

## 基于多分支自适应融合的多模态MR图像合成方法

邱宏轩<sup>1,2</sup>, 刘明洋<sup>3</sup>, 刘苏锐<sup>2</sup>, 王宇熙<sup>1,2</sup>, 彭博<sup>2</sup>, 戴亚康<sup>2</sup>, 于微波<sup>1</sup>

1. 长春工业大学电气与电子工程学院, 吉林 长春 130012; 2. 中国科学院苏州生物医学工程技术研究所, 江苏 苏州 215163; 3. 吉林建筑科技学院电气与机械工程学院, 吉林 长春 130114

**【摘要】**多模态磁共振图像(MR)为临床诊断与治疗提供全面且关键的信息,生成缺失模态图像有助于医学分析。针对现有合成方法在输入模态数量固定及合成图像的质量和结构保真度不佳等局限,本研究基于生成对抗网络框架提出一种多分支自适应融合与模态一致性引导的多模态磁共振图像合成方法(MMFC-GAN)。通过引入目标模态标签、零图像占位和多分支编码策略,实现对任意模态组合输入的高效医学图像合成。其中,目标模态标签引导解码器聚焦于所需特征,零图像占位处理输入模态不完整问题,而多分支编码确保各模态特征独立提取与灵活融合,增强模型对多模态图像输入的适应性与合成能力。此外,引入模态一致性引导机制,在潜在空间对不同模态的编码特征进行对齐,强化模态间的结构信息一致性,从而提升生成图像的结构保真度和合成质量。在BraTS2020和ISLES2015数据集上验证了本方法的有效性。实验结果表明,合成图像的峰值信噪比(PSNR)均超过24 dB,优于现有主流方法,且在视觉质量上达到最优水平。

**【关键词】**磁共振成像;多模态图像合成;生成对抗网络;特征一致性

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2026)04-0436-09

## Multi-branch adaptive fusion for multimodal magnetic resonance image synthesis

QIU Hongxuan<sup>1,2</sup>, LIU Mingyang<sup>3</sup>, LIU Surui<sup>2</sup>, WANG Yuxi<sup>1,2</sup>, PENG Bo<sup>2</sup>, DAI Yakang<sup>2</sup>, YU Weibo<sup>1</sup>

1. School of Electrical and Electronic Engineering, Changchun University of Technology, Changchun 130012, China; 2. Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China; 3. School of Electrical and Mechanical Engineering, Jilin University of Architecture and Technology, Changchun 130114, China

**Abstract:** Multimodal magnetic resonance imaging provides comprehensive and essential information for clinical diagnosis and treatment, and synthesizing missing modality images can significantly improve medical analysis. To address the limitations of existing image synthesis methods, including fixed input modalities, suboptimal image quality, and insufficient anatomical fidelity in synthesized images, a multi-branch modality-adaptive fusion and consistency-guided generative adversarial network is proposed for multimodal magnetic resonance image synthesis. Within this network, efficient medical image synthesis from arbitrary modality combinations is achieved through the incorporation of target modality labels, a zero-image placeholder strategy, and a multi-branch encoding mechanism. Specifically, target modality labels are used to guide the decoder to focus on desired features, a zero-image placeholder strategy is employed to address incomplete input modalities, and a multi-branch encoding mechanism is adopted to ensure independent feature extraction and flexible cross-modality fusion, thereby enhancing the model's adaptability to multimodal inputs and its synthesis performance. Furthermore, a modality consistency guidance mechanism is introduced to align encoded features from different modalities in the latent space, reinforcing cross-modal anatomical consistency and thus improving the anatomical fidelity and overall quality of synthesized images. The effectiveness of the proposed method is validated on the BraTS2020 and ISLES2015 datasets. Experimental results demonstrate that the synthesized images achieve a peak signal-to-noise ratio exceeding 24 dB, outperforming the existing image synthesis methods and exhibiting superior visual quality.

**Keywords:** magnetic resonance imaging; multimodal image synthesis; generative adversarial networks; feature consistency

**【收稿日期】**2025-11-21

**【基金项目】**国家自然科学基金(62471467, 62301557);江苏省重点研发计划(BE2022049-2);中国科学院磁共振技术联盟项目(2024GZL001);苏州市重点实验室项目(SZS2024007)

**【作者简介】**邱宏轩, 硕士, 研究方向: 医学图像处理, E-mail: 15584147870@163.com; 刘明洋, 硕士, 研究方向: 医学图像处理, E-mail: liumingyang06@163.com (邱宏轩和刘明洋为共同第一作者)

**【通信作者】**于微波, 硕士, 教授, 研究方向: 机器视觉与图像处理, E-mail: yuweibo@ccut.edu.cn; 戴亚康, 博士, 研究员, 研究方向: 医学图像处理, E-mail: daiyk@sibet.ac.cn

## 前言

磁共振成像(Magnetic Resonance Imaging, MRI)多模态数据在疾病诊疗中具有重要临床价值,不同模态的图像具有各自独特的成像原理与信息优势<sup>[1]</sup>,多模态MRI通过 $T_1$ 加权成像、 $T_2$ 加权成像、液体衰减反转恢复(Flair)以及弥散加权成像(DWI)等序列的互补信息,为病变表征提供更全面的数据支持<sup>[2]</sup>。但实际扫描常因设备差异、患者耐受度等因素导致模态缺失,直接影响计算机辅助诊断系统的鲁棒性<sup>[3]</sup>。如何在动态缺失场景下重建高质量的多模态图像,成为提升诊疗效率的核心挑战。

现有多模态图像合成方法普遍存在两大局限:其一,传统模型通常依赖固定模态组合(如 $T_1+T_2+Flair$ 合成 $T_1Gd$ ),模型架构和训练流程均基于完整模态数据设计<sup>[4]</sup>,难以应对临床中模态缺失情况多变的实际场景。一旦输入模态组合发生变化,往往需重新训练模型,缺乏灵活性和泛化能力;其二,不同模态图像在灰度分布和结构表现上存在较大差异,易导致模型在融合过程中出现结构表达偏差,影响合成图像的解剖一致性。传统方法常采用像素级损失优化局部强度差异<sup>[5]</sup>,难以有效建模和约束模态间的结构对应关系,导致生成图像在纹理细节和结构保真度方面表现欠佳。

