

中图法分类号: R737.9; TP751 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-14

论文引用格式: Zhang Zhen, Li Jiayue, Wang Xiaohua, Chen Xiaohe<sup>④</sup>. A Segmentation-Driven Model for Predicting Response to Neoadjuvant Therapy in Breast Cancer[J/OL]. Journal of Image and Graphics, XXXX:1-14. DOI: 10.11834/jig.260147. (张真, 李佳悦, 王晓华, 陈晓禾<sup>④</sup>. 面向乳腺癌新辅助治疗疗效预测的影像分割模型[J/OL]. 中国图象图形学报, XXXX:1-14. DOI: 10.11834/jig.260147. ) [DOI: 10.11834/jig.260147]

## 面向乳腺癌新辅助治疗疗效预测的影像分割模型

张真, 李佳悦\*, 王晓华, 陈晓禾<sup>④</sup>

<sup>①</sup>北京邮电大学信息与通信工程学院, 北京, 100876 <sup>②</sup>北方工业大学人工智能与计算机学院, 北京, 100144 <sup>③</sup>北京大学第三医院放射科, 北京, 100191 <sup>④</sup>中国石油大学(北京)人工智能学院, 北京, 102249

**摘要:** 目的 新辅助治疗(Neoadjuvant Therapy, NAT)已成为局部晚期乳腺癌术前治疗的标准策略,旨在缩小肿瘤体积、提高手术切除率并改善患者预后。目前NAT后的疗效评估主要依据病理完全缓解(pathological Complete Response, pCR)与非病理完全缓解(npCR)两种分类。然而,病理评估的固有滞后性以及乳腺肿瘤在磁共振成像(Magnetic Resonance Imaging, MRI)中呈现的复杂形态学表现与异质性增强模式,使得术前精准预测NAT疗效仍面临重大挑战。方法 该文提出基于传统U-Net架构的影像组学引导语义分割框架。该框架融合离散小波变换(Discrete Wavelet Transform, DWT)池化模块,显著增强了肿瘤边界等高频细节的刻画精度;创新性的引入组学增强Transformer模块,通过通道筛选与空间调制机制实现低层结构特征与高层纹理组学特征的有效融合,显著增强了模型对病理相关判别性特征的表征能力;并引入空间交叉注意力机制以优化肿瘤轮廓重建与瘤内异质性建模。结果 在三个公开乳腺MRI基准数据集上的广泛实验表明,所提模型在BreastDM和ISPY1数据集上均达到最优分割性能(Dice分别为88.68%和92.87%,HD95分别提升了19.7%和7.2%)。消融实验进一步验证了各关键模块的有效性,其中RAT模块单独引入即可使HD95指标相对降低3.4%,充分证明了影像组学先验知识在精细化边界建模中的关键作用。基于分割模型提取的深度特征可有效提升NAT疗效预测精度,且支持基于单次随机时间点DCE-MRI的稳健预测,展现出显著的临床转化潜力。结论 综上所述,该文所提模型为局部晚期乳腺癌患者提供了一种高效、精准的术前疗效评估工具,有望推动个体化治疗决策优化,提升乳腺癌诊疗的精准性与临床效率。

**关键词:** 乳腺肿瘤分割;磁共振成像;影像组学;新辅助治疗;U-Net;Transformer

## A Segmentation-Driven Model for Predicting Response to Neoadjuvant Therapy in Breast Cancer

Zhang Zhen, Li Jiayue\*, Wang Xiaohua, Chen Xiaohe<sup>④</sup>

<sup>①</sup>(School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China);  
(School of Artificial Intelligence and Computer Science, North China University of Technology, Beijing 100144, China); (Department of  
Radiology, Peking University Third Hospital, Beijing 100191, China); <sup>④</sup>(School of Artificial Intelligence, China University of Petroleum  
(Beijing), Beijing 102249, China)

收稿日期: 2026-03-24; 修回日期: 2026-06-03

\*通信作者: 李佳悦, 通信作者, 男, 讲师, 人工智能在交叉学科的应用。邮箱: lijy2017@ncut.edu.cn

基金项目: 国家自然科学基金(项目编号: 62506008); 北京市自然科学基金资助项目(项目编号: Z210013); 北京市教育委员会科学研究计划项目(项目编号: 110052972508-10); 北京智慧精准医学促进中心基金(项目编号: M20250363); 中国石油大学(北京)校基金(项目编号: 2462023YJRC005)

Supported by: The National Natural Science Foundation of China; Beijing Natural Science Foundation; The R&D Program of Beijing Municipal Education Commission; Beijing Huizhi Precision Medicine Promotion Center; Science Foundation of China University of Petroleum, Beijing.

©中国图象图形学报版权所有

**Abstract: Objective** Neoadjuvant Therapy (NAT) has been established as a standard preoperative treatment paradigm for locally advanced breast cancer, with the objectives of reducing tumor burden, enhancing surgical resectability, and maximizing the probability of breast-conserving surgery. Therapeutic efficacy is conventionally assessed by pathological Complete Response (pCR) versus non-pCR status, determined through postoperative histopathological examination. However, the inherent latency of pathological assessment substantially limits its utility for timely clinical decision-making and adaptive treatment modification. Furthermore, breast tumors characteristically exhibit complex morphological architectures and heterogeneous enhancement kinetics on Magnetic Resonance Imaging (MRI), particularly Dynamic Contrast-Enhanced MRI (DCE-MRI), which poses considerable challenges for precise tumor delineation and quantitative response evaluation. Consequently, developing a robust and effective computational framework capable of simultaneously achieving accurate tumor segmentation and extracting clinically discriminative features for early NAT response prediction represents a critical unmet need with substantial clinical significance. **Method** To address the aforementioned challenges, this study proposes a novel radiomics-guided semantic segmentation framework based on the classical U-Net architecture. The proposed model is designed to enhance feature representation, improve boundary delineation, and facilitate the extraction of discriminative features for downstream NAT response prediction. First, a Discrete Wavelet Transform (DWT)-based pooling module is introduced to replace conventional downsampling operations in the encoder. Unlike standard pooling methods that often result in the loss of high-frequency information, the DWT-based pooling decomposes feature maps into multiple frequency sub-bands, including low-frequency approximation components and high-frequency detail components. This multi-frequency decomposition enables the model to explicitly preserve edge information, fine structural details, and texture variations, which are essential for accurate tumor boundary delineation. As a result, the proposed module improves the model's ability to maintain spatial consistency and enhances its sensitivity to subtle structural changes. Second, a Radiomics-Augmented Transformer (RAT) module is incorporated to strengthen the interaction between deep semantic features and radiomics-inspired representations. Conventional convolutional neural networks primarily capture local spatial patterns, whereas radiomics features describe global statistical and texture characteristics that are closely related to tumor heterogeneity and pathological properties. By integrating these complementary sources of information, the proposed model achieves a more comprehensive representation of tumor characteristics. Specifically, the RAT module employs a channel-wise selection mechanism to adaptively fuse low-level structural features with high-level semantic and texture representations. This mechanism allows the network to dynamically emphasize informative channels while suppressing redundant or irrelevant features, thereby improving feature discriminability and robustness. Third, a Spatial Cross-Attention (SCA) mechanism is introduced in the decoding stage to refine tumor contour representation and enhance intra-tumoral heterogeneity modeling. By capturing long-range spatial dependencies and enabling interactions between features at different spatial locations, the SCA module effectively integrates global contextual information with local details. This design improves segmentation consistency, especially in regions with ambiguous boundaries or complex structures, and enhances the overall quality of the reconstructed feature maps. The overall architecture follows an encoder-decoder paradigm, in which hierarchical feature representations are progressively extracted and refined through the integration of the aforementioned modules. The encoder focuses on multi-scale feature extraction with enhanced detail preservation, while the decoder reconstructs high-resolution segmentation maps with improved boundary accuracy. Beyond segmentation, the deep features learned by the model are further utilized for NAT response prediction, enabling a unified framework that bridges image segmentation and clinical outcome analysis. Notably, the proposed method supports prediction based on a single randomly selected time point from DCE-MRI, thereby reducing the dependency on fully aligned temporal sequences and improving its applicability in real-world clinical scenarios. **Result** Comprehensive experiments were conducted on three publicly available breast MRI benchmark datasets to rigorously evaluate the effectiveness and cross-dataset generalization capability of the proposed method. Experimental results demonstrate that the proposed model consistently achieves state-of-the-art segmentation performance on both the BreastDM and ISPY1 datasets, with Dice similarity coefficients of 88.68% and 92.87%, respectively, indicating excellent spatial overlap with ground truth annotations. Regarding boundary precision, the proposed method yields substantial improvements in the 95th percentile Hausdorff Distance (HD95), with relative reductions of 19.7% and 7.2% over competing state-of-the-art approaches. These results underscore the superiority of the

proposed framework in capturing fine-grained boundary details while preserving global structural integrity. Systematic ablation studies further validate the individual and synergistic contributions of each component within the proposed framework. Notably, the Radiomics-Augmented Transformer (RAT) module plays a pivotal role in enhancing feature representation and boundary precision. When evaluated independently, the inclusion of the RAT module results in an additional 3.4% improvement in HD95, demonstrating its strong contribution to refining segmentation results and improving the modeling of tumor heterogeneity. Moreover, the combination of DWT-based pooling and SCA further enhances the overall performance, indicating the complementary nature of these modules. Furthermore, the deep features extracted by the proposed segmentation model demonstrate strong discriminative power for downstream NAT response prediction. Experimental results confirm that incorporating these learned features yields a substantial improvement in prediction accuracy over baseline approaches. Crucially, the demonstrated ability to achieve robust prediction based on a single randomly selected DCE-MRI time point highlights the exceptional flexibility and clinical practicality of the proposed method, substantially reducing the dependency on complete temporal imaging sequences and streamlining the data acquisition workflow in routine clinical practice.

