

# 融合梯度幅值引导门控的Transformer-CNN网络用于乳腺超声图像分割

董欣乐<sup>1</sup>, 张欣纯<sup>2</sup>, 黄梦思<sup>3</sup>, 胡嫣然<sup>2</sup>, 鲁意<sup>1</sup>

<sup>1</sup>云南大学信息学院计算机科学与技术, 云南 昆明

<sup>2</sup>云南大学信息学院物联网工程, 云南 昆明

<sup>3</sup>云南大学信息学院智能科学与技术, 云南 昆明

收稿日期: 2026年3月9日; 录用日期: 2026年4月9日; 发布日期: 2026年4月20日

## 摘要

乳腺超声图像中病灶的精准分割对乳腺癌的早期诊断与治疗规划具有重要意义。然而, 由于超声图像固有的噪声干扰、边界模糊以及病灶形态多样性, 传统基于卷积神经网络或Transformer的分割方法在全局语义建模与局部细节保持之间存在显著矛盾。为此, 本研究提出了一种梯度引导空间注意力门控机制, 并将其深度融合于一种新颖的Transformer-CNN并行混合架构中, 构建出GMG-TCNet (Gradient-guided Mixed Attention Network)。该网络通过自适应梯度门控动态提取并增强图像中的边缘先验信息, 结合空间注意力机制进一步抑制背景干扰, 实现了多尺度特征的精准对齐与融合。在四个公开乳腺超声数据集上的系统性实验表明, GMG-TCNet在mDice、mIoU、Precision等关键指标上均显著优于U-Net、Swin-UNet、TransUNet等主流模型, 尤其在边界分割精度上表现出明显优势。消融研究进一步验证了梯度引导机制与双分支结构的有效性。本研究不仅为乳腺超声图像分割提供了一种高性能解决方案, 也为先验知识与深度学习的协同建模提供了新思路。

## 关键词

梯度幅值引导, 注意力门控, Transformer-CNN混合网络, 乳腺超声图像分割

## A Transformer-CNN Network Fusing Gradient Magnitude-Guided Gating for Breast Ultrasound Image Segmentation

Xinle Dong<sup>1</sup>, Xinchun Zhang<sup>2</sup>, Mengsi Huang<sup>3</sup>, Yanran Hu<sup>3</sup>, Yi Lu<sup>1</sup>

<sup>1</sup>Department of Computer Science and Technology, Yunnan University, Kunming Yunnan

文章引用: 董欣乐, 张欣纯, 黄梦思, 胡嫣然, 鲁意. 融合梯度幅值引导门控的Transformer-CNN网络用于乳腺超声图像分割[J]. 计算机科学与应用, 2026, 16(4): 159-168. DOI: 10.12677/csa.2026.164118

<sup>2</sup>Department of Internet of Things Engineering, Yunnan University, Kunming Yunnan

<sup>3</sup>Department of Intelligence Science and Technology, Yunnan University, Kunming Yunnan

Received: March 9, 2026; accepted: April 9, 2026; published: April 20, 2026

## Abstract

Accurate segmentation of breast lesions in ultrasound images is crucial for early diagnosis and treatment planning of breast cancer. However, due to inherent noise, blurred boundaries, and morphological diversity of lesions, traditional segmentation methods based solely on Convolutional Neural Networks or Transformers face significant challenges in balancing global semantic modeling and local detail preservation. To address this, we propose a Gradient-guided Spatial Attention Gate and deeply integrate it into a novel Transformer-CNN parallel hybrid architecture, forming the GMG-TCNet. The network dynamically extracts and enhances edge prior information through an adaptive gradient gate, combined with a spatial attention mechanism to suppress background interference, achieving precise alignment and fusion of multi-scale features. Systematic experiments on four public breast ultrasound datasets demonstrate that GMG-TCNet significantly outperforms state-of-the-art models such as U-Net, Swin-UNet, and TransUNet in key metrics including mDice, mIoU, and Precision, particularly excelling in boundary segmentation accuracy. Ablation studies further validate the effectiveness of the gradient guidance mechanism and dual-branch structure. This study not only provides a high-performance solution for breast ultrasound image segmentation but also offers new insights into the synergistic modeling of prior knowledge and deep learning.

