

基于扩散循环一致性生成对抗网络的骨盆活跃骨髓区域分割方法

卓俐¹, 曾敏¹, 谭顺谦¹, 梁涛¹, 肖巍巍², 甄鑫¹

¹南方医科大学生物医学工程学院, 广东 广州 510515; ²中山大学肿瘤防治中心//华南肿瘤学国家重点实验室//肿瘤医学省部共建协同创新中心, 广东 广州 510060

摘要:目的 建立基于扩散循环一致性生成对抗网络的骨盆活跃骨髓(ABM)分割方法,突破传统解剖图谱方法个体化精度不足的技术瓶颈。方法 收集253例患者骨盆PET-CT数据,构建三阶段级联跨模态学习框架实现从CT到个体化ABM精准识别。首先通过循环一致性生成对抗网络建立CT-PET双向映射,采用9个残差模块学习跨模态特征关系。设计条件扩散模块基于1000步马尔可夫链实现渐进去噪,融合双向交叉注意力机制动态整合解剖与功能信息。最后构建多尺度渐进式特征金字塔分割网络,在4个尺度层级累积多模态特征实现ABM区域分割。采用峰值信噪比(PSNR)、结构相似性指数(SSIM)、归一化均方误差(NMSE)评估图像合成质量,Dice相似系数(DSC)、平均对称表面距离(ASSD)评估分割性能。结果 本方法均优于现有方法,PSNR达到 26.42 ± 0.63 dB,SSIM达到 0.894 ± 0.011 ,NMSE降至 0.0235 ± 0.0026 。在ABM分割任务中,平均Dice系数达到 0.777 ± 0.023 ,ASSD降至 3.52 ± 0.41 mm。结论 与传统方法相比,该方法显著提高了个体化分割精度,适用于直肠癌个体化骨髓保护放疗的临床应用。

关键词:直肠癌;活跃骨髓;扩散模型;生成对抗网络;图像分割

Diffusion cycle-consistent generative adversarial networks for pelvic active bone marrow segmentation

ZHUO Li¹, ZENG Min¹, TAN Shunqian¹, LIANG Tao¹, XIAO Weiwei², ZHEN Xin¹

¹School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China; ²Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China

Abstract: Objective To establish a pelvic active bone marrow (ABM) segmentation method based on diffusion cycle-consistent generative adversarial networks for improving individualized precision of conventional anatomical atlas-based methods. **Methods** We collected pelvic PET-CT data from 253 patients and constructed a 3-stage cascaded cross-modal learning framework for precise individualized ABM identification from CT images. The framework used cycle-consistent generative adversarial networks for bidirectional CT-PET mapping, conditional diffusion modules with 1000-step Markov chains for progressive denoising, and multi-scale progressive feature pyramid fusion networks for segmentation. The peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), normalized mean square error (NMSE), Dice similarity coefficient (DSC), and average symmetric surface distance (ASSD) were used for evaluation of the model performance for ABM segmentation. **Results** The proposed method outperformed the existing methods with a PSNR of 26.42 ± 0.63 dB, an SSIM of 0.894 ± 0.011 , and an NMSE of 0.0235 ± 0.0026 . For ABM segmentation, the average Dice coefficient of the model reached 0.777 ± 0.023 with an ASSD of 3.52 ± 0.41 mm. **Conclusion** Compared with the conventional methods, the propose method significantly improves individualized segmentation accuracy of the ABM and is thus suitable use in individualized bone marrow protection radiotherapy for rectal cancer.

Keywords: rectal cancer; active bone marrow; diffusion models; generative adversarial networks; image segmentation

局部进展期直肠癌是消化系统常见恶性肿瘤,全球每年新发约80万例,其中50%患者初诊时已属局部晚期^[1]。全程新辅助治疗(TNT)因能显著提高病理完全缓解率、降低远处转移风险而广泛应用^[2,3],但其3~4级急性血液学毒性发生率高达20%~30%,较常规治疗增加近3

倍^[4,5]。骨髓抑制作为主要剂量限制性毒性,导致约25%患者需中断治疗或减量,直接影响疗效^[6]。由于成人活跃造血骨髓50%以上分布在骨盆区域^[7],盆腔放疗成为骨髓抑制的主因,因此,精准识别和保护骨盆活跃骨髓(ABM)区域对于降低放疗相关骨髓毒性具有重要的临床意义。

正电子发射断层扫描(PET)基于¹⁸F-FDG示踪剂检测组织代谢活性,被认为是识别ABM区域的“金标准”^[8,9]。活跃造血骨髓因其高糖代谢需求,标准化摄取值(SUV)显著高于非活跃骨髓^[10],为个体化骨髓保护提供了可靠的功能成像基础。然而,在临床实践中基于PET的ABM识别面临的核心挑战是个体化精准度不足:传统基于解剖图谱配准的ABM识别方法仅依赖人

收稿日期:2025-06-13

基金项目:国家自然科学基金(82572381,82573772);国家自然科学基金青年基金(62106058);广东省自然科学基金(2024A1515012100) Supported by National Natural Science Foundation of China (82572381, 82573772); Supported by Natural Science Foundation for the Youth (NSFY) of China (62106058).

作者简介:卓俐,在读硕士研究生,E-mail: zhuoli0901@163.com

通信作者:甄鑫,博士,教授,E-mail: xinzhen@smu.edu.cn

群统计学数据,平均识别精度(Dice系数)仅为0.65~0.70,难以准确反映个体差异和病理状态下的骨髓分布变化^[10]。此外,PET图像空间分辨率相对较低(4 mm~6 mm),在精细解剖结构识别方面存在局限,影响放疗靶区的精确勾画^[11]。对于需要紧急开始放疗的患者,PET检查的预约和执行周期也可能延长治疗启动时间。

计算机断层扫描(CT)作为放疗计划的必备检查,具有优异的空间分辨率(<1 mm)和解剖结构显示能力^[12]。关键的科学问题在于:活跃骨髓的解剖特征与代谢功能间存在可建模的生物学关联。活跃造血骨髓具有高细胞密度和丰富的血管网络,这些生物物理特性在CT图像中表现为特定的密度分布模式,这些区域的骨骼微结构、血管密度和细胞组成与其代谢活性密切相关^[13-15]。造血细胞的高糖代谢需求使得18F-FDG摄取率与骨髓细胞密度呈强正相关,为建立CT-PET映射关系提供了生物学基础。

深度学习技术为建立这种解剖-功能映射关系提供了强有力的工具,通过学习大规模配对数据中CT特征与PET功能信息的复杂非线性关系,实现从单一CT图像推断ABM分布的跨模态预测。近年来,深度学习在医学图像跨模态合成领域取得重要进展,生成对抗网络(GAN)在医学图像转换中展现显著潜力^[16],Salehjahromi等基于GAN的CT到PET合成方法在肺癌诊断中达到了与真实PET相当的诊断准确率^[17],为跨模态医学图像合成的临床应用提供了重要验证。然而,骨盆区域复杂的解剖结构和个体代谢异质性对传统GAN方法的精度和稳定性提出更高要求。扩散模型作为新兴生成模型范式^[18],通过逐步去噪的迭代过程生成图像,在保持图像细节和解剖结构完整性方面表现出色。条件扩散模型能够将CT提供的解剖先验知识融入生成过程,显著提高合成PET图像的生物学真实性和诊断可靠性^[19]。结合循环一致性约束的双向映射机制,可确保跨模态转换过程中解剖信息的完整保持。

更重要的是,精确的ABM区域分割是实现个体化骨髓保护放疗的关键环节。传统方法的核心局限在于无法充分利用多模态信息的互补优势:CT图像提供高精度解剖结构,PET图像(或合成PET)提供功能代谢信息,两者的有效融合能够实现比单一模态更精确的ABM识别。基于深度学习的多模态分割网络通过自动学习不同模态间的关联模式,能够实现个体化的ABM区域精确识别,为制定最优的骨髓保护策略提供技术支持。

基于上述科学问题与临床需求,本文提出基于扩散循环一致性生成对抗网络的骨盆ABM分割方法。该方法通过三个层次的技术创新,构建了从CT图像到ABM精准分割的完整技术链条。在跨模态生成层面,创新性地融合扩散概率模型与循环一致性约束,实现了高保真且稳定的CT-PET跨模态映射;在特征融合层面,首创

双向交叉注意力条件扩散机制,通过动态权重自适应实现解剖与功能信息的深度融合;在精准分割层面,设计多尺度渐进式特征金字塔融合分割网络,实现从全局到局部的多尺度ABM区域精确识别。该方法有效提升了ABM识别的精确度和个体化水平,为制定精准的骨髓保护放疗策略提供了重要的技术基础。