**Conclusion** In conclusion, this study presents a novel radiomics-guided semantic segmentation framework that synergistically integrates multi-frequency feature decomposition, transformer-based radiomics-enhanced feature learning, and spatial cross-attention mechanisms within a unified end-to-end architecture. The proposed model achieves consistently superior performance in breast tumor segmentation, particularly in boundary delineation accuracy and fine-detail preservation, while simultaneously yielding highly discriminative features for NAT response prediction. By enabling reliable prediction from limited single-time-point imaging data, the method substantially reduces the dependency on complex multi-temporal imaging protocols and significantly enhances clinical applicability. Overall, the proposed framework provides a promising and practical computational tool for supporting personalized treatment planning and facilitating evidence-based clinical decision-making in breast cancer management.

**Key words:** Breast tumor segmentation; Magnetic Resonance Imaging; Neoadjuvant therapy; Radiomics; U-Net; Transformer

论文引用格式: [DOI:10.11834/jig.260147]

## 0 引言

根据《柳叶刀》最新发布的全球乳腺癌重大报告,乳腺癌已跃升为全球发病率最高的恶性肿瘤(Siegel等,2024)。国家癌症中心最新发布的全国癌症报告表明,我国乳腺癌年新发病例数达35.72万例,居我国女性恶性肿瘤发病率第二位,且增速位列全球首位(Volders等,2018)。面对乳腺癌发病率持续攀升的严峻形势,国家卫生健康委办公厅于2022年印发《乳腺癌筛查工作方案》,着力提升乳腺癌早期筛查的覆盖率与精准度(Montemurro等,2020)。

在临床实践中,局部晚期乳腺癌因肿瘤负荷大、侵袭范围广,往往需要更为积极的综合治疗策略,且术后复发转移风险显著升高。目前,新辅助治疗(Neoadjuvant Therapy, NAT)已成为中晚期乳腺癌术前治疗的重要策略(Cantini等,2023)。其主要目的在于手术前缩小肿瘤体积,减轻对周围组织的侵犯和淋巴结浸润程度,从而提高手术根治率,增加保乳

可能性,并降低术后复发或转移的风险,进而提高生活质量和患者的生存率。准确评估中晚期乳腺癌患者的NAT疗效、明确残留肿瘤及淋巴结的范围和边界对于手术方式的选择、治疗方案的优化以及预后判断均具有重要意义(Lobbes等,2013)。在NAT疗效评估中,影像学检查发挥着关键作用(Xiong等,2025)。其中磁共振成像(Magnetic Resonance Imaging, MRI)凭借其多方位、多序列成像能力和优越的软组织分辨率,能够更加准确地展示肿瘤的位置、大小与形态特征。同时,动态对比增强(Dynamic Contrast Enhanced, DCE)-MRI序列能进一步反映肿瘤的供血特点,对判断乳腺癌的组织学性质及NAT后的治疗反应具有重要价值。因此MRI被广泛应用于中晚期乳腺癌的NAT疗效评估。

目前乳腺癌患者接受NAT后的疗效评估可根据手术后的病理标本进行评判,常用标准有非病理完全缓解(non-pathological Complete Response, npCR)和病理完全缓解(pathological Complete Response, pCR)两种分类(El Adoui等,2019)。npCR表示术后病理标本中仍检出残余浸润性癌细

胞,提示需要进一步的辅助治疗。pCR是指乳腺原发灶和腋窝淋巴结中均未检测到浸润性癌细胞残留。由于不同患者对NAT的敏感性存在显著差异,若能在NAT前准确预判个体疗效反应,则可有效避免低敏感性患者因无效治疗而承受的毒副作用与治疗延误,并优化敏感性较高患者的用药剂量与疗程设计,从而最大化提升pCR的发生率并最大化治疗效果。尽管组织病理学检查仍是评估NAT是否达到pCR的金标准,但其依赖术后组织获取,导致该方法在临床应用中存在明显的滞后性(Duanmu等,2020)。

近年来,深度学习(Deep Learning, DL)(Shi, 2020)技术在NAT疗效预测中展现出巨大潜力。然而,由于乳腺MRI包含大量背景腺体和脂肪组织,直接对全局影像进行端到端预测极易引入假阳性干扰。因此,基于卷积神经网络(Convolutional Neural Networks, CNN)的精准特征提取与肿瘤分割在NAT疗效评估有重大意义(Shi, 2025)。CNN被广泛用于从多时相或多模态MRI中提取深层表征(Luo, 2014)。Ha等人(2019)首次将CNN引入该任务,采用一个10层网络结构实现了87.7%的预测准确率。为了进一步提升模型性能,后续研究逐渐转向多模态融合方法。Joo等人(2021)利用3D ResNet-50(Huang等,2017)从化疗前的DCE-MRI与T2加权图像中提取特征,并与临床数据融合,与单独使用临床数据相比显著提升了pCR的预测效果,融合模型的AUC达到88.8%,而仅使用临床信息的模型AUC为82.7%。同时有研究表明,影像组学特征能够有效量化肿瘤的内部纹理与形态异质性,并已被广泛应用于乳腺癌NAT疗效的预测研究中。Chitalia等人(2020)使用60个与纹理和形状相关的特征进行分类,取得了0.73的预测性能。在之后的工作中,Xia等人(2021)将特征数量扩展至1316个,构建了列线图模型,其预测性能达到了0.868。在最新的研究中,Xie等人(2024)基于肿瘤内部与周围血管的放射组学特征构建预测模型,最终融合模型的AUC达到0.82。Fan等人(2024)利用17个影像组学特征建立多任务模型,同时预测肿瘤缩小模式与pCR,在肿瘤缩小预测上的AUC为0.886,pCR预测的AUC为0.760。此外,在最新的语义分割网络架构研究中发现,通过引入基于U-Net(Ronneberger等,2015)与Transformer(Vaswani等,2017)融合架构以获取肿瘤

的精确区域,可进一步提升特征表达的稳定性和判别力。Li等人(2023)提出将DCE-MRI的深度学习特征与患者的临床特征进行联合建模,通过多模态融合策略显著增强模型的判别能力,其pCR预测AUC达到0.925。在最新的研究中,Huang等人(2023)利用治疗前后纵向MRI构建融合模型,引入反映疗程动态变化的特征,在多中心队列中取得了0.837~0.929的pCR预测性能,强调了纵向影像变化对NAT疗效评估的重要作用。尽管上述研究也意识到了在高度异质性的乳腺肿瘤NAT疗效预测任务中,单一维度的影响表征往往存在表达瓶颈,但选用的特征融合策略仅作向量拼接,难以在图像特征中对肿瘤分割以及疗效预测做准确的引导,或是依赖多时相或多模态影像,造成数据采集复杂、对时间配准要求严格。乳腺肿瘤的高度异质性使得人工影像组学特征鲁棒性不足,手动筛选组学特征存在很强的随机性,而只使用深度学习模型或是简单特征拼接又难以有效整合高层语义先验。因此,本文旨在解决以下三个关键挑战:(1)如何在不依赖多时相严格配准,即基于单次、随机时相DCE-MRI的情况下,实现病灶的精准分割与疗效的高精度预测?(2)如何将具备可解释性的影像组学先验知识与深度学习特征进行有效融合,以打破单一特征的表达瓶颈?(3)如何在模型设计中有效应对乳腺肿瘤边缘复杂且内部高度异质性的难题,从而提取具有判别性的个体化响应特征?为有效应对上述问题,该文基于U-Net架构提出了一种由组学引导的新型语义分割模型(Radiomic-Guided U-Net, RGU-Net)。乳腺病灶在MRI影像中表现出复杂的形态学特征与显著的内部异质性,通用的分割模型往往侧重于目标几何外廓的拟合,其在连续下采样过程中容易丢失反映病灶内部性质的微小纹理信息,且倾向于平滑处理具有病理意义的不规则轮廓。RGU-Net将高层组学先验知识融入模型,以指导低层特征的学习,从而在解剖结构复杂的病例中实现精确分割。该模型使用离散小波变换(Discrete Wavelet Transform, DWT)模块替代了传统的池化操作。DWT池化模块通过多尺度频域分解,用高频特征对低频特征调制,强化肿瘤的结构性差异与边界显著性,使潜在病灶特征更加显著。与此同时,通过构建组学增强Transformer(Radiomic-Augmented Transformer, RAT)模块,模型精巧地使用经过低层结构特征调制与通