为解决上述问题,本研究提出一种基于多分支自适应融合与模态一致性引导的多模态磁共振图像合成方法(Multi-branch Modality-adaptive Fusion and Consistency-guided Generative Adversarial Network, MMFC-GAN)。该方法通过多分支编码结构增强不同模态特征的独立建模与灵活融合能力,引入目标模态标签引导生成过程聚焦于目标模态关键特征,提升合成图像的保真度。结合零图像占位策略,使模型在不改变结构的前提下适应任意模态缺失情形,提升临床适用性。同时,引入模态一致性引导机制,通过特征一致性损失对齐多模态结构信息,增强生成图像的结构一致性与视觉质量。实验结果表明,该方法在脑肿瘤与脑卒中数据集上均表现出优异的结构保真与纹理还原能力,为多模态医学影像的高质量补全提供了可行路径。

## 1 相关工作

### 1.1 基于深度学习的图像合成方法

随着深度学习技术的不断发展,生成对抗网络(Generative Adversarial Networks, GAN)<sup>[6]</sup>在医学图像中的应用展现出显著优势<sup>[7-12]</sup>。基于GAN的生成模型通过对抗训练机制展现出高质量的图像重构能力:Pix2Pix利用条件生成对抗网络架构实现输入与

目标图像的像素级精准映射<sup>[13]</sup>。CycleGAN引入循环一致性损失构建无配对数据的图像生成模型,突破了传统方法对成对训练数据的刚性依赖,实现跨域的图像转换<sup>[14]</sup>。此类方法在磁共振(MR)图像合成中取得初步成效,如陈佳等<sup>[15]</sup>利用GAN结合自动编码器实现MR图像的有效合成,但其单模态输入机制难以应对多模态缺失场景,难以满足临床应用的多样化需求。

### 1.2 多模态特征融合方法

在医学图像生成领域,确保高质量输出与结构一致性是复杂脑组织及多样病变图像合成的核心挑战<sup>[16-17]</sup>。研究者通过改进网络架构与损失函数提升多模态合成质量。Hu等<sup>[18]</sup>引入差异损失以约束未扫描数据点,增强跨模态映射稳定性;Bi等<sup>[19]</sup>构建的PT-GAN框架通过生成器-判别器自编码结构实现自监督约束,显著提升多模态信息融合效率;Cao等<sup>[20]</sup>开发多分支编码器联合协作判别机制,强化目标特征学习能力。尽管这些方法在多模态融合方面取得进展,但固定的模态输入难以适应临床模态缺失情况多变的场景。MM-GAN通过多模态联合生成提升图像的整体质量与准确性,但其方法未充分考虑不同模态间在灰度与结构表征方面的差异,缺乏对模态间结构一致性的显式建模,容易在跨模态特征融合过程中出现结构对齐误差,从而导致生成图像在细节层面存在模糊或失真现象<sup>[21]</sup>。如何在确保解剖结构准确还原的前提下实现高质量图像合成,仍是当前多模态图像生成方法亟需解决的关键问题,亦是临床诊断和科研应用中的核心需求。

针对缺失模态图像的合成问题,本研究从两个层面进行创新:一是设计多分支编码架构,使每个模态拥有独立的编码分支,网络可以根据每个模态的特征进行独立处理,并在融合阶段通过进行信息整合;结合零图像占位策略使用零图像填充缺失模态,并结合目标模态索引,使得引导网络在有缺失模态时依然能够完成合理的特征融合,提升对不同模态输入组合的适应能力。二是引入模态一致性引导机制,通过特征一致性损失约束潜在空间映射关系<sup>[22]</sup>,强化模态间结构信息的一致性建模,从而实现缺失模态图像的高质量生成。

## 2 方法

本研究提出的MMFC-GAN网络包含一个生成网络(Generator network)和一个判别网络(Discriminator network)。生成网络由多分支编码器Multi\_ $E_n$ 、目标模态标签编码器 $E_t$ 及解码器 $D_n$ 构成;判别网络采用PatchGAN架构来构建编码器 $E_d$ ,通过局部图像块评估真实性。图1展示MMFC-GAN的整体网络框架。