1 材料和方法

1.1 基于扩散循环一致性生成对抗网络的骨盆活跃骨髓区分割研究

本文提出的基于扩散循环一致性生成对抗网络的骨盆活跃骨髓区分割方法构建了端到端的深度学习技术框架(图1)。该框架的核心创新在于提出了三阶段级联的跨模态学习范式,通过循环一致性约束、条件扩散生成和多尺度特征融合的协同机制,实现从单模态CT图像到高精度ABM区域语义分割的映射。系统架构首先通过循环一致性生成对抗网络建立CT与PET图像域间的双向同胚映射 $f: X_{CT} \leftrightarrow X_{PET}$,确保跨模态转换中解剖结构信息的完整保持,随后采用条件扩散概率模型 $p_{\theta}(x_{PET}^{opt}|x_{CT}, x_{PET}^{pseudo})$ 对合成PET到优化PET的映射关系进行建模,以CT作为条件约束,通过马尔可夫链蒙特卡洛采样从高斯噪声中逐步重构目标域特征表示,有效缓解传统生成对抗网络的模式崩塌问题,最终设计多尺度渐进式特征金字塔融合分割网络在不同空间分辨率层级实现CT解剖信息与合成PET功能信息的互补融合。

算法流程的技术创新体现在3个协同优化阶段:模态对齐阶段利用对抗学习机制训练双向生成器 $G_{CT \rightarrow PET}$ 和 $G_{PET \rightarrow CT}$,通过循环一致性损失约束映射的双射性,建立跨模态特征空间的拓扑等价关系;条件生成阶段基于变分推断框架,采用条件扩散模型对复杂的条件概率分布 $p(x_{PET}|x_{CT})$ 进行参数化建模,通过反向扩散过程 $x_{t-1} = \mu_{\theta}(x_t, t, c_{CT}) + \sigma_t \epsilon$ 实现高保真的跨模态图像合成;语义分割阶段构建多尺度特征金字塔网络,采用自底向上的特征聚合策略,通过残差连接实现不同模态特征的深度融合,最终输出像素级的ABM区域分割掩码。该技术路径的核心优势在于通过概率建模框架有效处理了跨模态转换中的不确定性问题,利用条件信息指导的扩散过程提升合成图像的生物学一致性,同时通过端到端的联合优化确保从图像合成到语义分割的全流程性能最优化。

1.2 循环一致性生成对抗网络模块

循环一致性生成对抗网络模块承担着建立CT与PET图像间双向映射关系的关键任务(图2A)。该模块通过循环一致性约束确保跨模态转换的可逆性,有效避免了传统模式的模式崩塌问题。网络包含两个对称的生成器-判别器对: $G_{CT \rightarrow PET}$ 和 $G_{PET \rightarrow CT}$ 分别负责双向图像转换,判别器 D_{PET} 和 D_{CT} 确保生成图像在目标域中的真实性。

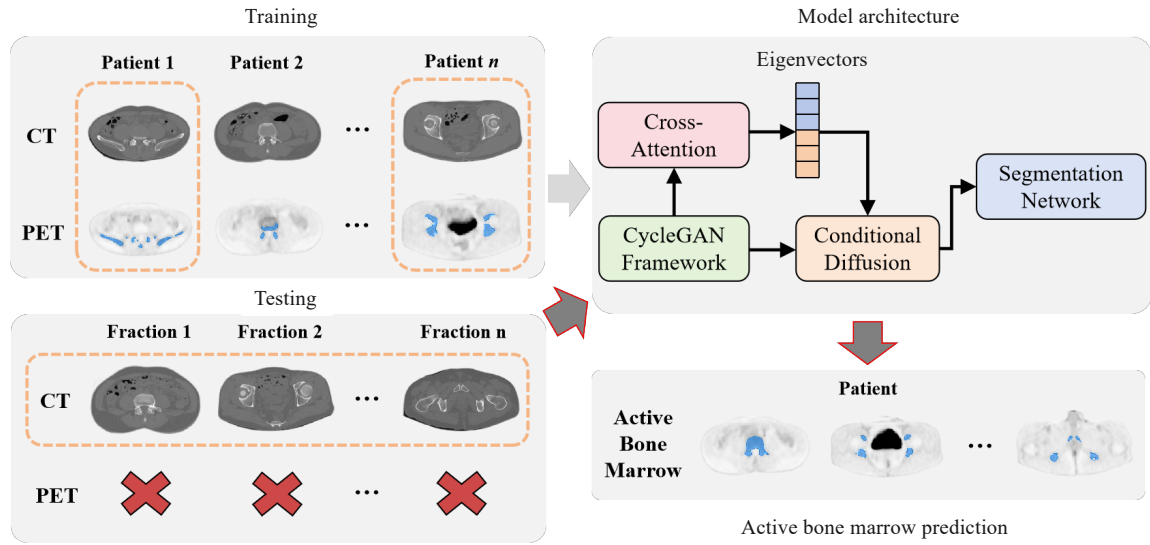


图1 基于扩散循环一致性生成对抗网络的骨盆活跃骨髓区分割方法整体架构图

Fig.1 Overall architecture of the pelvic active bone marrow segmentation method based on diffusion cycle-consistent generative adversarial network.

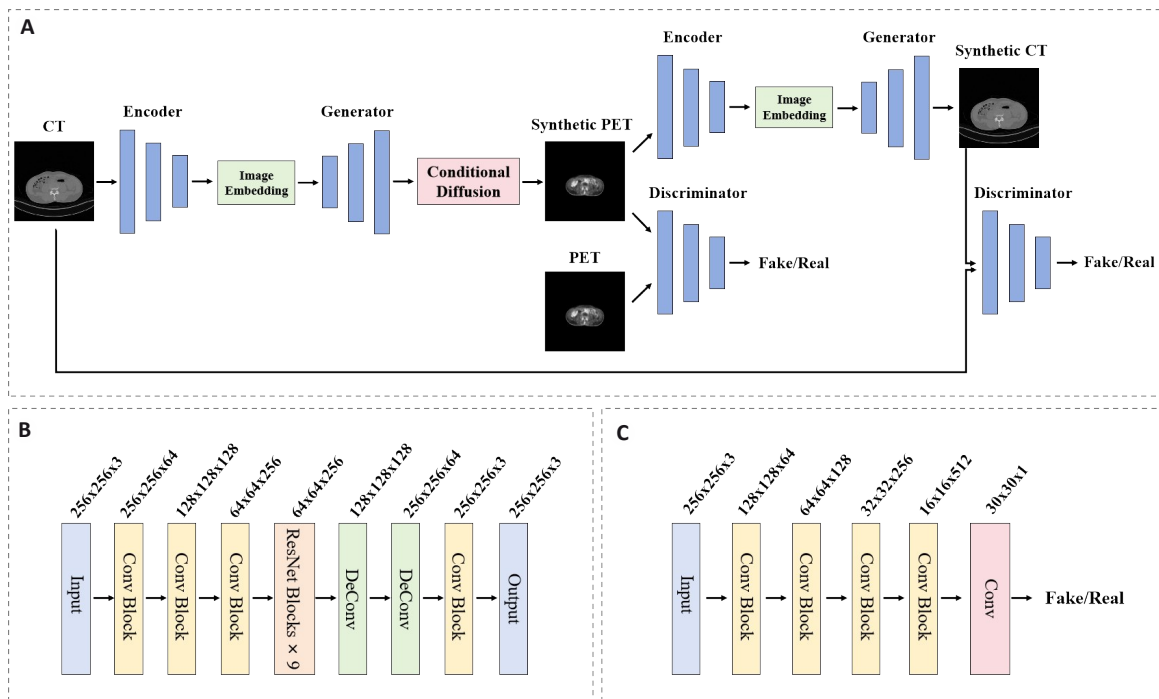


图2 循环一致性生成对抗网络模块架构图

Fig.2 Architecture diagram of the cycle-consistent generative adversarial network module. A: Overall architecture of the CycleGAN establishing bidirectional mapping. B: Generator network architecture containing 9 residual blocks. C: Discriminator network architecture employing spectral normalization.