道筛选优化的高层纹理组学特征,引导跨尺度注意力机制在深度特征中选择性强化与治疗反应相关的区域与通道,以应对乳腺肿瘤内部异质性强的问题,从而学习具有判别性的个体化响应特征表征。最终,通过引入空间交叉注意力(Spatial-wise Cross Attention, SCA)模块,模型进一步在高维特征空间中聚焦于关键响应区域,抑制与反应无关的背景干扰,实现单时相影像下更稳定、精准的疗效预测。尤其值得关注的是,RGU-Net能够利用单次随机时间点的DCE-MRI扫描进行NAT疗效预测,这极大地降低了数据收集的需求,并消除了严格对齐时间的必要性。该文在三个公开的乳腺MRI数据集上对RGU-Net进行了全面的验证,结果表明,该模型在分割准确性、鲁棒性和细节保留方面均展现出超越现有最先进模型的卓越性能。此外,从其分割结果中提取的特征显著改善了NAT疗效预测的准确性,这充分彰显了其在简化临床流程、实现个性化治疗规划方面的巨大临床应用潜力。

## 1 方法

影像组学(Wang等,2023)作为一种高通量计算范式,能够从数字化医学影像中系统性地提取海量定量特征,涵盖低阶特征(一阶统计特征、形状特征)以及刻画体素间复杂空间排列关系的高阶纹理特征

(Chang等,2025)。这些特征能够揭示传统影像学评估中肉眼难以捕获的微观异质性信息,而此类异质性已被广泛证实与肿瘤的生物行为、微环境特征及治疗敏感性密切相关。

为了应对乳腺MRI中肿瘤边界模糊且形状不规则、内部纹理异质性强以及多尺度特征表达不足等问题,该文提出了一种基于影像组学引导的深度语义分割模型RGU-Net。该方法通过引入影像组学先验知识与多尺度特征融合,有效提升了模型在结构细节感知和肿瘤区域判别方面的能力。得益于这一特征表征优化策略,模型在无需依赖多时项输入和时间对齐的条件下,即可在单一随机时相的DCE-MRI影像上实现稳定而精准的NAT疗效预测。接下来,该文将从RGU-Net整体框架、组学增强Transformer模块、离散小波变换池化模块和空间交叉注意力模块三个方面对该方法进行详细论述。

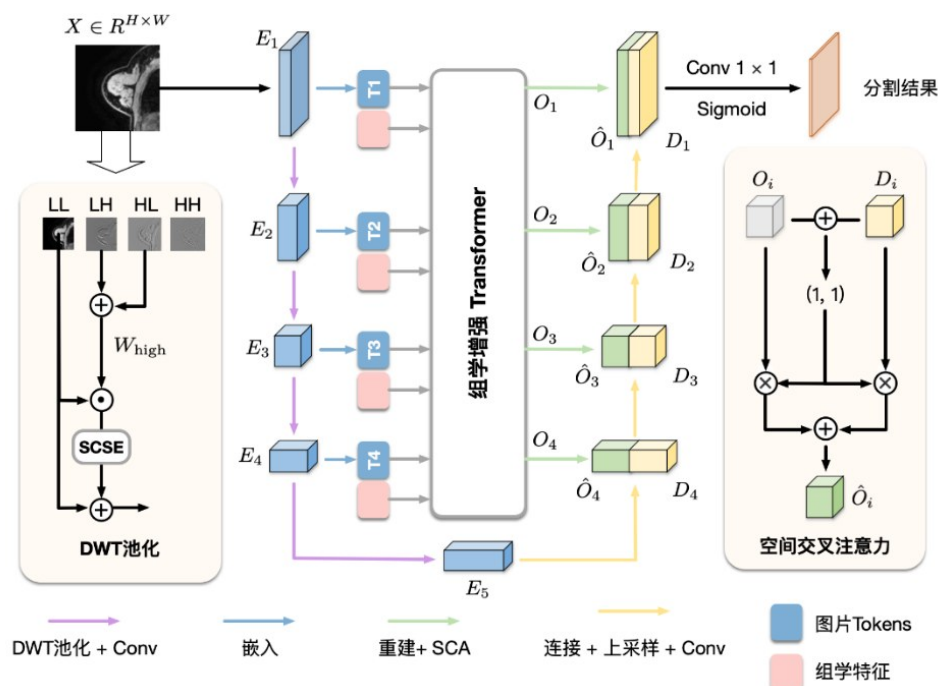
### 1.1 RGU-Net框架

RGU-Net是一种在经典U-Net架构基础上深度整合影像组学引导机制的语义分割框架,旨在系统地应对乳腺MRI肿瘤分割中的多重核心挑

图1:RGU-Net的整体架构

Fig 1: Overall architecture of RGU-Net

战,涵盖肿瘤形态的高度异质性、边界模糊不清以及传统影像组学特征可解释性与判别力不足等关键问题。



如图1所示,该文在传统U-Net及其通道交叉注意力框架(Wang等,2022)的基础上进行了方面的改进:首先,在输入图像的特征编码阶段,RAT模块在注意力计算中显式引入经调制与筛选优化的影像组学特征,从高层诊断视角增强低层特征的学习能力;其次,在编码器逐层下采样的过程中,将传统的下采样操作替换为DWT池化模块,以提升低层特征在通道维度上的表达能力,从而实现更精准的分割。最后,在解码器通过上采样和跳跃连接恢复分辨率时,引入SCA模块以弥补现有通道交叉注意力机制的不足,在训练过程中提供空间对齐与关注引导。综上所述,RGU-Net能够自适应地融合高层影像组学增强特征与低层特征,实现更灵活的多尺度特征交互,从而提升模型对全局肿瘤结构与局部细节的综合表征能力。

### 1.2 组学增强Transformer模块

RAT模块旨在将高层影像组学先验知识显式注入特征学习过程中,从诊断层面指导低层特征的细化与优化。RAT作用于U-Net各编码器层输出的多尺度特征图 $E_i$ ( $i = 1, 2, 3, 4$ )。每个特征图首先

被划分为若干不重叠的小块,这些小块经线性投影后形成图像标记序列 $T_i$ ( $1, 2, 3, 4$ ),从而编码局部化的语义与空间信息。

同时,从原始输入图像的ROI区域中提取多类别影像组学特征,包括一阶统计特征、形状特征以及基于灰度共生矩阵(Gray-Level Co-occurrence Matrix, GLCM)的纹理特征。组学特征筛选调制过程如图2所示。

具体而言,将ROI图像划分为 $4 \times 4$ 个不重叠的小块,每个小块分别提取10个一阶统计特征、10个形状特征和96个GLCM纹理特征。通过学习一阶统计特征与形状特征在不同空间位置的响应,生成一张空间注意力图 $A_{spa}$ ,用以衡量各小块在影像组学信息中的重要性:

$$A_{spa} = \sigma(f_s([\mathbf{F}_{state}; \mathbf{F}_{shape}])), \quad (1)$$

式中, $\mathbf{F}_{state}$ 和 $\mathbf{F}_{shape}$ 分别表示ROI小块的一阶统计特征与形状特征, $[\cdot]$ 表示特征拼接操作, $f_s(\cdot)$ 为两层卷积与归一化映射函数, $\sigma(\cdot)$ 为Sigmoid激活函数。由此获得每个ROI小块对应的GLCM权重,实现了纹理特征在空间层面的自适应调制,有效增强了肿瘤相关区域的响应,并抑制了非肿瘤区域的纹理干扰。

