

✧ 医学物理与工程学

Unified modality-aware denoising diffusion probabilistic model for detecting brain tumors in multiple sequences MRI

ZHAO Haoran¹, WANG Fei², LI Zhaotong³, GAO Song^{1*}

(1. Institute of Medical Technology, Peking University Health Science Center, Beijing 100191, China; 2. Department of Radiation Oncology, Peking University Cancer Hospital & Institute, Key Laboratory of Carcinogenesis and Translational Research [Ministry of Education], Beijing 100142, China; 3. Department of Medical Informatics, Medical School of Nantong University, Nantong 226001, China)

[Abstract] **Objective** To observe the value of unified modality-aware denoising diffusion probabilistic model (UMA-DDPM) for detecting brain tumors in multiple sequences MRI. **Methods** Totally 581 images of brain T1WI and T2WI from healthy individuals in IXI dataset were used for internal training, validation and testing, while T1WI and T2WI of 1 251 patients with brain tumors in BraTS21 public dataset were used for external testing. UMA-DDPM was proposed, including masking module, modality-aware module (MAM) and dual domain loss function, and its value for detecting brain tumors in multiple sequences brain MRI under “single mode” and “unified mode” was observed and compared with commonly used anomaly detection methods, respectively. **Results** UMA-DDPM could accurately identify abnormal areas in MRI, and the generated images had high similarity to original images, with better image quality than other models. The performance of single mode UMA-DDPM for detecting brain tumors in MRI was the best, while of unified mode UMA-DDPM was slightly lower, but still better than other models on the whole, and the modal adaptation capability and utilization of spatial and frequency information were improved. **Conclusion** UMA-DDPM could be used to effectively detect brain tumors in multiple sequences MRI under ensuring modal adaptation capability and image information utilization.

[Keywords] brain neoplasms; magnetic resonance imaging; artificial intelligence

DOI:10.13929/j.issn.1003-3289.2026.01.023

统一模态感知去噪扩散概率模型用于检出多序列 MRI 中的脑肿瘤

赵浩然¹, 王飞², 李兆同³, 高嵩^{1*}

(1. 北京大学医学部医学技术研究院, 北京 100191; 2. 北京大学肿瘤医院暨北京市肿瘤防治研究所放疗科恶性淋巴瘤发病机制及转化研究教育部重点实验室, 北京 100142; 3. 南通大学医学院医学信息学系, 江苏 南通 226001)

[摘要] **目的** 观察统一模态感知去噪扩散概率模型(UMA-DDPM)用于检出多序列 MRI 中的脑肿瘤的价值。**方法** 以 IXI 数据集中来自健康个体的 581 幅颅脑 T1WI 和 T2WI 对模型进行内部训练、验证及测试, 利用 BraTS21 公开数据集中 1 251 例脑肿瘤 T1WI 及 T2WI 数据进行外部测试; 以引入遮蔽模块、模态感知模块(MAM)和双域损失函数的 UMA-DDPM 分析颅脑多序列 MRI, 与常用异常检测方法比较, 评估其在“单模态”和“统一模态”下检出脑肿瘤的价值。**结果** UMA-DDPM 可精准识别 MRI 中的异常区域, 且生成图像与原始图像相似度较高, 质量亦优于其他模型。单模态

[基金项目] 2024 年度医学影像标准与创新研究项目(HX25A II 001038)、国家自然科学基金(82572374)。

[第一作者] 赵浩然(2000—), 男, 山西朔州人, 在读硕士。研究方向: 医学图像处理。E-mail: haoranzhao@hsc.pku.edu.cn

[通信作者] 高嵩, 北京大学医学部医学技术研究院, 100191。E-mail: gaoss@bjmu.edu.cn

[收稿日期] 2025-02-28 **[修回日期]** 2025-09-11

UMA-DDPM 检测 MRI 中的脑肿瘤的表现最优;统一模态 UMA-DDPM 性能略有下降,但整体仍优于其他模型,且其自适应能力及空间-频率信息利用率均有所提升。**结论** 基于 UMA-DDPM 可在保证模态自适应能力和图像信息利用率的情况下有效检出多序列 MRI 中的脑肿瘤。