生成器采用编码器-解码器架构(图2B)。编码器通过多个卷积块逐步下采样构建层次化特征表示 $\mathcal{F}_{enc} = f_{enc}(x_{input})$, 解码器通过反卷积操作和卷积块重建目标图像 $x_{output} = f_{dec}(\mathcal{F}_{enc})$ 。关键创新在于编码器-解码器间的多层残差模块(ResNet Block×9)。这一设计基于以下技术考量:9个残差块的深度设计能够学习足够复杂的跨模态特征变换关系,实验验证表明少于7个残差块会导致转换质量下降,多于11个残差块则会出现过

拟合现象;深层残差连接有效缓解了梯度消失问题,使得网络能够学习CT图像的解剖结构信息到PET图像代谢功能信息的复杂非线性映射;残差结构的恒等映射特性确保了跨模态转换既保持解剖结构一致性,又能有效传递代谢功能语义特征。

判别器网络采用多尺度卷积架构实现层次化判别(图2C)。与传统判别器的重要区别在于引入了多尺度判别策略:在不同分辨率层级(原始尺寸、1/2尺寸、1/4

尺寸)分别进行真实性判别,确保生成图像在全局结构和局部细节上都具备高度真实性。为提升训练稳定性,创新性地采用谱归一化替代传统的批归一化。谱归一化通过约束权重矩阵的最大奇异值使网络满足 1-Lipschitz 连续性,具有以下技术优势:保证梯度传播的稳定性,避免梯度爆炸或消失问题;提高网络对输入扰动的鲁棒性,增强模型的泛化能力;为后续扩散过程提供更可靠的初始化条件,生成的合成PET图像具有更好的特征分布特性,有利于扩散模型的训练收敛。

1.3 双向交叉注意力条件扩散模块

双向交叉注意力条件扩散模块代表了本研究最重要的技术创新,其核心在于提出了一种跨模态动态去噪方法(图3)。传统的扩散模型在去噪过程中通常依赖单一模态的信息,而本研究创新性地引入了双向交叉注意力机制,实现了CT和PET两种模态信息的动态融合。这一创新的关键在于充分利用CT图像丰富的解剖结构信息来指导PET图像的去噪重建过程,从而生成既保持解剖一致性又具有生物学真实性的合成PET图像。

扩散过程包含前向噪声添加和反向去噪重建两个阶段(图3A)。前向过程通过 T 步马尔可夫链将

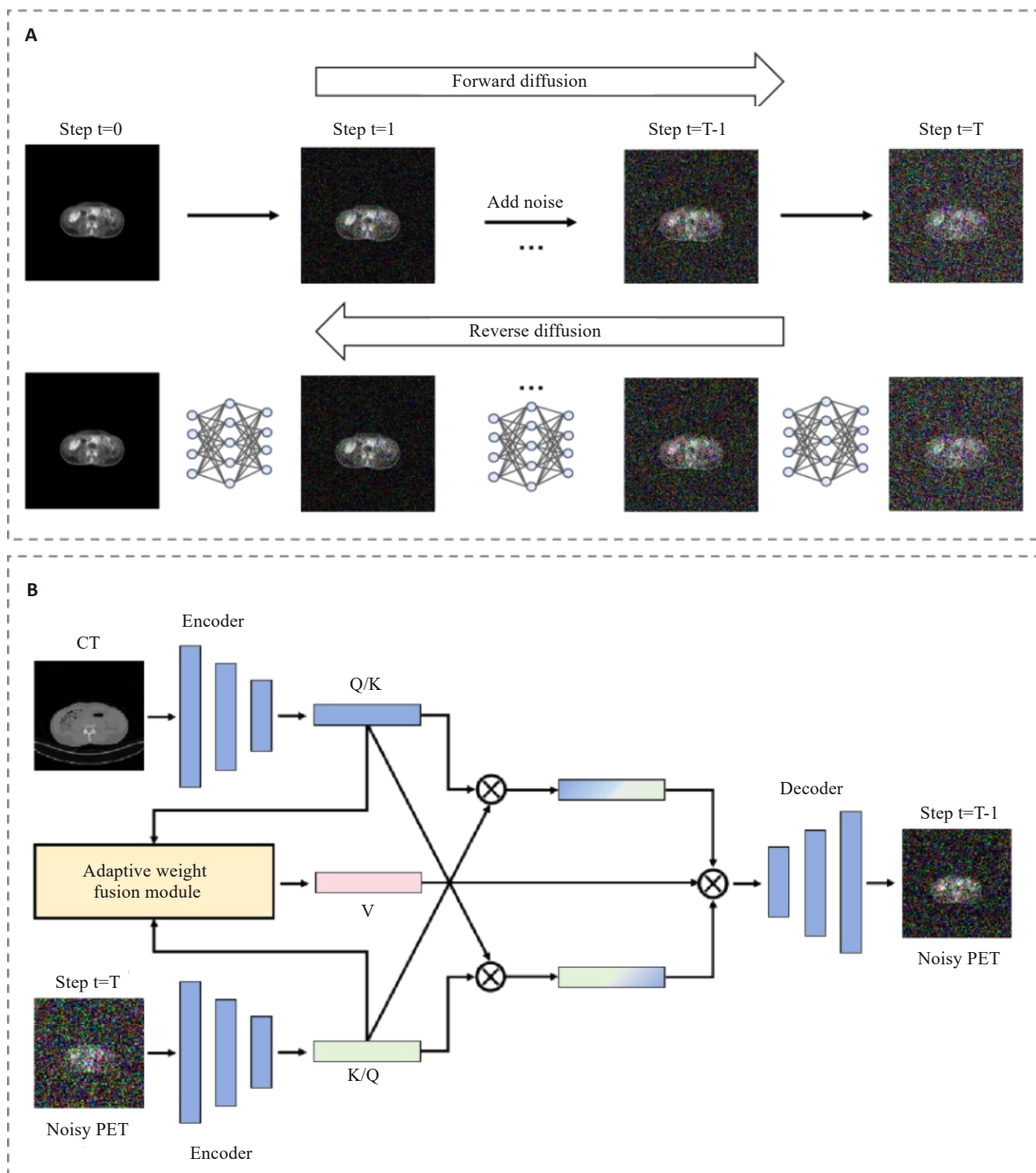


图3 条件扩散模块跨模态动态去噪方法

Fig. 3 Cross-modal dynamic denoising method of the conditional diffusion module. A: Forward noise addition and reverse denoising reconstruction processes based on Markov chain. B: Structure of the cross-modal dynamic denoising network integrating Adaptive Weight Fusion Module and Bidirectional Cross-Attention

CycleGAN生成的PET图像逐渐转换为噪声,反向过程以CT为条件从噪声中重建优化的合成PET图像。每一步都添加噪声,生成图像逐渐变得不可辨识。给定上一时刻的图像 x_{i-1} ,当前时刻的图像 x_i 服从以下分布:

$$q(x_i|x_{i-1}) = \mathcal{N}(x_i; \sqrt{1 - \beta_i} x_{i-1}, \beta_i I) \quad (1)$$

其中 β_i 控制噪声的增加, \mathcal{N} 表示正态分布, I 是单位矩阵。

反向去噪过程是本方法的核心创新所在(图3B)。其采用了独特的跨模态动态去噪策略:

$$p_\theta(x_{i-1}|x_i, c) = \mathcal{N}(x_{i-1}; \mu_\theta(x_i, t, c), \sigma_i^2 I) \quad (2)$$

其中 $\mu_\theta(x_i, t, c)$ 是生成的去噪均值,通过双向交叉注意力机制和动态权重调整策略计算得出, σ_i^2 是去噪过程中的噪声方差, c 表示CT图像的特征。在去噪过程中,CT特征和当前步噪声PET特征通过双向交叉注意力机制实现深度融合。具体而言,CT特征既作为Query去查询噪声PET特征信息,也作为Key/Value为噪声PET特征提供结构引导;同时,噪声PET特征也作为Query去获取CT的解剖结构信息,形成互补的双向特征交互。在去噪的初始阶段,系统更多地依赖CT图像提供的结构信息作为引导;随着去噪过程的进行,系统逐渐减少对CT特征的依赖,转而更多地信任自身的去噪中间结果。这一策略确保了最终生成的图像既保持了与CT图像的解剖一致性,又具备了PET图像特有的代谢功能特征。

动态自适应权重生成模块是该条件扩散系统的核心创新组件。该模块基于当前去噪步骤和实时去噪误差动态生成CT特征和前一步去噪结果的融合权重,实现了跨模态信息利用的自适应优化。模块采用指数衰减为基础模型,结合误差反馈校正机制,确保权重分配的合理性和鲁棒性。

CT特征的权重 W_{CT} 遵循指数衰减规律:

$$W_{CT} = W_0^{CT} \cdot e^{-\lambda \cdot step} \quad (3)$$

相应地,前一步去噪结果权重为:

$$W_{prev} = 1 - W_{CT} \quad (4)$$

为了增强系统对去噪质量的适应性,模块引入误差反馈校正机制:

$$\alpha_{correction} = 1 + (\epsilon_{denoise} \cdot \beta) \quad (5)$$

最终的自适应权重通过原始权重与误差校正因子的乘积获得:

$$W_{final} = W_{original} \cdot \alpha_{correction} \quad (6)$$

其中 W_0^{CT} 为初始权重, λ 是衰减系数, $\epsilon_{denoise}$ 为当前步骤的去噪误差, β 为校正系数。这种设计不仅保证权重过渡的平滑性,更重要的是通过实时误差反馈实现了权重分配的动态优化,使得系统能够根据去噪质量自适应调整跨模态信息的利用策略。