在此基础上,除空间维度的自适应调制外,为进一步区分不同GLCM纹理特征在分割任务中的重要性,

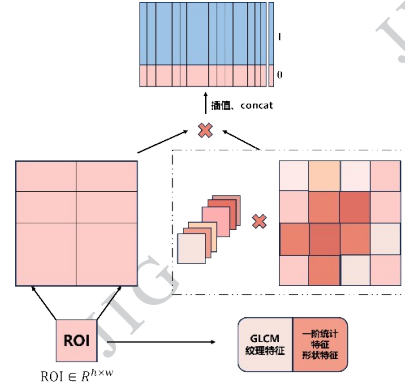


图2 组学特征的筛选调制过程

Fig 2 The feature selection and modulation process of radiomic features

RAT模块引入了可学习的通道权重矩阵 $\mathbf{W}_{GLCM}$ :

$$\mathbf{W}_{GLCM} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N], N = 96, \quad (2)$$

式中, $\mathbf{w}_c \in [0, 1]$ 表示第 $c$ 个GLCM通道的重要性。该向量用于筛选不同通道的纹理特征,使模型自适应强化与分割任务相关的通道,同时抑制无关通道干扰。最终得到联合调制后的GLCM响应,在训练过程中通过分割损失的反向传播动态优化,实现了对空间与通道双层次特征的重要性建模,从而促使模型自适应地选择并强化与分割任务最相关的纹理组学特征。

为了与图像标记序列对齐,筛选调制后的组学特征经过线性变换后拼接到标记序列末端,形成一个融合视觉结构与影像组学纹理先验的统一序列。该联合序列经层归一化处理,映射为包含图像查询分量 $Q_{vis,i}$ 与组学查询分量 $Q_{rad,i}$ 的增强查询向量 $Q_i$ 。键 $\mathbf{K}$ 和值 $\mathbf{V}$ 通过拼接所有尺度的查询 $Q_i$ 构建而成。在交叉注意力模块中,注意力权重 $A_i$ 的计算方式如下:

$$A_i = \sigma \left( \psi \left( \frac{[Q_{vis,i}; (\mathbf{W}_{GLCM} \cdot Q_{rad,i})]^T \cdot \mathbf{K} \cdot \mathbf{V}^T}{\sqrt{C_\Sigma}} \right) \right), \quad (3)$$

式中, $C_\Sigma$ 表示通道总数,本文设置为960, $\sigma$ 表示Softmax函数, $\psi$ 表示层归一化算子, $[\cdot]$ 表示特征拼接操作, $\mathbf{T}$ 表示矩阵转置。该机制确保影像组学先验能够直接影响注意力分布,其编码的病灶线索(如边缘清晰度和纹理均匀性)能够引导模型关注更

键的低层空间与通道特征,从而提升分割的精确性。

随后,经过注意力增强的特征将输入 Transformer 模块进行进一步细化。该模块包括残差连接、层归一化与前馈多层感知机(Multi-layer Perceptron, MLP)网络,其输出特征计算如下:

$$O_i = A_i^N + \text{MLP}(Q_i + A_i^N), \quad (4)$$

式中  $N$  为注意头的数量。最终,处理后的标记序列被重新映射为空间特征图并传递至解码器路径,使下游分割头能够生成准确且细节保留的预测结果。通过在多尺度特征中显式注入影像组学先验,RAT 模块增强了低层特征的语义相关性,改善了边界刻画能力,提升了肿瘤区域的整体可判别性。

### 1.3 离散小波变换池化

图1左侧展示了DWT池化模块在乳腺MRI切片上的应用。与传统池化方法往往会损失重要的空间和频率信息不同,DWT模块利用正交滤波器组将输入特征图分解为不同空间频率的子带分量。在特征提取过程中,频率筛选机制能够显著缓解深层卷积网络中常见的特征平滑现象,以避免网络在降低空间分辨率时,因过度拟合全局统计特性而抹除具有高度判别力的局部突变信号。这种频率层面的拆解为后续精细化勾画病灶边缘及挖掘异质性特征提供了更有力的支撑。通过低通算子  $L$  与高通算子  $H$  的卷积映射,Haar 函数计算生成四个不同的频率子带:低低频( $LL$ )、低高频( $LH$ )、高低频( $HL$ )和高高频( $HH$ ),其定义如下:

$$\text{DWT}_{\text{Haar}}(X) = \begin{pmatrix} LL & LH \\ HL & HH \end{pmatrix}, \quad (5)$$

式中,  $LL=(X*L^T L)_{\downarrow 2}$ ,  $LH=(X*L^T H)_{\downarrow 2}$ ,  $HL=(X*H^T L)_{\downarrow 2}$ ,  $HH=(X*H^T H)_{\downarrow 2}$ , \*表示卷积操作,  $T$  表示矩阵转置,  $\downarrow 2$  表示步长为2的下采样。 $LL$ 子带保留了粗糙的低频信息,有助于捕捉肿瘤的全局结构特征,如形状和位置。 $LH$ 与 $HL$ 子带编码了细粒度的水平和垂直高频成分,这对于强调关键分割线索(如边缘清晰度、纹理异质性以及肿瘤与周围组织之间的微小边界变化)至关重要,有助于增强通道特征表达能力,尤其在肿瘤边界不规则或对比度低的区域。 $HH$ 子带通常包含无结构的高频噪声,为了减少冗余并抑制噪声,该子带会被舍弃。保留的每个子带都会下采样至原始空间分辨率的四分之一( $C \times H/2 \times W/2$ )。

与传统仅使用  $LL$  子带进行下采样的做法不同,

该设计引入了选择性增强策略,以保留并放大有用的高频信息。具体来说,将  $LH$  和  $HL$  子带按元素相加,并通过非线性激活函数生成融合的高频特征图:

$$W_{\text{high}} = \sigma(LH + HL), \quad (6)$$

该高频特征图通过元素乘法调制  $LL$  子带,使高频边界信息能够选择性地增强全局特征表示。为了进一步提升特征的判别能力,并克服固定、不可训练的小波滤波器的限制,调制后的特征会通过空间与通道注意力(Spatial and Channel Squeeze and Excitation, SCSE)模块(Roy, 2018)进行精炼,该操作可自适应地根据任务相关性重新加权空间位置和通道:

$$O_{\text{dwt}} = \text{SCSE}(LL \cdot W_{\text{high}}) + LL, \quad (7)$$

最终输出不仅实现了下采样,同时增强了通道敏感性,尤其对边界和细微纹理变化更为敏感。与最大池化或平均池化相比,DWT池化在边界保留和分割精度上表现更优,使模型能够在复杂医学图像中更准确地刻画肿瘤边缘。

### 1.4 空间维度交叉注意力

SCA 模块旨在弥补纯通道注意力的局限,通过在编码器和解码器特征之间引入显式的空间对齐。通道注意力强调在特征通道上“关注什么”,而SCA则补充了在空间维度上“关注哪里”的能力,这对于精细化医学图像分割中的结构细节至关重要。

SCA 模块处理两个输入:解码器特征图  $D_i$  和编码器对应的跳跃连接特征  $O_i$ 。首先,将这两组特征通过逐元素相加融合,并经过 ReLU 激活,得到一个整合了低层与高层语义空间信息的中间特征表示。接着,将融合后的特征输入两个并行的  $1 \times 1$  卷积分支,每个分支后接 Sigmoid 激活函数,生成两个空间注意力图  $A_{i,1}^*$ ,  $A_{i,2}^*$ :

$$A_{i,1}^*, A_{i,2}^* = \sigma(\text{Conv}(\text{ReLU}(O_i + D_i))), \quad (8)$$

式中,  $\text{Conv}(\cdot)$  为卷积运算,  $\text{ReLU}(\cdot)$  为激活函数。这两个注意力图编码了不同区域的空间重要性,有效识别对精确分割关键的区域,例如病灶轮廓和异质性区域。随后,这些注意力图通过逐元素乘法对原始输入  $O_i$  和  $D_i$  进行空间加权,得到两组强调关键信息的特征图。

最后,将加权后的特征图通过逐元素相加融合,得到空间注意力输出  $\hat{O}_i$ :

$$\hat{O}_i = A_{i,1}^* \cdot O_i + A_{i,2}^* \cdot D_i, \quad (9)$$

该输出不仅保留了解码器和编码器的语义信息,还

引入了局部空间聚焦,使模型对细微解剖变化更为敏感。通过在网络中整合SCA,RGU-Net能够在全局上下文建模(由通道维度交叉注意力提供)的基础上,实现精确的空间对齐,从而提升分割性能,尤其是在肿瘤边界模糊或不规则的情况下效果更明显。

### 1.5 损失函数

本文采用由加权二元交叉熵(weighted Binary Cross-Entropy, wBCE)损失函数加权Dice损失构成的联合损失函数,并在此基础上引入稀疏正则项以实现GLCM特征自动选择与冗余抑制。前两项分别用于应对类别不平衡与边界模糊问题,其中加权wBCE通过引入类别权重平衡前景与背景像素的影响,其形式为:

$$L_{wBCE} = -\frac{1}{N} \sum_i [w_p y_i \log(p_i) + w_n (1 - y_i) \log(1 - p_i)], \quad (10)$$