[关键词] 脑肿瘤;磁共振成像;人工智能

[中图分类号] R739.41; R445.2 [文献标识码] A [文章编号] 1003-3289(2026)01-0113-06

MRI 对于诊断颅脑疾病具有重要作用^[1]。无监督异常检测 (unsupervised anomaly detection, UAD) 可自动于未标记影像学数据中检测异常,已成为传统依赖标注监督学习技术的潜在替代方案^[2];然而现有基于生成模型的 UAD 方法^[3]虽已取得一定进展,但在模态自适应能力^[4-5]和图像信息利用^[6-7]方面仍面临一定挑战。本研究提出自适应能力及空间-频率信息利用率均有所提升的统一模态感知去噪扩散概率模型 (unified modality-aware denoising diffusion probabilistic model, UMA-DDPM),观察其用于检出多序列 MRI 中的脑肿瘤的价值。

1 资料与方法

1.1 数据集 按 8:1:1 比例随机将 IXI 数据集^[8]中 581 幅来自健康个体的颅脑 T1WI 和 T2WI 分为内部训练集 ($n=465$)、内部验证集 ($n=58$) 及内部测试集 ($n=58$);以 BraTS21 公开数据集^[9]中的 1 251 例脑肿瘤 MRI 资料中的 T1WI 及 T2WI 数据进行外部测试。

1.2 数据预处理 首先将图像像素统一调整为 $192 \times 192 \times 100$,采用仿射变换技术将其与斯坦福研究所 (Stanford Research Institute, SRI) 24-Atlas 数据集对齐,确保图像空间方向标准化;之后采用 hybrid deep learning-based brain extraction tool (HD-BET)^[10]去除图像中的非脑组织;最后执行中心化操作,

并将图像分辨率降采样至 $96 \times 96 \times 50$ 体素。

1.3 UMA-DDPM 采用遮蔽模块、基于快速傅里叶变换 (fast Fourier transform, FFT) 对原始图像 X_0 进行掩盖处理后将输入 UMA-DDPM;计算输出图像空域和频率域损失,用于监督图像重建。见图 1。

1.3.1 遮蔽模块 遮蔽模块由频率遮蔽部分和空域遮蔽部分构成:频率遮蔽系通过 FFT 将图像转换为频率域以去除低频信号、保留高频信息,并利用二值掩码对频率域图像进行处理,强调高频成分;空域遮蔽则创建与原图尺寸相同的二值掩码,并将部分区域随机设置为 0(掩盖)或 1(保留),采用元素级乘法达到遮蔽效果。将经上述处理后的图像合并,得到双重遮蔽图像 \tilde{X}_0 作为下一步输入数据。见图 2。

1.3.2 正向扩散 首先,对 \tilde{X}_0 添加简单噪声 η ,模拟真实环境中的噪声干扰,获得加噪图像,见公式(1):

$$\tilde{X}_t = \sqrt{\alpha_t} \tilde{X}_0 + (1 - \alpha_t) \eta \quad (1)$$

其中, \tilde{X}_t 为 t 时刻的加噪图像块; α 为预设的计划噪声参数, $\bar{\alpha}_t$ 为累积的 α 值。

其后将 \tilde{X}_t 分割为 N 个分块 P_j , 其中 $j \in \{1, 2, \dots, N\}$;训练过程中,每次选择 1 个分块进行遍历迭代,且在选定分块区域创建 1 个与 \tilde{X}_0 大小相同的二值掩码