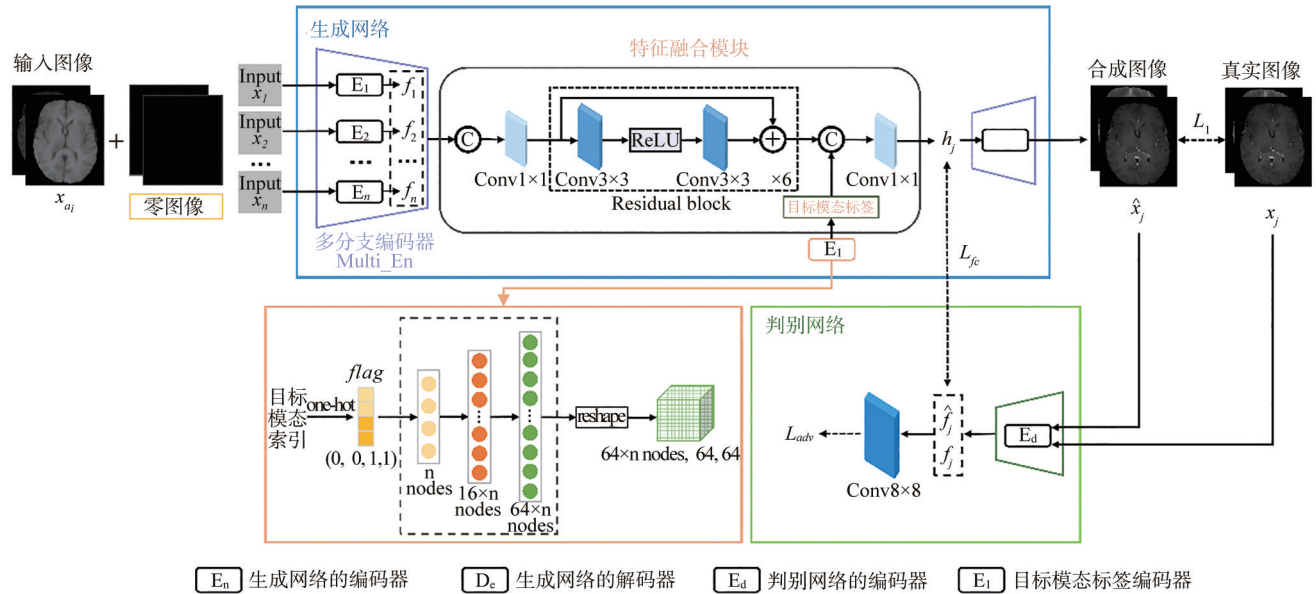


图1 MMFC-GAN整体网络框架  
Figure 1 Overall network framework of MMFC-GAN

### 2.1 网络结构

生成网络采用自适应多分支编码-解码架构,实现从任意输入模态组合到目标模态的灵活映射。生成网络的多分支编码器 Multi\_En 由多个独立分支组成,通过多分支结构能够独立建模不同模态的特征表示,确保每个模态的信息在融合前保持其原始特性。由于输入子集中存在模态数据缺失,无法保证所有编码器均能接收到有效输入,本研究采用零图像占位策略对缺失模态进行补全。具体而言,零图像定义为与真实输入图像在空间尺寸和通道数一致的  $H \times W \times 1$  全零矩阵,在训练阶段,当某模态数据缺失时,使用该零图像作为占位输入送入对应模态的编码器,以保持多分支结构的输入一致性,防止因输入缺失导致参数更新异常。此外,为增强模态间独立性,各模态编码器配备独立的权重矩阵与偏置向量,并通过输入模态索引引导图像进入对应的编码器。这种设计使每个分支能够针对特定模态的特征分布进行独立参数学习,从而保证所提取特征的模态特异性,有效提升特征提取的准确性与鲁棒性。每个分支编码器  $E_n$  都包含一个  $7 \times 7$  卷积层(步长1,填充3),与两个  $3 \times 3$  卷积层(步长2,填充1),每个  $H \times W \times 1$  的输入图像经编码器处理,得到  $H/4 \times W/4 \times 256$  的特征图。训练过程中,输入模态的组合从完整数据集  $x_k, \forall k \in [1, N]$  中随机选择输入模态组合  $x_a, \forall i \in [1, M]$ , 其中  $M(0 < M \leq N)$ ,  $N$  为完整模态数量。以增强模型对多样化输入的适应能力。

为实现多模态特征的有效融合,将各编码器提取

的特征沿通道维度拼接后,首先通过一个  $1 \times 1$  卷积层调整通道数进行降维,随后输入6个残差模块(每个模块包含两个  $3 \times 3$  卷积层),以提取更深层次的特征并增强模态间的关联性,获得多模态融合特征。为了提升模型在不同模态缺失场景下的生成能力,特征融合部分引入目标模态标签进行引导。首先将目标模态索引经过 one-hot 编码,生成一个  $1 \times 4$  的二进制 flag 向量,其中1表示模态可用,0表示模态缺失需补全。该 flag 向量被输送到目标模态标签编码器,该编码器由3个全连接层组成,节点数依次为  $N$ (输入层)、 $16N$ (隐藏层, ReLU 激活)、 $64N$ (输出层, Sigmoid 激活),最终输出一个与特征通道数相同的权重向量。该权重向量会被扩展为与特征图相同大小的权重图,并与多模态特征图进行逐通道逐元素相乘。由此可以根据模态是否缺失自动调整其特征的参与程度:缺失模态的特征会被减弱,可用模态的特征则被保留。通过 flag 向量的目标模态标签标识驱动权重自适应分布,确保解码器在生成过程中能够聚焦于有效模态的特征关联,从而实现缺失模态图像的高质量合成。将融合特征与目标模态标签拼接后,再经过  $1 \times 1$  卷积层,随后驱动解码器通过三层转置卷积将特征图恢复到原始图像尺寸,最终经 Tanh 激活输出归一化到  $[-1, 1]$  的合成图像  $\hat{X} = \{\hat{x}_j\}_{j=1}^{N-M}$ 。

判别网络采用 Pix2pix 中提出的 PatchGAN 结构来提升合成图像的真实性<sup>[13]</sup>。PatchGAN 将输入图像划分为  $30 \times 30$  的图像块,通过对每个图像块进行真伪概率判别,有效提升局部结构真实性的评估能力。

判别网络首先对合成图像和真实目标模态图像进行编码,生成各自的特征表示 $\hat{f}_j$ 和 $f_j$ ,并对其进行真实性评估。其用于特征提取的编码器 $E_d$ 的架构与生成网络的编码结构保持一致,当输入的生成图像和缺失的目标模态图像经过特征提取后,送入一层卷积核大小为 $8 \times 8$ (步长为2)的卷积层,通过扩大感受野强化对生成图像真实性的判别能力。

## 2.2 损失函数

为提升合成图像的结构保真度与合成质量,本文引入一种模态一致性引导机制,具体而言:通过特征一致性损失(Feature Consistency Loss,  $L_{fc}$ )强化模态间结构信息的一致性建模,在潜在空间对不同模态的编码特征进行对齐,增强合成图像与真实目标模态图像间的结构一致性。设输入图像特征经编码后的多模态信息融合特征为 $H = \{h_j\}_{j=1}^{N-M}$ 、合成图像特征 $\hat{F} = \{\hat{f}_j\}_{j=1}^{N-M}$ ,以及真实目标模态图像特征 $F = \{f_j\}_{j=1}^{N-M}$ ,特征一致性损失函数定义如下:

$$L_{fc} = \frac{1}{3(N-M)} \sum_j \left[ \|f_j - \hat{f}_j\|_2 + \|f_j - h_j\|_2 + \|\hat{f}_j - h_j\|_2 \right] \quad (1)$$

其 $\|\cdot\|_2$ 为 $L_2$ 范数。该损失通过量化输入特征与合成和真实特征间的差异,使合成图像在特征层面与目标模态对齐,从而强化了不同图像间信息的一致性。

为增强像素级的保真度,使用 $L_1$ 范数计算生成损失,通过计算真实目标模态图像 $X = \{x_j\}_{j=1}^{N-M}$ 与合成图像 $\hat{X} = \{\hat{x}_j\}_{j=1}^{N-M}$ 间的像素差异,生成损失如公式(2)所示:

$$L_1 = \frac{1}{N-M} \sum_j E \left[ \|x_j - \hat{x}_j\|_1 \right] \quad (2)$$

引入判别器 $D$ 强化生成图像的全局真实性。其对抗损失如式(3)所示:

$$L_{adv} = \frac{1}{N-M} \sum_j E \left[ D(\hat{x}_j)^2 + (D(x_j) - 1)^2 \right] \quad (3)$$