式中, $N$ 为像素总数, $y_i$ 为第*i*个像素的真实标签, $p_i$ 为第*i*个像素属于前景的预测概率, $w_p$ 与 $w_n$ 分别为前景类与背景类的权重系数。而加权Dice损失用于优化区域重叠度,其形式为:

$$L_{wDice} = 1 - \frac{2 \sum_i w_i p_i y_i + \epsilon}{\sum_i w_i (p_i + y_i) + \epsilon}, \quad (11)$$

式中, $\epsilon$ 保证分母不为0。为进一步强化模型对组学特征的有效利用,本文在联合损失中引入稀疏正则项,以促进通道权重矩阵的稀疏性,整体损失可表示为:

$$L_{total} = \alpha L_{wBCE} + \beta L_{wDice} + \lambda \|w\|_1, \quad (12)$$

式中, $\alpha, \beta$ 分别为上述两个损失函数在计算过程中的权重大小, $\lambda \|w\|_1$ 项用于约束组学特征的通道权重,使模型能够在训练过程中自动筛选出对分割最具贡献的特征通道,抑制冗余或噪声特征对模型判别的干扰,从而在提升分割精度的同时增强模型的稳定性与可解释性。

## 2 实验与结果

### 2.1 数据集

为验证RGU-Net的有效性,本研究选用了三个公开的乳腺MRI数据集进行评估。分割模型在BreastDM和ISPY1数据集上进行训练测试。为评价分割提取特征的性能,该文设计了一个下游分类任务:使用ISPY2数据集进行训练测试,将由RGU-Net

自动生成的感兴趣区域(Region of Interest, ROI)输入到一个在ISPY2真实ROI上预训练的分类器中,通过分类器的预测准确率间接衡量分割质量。分类准确率越高,表明分割效果越好。

(1) **BreastDM** (Zhao 等, 2023): 该数据集包含2018年1月至2021年9月期间收集的乳腺MRI检查影像。初始共纳入350例通过病理确诊的患者,经放射科医师排除既往乳腺手术、乳腺假体或DCE-MRI动态序列不完整的病例后,最终保留232例(其中恶性147例,良性85例)。患者年龄范围为22-79岁(中位数:50岁)。每例检查均包括1个增强前相位和8个连续的增强后相位,可全面反映乳腺病灶的动态特征。

(2) **ISPY1-Tumor-SEG-Radiomics 数据集** (Chitalia 等, 2021): 该数据集来源于The Cancer Imaging Archive (TCIA) (Clark 等, 2013), 最初共纳入230例II-III期乳腺癌患者。排除数据不完整的病例后,最终保留161例患者。该数据集提供了T1加权的DCE-MRI扫描图像,包括一次对比剂注射前和两次对比剂注射后的三个时相,并经过偏置场校正、重采样及z-score标准化等预处理操作。值得注意的是,该数据集包含经专家审核的高质量三维肿瘤手动标注,为分割任务的精确评估提供了可靠依据。

(3) **ISPY2/ACRIN-6698** (Newitt 等, 2021): ISPY2试验是一项多中心前瞻性研究,旨在评估乳腺癌患者对新辅助化疗的反应。在原始队列中,排除缺少中期治疗影像或病理完全缓解(pCR)标注的病例后,最终保留了229例患者。该数据集提供了多序列MRI扫描,包括T2加权像(T2WI)、DCE及弥散加权成像(DWI),共涵盖四个关键治疗时间点:治疗前基线(T0)、紫杉醇治疗三个周期后的早期阶段(T1)、紫杉醇与蒽环-环磷酰胺治疗之间的中期阶段(T2),以及术前治疗结束阶段(T3)。值得注意的是,所有患者均具有明确的pCR状态标注,为预测模型研究提供了高质量的参考标准。

上述三个数据集的系统性引入确保了模型评估的全面性、严谨性与临床相关性:BreastDM数据集提供了良性与恶性病例的多样组合,用于实现稳健的模型训练;ISPY1数据集包含高质量的三维人工分割标注,用于严格的分割性能评估;ISPY2数据集则提供了纵向的多时间点影像及确证的pCR结果。

值得指出的是,三个数据集在采集时相上本就处于未对齐的随机状态: BreastDM 病例收录完全随机, ISPY1-Tumor-SEG-Radiomics 仅按清晰度与显著性随机挑选, ISPY2 则固定提取可用病例最多的"治疗中期"单次图像。本模型无需完整时间序列与时序配准, 仅凭单次图像即可完成高精度预测, 切实验证了引言所述降低数据采集成本及配准难度的优势, 也提升了其在真实临床场景中的适用性。

在分割任务中, BreastDM 与 ISPY1 数据集提供了二维图像-标签对用于训练与测试, ISPY2 数据集用于推理阶段, 得到后续用于下游疗效预测任务的 ROI。跨数据集的设置主要受限于现有公共数据集的客观条件: 目前公开的乳腺肿瘤数据集中, 极少有同时具备"高质量肿瘤分割掩膜"、"完整的 NAT 阶段图像"的数据。因此, 利用在 ISPY1 上训练完毕的模型直接对 ISPY2 数据集进行推理。所有图像均被标准化为 160×160 像素, ROI 尺寸为 60×60, 并进行了强度归一化。仅在 BreastDM 数据集中应用了数据增强操作(水平/垂直翻转与随机旋转)。在下游分类任务中, 训练与验证阶段的 ROI 由 ISPY2 真实掩膜提取, 而测试阶段的 ROI 则由 RGU-Net 及其他对比模型预测的掩膜获得, 这些 ROI 统一调整为 112×112 像素。采用双路径融合分类模型: 影像组学通路通过多层感知机对 ROI 中提取的影像组学特征进行建模, 图像通路使用基于 ConvNeXt 的卷积网络从归一化 ROI 图像中提取深度特征。两路特征在切片维度进行聚合后拼接, 并通过全连接层完成二分类预测, 分类标签来源于 ISPY2 数据集提供的 CSV 文件。分类测试结果只依赖于模型分割出的 ROI 图像, 其他影响因素均保持一致。

## 2.2 实现细节与评价指标

在分割任务中, ISPY1 数据集采用五折交叉验证策略, 而较小的 BreastDM 数据集采用三折交叉验证。为了保证评价的客观性, ISPY1 数据集均按照切片划分, BreastDM 数据集的划分确保每个原始样本及其所有增强版本严格分配到同一子集。所有报告的性能指标均为各折的平均值。所有模型均使用 Python3.12 在 PyTorch2.5.1 框架下实现, 并在单块 NVIDIA RTX A6000 GPU(48GB 显存)上训练。训练进行了 200 个 epoch, 批量大小为 4, 采用 Adam 优化器, 初始学习率为 0.001, 并使用余弦退火学习率调度器。对于模型中的 Transformer 组件, 层数和注意

力头数均设置为 4, 四个编码器阶段特征图的分块大小分别配置为 8×8、4×4、2×2 和 1×1。在下游分类任务中, ISPY2 数据集按患者划分为训练集(143 例)和验证集(36 例), 用于在真实 ROI 上训练稳健的分类器。剩余 50 例作为测试集。需要注意的是, 在最终评估中, 这 50 例的 ROI 并非真实标注, 而是由各比较模型生成的分割掩码动态生成。本任务训练了 200 个 epoch, 批量大小为 10。该文使用 AdamW 优化器, 初始学习率为 1e-6, 权重衰减为 1e-2, 并采用余弦退火学习率调度器, 监控验证集准确率。

为了全面评估模型性能, 该文为分割和分类任务采用了不同的指标集合。在分割任务中, 该文使用三个标准指标评估预测掩码的空间精度: Dice 系数(Dice coefficient, Dice)和交并比(Intersection over Union, IoU)用于衡量空间重叠程度, 95 百分位的豪斯多夫距离(95th percentile of the Hausdorff Distance, HD95)用于评估边界精确度。在下游 NAT 反应预测任务中, 从多个角度评估分类性能。生成混淆矩阵(Confusion Matrix, CM)以提供详细结果, 并据此计算总体准确率(overall Accuracy, ACC)、敏感性(Sensitivity, SE)、特异性(Specificity, SP)、F1 分数(F1-score, F1)和受试者工作特征曲线下面积(Area Under the receiver operating characteristic Curve, AUC), 以评估模型的区分能力、类别特异性表现及整体有效性。

