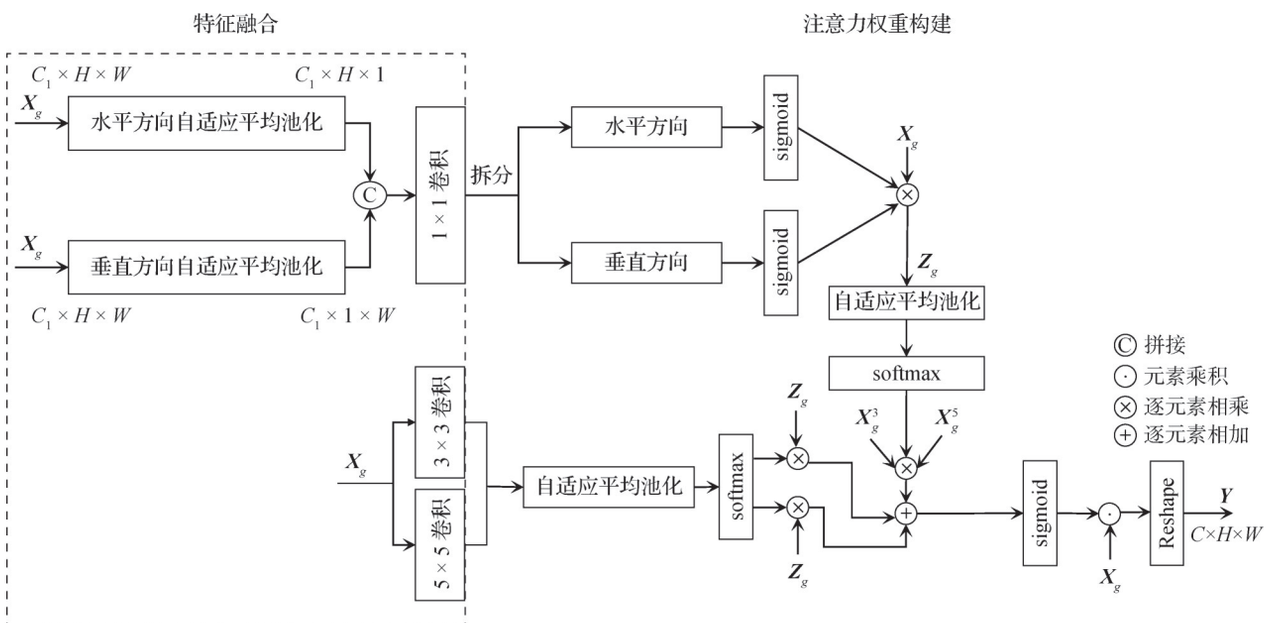


图2　DAFRM结构

Fig. 2　Structure of DAFRM

1.4　损失函数

基于人类视觉系统对边缘和结构信息的敏感性，本文通过引入平滑损失和结构相似性损失，显著提升模型对图像细节和边缘特征的保留能力。所设计的联合损失函数从像素级精度、边缘清晰度和结构一致性3个维度协同优化去噪效果。

像素损失用于衡量预测图像与真实图像像素之间的差异，通常可以通过真实值y与预测值$\hat{y}$之间的均方误差来定义，具体为

$$L_p = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{7}$$

式中，N为图像的像素数量，y_i为第i个像素位置上的真实值，$\hat{y}_i$为第i个像素位置上的预测值。

平滑损失更关注于图像的空间连续性和平滑度，它惩罚预测图像中像素值的突然变化，使得预测图像具有更平滑的外观。计算方法为

$$L_s = \frac{1}{N}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(\left\|\nabla_x\hat{y}(i,j)\right\|_2^2 + \left\|\nabla_y\hat{y}(i,j)\right\|_2^2\right) \tag{8}$$

式中，N为图像的像素数量，H, W为图像的高和宽，$\nabla_x\hat{y}(i,j)$和$\nabla_y\hat{y}(i,j)$分别是预测图像在位置(i,j)处沿水平方向和垂直方向的梯度。

无论是像素损失还是平滑损失，通常基于真实值与预测值之间的均方误差定义。其对异常值敏感，且仅衡量像素级差异，无法反映人类视觉系统的感知特性。相比之下，结构相似性指数（structural similarity index measurement，SSIM）（Horé 和 Ziou，2010）从亮度、对比度和结构3个维度评估图像相似性，更符合人类视觉感知特性。结构相似性损失函数计算为

$$L_{ss} = 1 - \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_x\sigma_y + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)} \tag{9}$$

式中，$\boldsymbol{x}$, $\boldsymbol{y}$分别为预测图像与真实图像，μ_x和μ_y分别为$\boldsymbol{x}$和$\boldsymbol{y}$的均值，σ_x和σ_y分别为$\boldsymbol{x}$和$\boldsymbol{y}$的标准差，σ_{xy}是$\boldsymbol{x}$和$\boldsymbol{y}$的协方差，C_1、C_2是常数，避免分母为0。因此，本文联合损失函数定义为

$$L = \alpha L_p + \beta L_s + \gamma L_{ss} \tag{10}$$

式中，α, β, γ分别为像素损失、平滑损失和结构相似性损失的预定义权重，经过多次实验结果对比，取$\alpha = 0.2, \beta = 0.7, \gamma = 0.1$。