最终综合采用特征一致性损失、生成损失及对抗损失的加权和形式,以协同约束生成网络的学习过程。定义如下:

$$L_{total} = L_1 + \lambda_1 L_{fc} + \lambda_2 L_{adv} \quad (4)$$

其中, $\lambda_1, \lambda_2$ 分别为特征一致性损失和对抗性损失的权重。

## 3 实验结果与分析

本研究实验在搭载Ubuntu操作系统的Linux服务器上完成,硬件配置为NVIDIA GeForce RTX 3090 GPU(24 GB显存),基于PyTorch框架实现算法开发。参数配置如下:初始化学习率设为 $2 \times 10^{-4}$ ,批处理规

模(batch size)为4,参数 $\lambda_1$ 和 $\lambda_2$ 的值分别设置为0.1和0.01。使用Adam优化器实现参数优化<sup>[23]</sup>。

### 3.1 实验数据

本研究选取BraTS2020数据集<sup>[24]</sup>和ISLES2015数据集<sup>[25]</sup>来评估所提出方法的有效性。BraTS2020数据集中每例包含4种MRI模态: $T_1, T_1Gd, T_2$ 和Flair,每个模态图像大小为 $240 \times 240 \times 155$ 。ISLES2015数据集采用SSIS子集,其中每例包含 $T_1, T_2, Flair$ 和DWI 4种MRI模态,每个模态图像大小为 $240 \times 240 \times 155$ 。均按8:2划分训练集与测试集。

实验采用统一预处理流程:使用体积图像的横轴面切片作为输入。将所有横轴面切片采用反射填充将图片尺寸统一调整为 $256 \times 256$ ,并将所有切片归一化处理。

### 3.2 评价指标

为评估模型性能,选用峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似度指数(Structural Similarity Index Measurement, SSIM)<sup>[26]</sup>以及感知图像相似度(Learned Perceptual Image Patch Similarity, LPIPS)<sup>[27]</sup>作为评价指标。三者从不同维度量化生成图像质量:PSNR基于均方误差度量像素级差异,SSIM通过亮度、对比度、结构三要素评估图像结构相似性,LPIPS计算图像块的感知相似性,更能反映人类感知。计算公式如下:

$$PSNR(x,y) = 10 \log_{10} \frac{\max^2(x,y)}{\frac{1}{N} \sum \|x - y\|_2^2} \quad (5)$$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\delta_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\delta_x^2 + \delta_y^2 + c_2)} \quad (6)$$

$$LPIPS(x,y) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|W_l \odot (\hat{x}_{hw}^l - \hat{y}_{hw}^l)\|_2^2 \quad (7)$$

### 3.3 实验结果与分析

**3.3.1 不同方法对比分析** 为了评估本研究方法MMFC-GAN的图像合成性能,选择两种先进的方法:PT-GAN<sup>[19]</sup>和MM-GAN<sup>[21]</sup>进行对比。所有方法均由3种模态图像合成单模态图像,共4种模态合成结果。在BraTS2020和ISLES2015数据集上有3种模态合成单模态任务的定量评估结果,如表1和表2所示,使用PSNR指标进行评测。

结果显示,MMFC-GAN在大多数模态中表现最佳,在BraTS2020数据集上的 $T_1, T_1Gd, Flair$ 和 $T_2$ 模态分别获得了26.93、29.76、27.18和27.29 dB的最佳PSNR分数,在ISLES2015数据集上的 $T_1, DWI, Flair$ 和 $T_2$ 模态分别获得了25.69、29.25、29.32和25.58 dB的最佳PSNR分数。相比之下,MMFC-GAN在跨数

表 1 BraTS2020 数据集 3 模态图像合成单模态图像对比方法的 PSNR 定量评价结果 (dB)

Table 1 PSNR quantitative results of comparative methods for three-modal to single-modal image synthesis on the BraTS2020 dataset (dB)

方法	T <sub>1</sub>	T <sub>1</sub> Gd	Flair	T <sub>2</sub>
PT-GAN	26.08	28.61	26.42	25.06
MM-GAN	25.89	29.18	23.85	25.41
MMFC-GAN	26.93	29.76	27.18	27.29

表 2 ISLES2015 数据集 3 模态图像合成单模态图像对比方法的 PSNR 定量评价结果 (dB)

Table 2 PSNR quantitative results of comparative methods for three-modal to single-modal image synthesis on the ISLES2015 dataset (dB)

方法	T <sub>1</sub>	DWI	Flair	T <sub>2</sub>
PT-GAN	24.01	28.78	28.73	24.29
MM-GAN	23.54	28.42	27.16	23.71
MMFC-GAN	25.69	29.25	29.32	25.58