## Keywords

Gradient Magnitude Guidance, Attention Gate, Transformer-CNN Hybrid Network, Breast Ultrasound Image Segmentation

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

乳腺癌是全球女性中最常见的恶性肿瘤之一，早期诊断对提高生存率至关重要。超声成像因其无辐射、低成本、实时性好等优势，成为乳腺筛查的首选影像学手段。然而，乳腺超声图像常受限于声学伪影、低对比度及组织异质性，使得病灶边界的视觉辨识与自动分割极具挑战。目前，基于深度学习的语义分割方法已在医学图像分析中取得显著进展。卷积神经网络凭借其局部感知与平移不变性，在纹理特征提取方面表现优异，但其感受野有限，难以建模长程依赖关系。而 Transformer 通过自注意力机制实现全局上下文建模，在捕获语义信息方面具有优势，却对局部细节敏感度不足。尽管已有研究尝试将 CNN 与 Transformer 融合，但多数方法仍存在特征对齐不精准、边界模糊等问题，尤其在处理形态多变、边缘不清的乳腺病灶时表现不佳。

为此，本研究提出一种梯度引导的混合注意力网络，通过引入图像梯度先验作为边缘增强信号，设计了一种可学习的空间注意力门控模块，并将其嵌入到 Transformer-CNN 双分支架构中。该方法不仅实现了全局与局部特征的互补融合，还通过梯度引导机制显式增强了边界区域的响应，从而提升了分割的

准确性与鲁棒性。

本研究的主要贡献包括：

1. 提出了梯度引导空间注意力门控机制，实现边缘感知的特征校准；
2. 构建了 Transformer-CNN 并行混合架构，协同建模全局语义与局部细节；
3. 在两个公开数据集上验证了方法的优越性与泛化能力。

## 2. 研究现状

乳腺癌早期诊断对提高患者存活率至关重要。传统的超声诊断依赖人工辨别，存在主观性强且效率低的问题。为此，国内外学者聚焦于开发计算机辅助诊断(CAD)方法，利用深度学习技术实现乳腺结节的自动分割。

国外在乳腺超声分割领域已形成较为成熟的研究体系。早期研究多以 U-Net [1]及其改进架构为基准，通过卷积神经网络提取病灶特征。近年来，Transformer [2]凭借全局建模优势在该领域表现突出，研究重点逐渐转向全局依赖建模与模型轻量化。然而，如何有效克服超声图像低对比度与噪声干扰，仍是现有研究的核心瓶颈。

相比之下，国内研究更多聚焦于算法的精细化优化与临床适配性。国内学者在构建“双分支”融合架构[3] [4]方面取得了显著进展，通过结合局部特征提取与全局长程依赖，提升了复杂背景下对病变区域的定位精度。此外，针对临床筛查需求，国内研究在模型轻量化及移动端部署方面也进行了积极探索[5]。尽管如此，在多尺度病灶的协同提取与特征高效融合等方面仍有待优化。

近年来，针对医学图像中边界模糊的问题，研究者提出了多种边界增强策略，主要包括基于损失函数的约束方法与基于注意力机制[6]的结构改进方法。典型的边界损失(Boundary Loss)通过在优化目标中显式引入轮廓约束，从而提升分割结果的边界一致性。然而，该类方法依赖于损失层面的间接监督，难以在特征提取阶段对边缘信息进行显式建模。另一类方法通过引入边缘注意力模块(Edge Attention/Spatial Attention)，利用额外分支或注意力权重强化边界区域响应，但其边缘信息通常来源于浅层特征或学习得到的隐式表示，缺乏稳定的物理先验支撑。

相比之下，本文提出的 GSA-Gate 通过引入梯度幅值作为显式边缘先验，将传统图像梯度信息与深度特征进行融合建模。一方面，梯度幅值能够直接反映像素级的强度变化，具有明确的边界指示意义；另一方面，通过可学习的门控网络对梯度信息进行自适应调制，使其能够动态适配不同层级特征。这种“显式先验 + 可学习调制”的机制，在边界增强的稳定性与表达能力之间取得了更优平衡。

在 CNN-Transformer 混合架构方面，现有方法(如 TransUNet [7]、Swin-UNet [8]等)多采用串行融合或编码器-解码器嵌入方式，主要侧重于全局语义建模能力的提升。然而，这类方法通常依赖隐式特征学习进行多尺度融合，缺乏对局部结构信息(尤其是边界)的针对性建模。本文提出的并行双分支结构通过在特征提取阶段同时建模局部纹理(CNN)与全局依赖(Transformer)，并结合加权融合机制实现动态信息分配，同时借助 GSA-Gate 对边界区域进行显式强化，从而在结构设计层面实现了全局-局部-边界的协同优化。