2 实　验

2.1 实验数据集与评估方法

本实验中数据来源于上海市第九人民医院，包括不同X射线剂量下，切片厚度为0.2 mm的口腔CBCT图像。图像涵盖70 kV 6 mA、75 kV 10 mA、80 kV 10 mA、85 kV 10 mA和100 kV 10 mA这5种剂量下的CBCT图像各500幅。在实验中，选择最低剂量70 kV 6 mA的图像作为LDCT数据，100 kV 10 mA的图像作为NDCT数据，LDCT与NDCT一一对应。为进行模型训练和评估，从数据集中选取450对高低剂量图像对作为训练集，50对高低剂量图像作为测试集。

实验中采用低剂量CT图像质量评估常用的3个评价指标：PSNR、SSIM和均方根误差（root mean squared error，RMSE）。这3个评价指标从各个方面定性定量地分析去噪后图像的质量。

2.2 实验设置细节

本文实验均于Python 3.9，PyTorch 1.12.1，CUDA11.6实验环境下运行。为公平起见，所有网络模型的学习率设置为0.000 1，采用Adam优化器进行训练，实验均在24 G的NVIDIA RTX 3090 Ti上完成。

2.3 实验结果

2.3.1 视觉效果分析

将7种不同的低剂量CT图像去噪方法与本文方法进行比较，包括EDCNN（edge enhancement-based densely connected convolutional neural network）（Liang等，2020）、UNAD（universal anatomy-initialized noise distribution learning framework）（Gu等，2024）、CNN_OCT（octave convolution）（Won等，2020）、U-Net（Ronneberger等，2015）、RED-CNN（residual encoder-decoder convolutional neural network）（Chen等，2017）、WGAN（Yang等，2018）和TED-Net（Wang等，2021）。上述方法都是对2016-NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge公开数据集进行的评估。

为保证对比结果的公平性，未对本文之外的初始模型参数作任何修改，使得各方法在口腔低剂量CBCT图像去噪上具有相对公平的比较基准。

图3和图4分别展示了不同模型对口腔低剂量CBCT图像数据集中有代表性的两幅切片的去噪结果。在图3中选取牙齿整体放大区域为感兴趣区域（region of interset，ROI），并分别用黄色框和红色框标记出ROI 1和ROI 2。为方便进一步详细观察，图3底部提供了ROI 1和ROI 2的放大图像，并在图5展示了不同区域评价指标的柱状图可视化，以便于直观地比较各模型的性能。通过观察图中牙齿整体区域放大图以及箭头所指牙髓区域，可以看出各模型在抑制图像噪声和伪影方面的不同表现。尽管各模型在肺部低剂量CT去噪中都有其独特的优势，但当应用于口腔低剂量CBCT图像时，却面临诸多挑战。

特别地，EDCNN在设计时更关注传统图像处理中易被忽略的图像边缘信息，但在处理口腔低剂量CBCT图像时，这种对边缘的过度关注反而导致牙齿连接区域过度平滑，一定程度上影响图像的真实性，如图3(b)和图4(b)所示。UNAD虽然有效去除噪声，但相较于肺部低剂量CT去噪的优异表现，处理后所得到的口腔低剂量CBCT图像太过模糊，出现部分细节丢失现象，整体图像质量较差，如图3(c)和图4(c)所示。RED-CNN网络所得图像尽管在某种程度上去除噪声，但部分区域出现结构变形，影响图像的准确性，如图3(f)所示。如图3(d)和图3(e)所示，CNN_OCT和U-Net在消除大部分噪声和伪影的同时，较好地保留牙齿区域的结构特征，但是生成的图像在视觉呈现以及评价指标值上仍存在一定不足。