## 2.3 结果与讨论

分割性能如表 1、表 2 所示, 该文在 BreastDM 和 ISPY1 数据集上对 RGU-Net 与多种先进模型进行了全面比较, 包括基于 CNN 的架构(U-Net、U-Net++(Zhou 等, 2020)、AAU-Net(Chen 等, 2023))、基于 Transformer 的架构(TransUNet(Chen 等, 2024)、Swin-UNet(Cao 等, 2021)、MTUNet(Wang 等, 2021))。Transformer 类模型(如 TransUNet 和 MTUNet)在较大的 ISPY1 数据集上展现出竞争力, 但在样本量受限的 BreastDM 数据集上性能出现显著退化, 深刻揭示了纯 Transformer 架构对训练数据规模的高度依赖性。相比之下, CNN 类架构(如 U-Net 及其变体)在有限数据条件下展现出更强的泛化鲁棒性, 但受限于局部感受野, 其性能天花板相对有限。值得注意的是, RGU-Net 在两个数据集的全部三项评价指标上均取得了最优表现, 充分验证了影像组学引导策略与多尺度特征融合机制的协同优势。分

表1 BreastDM上的分割结果

Table 1 Segmentation results on BreastDM

	Method	Dice	HD95	IoU
BreastDM	RGU-Net	88.68 ± 0.52	2.39 ± 0.11	79.67 ± 0.84
	U-Net	85.94 ± 0.26	3.30 ± 0.27	75.35 ± 0.41
	U-Net++	85.57 ± 0.17	3.05 ± 0.05	74.77 ± 0.27
	Swin-UNet	80.43 ± 0.63	8.71 ± 0.24	67.35 ± 0.87
	TransUNet	82.69 ± 2.18	5.05 ± 0.78	70.55 ± 3.13
	AAU-Net	86.59 ± 1.16	3.21 ± 0.33	76.36 ± 1.80
	MTUNet	74.22 ± 8.50	13.26 ± 11.95	59.48 ± 10.38

表2 ISPY1上的分割结果

Table 2 Quantitative segmentation results on ISPY1

	Method	Dice	HD95	IoU
ISPY1	RGU-Net	92.87 ± 0.33	2.57 ± 0.34	86.68 ± 0.56
	U-Net	91.40 ± 0.61	2.77 ± 0.59	84.17 ± 1.03
	U-Net++	91.70 ± 0.28	5.07 ± 0.41	84.68 ± 0.53
	Swin-UNet	81.57 ± 0.69	8.95 ± 0.66	69.41 ± 0.81
	TransUNet	91.72 ± 0.15	2.78 ± 0.29	84.71 ± 0.25
	AAU-Net	78.53 ± 5.39	7.51 ± 3.09	64.91 ± 7.29
	MTUNet	61.73 ± 9.90	25.86 ± 6.30	45.41 ± 10.83

割效果在图3中展示。图3共8列,列从左到右依次显示TransUNet、UNet、UNet++、AAUNet、Swin-UNet、MTUNet的输出,以及真实标注和该文提出的RGU-Net的预测结果。对于每个预测掩码,黄色表示真正例(True Positives, TP, 即正确分割部分),红色表示假正例(False Positives, FP, 即过度分割部分),而从真实标注可见的绿色表示假负例(False Negatives, FN, 即欠分割部分)。

消融实验结果如表3所示。提出的模块都对基线模型——通道交叉融合Transformer(Channel-wise Cross-fusion Transformer, CCT)架构作出了不同的贡

表3 ISPY1数据集上的消融实验结果

Table 3 Ablation study results on the ISPY1 dataset

	Method	Dice	HD95	IoU
ISPY1	CCT	90.34 ± 0.23	2.74 ± 0.26	82.39 ± 0.37
	DWT	90.72 ± 0.56	2.64 ± 0.22	83.02 ± 0.82
	SCA	91.70 ± 0.28	2.66 ± 0.31	84.68 ± 0.53
	RGU-Net	92.87 ± 0.33	2.57 ± 0.34	86.68 ± 0.56

献。表中DWT以及SCA分别代表CCT+DWT Pooling和CCT+DWT Pooling+SCA。DWT池化模块保留了传统最大池化中通常丢失的高频细节,从而提升了边缘清晰度和细微结构的表达能力。SCA模块增强了模型对肿瘤核心区域的关注,同时保持了全局上下文信息。RAT模块通过引入经调制筛选后的高层影像组学纹理先验作为语义锚点,实现了对低层特征的精细化引导,带来了最显著的性能提升,Dice提升约1.3%,IoU提升约2.4%,尤其是边界距离指标HD95实现了3.4%的相对下降,实现了肿瘤边缘的精确分割。这些改进表明,RAT引入的高层纹理先验能够有效强化语义对齐能力,使模型在结构一致性与边界精度方面均表现出更加稳定和可靠的提升。

NAT反应预测性能:如表4所示,RGU-Net在ISPY2数据集上实现了最高的AUC(0.838)和ACC(0.780),同时在SP为0.824、SE为0.688的情况下实现了平衡的性能表现,代表了所有模型中最具临床意义的性能权衡。相比之下,Swin-UNet虽然获得

表4 在ISPY2数据集上的分类结果

Table 4 Classification results on the ISPY2 dataset

	Method	ACC	AUC	F1	SE	SP
ISPY2	RGU-Net	0.780	0.838	0.667	0.688	0.824
	MTUNet	0.700	0.754	0.400	0.312	0.882
	SwinUNet	0.700	0.693	0.118	1.000	0.062
	TransUNet	0.640	0.756	0.500	0.562	0.676
	U-Net++	0.700	0.719	0.651	0.873	0.618
	U-Net	0.680	0.778	0.467	0.438	0.794
	AAU-Net	0.660	0.743	0.564	0.688	0.647
	ResNet50	0.604	0.749	0.387	0.400	0.697

了最高的敏感性,但以极低的特异性为代价,这表明其过度依赖于在pCR与npCR病例中共同存在的非特异性特征,导致了大量假阳性结果。尽管ResNet50在图像级任务中是一种性能强大的分类器,但在本任务中表现不佳,原因在于其易受背景干扰且无法有效分离病灶特异性特征,从而导致对患者的漏检。

如图4所示,该文进一步通过混淆矩阵展示不同模型的性能。这些结果突显了RGU-Net在融

合影像组学先验、多尺度频率特征以及跨尺度空间对齐方面的能力,从而实现了稳定、准确且可解释的分割与分类。该模型在不同类型的肿瘤(包括浸润性、多灶性和隐匿性病变)上均表现出色,表明其在新辅助治疗反应评估和个体化治疗方案制定中具有潜在的临床应用价值。

#### 2.4 特征融合可解释性

以往的影像组学模型在特征处理上多依赖人工手动筛选或无差别全量输入,存在较强的人为干预与主观盲目性;而本研究在损失函数中引入了L1稀疏正则化项,使模型能够在训练过程中通过梯度反向传播,自适应地筛选出对肿瘤分割结果影响最大的关键特征。具体而言,原始纹理特征在进入RAT模块前共计96维,经L1稀疏约束训练后,大量与任务无关维度的权重被自动压缩至接近于零,仅保留少量具有显著判别性的关键纹理维度参与后续调制,这一筛选过程完全由梯度驱动,无需任何人工先验介入。

在特征融合阶段,筛选后的纹理特征再由一阶统计特征与形状特征基于概率分布进行调制,并分别与编码器各层形状为[64, 160]、[128, 80]、[256, 40]、[512, 20]的图像特征进行跨尺度交互。参与调