编码器部分通过连续的卷积和池化操作提取多尺度特征,解码器部分通过上采样和跳跃连接重建图像细节。在网络的瓶颈层引入双向交叉注意力模块,实现

CT和PET特征的深度融合。这一设计不仅保证了特征提取的有效性,更重要的是通过多层次的双向特征交互实现了跨模态信息的充分利用。

1.4 多尺度渐进式特征金字塔融合分割网络

多尺度渐进式特征金字塔融合分割网络(MP-FPSN)是本研究在分割技术方面的重要创新,其设计理念突破了传统特征拼接方法的局限性(图4)。传统多模态分割方法通常采用通道维度直接拼接的方式融合不同模态特征,这种做法存在3个关键问题:(1)特征维度线性增长导致参数爆炸和计算复杂度急剧上升;(2)简单拼接无法有效建模模态间的交互关系,容易产生特征冗余;(3)不同模态特征的语义层次差异被忽略,影响融合效果。

该网络的核心创新在于提出了渐进式特征融合策略,通过层次化的特征累积过程实现了多模态信息的有效整合。与传统方法简单拼接不同模态特征的做法相比,本方法采用特征相加和金字塔式累积的方式,不仅将参数量降低约40%,计算复杂度减少35%,更重要的是通过逐元素相加操作保持了各模态特征的完整性和互补性,实现了"1+1>2"的融合效果。

网络采用双分支编码器结构,CT和sPET图像分别通过独立的编码路径提取特征,这一设计确保了各模态特征的专业化提取。编码器在四个不同尺度(1/2、1/4、1/8、1/16)上提取特征,每个尺度都通过自适应特征对齐模块使用 1×1 卷积将不同模态的特征维度统一。

渐进式融合策略的核心创新体现在其"由粗到精"的层次化累积过程。该策略从网络的最深层开始,首先在1/16尺度上将CT和sPET特征直接相加:

$$F_{fused}^{(16)} = \tilde{F}_{CT}^{(16)} + \tilde{F}_{sPET}^{(16)} \quad (7)$$

然后将融合结果上采样并与上一层级的特征再次融合:

$$F_{up}^{(i)} = \text{Upsample}(F_{fused}^{(2i)}) \quad (8)$$

$$F_{fused}^{(i)} = \text{Conv}_{3 \times 3}(F_{up}^{(i)} + \tilde{F}_{CT}^{(i)} + \tilde{F}_{sPET}^{(i)}) \quad (9)$$

其中 $i \in \{8, 4, 2\}$,这一过程层层向上累积,形成丰富的多模态特征金字塔表示。关键创新在于每次特征融合后都使用 3×3 卷积进行特征细化,并引入残差连接确保信息流的完整性。这种设计的优势在于:(1)逐元素相加操作天然保持了各模态的完整信息,避免了拼接可能导致的冗余和维度爆炸问题;(2)金字塔式累积使得浅层特征能够同时获得深层语义信息和多模态融合信息,增强了特征表达能力;(3)渐进式融合策略实现了从全局到局部的多尺度信息整合,显著提升了分割边界的精确性。

解码器部分采用标准的上采样结构,但在跳跃连接的设计上进行了重要改进。传统的跳跃连接通常使用单模态的特征,而本方法创新性地使用融合后的金字塔

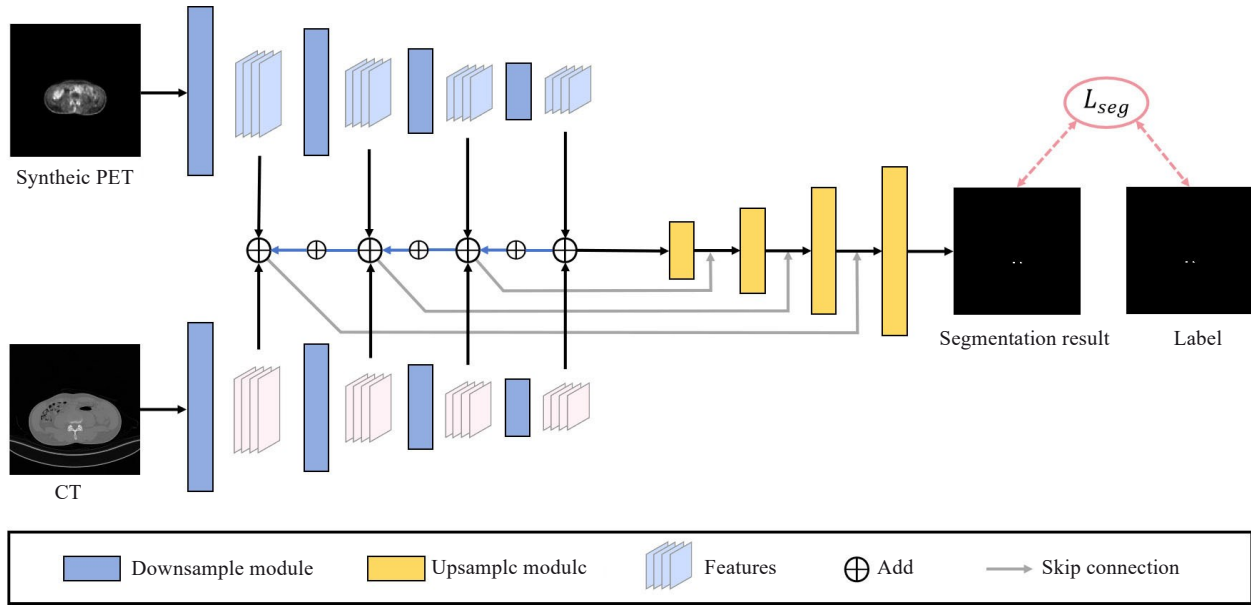


图4 多尺度渐进式特征金字塔融合分割网络架构

Fig.4 Architecture of multi-scale progressive feature pyramid fusion segmentation network.

特征作为跳跃连接。这一改进的技术优势包括:(1)解码器能够充分利用多模态融合的优势,在重建过程中同时考虑解剖结构和功能信息;(2)跳跃连接传递的不再是原始单模态特征,而是经过渐进式融合优化的多模态表示,大幅提升了细节恢复能力;(3)多尺度融合特征的跳跃连接为解码器提供了更丰富的语义层次信息,使得最终分割结果在保持边界精确性的同时具备更强的语义一致性。

1.5 损失函数设计

1.5.1 循环一致性生成对抗网络损失函数 循环一致性生成对抗网络的损失函数设计充分考虑了跨模态转换的特殊性,采用对抗损失和循环一致性损失相结合的策略。对抗损失确保生成图像在目标域中的真实性,其表达式为:

$$L_{GAN}(G, D, X, Y) = \mathbb{E}_{y \sim P_{data}(y)} [\log D(y)] + \mathbb{E}_{x \sim P_{data}(x)} [\log (1 - D(G(x)))] \quad (10)$$

循环一致性损失是本模块的核心约束,通过确保双向转换的一致性来维持图像内容的完整性:

$$L_{cyc}(G_{CT \rightarrow PET}, G_{PET \rightarrow CT}) = \mathbb{E}_{x \sim P_{data}(x)} [\|G_{PET \rightarrow CT}(G_{CT \rightarrow PET}(x)) - x\|_1] + \mathbb{E}_{y \sim P_{data}(y)} [\|G_{CT \rightarrow PET}(G_{PET \rightarrow CT}(y)) - y\|_1] \quad (11)$$

总的损失函数将两者结合,其中循环一致性损失的权重系数 λ_{cyc} 设置为10,以确保转换过程的稳定性:

$$L_{CycleGAN} = L_{GAN}(G_{CT \rightarrow PET}, D_{PET}, CT, PET) + L_{GAN}(G_{PET \rightarrow CT}, D_{CT}, PET, CT) + \lambda_{cyc} L_{cyc} \quad (12)$$

1.5.2 条件扩散模块损失函数 条件扩散模块的损失函数设计体现了跨模态动态去噪的特殊要求。基础的去噪匹配损失确保网络能够准确预测各时间步的噪声:

$$L_{diff} = \mathbb{E}_{t, x_0, \epsilon, c} [\|\epsilon - \epsilon_0(x_t, t, c)\|^2] \quad (13)$$

其中 x_t 表示在时间步 t 上的噪声图像, ϵ 是真实的噪声,表示从清晰图像 x_0 添加噪声后得到的扰动。 $\epsilon_0(x_t, t, c)$ 是噪声预测网络的输出,网络根据当前图像 x_t 、时间步 t 和条件信息 c (CT)来预测噪声。为了进一步提升合成图像的质量和生物学真实性,引入感知损失来约束高层特征的一致性:

$$L_{perceptual} = \sum_j \frac{1}{C_j H_j W_j} \|\Phi_j(x_{pred}) - \Phi_j(x_{gt})\|_2^2 \quad (14)$$

条件扩散模块的总损失函数为: $L_{diffusion} = L_{diff} + \lambda_p L_{perceptual}$,其中感知损失权重 λ_p 设置为0.1。