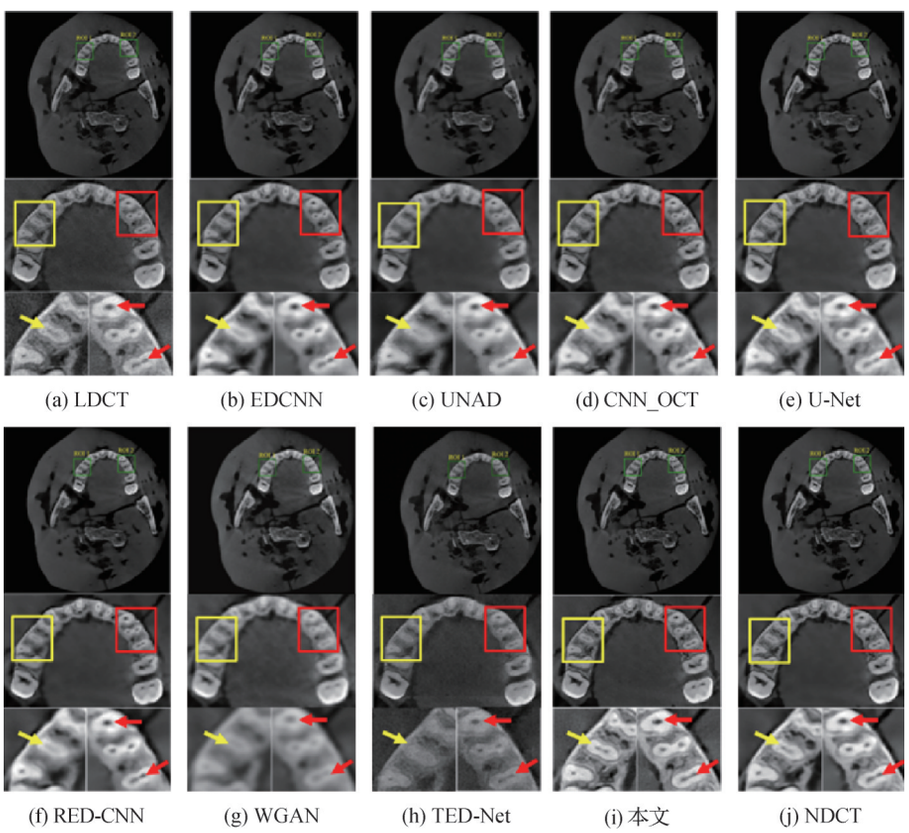


图3　不同模型对切片1的去噪结果

Fig. 3　Visualization of denoising results of slice 1 by different models

((a) LDCT;(b) EDCNN;(c) UNAD;(d)CNN_OCT;(e)U-Net;(f)RED-CNN;(g)WGAN;(h)TED-Net;(i)ours;(j)NDCT)

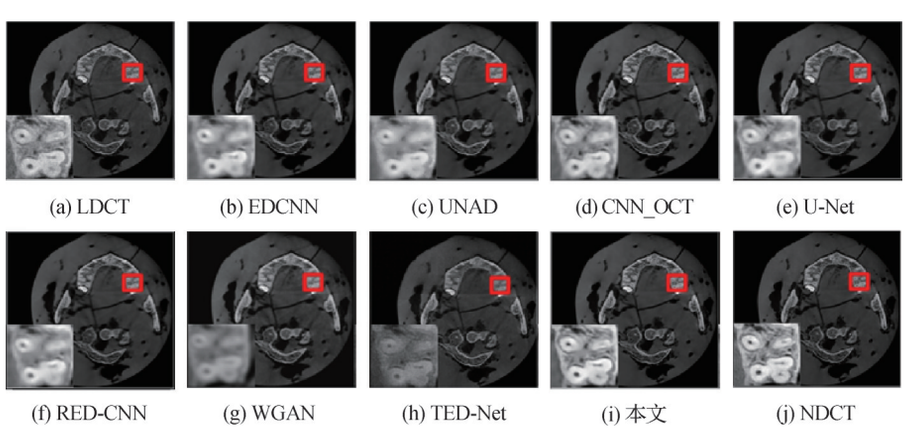


图4　不同模型对切片2的去噪结果

Fig. 4　Visualization of denoising results of slice 2 by different models

((a) LDCT;(b) EDCNN;(c) UNAD;(d)CNN_OCT;(e)U-Net;(f)RED-CNN;(g)WGAN;(h)TED-Net;(i)ours;(j)NDCT)

此外，由于口腔CBCT图像数据集相对较小，WGAN和TED-Net难以有效捕捉局部特征信息，导致所生成的去噪图像质量较差，如图3(g)和图3(h)所示。在深入分析不同网络模型去噪结果后发现，它们所生成的口腔低剂量CT图像普遍存在过度平滑现象。相比之下，本文方法在保留细节的同时有效去除噪声和伪影，同时，生成的图像更符合人眼视觉感知，与NDCT图像相似度最高，如图3(i)和图4(i)所示，为后续牙齿分割(秦俊 等，2023)以及疾病诊断提供可靠的影像数据基础。