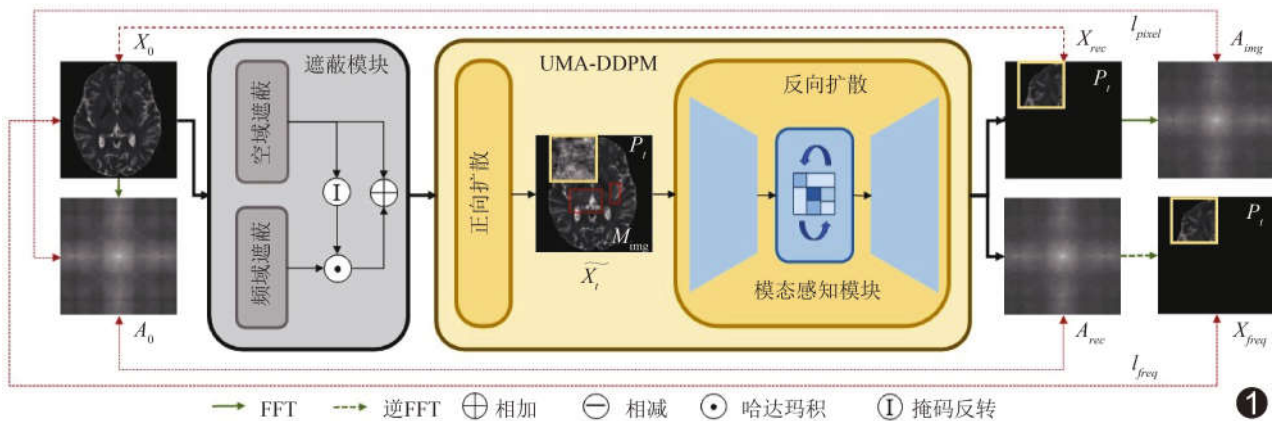


图 1 UAM-DDPM 结构示意图 (X_0 为原始图像; A_0 为原始图像的频率幅值图; \tilde{X}_t 为 t 时刻的加噪图像块; A_{rec} 为模型直接输出的重建频率幅值图; X_{rec} 为模型直接输出的重建图像; l_{freq} 为频率域损失; l_{pixel} 为像素域损失; A_{img} 为 X_{rec} 经 FFT 得到的频率幅值图; X_{freq} 为 A_{rec} 经逆 FFT 得到的重建图像)

$M_{patch}^j \in R^{W \times H}$ (R 为实数域、 W 为图像宽度、 H 为图像高度), 将选定区域内像素值设为 1、未选定区域像素值设为 0。获得掩码乘法定义的合并图像 \hat{X}_{ij} , 见公式(2):

$$\hat{X}_{ij} = \tilde{X}_i \odot M_{patch}^j + \tilde{X}_0 \odot \bar{M}_{patch}^j \quad (2)$$

其中, \odot 为元素级乘法; \bar{M}_{patch}^j 是 M_{patch}^j 的逆掩码。

最后, 将 \hat{X}_{ij} 输入扩散模型进行反向去噪, 恢复非掩码区域对应的分块。

1.3.3 模态感知模块 (modality-aware module, MAM) UMA-DDPM 的核心为嵌入 U-Net 瓶颈层的 MAM, 主要由模态感知自提示 (modality-aware self-prompting, MASP) 机制、2 个串联残差块和 1 个 Mamba 模块组成。见图 3。