据集和不同模态方面表现出更强的鲁棒性和合成能力。不同对比方法在 BraTS2020 和 ISLES2015 数据集上的可视化结果如图 2 所示。本研究方法在标记

的病灶区域较对比方法优势显著,边界更清晰锐利、纹理细节更真实,且有效抑制了伪影,验证了对病灶核心区域的精准重建能力。

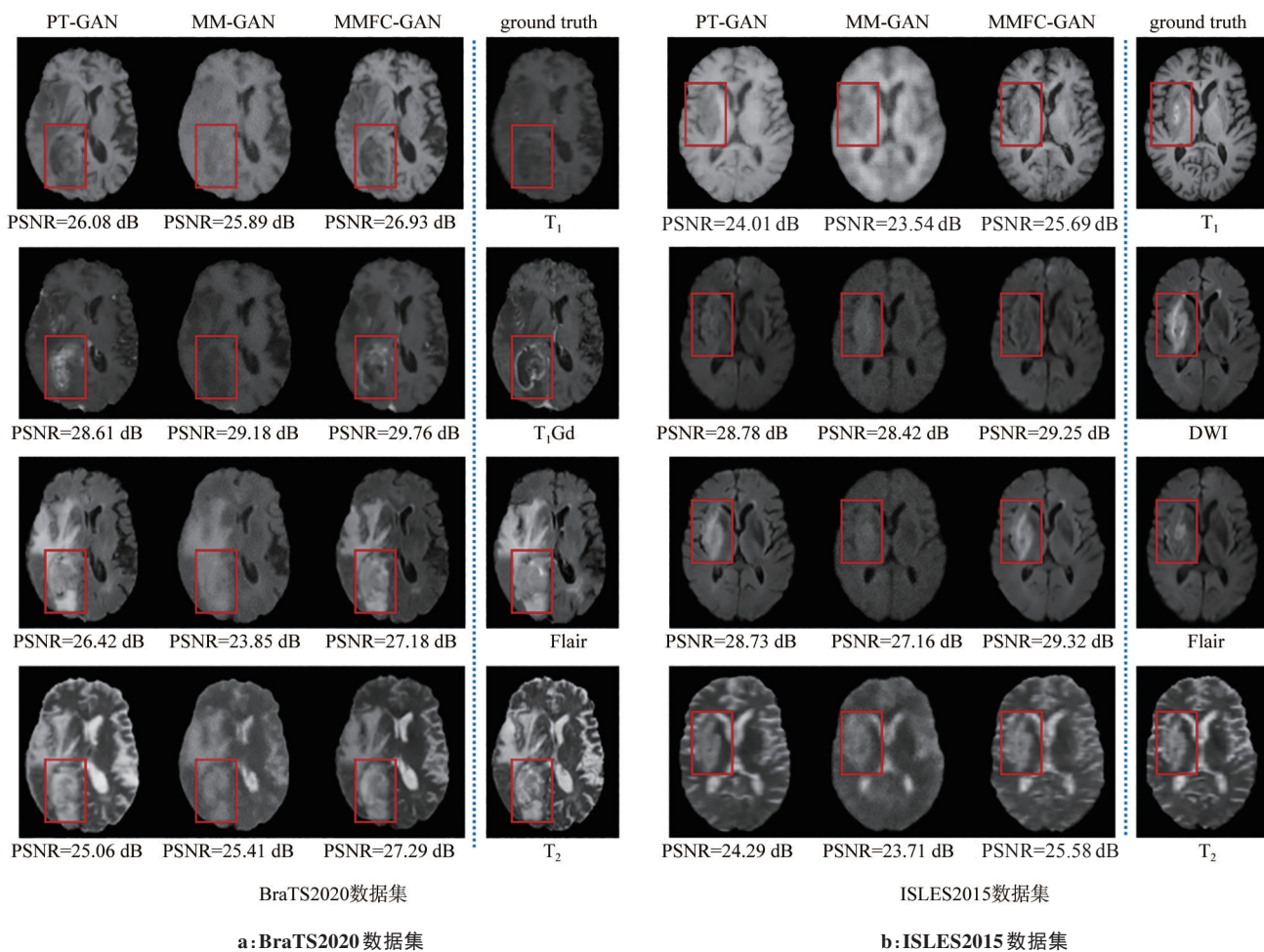


图 2 3 模态图像合成单模态图像任务不同方法的可视化对比结果

Figure 2 Visual comparison of different methods for three-modal to single-modal image synthesis

3.3.2 不同任务对比分析 为了验证 MMFC-GAN 的泛化性,本文在 BraTS2020 和 ISLES2015 数据集上构建 3 类跨模态合成任务(单→3 模态、双→双模态、3→单模态)。定量评估结果见表 3~4,图 3~4 展示了

MMFC-GAN 在不同任务上合成图像的可视化效果。

在 BraTS2020 和 ISLES2015 数据集上的实验结果表明,多模态输入显著提升合成图像质量。在 BraTS2020 数据集中,T<sub>1</sub>+Flair+T<sub>2</sub>多模态生成 T<sub>1</sub>Gd 模态图像 PSNR

表3 MMFC-GAN在BraTS2020数据集对不同生成任务进行定量评估结果(PSNR/SSIM/LPIPS)

Table 3 Quantitative evaluation results of MMFC-GAN for different generation tasks on the BraTS2020 dataset (PSNR/SSIM/LPIPS)

输入模态				输出模态			
T <sub>1</sub>	T <sub>1</sub> Gd	Flair	T <sub>2</sub>	T <sub>1</sub>	T <sub>1</sub> Gd	Flair	T <sub>2</sub>
√				-	29.15 dB/0.9544/0.0815	25.61 dB/0.9299/0.0958	25.98 dB/0.9361/0.0802
	√			26.20 dB/0.9544/0.0644	-	25.75 dB/0.9269/0.0998	25.92 dB/0.9334/0.0822
		√		25.71 dB/0.9448/0.0736	28.45 dB/0.9470/0.0972	-	25.84 dB/0.9288/0.0830
			√	25.69 dB/0.9500/0.0730	28.71 dB/0.9515/0.0902	25.85 dB/0.9299/0.0958	-
		√	√	26.20 dB/0.9545/0.0601	28.77 dB/0.9515/0.0751	-	-
√			√	26.66 dB/0.9583/0.0540	-	26.93 dB/0.9372/0.0779	-
	√			26.76 dB/0.9576/0.0534	-	-	26.73 dB/0.9405/0.0662
√			√	-	29.46 dB/0.9564/0.0690	26.61 dB/0.9381/0.0779	-
√		√		-	29.27 dB/0.9553/0.0698	-	26.69 dB/0.9422/0.0668
√	√			-	-	26.32 dB/0.9331/0.0816	26.19 dB/0.9384/0.0683
	√	√	√	26.93 dB/0.9611/0.0510	-	-	-
√		√	√	-	29.76 dB/0.9591/0.0701	-	-
√	√		√	-	-	27.18 dB/0.9434/0.0791	-
√	√	√		-	-	-	27.29 dB/0.9473/0.0635

表4 MMFC-GAN在ISLES2015数据集对不同生成任务进行定量评估结果(PSNR/SSIM/LPIPS)

Table 4 Quantitative evaluation results of MMFC-GAN for different generation tasks on the ISLES2015 dataset (PSNR/SSIM/LPIPS)