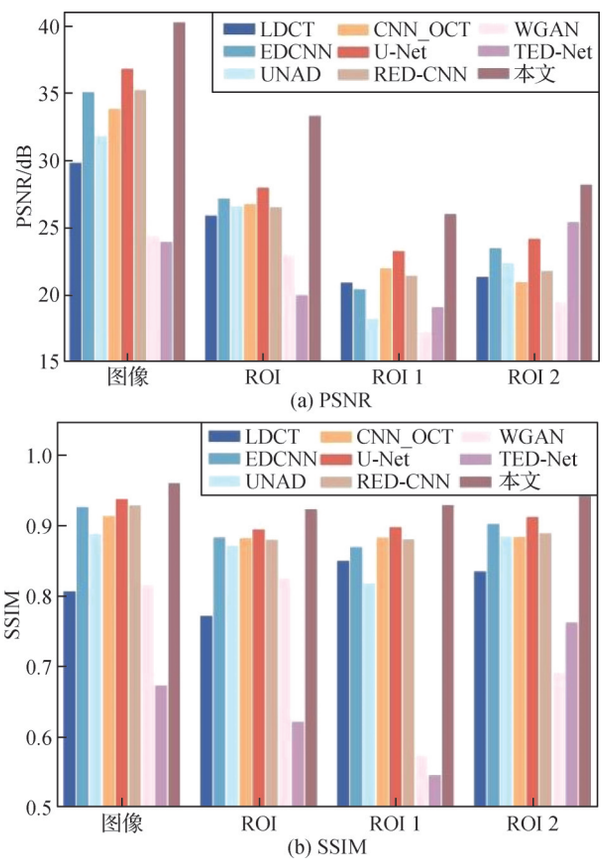


图5　不同模型对切片1不同区域评价指标的柱状图可视化

Fig. 5　Different models visualized the histograms of the evaluation indicators of different regions of slice 1 ((a) PSNR; (b) SSIM)

为了全面评估本文模型在不同成像条件下的泛化能力，将本文方法在70 kV 6 mA、75 kV 10 mA、80 kV 10 mA和85 kV 10 mA不同剂量下的口腔CBCT数据集上展开测试。图6直观地展示在不同剂量条件下，经模型处理前后的CBCT图像对比，图中红色箭头所指区域为牙髓区域。

从图6可以看出，在最低剂量70 kV 6 mA条件下，原始图像存在明显的噪声和伪影，牙髓腔呈现模糊化特征，牙体—牙周组织边界分辨不清。然而，经本文模型去噪处理后，各剂量下的CBCT图像质量均得到显著提升，这充分表明了本文模型对不同剂量条件下CBCT图像进行去噪处理的有效性。

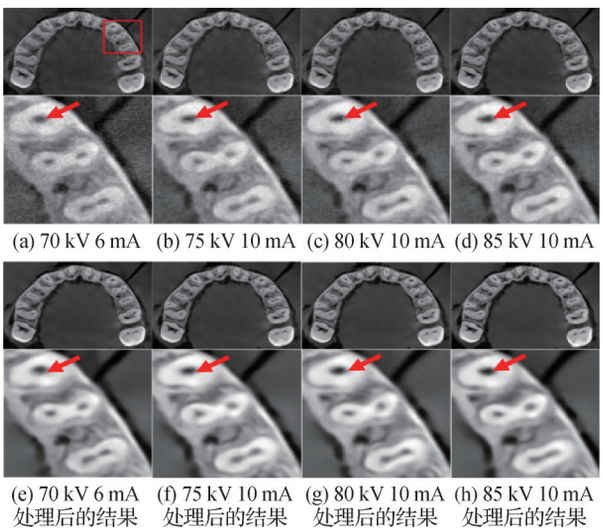


图6　不同剂量CBCT去噪结果可视化

Fig. 6　Visualization of denoising results ((a) 70 kV 6 mA; (b) 75 kV 10 mA; (c) 80 kV 10 mA; (d) 85 kV 10 mA; (e) result of 70 kV 6 mA; (f) result of 75 kV 10 mA; (g) result of 80 kV 10 mA; (h) result of 85 kV 10 mA)

2.3.2　定量结果分析

对各模型在低剂量口腔CBCT数据集上的去噪结果进行定量分析，如表1所示，列出了不同网络模型去噪后图像的PSNR、RMSE和SSIM指标的平均值及标准差。评价指标中，较高的PSNR和SSIM值以及较低的RMSE值均表明降噪图像的质量更优。表1的实验结果表明，本文模型在降噪性能上显著优于其他对比模型。与降噪前的LDCT图像相比，本文模型的PSNR指标提升8.890 6 dB，SSIM指标提升0.157 4。与次优模型相比，PSNR和SSIM指标分别提升2.524 0 dB和0.025 6。

此外，为了直观展示不同模型性能差异，图7通过箱线图呈现了各对比模型在PSNR和SSIM两个客观评价指标上的分布特征。从统计学角度分析，箱体长短反映数据离散程度，长度越短则表明数据分布越集中。从图7可以看出，本文模型PSNR指标值以37.0～38.5 dB的核心分布区间覆盖高信噪比区域，箱体上沿突破40.6 dB，较次优模型整体上移3.1 dB且75%的数据点高于37.0 dB。此外，在