1.5.3 分割网络损失函数 针对ABM区域分割任务的特殊性,采用Dice损失和交叉熵损失相结合的策略。Dice损失用于处理类别不平衡问题:

$$L_{dice} = 1 - \frac{2 \sum_{i=1}^N p_i g_i + \epsilon}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2 + \epsilon} \quad (15)$$

交叉熵损失确保像素级分类准确性:

$$L_{ce} = -\sum_{i=1}^N g_i \log(p_i) \quad (16)$$

其中 p_i 是预测概率, g_i 是真实标签, ϵ 是平滑项。分割网络的总损失函数为:

$$L_{seg} = L_{dice} + L_{ce} \quad (17)$$

1.6 实验数据

本研究使用的医学图像数据采集于中山大学肿瘤防治中心核医学科和放疗科,所有采集数据均进行匿名化与隐私信息脱敏处理,严格遵循医学伦理要求对患者信息保密。最终收集共253例初诊肿瘤患者的骨盆PET-CT图像数据,患者年龄为18~70岁,男女比例约为1.1:1。图5展示了典型的PET-CT图像数据及其对应的ABM标注示例,其中PET图像清晰显示骨盆区域的

代谢活跃骨髓分布,CT图像提供了详细的解剖结构信息。CT图像采集参数为层厚2.5 mm,图像尺寸512×512像素,PET图像重建后层厚为2.5 mm,图像矩阵为256×256。

数据预处理流程包括PET-CT图像精确配准以确保空间对应关系,使用Otsu阈值分割算法去除扫描床等冗余信息并提取骨盆感兴趣区域。针对CT图像采用窗宽窗位技术进行强度归一化至 $[-1, 1]$,PET图像基于

标准化摄取值(SUV)进行归一化处理,确保数据分布一致性。由经验丰富的核医学医师组成专家团队负责ABM区域的精确标注,以PET图像中代谢活跃区域为依据,结合解剖学知识进行手工勾画和质量控制。

数据集按照7:3的比例随机分配,其中177例用于训练集,76例用于测试集。数据增强策略包括随机旋转($\pm 10^\circ$)、水平翻转和强度抖动等操作以提高模型的泛化能力。

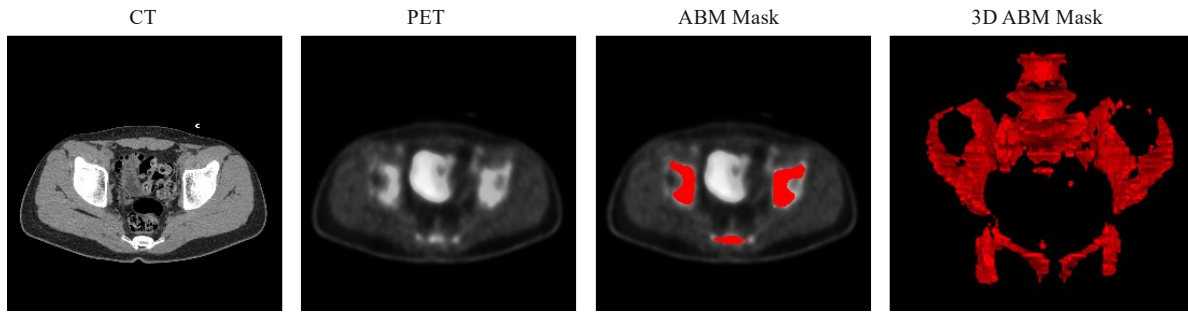


图5 典型PET-CT图像数据及ABM标注示例

Fig.5 Typical PET-CT image data and ABM annotation examples.

1.7 评价指标

本研究采用以下评价指标来全面评估基于扩散循环一致性生成对抗网络的骨盆活跃骨髓区分割方法的性能。

1.7.1 图像合成质量评价指标 峰值信噪比(PSNR)用于量化合成PET图像与真实PET图像间的像素级差异:

$$PSNR = 10 \log_{10} \frac{MAX^2}{MSE} \quad (18)$$

其中MAX为图像的最大像素值,MSE为均方误差。PSNR值越高表示图像质量越好。

结构相似性指数(SSIM)从亮度、对比度和结构3个维度评估图像相似性:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (19)$$

其中 μ_x, μ_y 分别为图像 x, y 的均值, σ_x, σ_y 为标准差, σ_{xy} 为协方差, c_1, c_2 为稳定常数。SSIM取值范围为 $[0, 1]$,值越接近1表示结构相似性越高。

归一化均方误差(NMSE)评估合成图像与真实图像间的平均像素差异:

$$NMSE = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N y_i^2} \quad (20)$$

其中 y_i 为真实像素值, \hat{y}_i 为合成像素值,N为像素总数。NMSE值越小表示图像合成质量越好。

1.7.2 分割精度评价指标 Dice相似系数(DSC)评估预测分割结果与真实标注的重叠程度:

$$DSC = \frac{2|P \cap G|}{|P| + |G|} \quad (21)$$

其中P为预测分割结果,G为真实标注。DSC取值范围为 $[0, 1]$,值越接近1表示分割精度越高。

平均对称表面距离(ASDD)评估预测轮廓与真实轮廓间的平均距离:

$$ASDD = \frac{1}{2} \left(\frac{1}{|S_p|} \sum_{p \in S_p} d(p, S_G) + \frac{1}{|S_G|} \sum_{g \in S_G} d(g, S_p) \right) \quad (22)$$

其中 S_p, S_G 分别为预测和真实分割的表面点集, $d(p, S_G)$ 和 $d(g, S_p)$ 分别表示点到集合的最短距离。ASDD值越小表示分割边界越准确。

1.8 实验设计

本研究中所有方法与模型均基于Python 3.9环境实现,采用PyTorch作为深度学习框架。实验硬件平台为配备2张NVIDIA RTX A6000 GPU(每张48GB显存)的Linux服务器,CUDA版本为12.4,采用分布式并行训练策略以提升训练效率。

针对基于扩散循环一致性生成对抗网络的骨盆活跃骨髓区分割模型,训练策略采用分阶段优化方案。首先循环一致性生成对抗网络模块与多尺度渐进式特征金字塔融合分割网络进行联合训练,建立稳定的跨模态映射关系和初步的分割能力,共执行150个epoch,随后引入条件扩散模块进行联合训练200个epoch,总训练周期为350个epoch。这种渐进式策略有效避免了复杂多模块网络初期训练的不稳定性,确

保各组件的协同发展。

优化器配置方面,全部模块统一采用 Adam 优化器,动量参数 $\beta_1=0.9, \beta_2=0.999$,权重衰减系数设置为 1×10^{-4} 以防止过拟合。批处理大小设置为 4。学习率调度策略采用阶梯式衰减机制,初始学习率设置为 1×10^{-4} ,在第 150 和 300 个 epoch 处分别衰减至原值的 0.1 倍,即在第 2 阶段开始时和训练后期进行两次衰减,以确保模型在不同训练阶段的稳定收敛。

损失函数权重配置基于消融实验确定:循环一致性损失权重 λ 设置为 10,保证跨模态转换的双向一致性;感知损失权重设置为 0.1,平衡像素级重建和高层语义特征保持;分割任务中 Dice 损失和交叉熵损失权重比例设置为 1:1,有效处理类别不平衡问题。扩散过程的时间步数 T 设置为 1000 步,在生成质量和计算效率间取得平衡。动态自适应权重生成模块中的衰减系数 λ 设置为 0.1,确保跨模态信息融合的渐进式过渡。

2 结果

2.1 图像合成质量评估

本研究对所提出方法与现有代表性图像生成方法在图像合成任务中的性能进行了定量评估。比较方法包括 U-net、Pix2pix、CycleGAN、WGAN、DCGAN 和 DDPM。评估指标包括峰值信噪比(PSNR)、结构相似性指数(SSIM)和归一化均方误差(NMSE)。

各方法的定量评估结果见表 1。本方法的 PSNR 为 26.42 ± 0.63 dB, SSIM 为 0.894 ± 0.011 , NMSE 为 0.0235 ± 0.0026 。在所有比较方法中,本方法在 PSNR 和 SSIM 指标上数值最高, NMSE 指标数值最低。训练时间方面,本方法平均耗时 239.8 s/epoch。

图 6 展示了不同方法生成图像的视觉对比结果。图中显示了 3 个不同病例的骨盆横断面图像。第一列为真实 PET 图像,其余各列为不同方法生成的合成 PET 图像。每种方法下方显示与真实 PET 的差异热力图。

表 1 各图像合成方法的性能比较

Tab.1 Quantitative comparison of different methods for image synthesis quality

Methods	PSNR	SSIM	NMSE	Average time (Epoch/s)
U-net ^[20]	21.27±1.31	0.787±0.031	0.0438±0.0062	57.1
Pix2pix ^[16]	22.61±1.04	0.815±0.026	0.0359±0.0045	48.9
CycleGAN ^[21]	24.33±0.91	0.861±0.022	0.0297±0.0041	53.4
WGAN ^[22]	20.42±1.52	0.771±0.035	0.0481±0.0073	65.3
DCGAN ^[23]	19.08±1.89	0.738±0.044	0.0541±0.0095	44.8
DDPM ^[18]	23.71±0.97	0.849±0.024	0.0324±0.0042	197.2
Ours	26.42±0.63	0.894±0.011	0.0235±0.0026	239.8