MASP 机制根据不同模态输入图像自动调整其提取特征方式, 以确保对每个模态特征进行有效处理,

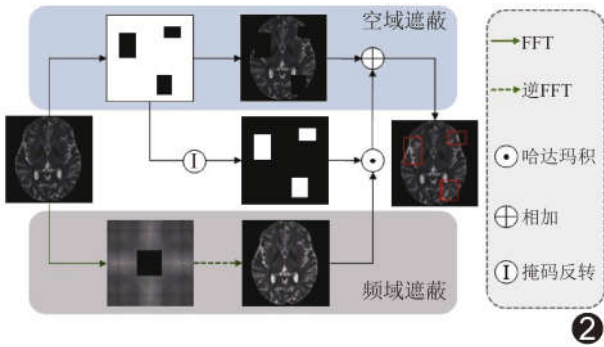


图 2 遮蔽模块示意图

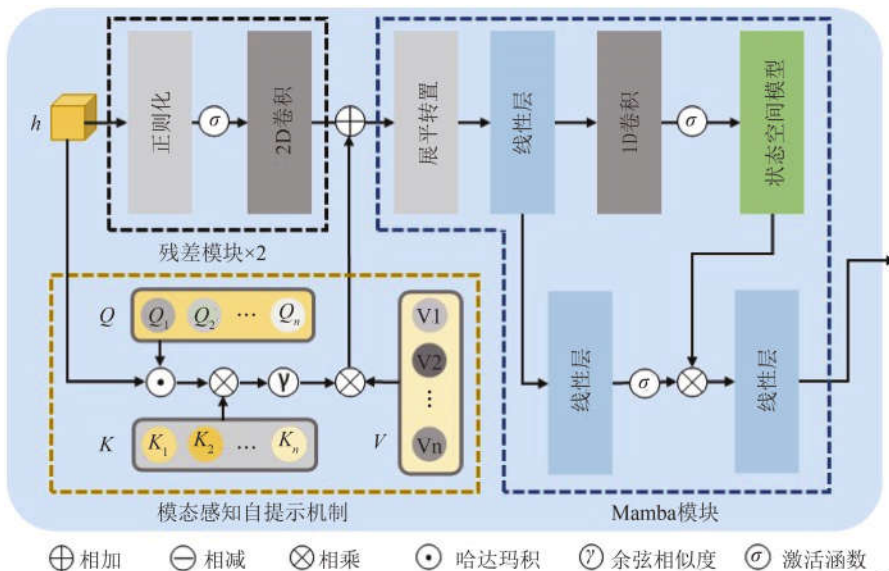


图 3 MAM 结构示意图 (h 为瓶颈特征; Q 为查询矩阵; K 为键矩阵; V 为值矩阵)

并主要基于 3 个可学习参数矩阵(查询矩阵 Q 、键矩阵 K 、值矩阵 V)实现模态自感知; 对其中的瓶颈特征 h 首先与 Q 进行 Hadamard 积运算, 再与 K 进行余弦相似度计算, 以衡量图像特征与预设键的匹配程度, 见公式(3):

$$A = \{ \gamma(h \odot Q, K_1), \gamma(h \odot Q, K_2), \dots, \gamma(h \odot Q, K_n) \} \quad (3)$$

其中, A 为相似度权重向量, 代表每个模态在 V 中的权重; \odot 表示 Hadamard 积; $\gamma(\cdot)$ 为余弦相似度函数; K_1, K_2, \dots, K_n 为 K 的分块。

之后对 A 与 V 进行矩阵乘法运算, 获得适应于 h 的模态提示 V_p , 用于指导 U-Net 解码器理解和处理不同模态; 经过归一化、Sigmoid 线性单元 (Sigmoid linear unit, SiLU) 激活函数和 2D 卷积块组成的 2 个残差网络等步骤后, 产生残差特征 h_{res} , 并将其加到 V_p 上, 即 $h_{res} + V_p$, 从而获得模态感知中心特征 h_p 。

Mamba 模块通过高效并行处理和深层特征融合提高模型表达能力, 使其更精确地去噪和检测异常。将 h_p 展平并转置后, 经 Mamba 模块的 2 个平行分支进行处理: 第 1 个分支通过线性层和 1D 卷积层捕捉局部依赖关系, 并通过 SiLU 激活函数和状态空间模型 (state space model, SSM) 层进一步提取复杂模式; 第 2 个分支仅使用线性层和 SiLU 激活函数。最后, 将 2 个分支特征通过 Hadamard 合并, 并通过线性层恢复到原始形状, 最终用于图像去噪和重建。