表 1　不同模型降噪结果图像评价指标

Table 1　Image evaluation indexes of noise reduction results of different models

模型	PSNR/dB ↑	RMSE ↓	SSIM ↑
LDCT	29.512 4 ± 0.215 0	0.033 5 ± 0.000 8	0.795 0 ± 0.007 3
EDCNN	34.618 9 ± 0.359 8	0.018 6 ± 0.000 8	0.916 9 ± 0.006 0
UNAD	31.443 0 ± 0.275 6	0.026 8 ± 0.000 9	0.876 6 ± 0.007 3
CNN_OCT	33.555 5 ± 0.283 0	0.021 0 ± 0.000 7	0.904 5 ± 0.006 0
U-Net	35.879 0 ± 0.532 8	0.016 1 ± 0.001 0	0.926 8 ± 0.006 3
RED-CNN	34.863 0 ± 0.365 1	0.018 1 ± 0.000 8	0.920 1 ± 0.005 5
WGAN	24.412 1 ± 0.077 6	0.060 2 ± 0.000 5	0.800 8 ± 0.009 8
TED-Net	24.050 6 ± 0.081 5	0.062 7 ± 0.000 6	0.654 8 ± 0.001 0
本文	**38.403 0 ± 0.913 9**	**0.012 1 ± 0.001 2**	**0.952 4 ± 0.005 0**

注：加粗字体表示各列最优结果，↑表示数值越大指标越好；↓表示数值越小指标越好。

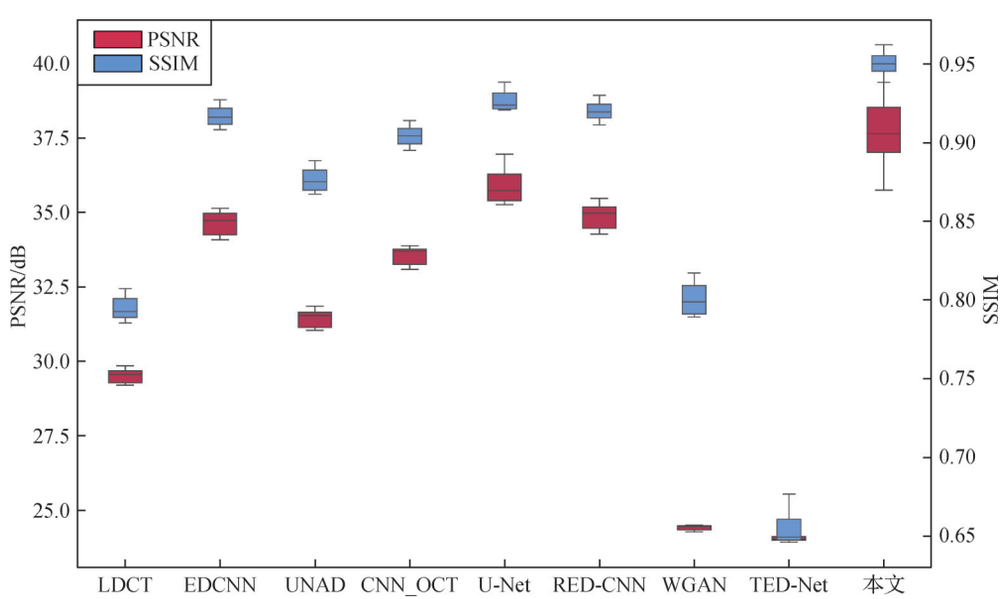


图 7　不同模型降噪结果评价指标箱线图

Fig. 7　Boxplots of evaluation indicators of noise reduction results of different models

SSIM 指标维度上，本文模型以 0.938～0.962 的无离群紧凑分布占据绝对优势，其指标稳定居于 0.938 以上。上述这些统计学特征进一步验证了本模型在噪声抑制和结构保真度的优异性能。

为了系统验证模型在不同剂量条件下的鲁棒性，表 2 对比了 75 kV 10 mA 至 85 kV 10 mA 区间内各剂量组别的客观评价指标。实验数据显示，在 75 kV 10 mA 条件下，本文模型的 PSNR 指标值由原始图像的 29.167 2 dB 提升至 35.466 4 dB，SSIM 指标值从 0.829 1 提升至 0.936 2，对于 80 kV 10 mA 和 85 kV 10 mA 剂量条件下的 CBCT 图像，去噪后的各项指标也呈现明显改善。通过对比分析可以发现，无论在低剂量还是较高剂量条件下，本文模型均能有效提高图像的 PSNR 和 SSIM 值，同时降低 RMSE 值，证明本文模型在不同剂量条件下的良好泛化能力。

2.4　消融实验

与基于传统的编码—解码器架构的降噪网络相比，本文在其基础上增加多尺度特征融合和方向注意力特征细化模块，并设计组合损失函数来约束网络。为探索不同因素对网络的影响，设计了消融实验，包括网络模块消融实验和损失函数消融实验两

表2　模型处理前后不同剂量图像评价指标

Table 2　Evaluation indexes of different dose images before and after model treatment