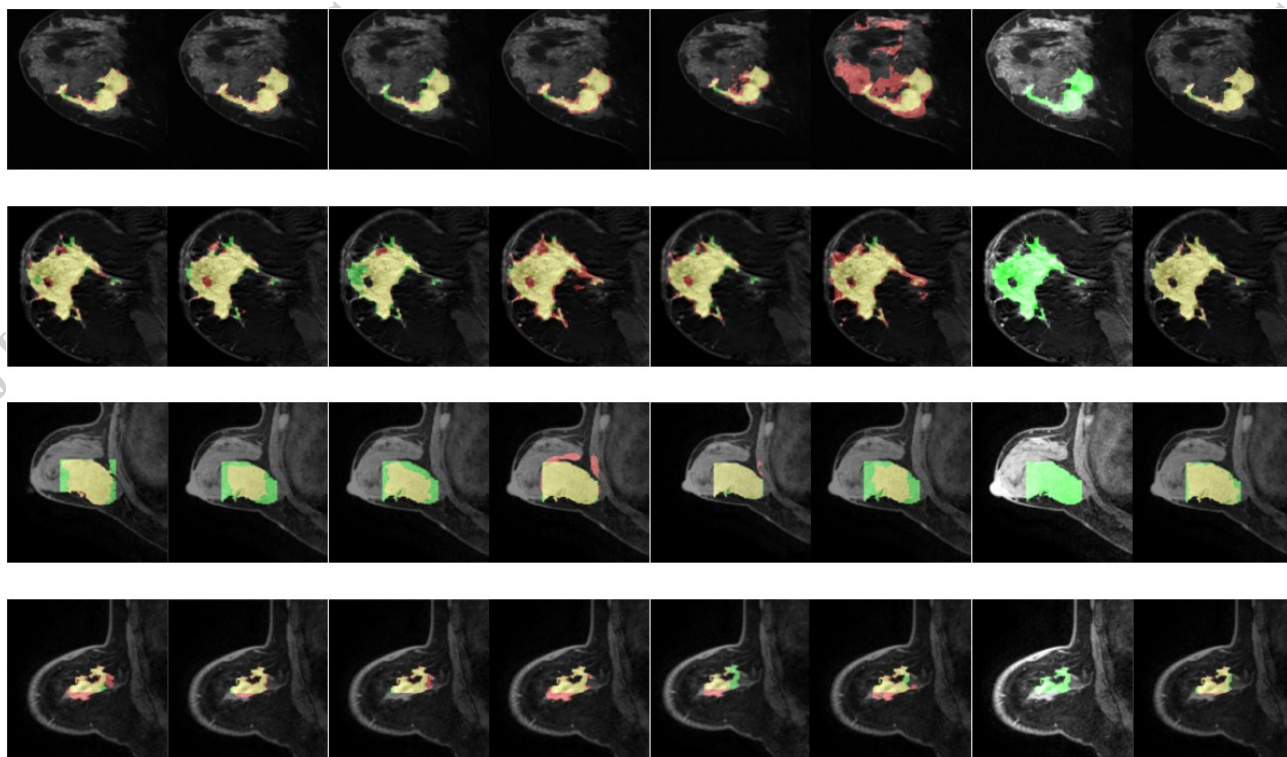


图3 分割结果的视觉比较。

Fig 3 Visual comparison of segmentation results.

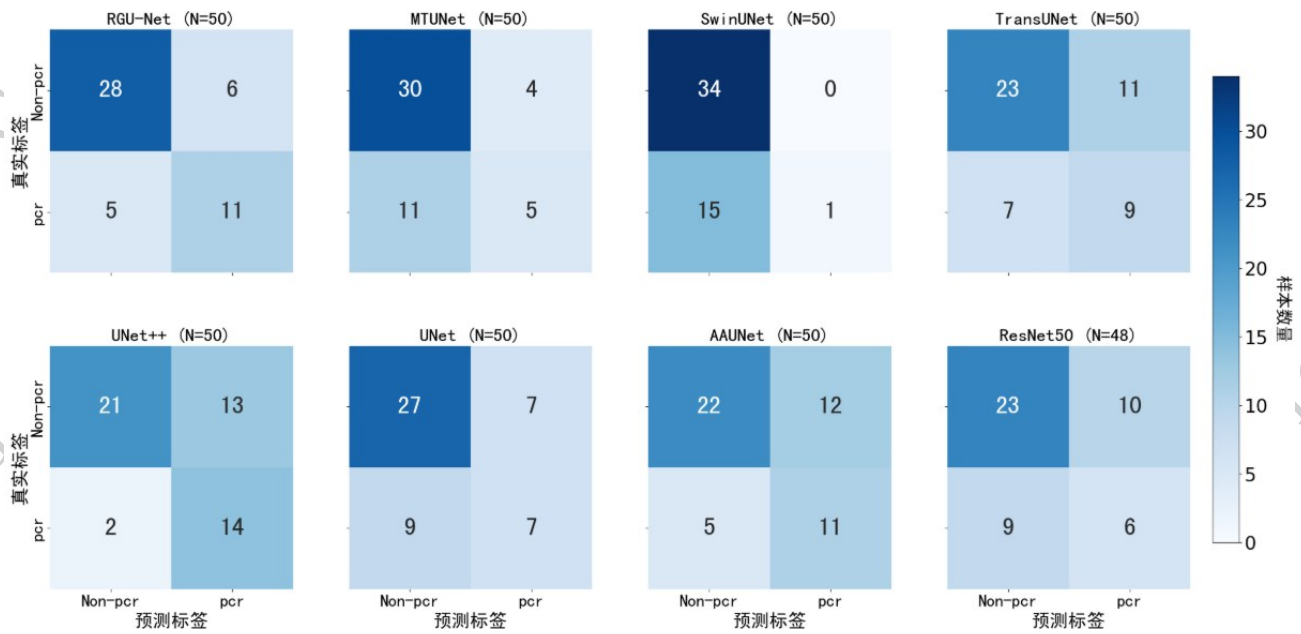


图4 混淆矩阵展示不同模型在ISPY2数据集上的分类性

Fig 4 Confusion matrices of different models on the ISPY2 dataset

制的组学特征维度仅为个位到十位数量级,而受其引导的深度图像特征在四个尺度上累计已达到数百万维,如此悬殊体量下少量组学先验仍能稳定作用于大规模图像特征,恰恰说明这些被筛选保留的维度承载了高度浓缩的语义信息,而非冗余噪声,这一以小驭大的现象本身即是对特征融合可解释性的有力佐证。

从实验层面说明该设计的有效性,我们在ISPY1数据集上开展了消融实验:在已包含CCT、DWT和SCA三个模块的基础上,进一步引入组学特征引导即构成完整的RGU-Net,其Dice由91.70%提升至92.87%,IoU由84.68%提升至86.68%,HD95由2.66降至2.57。该结果表明,经L1稀疏筛选并参与调制的组学特征确实承载了可被模型利用的判别性信息,为本文所主张的可解释性特征融合提供了直接的定量证据。

### 3 结论

该文提出了一种基于影像组学引导的乳腺癌NAT疗效预测框架——RGU-Net。该框架创新性地融合了DWT频域池化、RAT组学增强Transformer以及SCA空间交叉注意力三大核心模块,系统性地提升了肿瘤边界精细刻画能力与上下文感知的多尺度特征学习能力。在三个公开乳腺MRI基准数据集上

的系统验证表明, RGU-Net在分割精度、边界准确性与跨数据集泛化性方面均显著优于现有最先进模型。此外,基于分割结果提取的深度特征在NAT反应预测任务中取得了显著的性能提升,充分验证了该框架在个体化乳腺癌精准治疗中提供可靠、可解释临床决策支持的巨大应用潜力。未来工作将进一步探索多模态影像数据融合与外部多中心队列验证,以推动模型向临床实际部署迈进。

### 参考文献

- Cantini L, Trapani D, Guidi L, Boscolo Bielo L, Scafetta R, Koziej M, et al. 2024. Neoadjuvant therapy in hormone receptor-positive/HER2-negative breast cancer. *Cancer Treatment Reviews*, 123: 102669 [DOI: 10.1016/j.ctrv.2023.102669]
- Cao H, Wang Y Y, Chen J, Jiang D S, Zhang X P, Tian Q, et al. 2021. Swin-Unet: Unet-like pure transformer for medical image segmentation[EB/OL].[2026-05-17]. <https://arxiv.org/abs/2105.05537>
- Chang L C, Liu J M, Zhu J L, Wang Z H, Huang X M, Li Y H, et al. 2025. Advancing precision medicine: the transformative role of artificial intelligence in immunogenomics, radiomics, and pathomics for biomarker discovery and immunotherapy optimization. *Cancer Biology & Medicine*, 22(1): 33-47 [DOI: 10.20892/j.issn.2095-3941.2024.0376]
- Chen G P, Li L, Dai Y, Zhang J and Yap M H. 2023. AAU-Net: an adaptive attention U-Net for breast lesions segmentation in ultra-