输入模态				输出模态			
T <sub>1</sub>	DWI	Flair	T <sub>2</sub>	T <sub>1</sub>	DWI	Flair	T <sub>2</sub>
√				-	28.15 dB/0.9333/0.0686	28.09 dB/0.9119/0.0956	24.93 dB/0.8924/0.0998
	√			23.89 dB/0.8972/0.0813	-	27.75 dB/0.9206/0.0941	23.96 dB/0.8733/0.1106
		√		23.83 dB/0.9119/0.0721	29.08 dB/0.9458/0.0845	-	24.55 dB/0.8854/0.0967
			√	23.69 dB/0.8936/0.0886	27.50 dB/0.9179/0.0917	27.25 dB/0.9032/0.1140	-
		√	√	24.59 dB/0.9192/0.0703	29.25 dB/0.9468/0.0549	-	-
√			√	24.47 dB/0.9112/0.0739	-	28.72 dB/0.9339/0.0974	-
	√			24.07 dB/0.9161/0.0711	-	-	25.19 dB/0.8976/0.1052
√			√	-	28.38 dB/0.9354/0.0676	28.78 dB/0.9297/0.1021	-
√		√		-	29.13 dB/0.9468/0.0474	-	25.50 dB/0.9023/0.0992
√	√			-	-	28.77 dB/0.9353/0.0883	25.02 dB/0.8949/0.1011
	√	√	√	25.69 dB/0.9197/0.0676	-	-	-
√		√	√	-	29.25 dB/0.9485/0.0560	-	-
√	√		√	-	-	29.32 dB/0.9396/0.0850	-
√	√	√		-	-	-	25.58 dB/0.9022/0.1001

达29.76 dB,相比于单模态输入(如T<sub>1</sub>/T<sub>2</sub>/Flair生成T<sub>1</sub>Gd模态图像,PSNR值范围为28.45~29.15 dB)显著提高了图像质量,在标记的病灶区域也更加清晰可信;T<sub>1</sub>Gd与T<sub>1</sub>/T<sub>2</sub>/Flair等组合也表现出更高的PSNR值,进一步验证了多模态图像合成的有效性。类似地,在ISLES2015数据集中,T<sub>1</sub>+Flair+T<sub>2</sub>多模态合成任务的

PSNR值达到了29.25 dB,高于T<sub>1</sub>单模态任务的PSNR值(24.93~28.15 dB)。DWI与T<sub>1</sub>/T<sub>2</sub>/Flair组合的多模态(如T<sub>1</sub>+DWI+T<sub>2</sub>任务,PSNR值为29.32 dB)同样展示出更高的PSNR值,在SSIM和LPIPS评价上也表现最佳。以上结果表现出多模态融合在提升图像质量方面的显著优势。

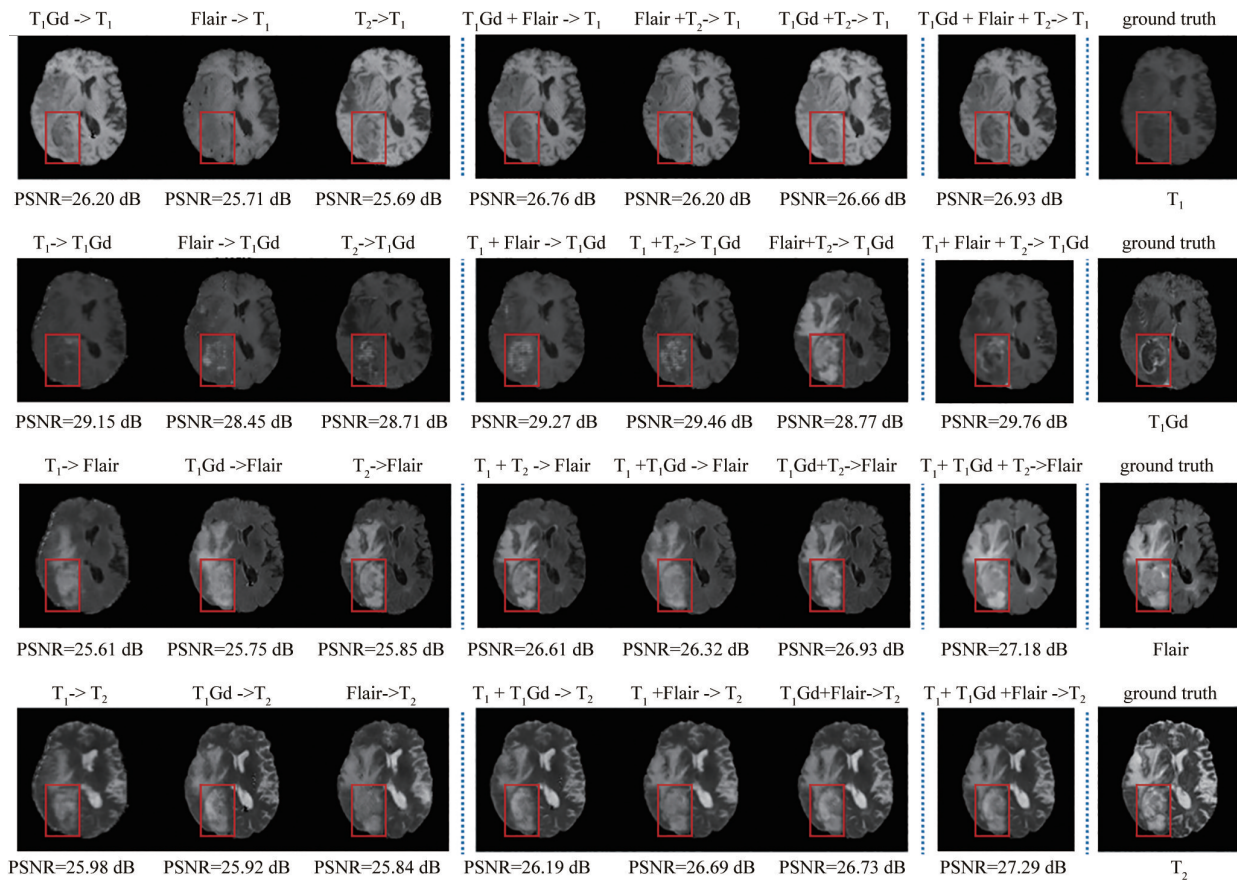


图3 MMFC-GAN在BraTS2020数据集上对不同任务的可视化结果

Figure 3 Visualization results of MMFC-GAN for different tasks on the BraTS2020 dataset