处理前后	剂量	PSNR/dB ↑	RMSE ↓	SSIM ↑
处理前	75 kV 10 mA	29.167 2 ± 0.160 5	0.034 8 ± 0.000 6	0.829 1 ± 0.007 6
	80 kV 10 mA	30.217 8 ± 0.153 4	0.030 8 ± 0.000 5	**0.852 3 ± 0.006 5**
	85 kV 10 mA	**30.428 6 ± 0.269 5**	**0.030 1 ± 0.000 9**	0.845 3 ± 0.007 6
处理后	75 kV 10 mA	35.466 4 ± 0.615 1	0.016 9 ± 0.001 2	0.936 2 ± 0.006 2
	80 kV 10 mA	34.656 5 ± 0.502 1	0.018 5 ± 0.001 1	0.923 8 ± 0.006 7
	85 kV 10 mA	**36.194 6 ± 0.896 3**	**0.0156 ± 0.001 6**	**0.938 1 ± 0.007 3**

注：加粗字体表示各项指标在模型处理前后的最优结果。↑表示数值越大指标越好；↓表示数值越小指标越好。

部分。表3为消融实验定量结果，图8对模块消融实验以及损失函数消融实验结果进行可视化对比展示。网络模块消融实验讨论MFFM与DAFRM的有效性，其中M1表示未添加任何模块的编解码基线模型，M2表示添加MFFM，M3表示添加DAFRM，M4表示添加两个模块，即为本文方法。损失函数消融实验部分分别评估单独使用与联合使用联合损失函数对网络去噪结果的影响，L_p表示单独使用像素损失，L_{ps}表示同时使用像素损失以及平滑损失，L_{pss}表示使用像素损失、平滑损失和结构相似性损失作为最终损失函数。

表3　消融实验结果

Table 3　Experimental results of ablation

模型	PSNR/dB ↑	RMSE ↓	SSIM ↑
LDCT	29.512 4	0.033 5	0.795 0
M1 + L_{pss}	33.806 6	0.020 4	0.843 7
M2 + L_{pss}	35.439 0	0.016 9	0.923 2
M3 + L_{pss}	35.603 4	0.016 6	0.929 8
M4 + L_p	34.657 3	0.018 5	0.909 2
M4 + L_{ps}	35.716 2	0.016 4	0.925 3
M4 + L_{pss}（本文）	**38.403 0**	**0.012 1**	**0.952 4**