- sound images. *IEEE Transactions on Medical Imaging*, 42(5): 1289-1300 [DOI: 10.1109/TMI.2022.3226268]
- Chen J N, Mei J R, Li X H, Lu Y J, Yu Q H, Wei Q, et al. 2024. TransUNet: rethinking the U-Net architecture design for medical image segmentation through the lens of transformers. *Medical Image Analysis*, 97: 103280 [DOI: 10.1016/j.media.2024.103280]
- Chitalia R D, Rowland J, McDonald E S, Clark A S, DeMichele A and Kontos D. 2020. Imaging phenotypes of breast cancer heterogeneity in preoperative breast dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) scans predict 10-year recurrence. *Clinical Cancer Research*, 26(4): 862-869 [DOI: 10.1158/1078-0432.CCR-18-4067]
- Chitalia R D, Pati S, Bhalariao M, Gastounioli A, Jahani N, Cohen E, et al. 2021. Expert tumor annotations and radiomic features for the ISPY1/ACRIN 6657 trial data collection [DB/OL]. [2026-05-17]. <https://www.cancerimagingarchive.net> [DOI: 10.7937/TCIA.XC7A-QT20]
- Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. 2013. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *Journal of Digital Imaging*, 26(6): 1045-1057 [DOI: 10.1007/s10278-013-9622-7]
- Duanmu H Y, Huang P B, Brahmavar S, Lin S, Ren T, Kong J, et al. 2020. Prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer using deep learning with integrative imaging, molecular and demographic data//*Medical Image Computing and Computer Assisted Intervention - MICCAI 2020*. Lima: Springer: 242-252 [DOI: 10.1007/978-3-030-59710-8\_24]
- El Adoui M, Drisis S and Benjelloun M. 2019. Predict breast tumor response to chemotherapy using a 3D deep learning architecture applied to DCE-MRI data//*International Work-Conference on Bioinformatics and Biomedical Engineering*. Granada: Springer: 33-40 [DOI: 10.1007/978-3-030-17935-9\_4]
- Fan M, Wang K L, Pan D, Zheng B and Li H. 2024. Radiomic analysis reveals diverse prognostic and molecular insights into the response of breast cancer to neoadjuvant chemotherapy: a multicohort study. *Journal of Translational Medicine*, 22(1): 637 [DOI: 10.1186/s12967-024-05487-y]
- Ha R, Chin C, Karcich J, Liu M Z, Chang P, Mutuzan E, et al. 2019. Prior to initiation of chemotherapy, can we predict breast tumor response? Deep learning convolutional neural networks approach using a breast MRI tumor dataset. *Journal of Digital Imaging*, 32(5): 693-701 [DOI: 10.1007/s10278-018-0144-1]
- Huang G, Liu Z, Van Der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE: 4700-4708 [DOI: 10.1109/CVPR.2017.243]
- Huang Y H, Zhu T, Zhang X L, Huang Y, Li Y Y, Chen Y, et al. 2023. Longitudinal MRI-based fusion novel model predicts pathological complete response in breast cancer treated with neoadjuvant chemotherapy: a multicenter, retrospective study. *eClinicalMedicine*, 58: 101899 [DOI: 10.1016/j.eclinm.2023.101899]
- Joo S, Ko E S, Kwon S, Jeon E, Jung H, Kim J Y, et al. 2021. Multimodal deep learning models for the prediction of pathologic response to neoadjuvant chemotherapy in breast cancer. *Scientific Reports*, 11(1): 18800 [DOI: 10.1038/s41598-021-98408-8]
- Li Y T, Fan Y H, Xu D H, Qiang L, Zhang Y, Zhang H J, et al. 2023. Deep learning radiomic analysis of DCE-MRI combined with clinical characteristics predicts pathological complete response to neoadjuvant chemotherapy in breast cancer. *Frontiers in Oncology*, 12: 1041142 [DOI: 10.3389/fonc.2022.1041142]
- Lobbes M B I, Prevos R, Smidt M, Tjan-Heijnen V C G, Van Goethem M, Schipper R, et al. 2013. The role of magnetic resonance imaging in assessing residual disease and pathologic complete response in breast cancer patients receiving neoadjuvant chemotherapy: a systematic review. *Insights into Imaging*, 4(2): 163-175 [DOI: 10.1007/s13244-013-0219-y]
- Luo M, Huang J and Yang F. 2014. MRI brain tumor segmentation method based on multi-modal 3D-CNNs feature extraction. *Science Technology and Engineering*, 14(31): 78-83 (罗蔓, 黄靖, 杨丰. 2014. 基于多模态3D-CNNs特征提取的MRI脑肿瘤分割方法. *科学技术与工程*, 14(31): 78-83)
- Montemurro F, Nuzzolese I, Ponzzone R and Agostinetto E. 2020. Neoadjuvant or adjuvant chemotherapy in early breast cancer? *Expert Opinion on Pharmacotherapy*, 21(9): 1071-1082 [DOI: 10.1080/14656566.2020.1746273]
- Newitt D C, Partridge S C, Zhang Z, Gibbs J E, Chenevert T L, Rosen M A, et al. 2021. ACRIN 6698/I-SPY2 breast DWI [DB/OL]. [2026-05-17]. <https://www.cancerimagingarchive.net> [DOI: 10.7937/TCIA.KK02-6D95]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*. Munich: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4\_28]
- Roy A G, Navab N and Wachinger C. 2018. Concurrent spatial and channel squeeze and excitation in fully convolutional networks//*Medical Image Computing and Computer Assisted Intervention - MICCAI 2018*. Granada: Springer: 421-429 [DOI: 10.1007/978-3-030-00928-1\_48]
- Siegel R L, Giaquinto A N and Jemal A. 2024. Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians*, 74(1): 12-49 [DOI: 10.3322/caac.21820]
- Shi J, Wang L L, Wang S S, Chen Y X, Wang Q, Wei D M, et al. 2020. Applications of deep learning in medical imaging: a survey. *Journal of Image and Graphics*, 25(10): 1953 - 1981 (施俊, 汪琳琳, 王珊珊, 陈艳霞, 王乾, 魏冬铭等. 2020. 深度学习在医学影像中的应用综述. *中国图象图形学报*, 25(10): 1953 - 1981) [DOI: 10.11834/jig.200255]
- Shi J, Wang T T, Zhu Z Q, Zhao M F, Wang B X and An H. 2025. © 中国图象图形学报版权所有

- Deep learning-based medical image segmentation methods. *Journal of Image and Graphics*, 30(6): 2161 - 2186 (石军, 王天同, 朱子琦, 赵敏帆, 王炳勋, 安虹. 2025. 基于深度学习的医学图像分割方法综述. *中国图象图形学报*, 30(6): 2161 - 2186) [DOI:10.11834/jig.240467]
- Waswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. 2017. Attention is all you need [EB/OL]. [2026-05-17]. <https://arxiv.org/abs/1706.03762>
- Volders J H, Negenborn V L, Spronk P E, Krekel N M, Schoonmade L J, Meijer S, et al. 2018. Breast-conserving surgery following neoadjuvant therapy - a systematic review on surgical outcomes. *Breast Cancer Research and Treatment*, 168(1): 1-12 [DOI: 10.1007/s10549-017-4598-5]
- Wang H N, Cao P, Wang J Q and Zaiane O R. 2022. UCTransNet: rethinking the skip connections in U-Net from a channel-wise perspective with transformer // Proceedings of the 36th AAAI Conference on Artificial Intelligence. Vancouver: AAAI Press: 2441-2449 [DOI: 10.1609/aaai.v36i3.20144]
- Wang H Y, Xie S, Lin L F, Iwamoto Y, Han X H, Chen Y W, et al. 2021. Mixed transformer U-Net for medical image segmentation [EB/OL]. [2026-05-17]. <https://arxiv.org/abs/2111.04734>
- Wang Y Y and Yu J Q. 2023. Advances in imaging and radiomics of axillary lymph node metastasis in breast cancer. *Radiologic Practice*, (05): 662-666 (汪媛媛, 余建群. 2023. 乳腺癌腋窝淋巴结转移的影像及影像组学研究进展. *放射学实践*, (05): 662-666) [DOI:10.13609/j.cnki.1000-0313.2023.05.024]
- Xia B Q, Wang H, Wang Z, Dong M, Luo M L, Zheng B, et al. 2021. A combined nomogram model to predict disease-free survival in triple-negative breast cancer patients with neoadjuvant chemotherapy. *Frontiers in Genetics*, 12: 783513 [DOI: 10.3389/fgene.2021.783513]
- Xie T W, Gong J, Zhao Q F, Wang Z, Chen Z H, Zhang K, et al. 2024. Development and validation of peritumoral vascular and intratumoral radiomics to predict pathologic complete responses to neoadjuvant chemotherapy in patients with triple-negative breast cancer. *BMC Medical Imaging*, 24(1): 136 [DOI: 10.1186/s12880-024-01311-7]
- Xiong X, Zheng L W, Ding Y, Chen Y F, Cai Y W, Wang L P, et al. 2025. Breast cancer: pathogenesis and treatments. *Signal Transduction and Targeted Therapy*, 10: 49 [DOI: 10.1038/s41392-024-02108-4]
- Zhao X M, Liao Y H, Xie J H, He X X, Zhang S Q and Zhang X X. 2023. BreastDM: a DCE-MRI dataset for breast tumor image segmentation and classification. *Computers in Biology and Medicine*, 164: 107255 [DOI: 10.1016/j.combiomed.2023.107255]
- Zhou Z W, Siddiquee M M R, Tajbakhsh N and Liang J M. 2020. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 39(6): 1856-1867 [DOI: 10.1109/TMI.2019.2959609]