1.3.4 损失函数 引入频率域和空域的双域损失函数, 以平衡模型在频率域与空域的重建性能。

计算模型直接输出的重建频率幅值图 A_{rec} 与原始图像的频率幅值图 A_0 的频率域损失 l_{freq} 及像素域损失 l_{pixel} , 见公式(4)、(5):

$$l_{freq} = \| A_0 - A_{img} \|_1 + \| A_0 - A_{rec} \|_1 \quad (4)$$

$$l_{pixel} = \| X_0 - X_{freq} \|_1 + \| X_0 - X_{rec} \|_1 \quad (5)$$

其中, X_{rec} 为模型直接输出的重建图像, A_{img} 为 X_{rec} 经 FFT 得到的频率幅值图, X_{freq} 为 A_{rec} 经逆 FFT 得到的重建图像, $\| \cdot \|_1$ 为 L1 范数。最终得到的损失函数 l_{DDPM} , 见公式(6):

$$l_{DDPM} = \alpha l_{freq} + \beta l_{pixel} \quad (6)$$

其中, α 和 β 是权重因子。

1.4 检测异常 检测异常时,首先计算 UMA-DDPM 内部测试阶段输出图像与原始图像的差值,即重建误差,以之作为像素级的异常得分 $\Delta S: \Delta S = X_0 - X_{rec1}$; 之后,基于内部验证集进行搜索,迭代计算不同阈值下的戴斯相似系数(Dice similarity coefficient, DSC),最终选取可使 DSC 最大化的阈值对 ΔS 进行二值化。

1.5 评估模型 以 DSC、精确率-召回率曲线下面积(area under the precision-recall curve, AUPRC)、真阴性率(true negative rate, TNR)、真阳性率(true positive rate, TPR)、结构相似性指数(structural similarity index, SSIM)和峰值信噪比(peak signal-to-noise ratio, PSNR)评估 UMA-DDPM 在“单模态”和“统一模态”下检测异常的性能,前者指遵循“单对单”范式(采用单一序列数据进行训练,并在其相同序列中进行测试),后者指遵循“全对单”范式(采用多序列数据联合训练,在任意单一序列中进行测试)。DSC 用于量化检测异常区域与实际异常区域的重叠程度; AUPRC 用于评估模型在不同分类阈值下的表现; TNR 和 TPR 分别衡量模型识别非异常区域和检测异常像素的准确性; SSIM 评估生成图像与原始图像的相似性; PSNR 通过均方误差衡量生成图像的失真程度和噪声水平。

在“单模态”下将 UMA-DDPM 与其他常用检测异常方法进行对比,包括 1 种基于生成对抗网络(generative adversarial networks, GAN)的方法(即 f-AnoGAN^[2])、3 种基于自编码器(autoencoder, AE)的方法[变分 AE(variational AE, VAE)^[4]、稀疏 VAE(sparse VAE, SVAE)^[5]和去噪 AE(denoising AE, DAE)^[11]]及 3 种基于扩散模型的方法[异常检测去噪扩散概率模型(anomaly detection denoising diffusion probabilistic model, AnoDDPM)^[6]、分块去噪扩散概率模型(patch denoising diffusion probabilistic model, pDDPM)^[7]和掩码去噪扩散概率模型(masked denoising diffusion probabilistic model, mDDPM)^[12]]。

1.6 消融实验 以去除 MAM 及双域损失函数的 DDPM 模型为基线架构,逐一添加上述二者,于统一条件下评估其对 UMA-DDPM 的贡献度。

2 结果

2.1 检测效能 UMA-DDPM 可精准识别 MRI 中的异常区域,且生成图像与原始图像相似度较高,图像质量亦优于其他模型,见图 4。单模态 UMA-DDPM 检测 MRI 中的脑肿瘤的表现最优;统一模态 UMA-DDPM 的性能较之略有下降,但整体仍优于其他模型,且其自适应能力及空间-频率信息利用率均有所提升,见表 1。