2.2 ABM 区域分割精度分析

我们进一步评估了模型在骨盆活跃骨髓(ABM)区域的分割性能,涵盖 3 个主要解剖结构:骶髂脊柱、髌骨、下盆腔。采用 Dice 相似系数(DSC)评估分割区域的重叠程度,采用平均对称表面距离(ASSD)评估边界精确性。

分割性能评估结果见表 2。本方法在 3 个解剖区域的 DSC 分别为:骶髂脊柱 0.806 ± 0.019 、髌骨 $0.771 \text{ mm} \pm 0.023$ 、下盆腔 0.754 ± 0.026 ,对应的 ASSD 分别为 $3.02 \text{ mm} \pm 0.33 \text{ mm}$ 、 $3.58 \text{ mm} \pm 0.41 \text{ mm}$ 和 $3.95 \text{ mm} \pm 0.49 \text{ mm}$ 。整体平均 DSC 为 0.777 ± 0.023 ,平均 ASSD 为 $3.52 \text{ mm} \pm 0.41 \text{ mm}$ 。图 7 展示了典型的 ABM 分割结果,其中绿色区域表示真实标注,蓝色区域表示模型预测结果。

2.3 条件扩散模块实验分析

本研究评估了条件扩散模块在不同扩散步数(Diffusion Steps T)下对图像合成与分割性能的影响。实验设置扩散步数分别为 100、250、500、1000、1500 和 2000 步,评估指标包括 PSNR、SSIM、NMSE 和 Dice

系数。

不同扩散步数的实验结果见表 3。当扩散步数为 1000 时,各项指标达到最优值:PSNR 为 26.42 ± 0.63 dB, SSIM 为 0.894 ± 0.011 , NMSE 为 0.0235 ± 0.0026 , Dice 为 0.777 ± 0.023 。继续增加扩散步数至 1500 和 2000 步时,性能提升有限但计算时间显著增加。此外,本研究比较了不同权重策略对模型性能的影响,包括固定权重、线性衰减、余弦衰减、指数衰减以及指数衰减结合误差反馈机制。权重策略实验结果见表 4,指数衰减结合误差反馈机制在各项指标上数值最优。

2.4 消融实验结果

为验证各技术组件的有效性,本研究设计了渐进式消融实验。实验从仅使用 CT 图像的基线方法开始,逐步加入 CycleGAN 模块、条件扩散模块和多尺度特征融合网络,最后采用端到端联合优化策略。所有实验均在相同的数据集和硬件环境下进行,以确保对比的公平性。

消融实验结果见表 5。基线方法的 Dice 系数为

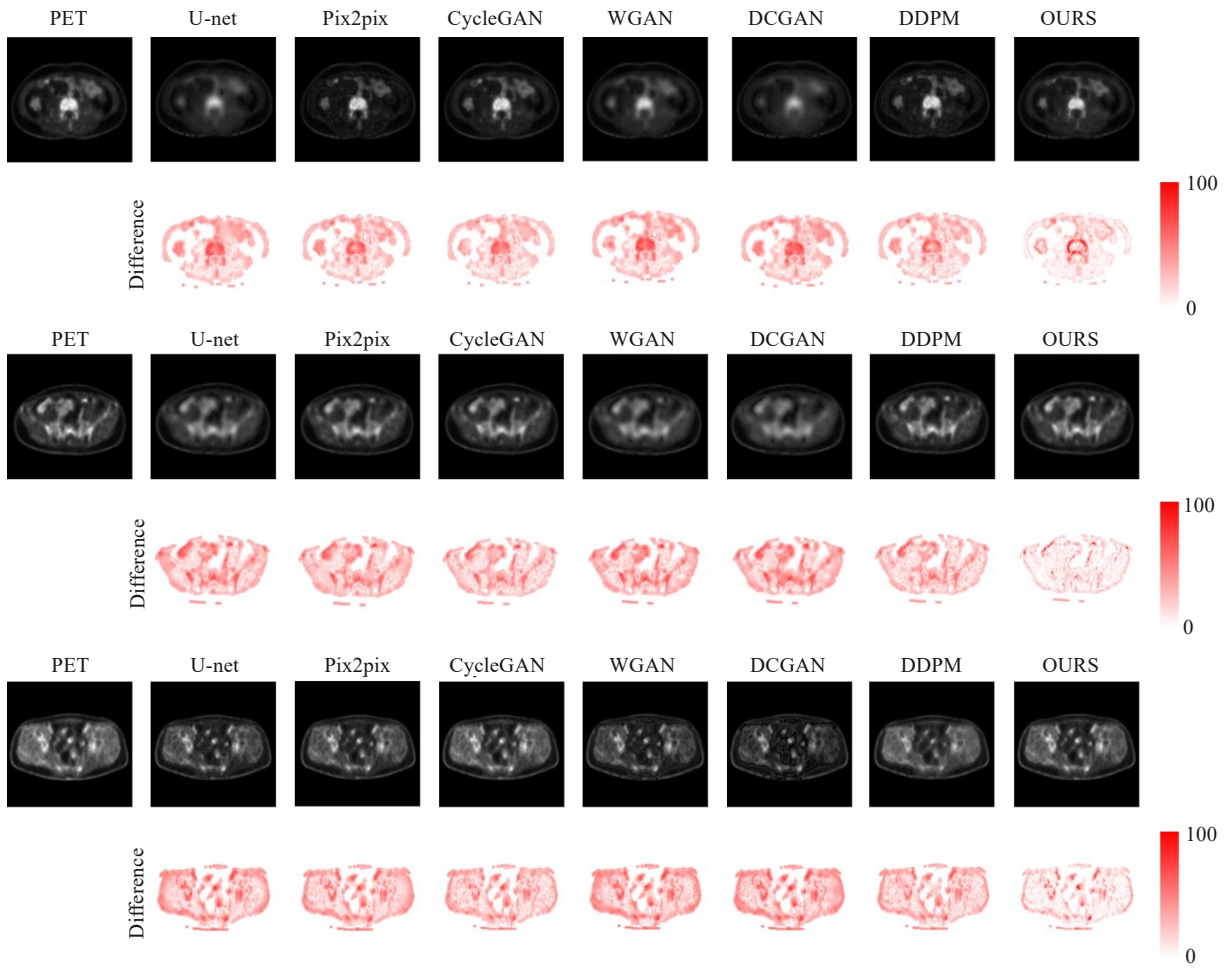


图6 不同方法生成图像的可视化对比

Fig.6 Visual comparison of synthesized images with different methods.

表2 多种方法在不同ABM区域的分割性能比较

Tab.2 Comparison of segmentation performance of different methods in different ABM regions.

Methods	Dice				ASSD			
	Lumbosacral spine	Ilium	Lower pelvis	Average	Lumbosacral spine	Ilium	Lower pelvis	Average
GAN+U-net ^[20]	0.696±0.031	0.652±0.039	0.638±0.035	0.662±0.035	4.21±0.48	5.08±0.63	5.61±0.71	4.97±0.61
SynSeg-Net ^[24]	0.718±0.027	0.675±0.034	0.661±0.032	0.685±0.031	3.89±0.44	4.68±0.58	5.15±0.64	4.57±0.55
Robust-Mseg ^[25]	0.731±0.024	0.693±0.030	0.675±0.028	0.700±0.027	3.67±0.40	4.35±0.52	4.83±0.58	4.28±0.50
SCM ^[26]	0.715±0.029	0.677±0.033	0.661±0.031	0.684±0.031	3.87±0.43	4.62±0.56	5.10±0.62	4.53±0.54
PEMMA ^[27]	0.748±0.022	0.709±0.026	0.688±0.027	0.715±0.025	3.45±0.38	4.12±0.49	4.56±0.55	4.04±0.47
Ours	0.806±0.019	0.771±0.023	0.754±0.026	0.777±0.023	3.02±0.33	3.58±0.41	3.95±0.49	3.52±0.41

0.664±0.036, ASSD 为 4.97 mm±0.62 mm。加入 CycleGAN 模块后,利用生成的合成 PET 图像提供跨模态信息, Dice 系数提升至 0.702±0.028 (提升 5.7%), ASSD 降至 4.28 mm±0.51 mm。条件扩散模块的引入显著改善了图像合成质量, PSNR 从 24.35dB 提升至 26.42 dB,相应地分割性能进一步改善, Dice 系数达到 0.726±0.031。