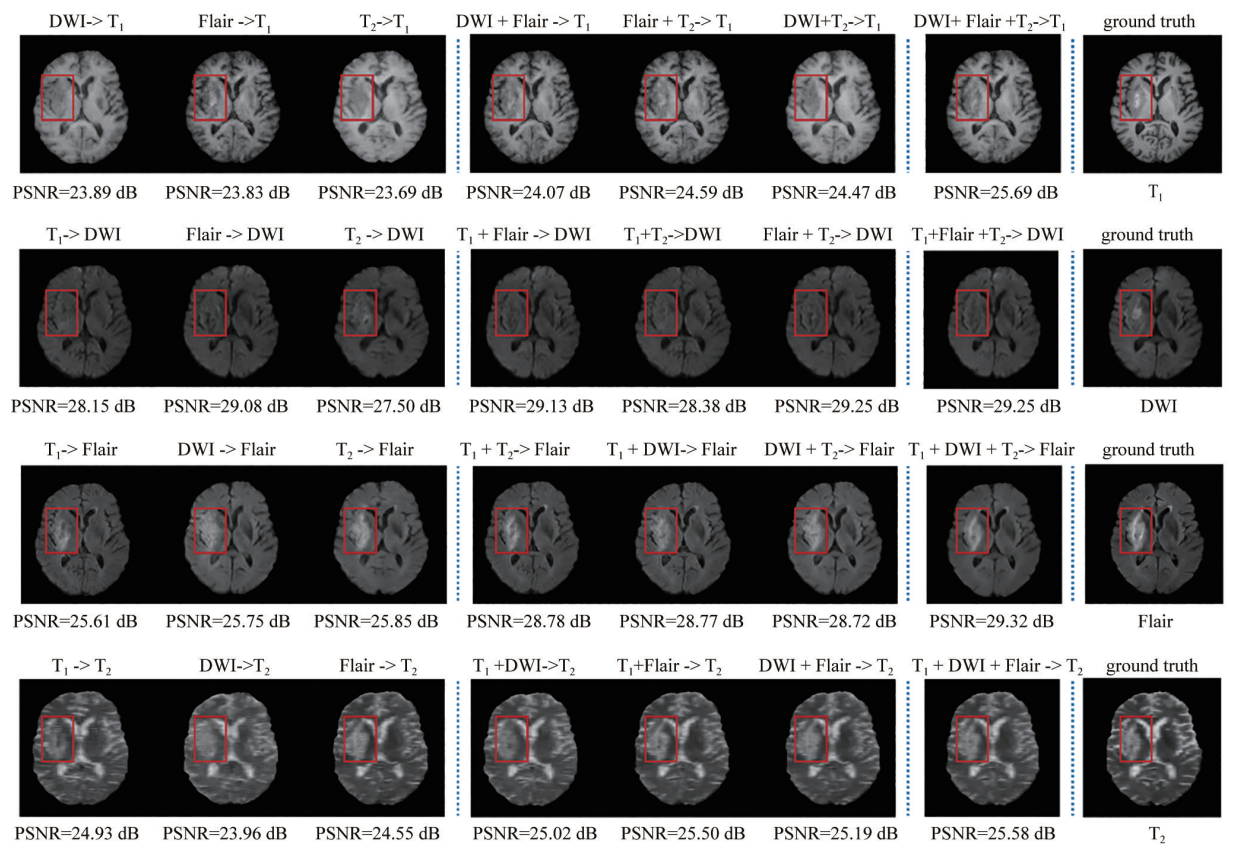


图4 MMFC-GAN在ISLES2015数据集上对不同任务的可视化结果

Figure 4 Visualization results of MMFC-GAN for different tasks on the ISLES2015 dataset

**3.3.3 组合损失的定量分析** 为系统评估损失函数组合对多模态医学图像合成的影响,本研究在BraTS2020与ISLES2015数据集上进行损失函数组合的消融实验。联合像素重建损失 $L_1$ 、特征一致性损失 $L_{fc}$ 和对抗损失 $L_{adv}$ 的优化效果源于多层次约束的互补机制: $L_1$ 像素级绝对误差约束全局结构对齐, $L_{fc}$ 迫使图像间潜在空间特征分布匹配以保留高频细节,而 $L_{adv}$ 通过对抗训练驱动生成图像服从目标模态的数据分布特性以增强局部真实性。三者构建起“像素-特征-分布”

的层次化优化链,实现合成图像在解剖结构、细节保真度与模态特异性上的多维度提升。实验表明,在BraTS2020数据集上,联合损失策略使合成 $T_1$ Gd模态图像的PSNR达到29.47 dB,SSIM与LPIPS分别提升至0.957和0.072;在ISLES2015数据集上该组合同样实现最优性能。具体实验结果如表5所示。3项指标的一致性优化证实了合成图像在PSNR、SSIM和LPIPS三个维度上均接近真实图像分布,表现出引入特征一致性损失的有效性。

表5 在BraTS2020数据集和ISLES2015数据集上3模态输入到单模态输出的组合损失函数消融实验结果  
Table 5 Results of the ablation study on the combined loss function for three-modal input to single-modal output on the BraTS2020 and ISLES2015 datasets

方法	$L_1$	$L_{fc}$	$L_{adv}$	BraTS2020数据集				ISLES2015数据集			
				$T_1$	$T_1$ Gd	Flair	$T_2$	$T_1$	DWI	Flair	$T_2$
PSNR/dB	√		√	26.64	29.08	26.60	26.47	24.23	28.86	28.59	25.00
	√	√		26.11	28.93	26.45	26.58	24.31	28.93	28.61	25.27
	√	√	√	27.03	29.47	27.08	27.06	24.43	29.00	28.86	25.46
SSIM	√		√	0.956	0.952	0.933	0.939	0.956	0.952	0.933	0.939
	√	√		0.953	0.952	0.934	0.940	0.953	0.952	0.934	0.940
	√	√	√	0.960	0.957	0.942	0.947	0.960	0.957	0.942	0.947
LPIPS	√		√	0.057	0.075	0.078	0.071	0.070	0.058	0.087	0.093
	√	√		0.059	0.074	0.079	0.072	0.077	0.063	0.098	0.107
	√	√	√	0.056	0.072	0.077	0.067	0.069	0.057	0.086	0.092

## 4 结论

本研究提出了一种多模态磁共振图像合成方法MMFC-GAN,通过多分支自适应融合和模态一致性引导机制,针对临床中常见的模态缺失问题提供有效解决方案。在BraTS2020与ISLES2015数据集上的实验验证表明,该方法能够充分融合多模态信息,生成结构一致且质量高的合成图像。网络设计方面,采用多分支编码器架构,实现对各模态图像信息的有效整合;结合零图像占位策略,灵活处理缺失模态,支持从单模态到全模态的多种输入组合,突破了传统方法对固定模态输入的限制。此外,模态一致性引导机制通过特征一致性损失,在潜在空间对不同模态特征进行约束,显著提升了合成图像的质量。本研究为缺失模态条件下的临床辅助诊断提供了有力支持。未来工作将聚焦于开发能够在合成过程中关注病灶区域的网络结构,并探索方法向三维图像合成的扩展,以进一步提升模型的临床应用价值和多模态融合效果。