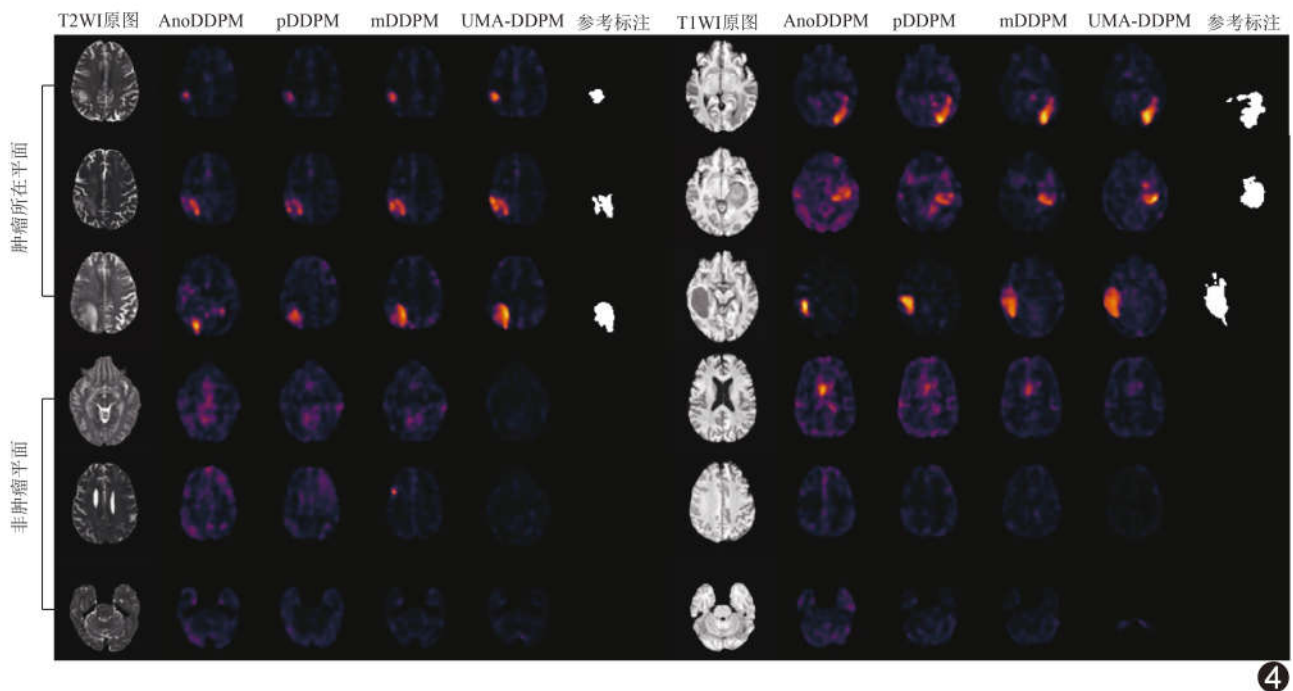


图 4 UMA-DDPM 与其他基于 DDPM 方法检测 BraTS21 数据集 MRI 中的脑肿瘤

表 1 不同模型检测 MRI 中的脑肿瘤的性能(%)

模型	T1WI						T2WI					
	DSC	AUPRC	TNR	TPR	SSIM	PSNR	DSC	AUPRC	TNR	TPR	SSIM	PSNR
f-AnoGAN	12.17	11.38	12.25	10.90	80.99	29.48	22.74	18.11	19.37	18.29	80.15	29.16
VAE	16.43	12.69	17.97	13.02	82.82	26.98	28.79	25.43	32.85	24.53	82.12	28.45
SVAE	19.67	14.90	18.90	14.32	82.92	29.48	33.61	30.42	35.95	35.38	82.39	28.75
DAE	25.23	22.32	23.33	19.41	83.59	28.34	37.25	34.74	37.14	33.80	82.91	29.78
AnoDDPM	23.00	21.02	20.58	19.73	82.84	30.53	36.88	35.48	29.65	27.20	84.13	26.20
pDDPM	26.20	24.18	27.69	22.80	83.25	28.25	41.01	38.11	42.06	36.08	84.47	27.69
mDDPM	28.92	23.33	27.90	22.28	84.43	29.86	42.68	40.09	44.53	38.86	85.92	28.69
单模态 UMA-DDPM	30.14	24.26	29.01	23.35	85.70	31.44	44.85	43.11	50.64	47.30	86.45	30.67
统一模态 UMA-DDPM	29.06	23.39	28.52	22.23			43.10	41.36	51.28	45.34		