采用多尺度渐进式特征金字塔融合网络替代简单特征拼接后,分割边界精度显著提升, Dice 系数达到

0.758±0.025, ASSD 降至 3.76 mm±0.44 mm (较上一步改善 8.5%)。完整方法通过端到端联合优化实现各组件间的协同效应,最终达到 Dice 系数 0.777±0.023, ASSD 为 3.52 mm±0.41 mm,相比基线方法分别提升 17.0% 和 29.2%。

3 讨论

直肠癌全程新辅助治疗中骨髓毒性个体化防护是当前临床面临的关键挑战^[4,5]。深度学习在自动分割领

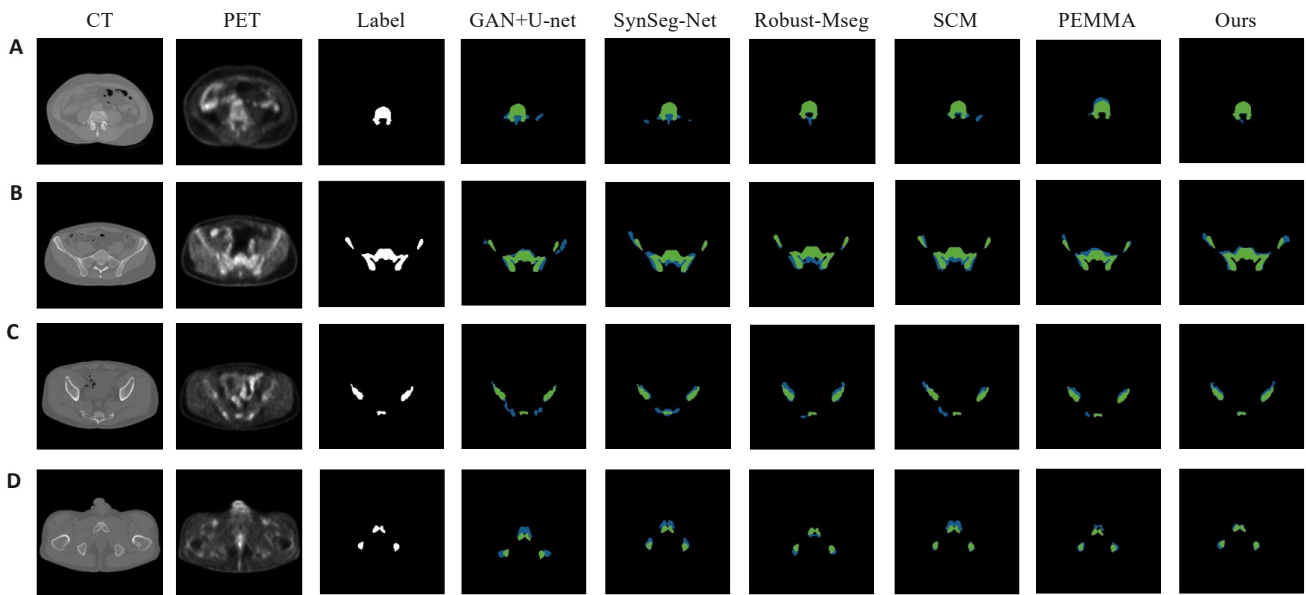


图7 ABM分割结果的可视化示例

Fig.7 Visual examples of ABM segmentation results. (A) Lumbosacral spine region; (B) Ilium region; (C) Lower pelvis region; (D) Representative case with ground truth (green) and predictions (blue).

表3 不同扩散步数下的图像生成效果及分割预测对比

Tab.3 Comparison of image generation effects and segmentation prediction under different diffusion steps

Diffusion steps T	PSNR (dB)	SSIM	NMSE	Dice	Average time (Epoch/s)
100	22.28±1.29	0.797±0.043	0.0449±0.0076	0.548±0.049	113.5
250	23.94±1.05	0.827±0.036	0.0383±0.0063	0.625±0.041	142.6
500	25.58±0.86	0.861±0.029	0.0305±0.0050	0.715±0.020	184.0
1000	26.42±0.63	0.894±0.011	0.0235±0.0026	0.777±0.023	239.8
1500	26.21±0.70	0.888±0.025	0.0238±0.0039	0.768±0.030	337.1
2000	26.03±0.76	0.884±0.028	0.0245±0.0043	0.761±0.033	431.0

表4 不同权重策略对模型性能的影响

Tab.4 Impact of different weighting strategies on model performance

Weight strategy	Weight function	PSNR (dB)	SSIM	NMSE	Dice
Fixed weight (0.3:0.7)	$W_{CT} = 0.3$	24.91±0.92	0.851±0.026	0.0339±0.0048	0.728±0.037
Fixed weight (0.5:0.5)	$W_{CT} = 0.5$	25.07±0.84	0.862±0.022	0.0323±0.0044	0.745±0.033
Fixed weight (0.7:0.3)	$W_{CT} = 0.7$	24.73±0.99	0.845±0.029	0.0357±0.0055	0.719±0.040
Linear decay	$W_{CT} = 1 - \frac{t}{T}$	25.58±0.75	0.871±0.018	0.0287±0.0036	0.763±0.026
Cosine decay	$W_{CT} = \frac{1 + \cos(\frac{\pi t}{T})}{2}$	25.84±0.70	0.878±0.016	0.0271±0.0033	0.769±0.024
Exponential decay	$W_{CT} = e^{-\lambda t}$	26.08±0.66	0.885±0.014	0.0253±0.0030	0.772±0.023
Exponential decay+error feedback	$W_{CT} = e^{-\lambda t} (1 + \beta \epsilon)$	26.42±0.63	0.894±0.011	0.0235±0.0026	0.777±0.023

域的快速发展为解决复杂医学图像问题提供了新的技术路径^[28],但现有方法在骨髓保护的个性化精度方面仍存在明显不足。为此,我们构建了三阶段级联跨模态学习框架,通过扩散循环一致性生成对抗网络实现从CT到个性化ABM精准识别的技术突破。实验结果显示,该方法在图像合成PSNR达到(26.42±0.63) dB,SSIM

达到0.894±0.011;ABM分割平均Dice系数达到0.777±0.023,较传统方法提升超过10%。

现有跨模态图像合成方法在医学图像转换中呈现出明显的性能差异。传统GAN方法局限性明显,WGAN^[22]和DCGAN^[23]的PSNR分别仅为(20.42±1.52) dB和(19.08±1.89) dB,主要受训练不稳定和模式崩塌

表5 消融实验结果

Tab.5 Ablation study results

Configuration	Input modality	Dice	ASSD	PSNR (dB)	SSIM	NMSE
Baseline	CT only	0.664±0.036	4.97±0.62	-	-	-
+CycleGAN	CT+pseudo-PET	0.702±0.028	4.28±0.51	24.35±0.81	0.861±0.017	0.0281±0.0033
+Conditional diffusion	CT+enhanced-PET	0.726±0.031	4.11±0.55	26.42±0.63	0.894±0.011	0.0235±0.0026
+Multi-scale fusion	CT+enhanced-PET+advanced fusion	0.758±0.025	3.76±0.44	26.42±0.63	0.894±0.011	0.0235±0.0026
Complete method	Full pipeline with joint optimization	0.777±0.023	3.52±0.41	26.42±0.63	0.894±0.011	0.0235±0.0026

问题困扰。CycleGAN^[21]虽然在无配对数据转换中表现良好,但其PSNR为(24.33±0.91) dB,循环一致性约束仅通过L1损失实现,难以保证复杂医学图像的细节一致性。扩散模型DDPM^[18]在图像质量上有所改善[PSNR为(23.71±0.97) dB],但仍缺乏跨模态转换的有效约束机制。针对这些技术局限,我们提出的扩散循环一致性生成对抗网络通过融合扩散概率模型与循环一致性约束,有效解决了传统方法的技术缺陷,PSNR较最佳对比方法CycleGAN提升8.6%。近期研究表明,扩散模型在放疗合成CT生成领域已被证明优于传统GAN方法^[29,30]。该网络的核心创新在于双向交叉注意力条件扩散机制,采用跨模态动态去噪策略实现CT解剖信息与PET代谢信息的深度融合。在去噪初期依赖CT结构信息引导,随过程推进逐渐转向自身去噪结果,确保生成图像既保持解剖一致性又具备代谢功能特征。1000步扩散过程在质量和效率间达到最优平衡,值得注意的是,当扩散步数超过1000步时,模型性能出现轻微下降。这一现象可能归因于过度扩散导致的信息损失和计算误差累积,这与医学图像生成领域的相关研究发现一致^[31]。动态自适应权重生成模块采用指数衰减结合误差反馈机制,相比固定权重方法PSNR提升1.51dB,验证了动态权重调整的有效性。