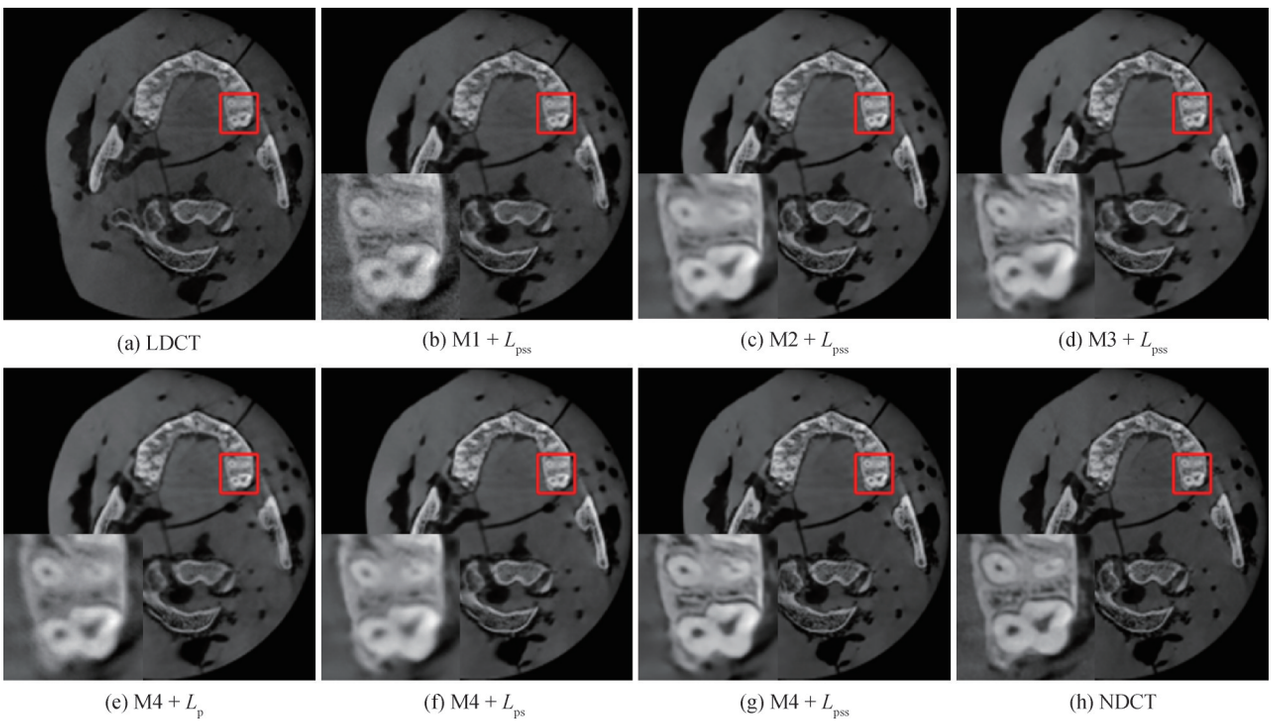
注：加粗字体表示各列最优结果。

从表3可以看出，本文设计的各个模块以及添加的损失函数对去噪后图像的评价指标均有不同程度的贡献。当单独引入MFFM和DAFRM时，图像的评价指标均有显著提升，其中，SSIM指标分别提高9.42%和10.21%。如图8(c)所示，MFFM通过多尺度特征融合有效捕捉牙齿边缘和根管等微小结构，提升细节保留能力。DAFRM通过方向感知机制增强牙釉质—牙本质界面及牙髓区域的特征表达，优化对比度和清晰度，如图8(d)所示。此外，平滑损失函数和结构相似性损失函数的加入使SSIM指标较单独使用RMSE损失函数提高4.7%。从表3和图8(g)可以看出，结合MFFM和DAFRM并使用组合损失函数的本文模型在评价指标以及视觉效果上均取得最优结果，充分验证了MFFM与DAFRM的协同作用以及组合损失函数的有效性。

3　结　论

本文提出一种针对口腔低剂量CBCT图像去噪的多尺度注意力编解码网络，旨在解决因口腔牙齿结构特殊性导致的图像过度平滑和结构特征模糊等问题。通过引入MFFM和DAFRM，该模型能够有效捕捉和整合不同尺度的特征信息，并聚焦于关键细节区域，从而在去噪过程中更好地保留牙齿边缘和牙髓等细微结构。与其他算法相比，本文算法去噪后的图像具有较高的客观评价指标以及较好的视觉效果。这为后续的牙髓疾病诊断以及牙齿正畸提供更为准确、可靠的图像依据。

然而，本文仍存在一定的局限性。一方面，使用的全监督学习高度依赖配准良好的高低剂量图像对，而实际临床场景中受患者位移等因素影响，精准配准数据获取难度较大；另一方面，当前实验主要基于单中心单设备数据集，模型在多厂商设备兼容性及复杂病理特征的去噪泛化能力仍需进一步验证。鉴于此，后续研究可重点探索自监督/弱监督学习框架在低剂量口腔CBCT去噪中的应用，通过噪声建模或域适应技术缓解数据依赖问题。同时构建多中



(a) LDCT (b) M1 + L_{pss} (c) M2 + L_{pss} (d) M3 + L_{pss}

(e) M4 + L_{p} (f) M4 + L_{ps} (g) M4 + L_{pss} (h) NDCT

图8　消融实验结果

Fig. 8　Experimental results of ablation

((a) LDCT; (b) M1 + L_{pss}; (c) M2 + L_{pss}; (d) M3 + L_{pss}; (e) M4 + L_{p}; (f) M4 + L_{ps}; (g) M4 + L_{pss}; (h) NDCT)

心多设备的口腔CBCT数据集，系统纳入包含不同病理特征的临床案例，从而增强模型对复杂临床场景的适应性。

参考文献(References)

Aharon M, Elad M and Bruckstein A. 2006. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing, 54(11): 4311-4322 [DOI: 10.1109/TSP.2006.881199]

Brenner D J and Hall E J. 2007. Computed tomography—an increasing source of radiation exposure. New England Journal of Medicine, 357(22): 2277-2284 [DOI: 10.1056/NEJMra072149]

Chen H, Zhang Y, Kalra M K, Lin F, Chen Y, Liao P X, Zhou J L and Wang G. 2017. Low-dose CT with a residual encoder-decoder convolutional neural network. IEEE Transactions on Medical Imaging, 36(12): 2524-2535 [DOI: 10.1109/TMI.2017.2715284]

Diwakar M, Kumar P and Singh A K. 2020. CT image denoising using NLM and its method noise thresholding. Multimedia Tools and Applications, 79(21): 14449-14464 [DOI: 10.1007/s11042-018-6897-1]

Feruglio P F, Vinegoni C, Gros J, Sbarbati A and Weissleder R. 2010. Block matching 3D random noise filtering for absorption optical projection tomography. Physics in Medicine and Biology, 55(18): 5401-5415 [DOI: 10.1088/0031-9155/55/18/009]

Gu L R, Deng W J and Wang G L. 2024. UNAD: universal anatomy-initialized noise distribution learning framework towards low-dose CT denoising//Proceedings of 2024 IEEE International Conference on Acoustics, Speech and Signal Processing. Seoul, Korea(South): IEEE: 1671-1675 [DOI: 10.1109/ICASSP48485.2024.10446919]

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

Horé A and Ziou D. 2010. Image quality metrics: PSNR vs. SSIM//Proceedings of the 20th International Conference on Pattern Recognition. Istanbul, Turkey: IEEE: 2366-2369 [DOI: 10.1109/ICPR.2010.579]

Huang Z Z, Zhang J P, Zhang Y and Shan H M. 2022. DU-GAN: generative adversarial networks with dual-domain U-Net-based discriminators for low-dose CT denoising. IEEE Transactions on Instrumentation and Measurement, 71: #4500512 [DOI: 10.1109/TIM.2021.3128703]