注:评估 DSC、AUPRC、TNR 及 TPR 均基于 Brats21 数据集,评估 SSIM 和 PSNR 基于 IXI 数据集

2.2 消融实验 消融实验结果显示, MAM 可提升 UMA-DDPM 检测 MRI 中的脑肿瘤的性能;加入双域损失函数后,上述 2 个参数均获得进一步提升,即模态自适应能力和空间-频率信息利用率均对模型检测能力有所贡献。见表 2 和图 5。

3 讨论

本研究提出的 UMA-DDPM 在模态自适应能力和空间-频率信息利用率方面有所创新,其检测多序列 MRI 中的脑肿瘤的性能优于既往常用方法:相比基于 GAN 的 f-AnoGAN, UMA-DDPM 在检测精度和恢复图像质量方面具有明显优势;相比基于 AE 的 VAE、SVAE 和 DAE,UMA-DDPM 能更有效地处理复杂异常结构,且在较大噪声样本中的表现尤为突出;同时,与基于扩散模型的 AnoDDPM、pDDPM 和 mDDPM 相比,

表 2 UMA-DDPM 消融实验结果

MR 序列	遮蔽模块	MAM 模块	双域损失函数	DSC(%)	AUPRC(%)
T1W	✓			28.92	23.33
	✓	✓		29.49	23.49
	✓	✓	✓	30.14	24.26
T2W	✓			42.68	40.09
	✓	✓		43.55	41.14
	✓	✓	✓	44.85	43.11

注:✓表示对应组合中包含该部分

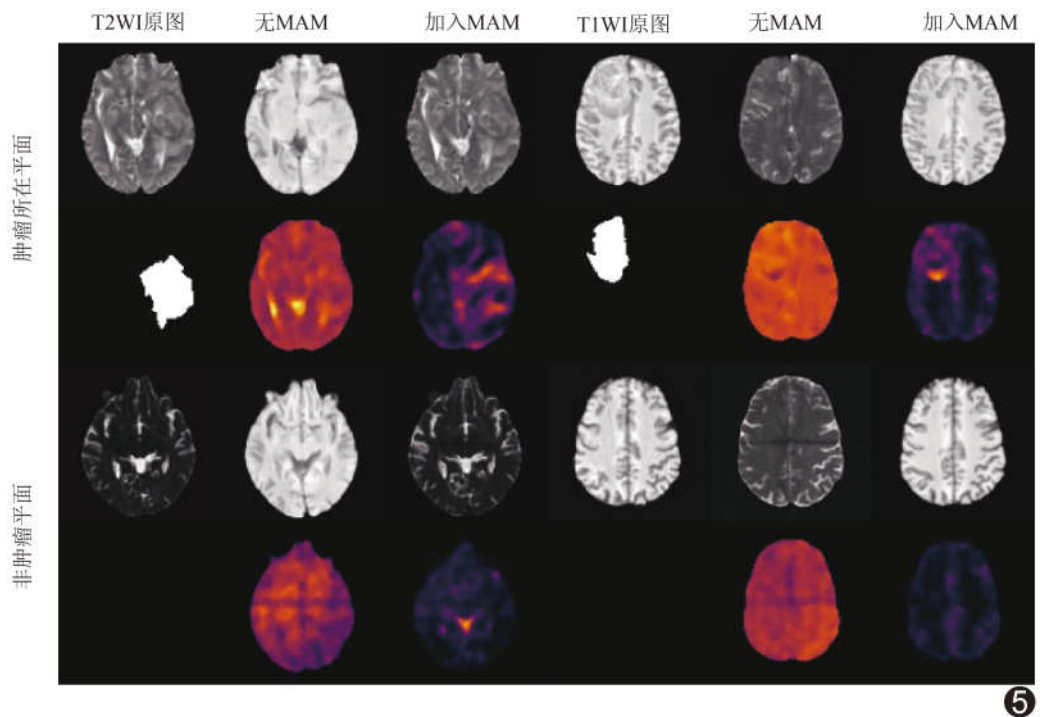


图 5 UMA-DDPM 检测 BraTS21 数据集 MR 图像中脑肿瘤的 MAM 消融实验结果 无 MAM 时,模型倾向于将 T1WI 错误地重建为 T2WI;加入 MAM 后,模型感知和重建无需依赖外部信息,并能准确识别异常区域