在分割方面,现有方法同样面临技术瓶颈。早期GAN+U-net^[20]平均Dice系数仅为0.662±0.035,因缺乏深度多模态特征融合机制。改进方法如SynSeg-Net^[24]、Robust-Mseg^[25]等,Dice系数分别达到0.685±0.031、0.700±0.027,虽有提升但仍低于0.72的临床可接受阈值。这些方法多采用简单特征拼接策略,容易导致维度爆炸和信息冗余。为克服这些限制,设计的多尺度渐进式特征金字塔融合分割网络(MP-FPSN)突破了传统特征拼接的局限性,采用“由粗到精”的渐进式融合策略,从1/16尺度开始逐层向上累积,使浅层特征同时获得深层语义信息和多模态融合信息。与传统拼接方法相比,这种设计采用特征相加和金字塔式累积,参数量降低约40%,计算复杂度减少35%。

性能对比显示,该方法相比最强基线PEMMA^[27]取得显著优势。PEMMA平均Dice系数达到0.715±0.025,

而我们的方法在Dice系数上比PEMMA提升8.7%,ASSD改善13.9%。这一提升主要归因于3个技术创新的协同作用:扩散循环一致性生成对抗网络的高保真图像生成、双向交叉注意力机制的深度跨模态融合、以及多尺度渐进式特征金字塔网络的精细分割性能。消融实验显示,完整方法(CG+CDM+MP-FPSN)相比基线提升17.0%。在最具挑战性的下盆腔区域,该方法Dice系数达到0.754±0.026,而传统GAN+U-net仅为0.638±0.035,展现出优秀的鲁棒性。方法实现的0.777平均Dice系数首次突破0.75的临床高精度阈值,相比传统基于解剖图谱方法(精度0.65-0.70)有显著提升^[10]。

但本研究仍存在一些局限性。首先是数据集多样性问题,当前基于单中心253例患者数据的研究,在处理复杂病理状态时可能面临泛化能力限制。未来研究应建立多中心的大规模数据集。其次是计算效率挑战,扩散过程的高计算复杂度导致训练时间达239.8 s/epoch,影响了临床应用的便捷性。未来研究应重点发展知识蒸馏、模型剪枝和量化等模型压缩技术,在保持精度的前提下显著提升推理速度。

综上所述,通过扩散循环一致性生成对抗网络、双向交叉注意力条件扩散机制和多尺度渐进式特征金字塔融合分割网络的协同创新,构建的三阶段级联跨模态学习框架在技术指标和临床应用方面均实现了重要突破。本研究为个体化精准骨髓保护提供了切实可行的技术路径,有望在直肠癌精准放疗中发挥重要作用,推动放疗领域向个体化精准医疗方向发展。

Declaration of interests: The authors declare no competing interests.

参考文献:

- [1] Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020 [J]. *CA A Cancer J Clinicians*, 2020, 70(3): 145-64.
- [2] Conroy T, Bosset JF, Etienne PL, et al. Neoadjuvant chemotherapy with FOLFIRINOX and preoperative chemoradiotherapy for patients with locally advanced rectal cancer (UNICANCER-PRODIGE 23): a multicentre, randomised, open-label, phase 3 trial [J]. *Lancet Oncol*, 2021, 22(5): 702-15.
- [3] Garcia-Aguilar J, Patil S, Gollub MJ, et al. Organ preservation in patients with rectal adenocarcinoma treated with total neoadjuvant

- therapy[J]. *J Clin Oncol*, 2022, 40(23): 2546-56.
- [4] Jin J, Tang Y, Hu C, et al. Multicenter, randomized, phase III trial of short-term radiotherapy plus chemotherapy versus long-term chemoradiotherapy in locally advanced rectal cancer (STELLAR) [J]. *J Clin Oncol*, 2022, 40(15): 1681-92.
- [5] Bahadoer RR, Dijkstra EA, van Etten B, et al. Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial[J]. *Lancet Oncol*, 2021, 22(1): 29-42.
- [6] Diefenhardt M, Ludmir EB, Hofheinz RD, et al. Association of treatment adherence with oncologic outcomes for patients with rectal cancer: a post hoc analysis of the CAO/ARO/AIO-04 phase 3 randomized clinical trial [J]. *JAMA Oncol*, 2020, 6(9): 1416-21.
- [7] Hayman JA, Callahan JW, Herschtal A, et al. Distribution of proliferating bone marrow in adult cancer patients determined using FLT-PET imaging[J]. *Int J Radiat Oncol Biol Phys*, 2011, 79(3): 847-52.
- [8] McGuire SM, Menda Y, Ponto LL, et al. A methodology for incorporating functional bone marrow sparing in IMRT planning for pelvic radiation therapy [J]. *Radiother Oncol*, 2011, 99(1): 49-54.
- [9] Li N, Noticewala SS, Williamson CW, et al. Feasibility of atlas-based active bone marrow sparing intensity modulated radiation therapy for cervical cancer [J]. *Radiother Oncol*, 2017, 123(2): 325-30.
- [10] Yusufaly T, Miller A, Medina-Palomo A, et al. A multi-atlas approach for active bone marrow sparing radiation therapy: implementation in the NRG-GY006 trial [J]. *Int J Radiat Oncol Biol Phys*, 2020, 108(5): 1240-7.
- [11] Czernin J, Allen-Auerbach M, Nathanson D, et al. PET/CT in oncology: current status and perspectives [J]. *Curr Radiol Rep*, 2013, 1(3): 177-90.
- [12] Yang TJ, Oh JH, Apte A, et al. Clinical and dosimetric predictors of acute hematologic toxicity in rectal cancer patients undergoing chemoradiotherapy [J]. *Radiother Oncol*, 2014, 113(1): 29-34.
- [13] Zhu JD, Wu H, Chen YL, et al. The correlation between the change of Hounsfield units value and Modic changes in the lumbar vertebral endplate [J]. *BMC Musculoskelet Disord*, 2021, 22(1): 509.
- [14] Meyer HJ, Pönisch W, Monecke A, et al. Can diagnostic low-dose whole-body CT reflect bone marrow findings in systemic mastocytosis [J]. *Anticancer Res*, 2020, 40(2): 1015-22.
- [15] Goodsitt MM, Shenoy A, Shen JC, et al. Evaluation of dual energy quantitative CT for determining the spatial distributions of red marrow and bone for dosimetry in internal emitter radiation therapy [J]. *Med Phys*, 2014, 41(5): 051901.
- [16] Isola P, Zhu JY, Zhou TH, et al. Image-to-image translation with conditional adversarial networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 5967-76.
- [17] Salehjahromi M, Karpinets TV, Sujit SJ, et al. Synthetic PET from CT improves diagnosis and prognosis for lung cancer: Proof of concept [J]. *Cell Rep Med*, 2024, 5(3): 101463.
- [18] Kazerouni A, Aghdam EK, Heidari M, et al. Diffusion models in medical imaging: a comprehensive survey [J]. *Med Image Anal*, 2023, 88: 102846.
- [19] DhariwalPrasanna, NicholAlex. Diffusion models beat GANs on image synthesis [C]. Proceedings of the 35th International Conference on Neural Information Processing Systems, 2021: 8780-94.
- [20] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [M]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-41.
- [21] Zhu JY, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017, Venice, Italy. IEEE, 2017: 2242-51.
- [22] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks [C]. Proceedings of the 34th International Conference on Machine Learning, 2017: 214-23.
- [23] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [J]. arXiv preprint arXiv:1511.06434, 2015.
- [24] Huo YK, Xu ZB, Moon H, et al. SynSeg-net: synthetic segmentation without target modality ground truth [J]. *IEEE Trans Med Imaging*, 2018: 10.1109/TMI.2018.2876633.
- [25] Chen C, Dou Q, Jin YM, et al. Robust multimodal brain tumor segmentation via feature disentanglement and gated fusion [M]//Medical Image Computing and Computer Assisted Intervention - MICCAI 2019. Cham: Springer International Publishing, 2019: 447-56.
- [26] Diao ZS, Jiang HY, Shi TY, et al. Siamese semi-disentanglement network for robust PET-CT segmentation [J]. *Expert Syst Appl*, 2023, 223: 119855.
- [27] Saadi N, Saeed N, Yaqub M, et al. *PEMMA*: parameter-efficient multi-modal adaptation for medical image segmentation [M]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2024. Cham: Springer Nature Switzerland, 2024: 262-71.
- [28] Ponnusamy R, Zhang M, Wang Y, et al. Automatic segmentation of bone marrow lesions on MRI using a deep learning method [J]. *Bioengineering (Basel)*, 2024, 11(4): 374.
- [29] Sizikova E, Badal A, Delfino JG, et al. Synthetic data in radiological imaging: current state and future outlook [J]. *Bjrtartificial Intell*, 2024, 1: ubae007.
- [30] Hu C, Cao N, Li XH, et al. CBCT-to-CT synthesis using a hybrid U-Net diffusion model based on transformers and information bottleneck theory [J]. *Sci Rep*, 2025, 15: 10816.
- [31] Shi YQ, Abulizi A, Wang H, et al. Diffusion models for medical image computing: a survey [J]. *Tsinghua Sci Technol*, 2025, 30(1): 357-83.

(编辑:经 媛)