## 【参考文献】

- [1] Ting H, Liu MH. Multimodal transformer of incomplete MRI data for brain tumor segmentation[J]. IEEE J Biomed Health Inform, 2024, 28(1): 89-99.
- [2] 张裕辉,杨颖思,范伟雄,等. 基于多模态常规MRI的影像组学预测胶质瘤术后早期复发的价值[J]. 中国医学物理学杂志, 2025, 42(2): 208-212.  
Zhang YH, Yang YS, Fan WX, et al. Predictive value of multi-modal conventional MRI radiomics for early postoperative glioma recurrence [J]. Chinese Journal of Medical Physics, 2025, 42(2): 208-212.
- [3] 徐盼盼,张东,袁达龙. 基于生成对抗网络的多模态MR图像缺失模态合成[J]. 中国医学物理学杂志, 2023, 40(7): 827-832.  
Xu PP, Zhang D, Yuan DL. Incomplete multimodality MR image synthesis based on generative adversarial network[J]. Chinese Journal of Medical Physics, 2023, 40(7): 827-832.
- [4] Albekairi M, Mohamed MVO, Kaaniche K, et al. Multimodal medical image fusion combining saliency perception and generative adversarial network[J]. Sci Rep, 2025, 15(1): 10609.
- [5] Aliniya P, Nicolescu M, Nicolescu M, et al. Hybrid region and pixel-level adaptive loss for mass segmentation on whole mammography images[C]//Advances in Visual Computing. Cham: Springer Nature Switzerland, 2023: 3-17.
- [6] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Commun ACM, 2020, 63(11): 139-144.
- [7] 常晓,蔡昕,杨光,等. 生成对抗网络在医学图像转换领域的应用[J]. 波谱学杂志, 2022, 39(3): 366-380.  
Chang X, Cai X, Yang G, et al. Applications of generative adversarial networks in medical image translation [J]. Chinese Journal of Magnetic Resonance, 2022, 39(3): 366-380.

- [8] Heng Y, Yinghua M, Khan FG, et al. Survey: application and analysis of generative adversarial networks in medical images[J]. *Artif Intell Rev*, 2024, 58(2): 39.
- [9] Skandarani Y, Jodoin PM, Lalonde A. GANs for medical image synthesis: an empirical study[J]. *J Imaging*, 2023, 9(3): 69.
- [10] Jeong JJ, Tariq A, Adejumo T, et al. Systematic review of generative adversarial networks (GANs) for medical image classification and segmentation[J]. *J Digit Imaging*, 2022, 35(2): 137-152.
- [11] Lee TH, Munasinghe V, Li YM, et al. GAN-based medical image registration for augmented reality applications[C]//2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS). Piscataway, NJ, USA: IEEE, 2022: 279-282.
- [12] Zhang PH, Li JY, Wang YN, et al. Domain adaptation for medical image segmentation: a meta-learning method[J]. *J Imaging*, 2021, 7(2): 31.
- [13] 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). Piscataway, NJ, USA: IEEE, 2017: 5967-5976.
- [14] 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). Piscataway, NJ, USA: IEEE, 2017: 2242-2251.
- [15] 陈佳, 骆爽, 何儒汉, 等. 基于生成对抗网络的MR图像合成研究[J]. *计算机应用研究*, 2021, 38(5): 1563-1566.  
Chen J, Luo S, He RH, et al. Research on MR image synthesis based on generative adversarial network [J]. *Application Research of Computers*, 2021, 38(5): 1563-1566.
- [16] Wolterink JM, Dinkla AM, Savenije MHF, et al. Deep MR to CT synthesis using unpaired data [C]//Simulation and Synthesis in Medical Imaging. Cham: Springer International Publishing, 2017: 14-23.
- [17] 李思宁, 毕佳俊, 于瀛, 等. 失眠相关脑神经及血管改变的功能性磁共振研究进展[J]. *神经解剖学杂志*, 2025, 1(2): 251-254.  
Li SN, Bi JJ, Yu Y, et al. Research advances in functional magnetic resonance imaging on neurologic and vascular aspects of insomnia[J]. *Chinese Journal of Neuroanatomy*, 2025, 1(2): 251-254.
- [18] Hu C, Li C, Wang HF, et al. Self-supervised learning for MRI reconstruction with a parallel network training framework [C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2021. Cham: Springer International Publishing, 2021: 382-391.
- [19] Bi ZW, Cao B, Zuo WM, et al. Learning a prototype discriminator with RBF for multimodal image synthesis[J]. *IEEE Trans Image Process*, 2022, 31: 6664-6678.
- [20] Cao B, Zhang H, Wang NN, et al. Auto-GAN: self-supervised collaborative learning for medical image synthesis[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence and the Thirty-Second Conference on Innovative Applications of Artificial Intelligence and the Tenth Symposium on Educational Advances in Artificial Intelligence. Palo Alto, CA, USA: AAAI Press, 2020: 10486-10493.
- [21] Sharma A, Hamarneh G. Missing MRI pulse sequence synthesis using multi-modal generative adversarial network [J]. *IEEE Trans Med Imaging*, 2020, 39(4): 1170-1183.
- [22] Sun FC, Kong T, Huang WB, et al. Feature pyramid reconfiguration with consistent loss for object detection [J]. *IEEE Trans Image Process*, 2019, 28(10): 5041-5051.
- [23] Kingma DP, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2017-01-30). <https://arxiv.org/abs/1412.6980>.
- [24] Menze BH, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS)[J]. *IEEE Trans Med Imaging*, 2015, 34(10): 1993-2024.
- [25] Maier O, Menze BH, von der Gabelntz J, et al. ISLES 2015 - a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI[J]. *Med Image Anal*, 2017, 35: 250-269.
- [26] Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Trans Image Process*, 2004, 13(4): 600-612.
- [27] Zhang R, Isola P, Efros AA, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 586-595.

(编辑:薛泽玲)