UMA-DDPM 亦具有整体优越性。分析原因,UMA-DDPM 引入了 MAM 和频率域空域的双域损失函数,可灵活适应不同序列图像特征,避免现有方法在切换

模态中的性能损失,同时又能更好地恢复和重建图像信息,确保其检测异常的精度;同时,其每秒浮点运算次数(floating point operations per second, FLOPS)约 1.4 GFLOPS,内存需求为 4.2 GB,虽较传统模型有所增加,但仍在多数硬件配置可接受范围内。

综上所述,本研究提出的 UMA-DDPM 可在保证模态自适应能力和图像信息利用率的情况下,有效检出多序列 MRI 中的脑肿瘤。但本研究基于无监督学习提出 UMA-DDPM,可能出现对极少数异常类型识别困难;仅基于 T1WI 和 T2WI 训练模型,且未针对不同类型脑肿瘤进行分组分析。未来将进一步优化模型训练过程,尤其针对更复杂医学影像的数据集,纳入更多类型数据、开发更精细的检测模型,进一步提高模型的准确性和鲁棒性。

利益冲突:全体作者声明无利益冲突。

作者贡献:赵浩然研究设计、撰写文章;王飞查阅文献、数据分析;李兆同查阅文献、图像处理;高嵩指导、审阅文章。

[参考文献]

- [1] ABD-ELLAH M K, AWAD A I, KHALAF A A M, et al. A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned[J]. Magn Reson Imaging, 2019, 61:300-318.
- [2] SCHLEGL T, SEEBÖCK P, WALDSTEIN S M, et al. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks[J]. Med Image Anal, 2019, 54:30-44.
- [3] 安政,韩乐,史明,等.基于双流判别器的生成对抗皮肤病变分割网络[J].中国医学影像技术,2024,40(12):1914-1919.
- [4] BAUR C, DENNER S, WIESTLER B, et al. Autoencoders for unsupervised anomaly segmentation in brain MR images: A comparative study[J]. Med Image Anal, 2021, 69:101952.
- [5] BEHRENDT F, BENGIS M, BHATTACHARYA D, et al. Capturing inter-slice dependencies of 3D brain MRI-scans for unsupervised anomaly detection[EB/OL]. (2022-05-09) [2023-05-13]. <https://openreview.net/pdf?id=db8wDgKH4p4>.
- [6] WYATT J, LEACH A, SCHMON S M, et al. Anoddp: Anomaly detection with denoising diffusion probabilistic models using simplex noise//2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR) workshops [C]. New Orleans: IEEE, 2022:650-656.
- [7] BEHRENDT F, BHATTACHARYA D, KRÜGER J, et al. Patched diffusion models for unsupervised anomaly detection in brain MRI[EB/OL]. (2023-03-07) [2025-02-28]. <https://doi.org/10.48550/arXiv.2303.03758>.
- [8] IXI dataset-brain development[EB/OL]. (2024-03-12) [2025-02-28]. <https://brain-development.org/ixi-dataset/>.
- [9] MENZE B H, JAKAB A, BAUER S, et al. The multimodal brain tumor image segmentation benchmark (BRATS)[J]. IEEE Trans Med Imaging, 2014, 34(10):1993-2024.
- [10] ISENSEE F, SCHELL M, PFLUEGER I, et al. Automated brain extraction of multisequence MRI using artificial neural networks[J]. Hum Brain Mapp, 2019, 40(17):4952-4964.
- [11] IQBAL H, KHALID U, HUA J, et al. Unsupervised anomaly detection in medical images using masked diffusion model[EB/OL]. (2023-05-31) [2025-02-28]. <https://arxiv.org/pdf/2305.19867v1>.
- [12] KASCENAS A, PUGEAULT N, O'NEIL A Q. Denoising autoencoders for unsupervised anomaly detection in brain MRI//International conference on medical imaging with deep learning[C]. Switzerland: PMLR, 2022:653-664.