Kim B, Han M, Shim H and Baek J. 2019. A performance comparison of convolutional neural network-based image denoising methods: the effect of loss functions on low-dose CT images. Medical Physics, 46(9): 3906-3923 [DOI: 10.1002/mp.13713]

Liang T F, Jin Y, Li Y D and Wang T. 2020. EDCNN: edge enhancement-based densely connected network with compound loss

for low-dose CT denoising//Proceedings of the 15th IEEE International Conference on Signal Processing (ICSP). Beijing, China: IEEE: 193-198 [DOI: 10.1109/ICSP48669.2020.9320928]

Marcos L, Alirezaie J and Babyn P. 2021. Low dose CT image denoising using boosting attention fusion GAN with perceptual loss//Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Mexico: IEEE: 3407-3410 [DOI: 10.1109/EMBC46164.2021.9630790]

Pan X C, Sidky E Y and Vannier M. 2009. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? Inverse Problems, 25(12): #123009 [DOI: 10.1088/0266-5611/25/12/12300]

Qin J, Lu T L, Ji B and Li Y Q. 2024. Tooth segmentation network for low-dose CT. Journal of Image and Graphics, 29(3): 686-696 (秦俊, 卢婷岚, 纪柏, 李雨晴. 2024. 面向低剂量CT的牙齿分割网络. 中国图象图形学报, 29(3): 686-696) [DOI: 10.11834/jig.230540]

Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]

Wang D Y, Fan F L, Wu Z, Liu R, Wang F and Yu H Y. 2023. CTformer: convolution-free Token2Token dilated vision transformer for low-dose CT denoising. Physics in Medicine and Biology, 68(6): #065012 [DOI: 10.1088/1361-6560/acc000]

Wang D Y, Wu Z and Yu H Y. 2021. TED-Net: convolution-free T2T vision transformer-based encoder-decoder dilation network for low-dose CT denoising//Proceedings of the 12th International Workshop on Machine Learning in Medical Imaging. Strasbourg, France: Springer: 416-425 [DOI: 10.1007/978-3-030-87589-3_43]

Wei X Q and Wang Y J. 2021. Application progresses of cone beam CT in oral and maxillofacial diseases. Chinese Journal of Medical Imaging Technology, 37(4): 603-607 (韦雪琼, 王远军. 2021. 锥形束CT用于口腔颌面疾病进展. 中国医学影像技术, 37(4): 603-607) [DOI: 10.13929/j.issn.1003-3289.2021.04.030]

Willemink M J and Noël P B. 2019. The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence. European Radiology, 29(5): 2185-2195 [DOI: 10.1007/s00330-018-5810-7]

Won D K, An S, Park S H and Ye D H. 2020. Low-dose CT denoising using octave convolution with high and low frequency bands//Proceedings of the 3rd International Workshop on Predictive Intelligence in Medicine. Lima, Peru: Springer: 68-78 [DOI: 10.1007/978-3-030-59354-4_7].

Yang Q S, Yan P K, Zhang Y B, Yu H Y, Shi Y Y, Mou X Q, Kalra M K, Zhang Y, Sun L and Wang G. 2018. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. IEEE Transactions on Medical Imaging, 37(6): 1348-1357 [DOI: 10.1109/TMI.2018.2827462]

You C Y, Yang Q S, Shan H M, Gjesteby L, Li G, Ju S H, Zhang Z Y, Zhao Z, Zhang Y, Cong W X and Wang G. 2018. Structurally-sensitive multi-scale deep neural network for low-dose CT denoising. IEEE Access, 6: 41839-41855 [DOI: 10.1109/ACCESS.2018.2858196]

Zhang H, Ma J H, Wang J, Liu Y, Lu H B and Liang Z R. 2014. Statistical image reconstruction for low-dose CT using nonlocal means-based regularization. Computerized Medical Imaging and Graphics, 38(6): 423-435 [DOI: 10.1016/j.compmedimag.2014.05.002]

Zhao Y, Li J C, Cheng B D, Niu N J, Wang L G, Gao G W and Shi J. 2024. Applications and challenges of deep learning in dental imaging. Journal of Image and Graphics, 29(3): 586-607 (赵阳, 李俊诚, 成博栋, 牛娜君, 王龙光, 高广谓, 施俊. 2024. 深度学习在口腔医学影像中的应用与挑战. 中国图象图形学报, 29(3): 586-607) [DOI: 10.11834/jig.230062]

作者简介

张立, 男, 硕士研究生, 主要研究方向为医学图像处理和深度学习。E-mail: jyouzdl@163.com

陈胜, 通信作者, 男, 副教授, 主要研究方向为医学图像处理、模式识别、计算机辅助诊断。E-mail: chnshn@163.com

华旖筠, 女, 硕士研究生, 主要研究方向为模式识别与智能控制。E-mail: huayiyun_14@163.com