## 3. 数据集

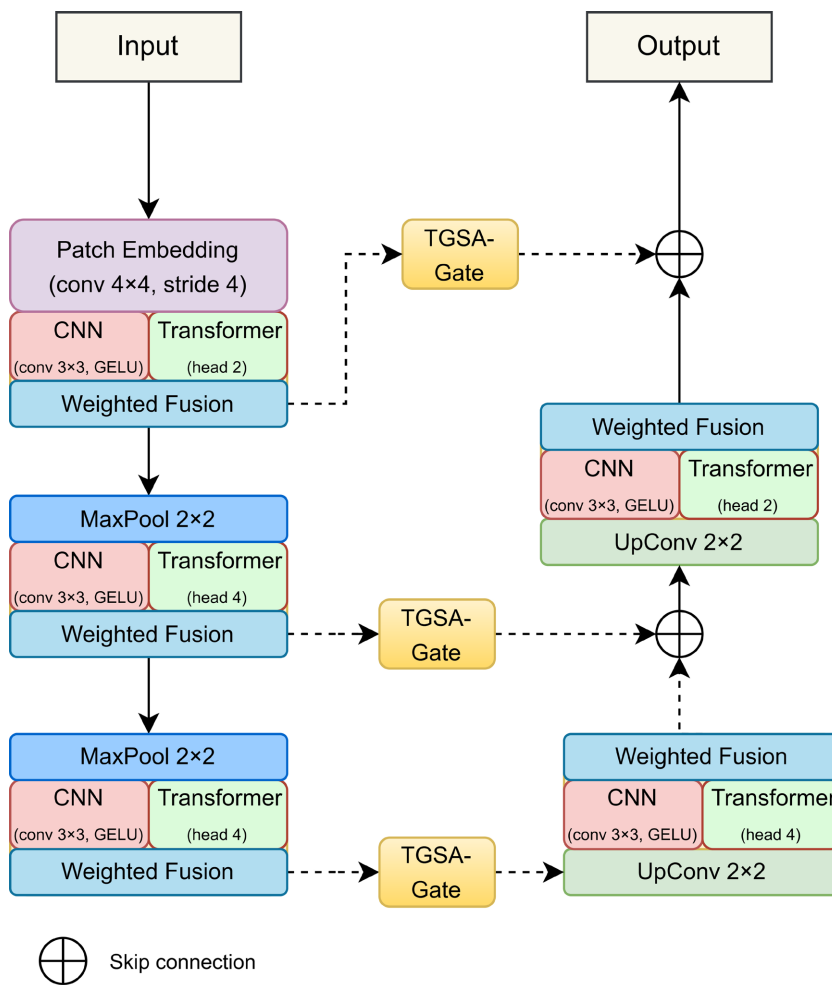
本研究选取了两个在图像质量上具有显著差异的公开数据集来验证我们提出的方法，总共包括 1591 张图片。表 1 简单总结了这两个数据集中样本的分布情况。Dataset1 包含了来自 BUSI 的 780 张图片，这些图像由埃及巴尼苏韦夫大学医院采集。Dataset2 包含了来自 BUS\_UC 的 811 张图片，其原始数据来源于公开的超声病例库，并由经验丰富的放射科医师提供了像素级分割标注。

**Table 1.** Dataset distribution  
**表 1.** 数据集分布情况

病变情况	数据集	
	Dataset1	Dataset2
良性	437	358
恶性	210	453
正常	133	-
总计	780	811

### 4. 网络架构设计

本研究提出梯度引导混合注意力网络(Gradient-guided Mixed Attention Network, GMG-TCNet), 整体架构如图 1 所示。GMG-TCNet 采用双分支编码器 - 解码器结构, 主要由 Transformer 分支、CNN 分支和梯度空间注意力门控模块(GSA-Gate)三部分组成。



**Figure 1.** Overall architecture of GMG-TCNet  
**图 1.** GMG-TCNet 整体架构图

## 4.1. 双分支编码器

输入图像首先经过  $4 \times 4$  卷积(步长为 4)进行补丁嵌入(Patch Embedding), 将图像划分为不重叠的  $16 \times 16$  像素块, 并映射到特征空间。随后特征分别输入两个并行分支: CNN 分支采用经典卷积结构, 包含 3 层下采样模块, 每层由两个  $3 \times 3$  卷积层、批归一化层(BatchNorm)和 GELU 激活函数组成, 使用最大池化(MaxPool  $2 \times 2$ )进行下采样, 该分支专注于提取局部纹理特征和边缘细节, 感受野逐步扩大, 能够有效捕捉乳腺病灶的微结构特征。Transformer 分支基于多头自注意力机制(Multi-head Self-Attention), 各层注意力头数分别为 2、4、4, 该分支通过全局上下文建模, 建立图像各区域间的长距离依赖关系, 特别适合学习器官的整体形态结构和解剖学先验, Transformer 分支的具体超参数设置如表 2 所示。双分支编码器的具体实现细节如图 2 所示。

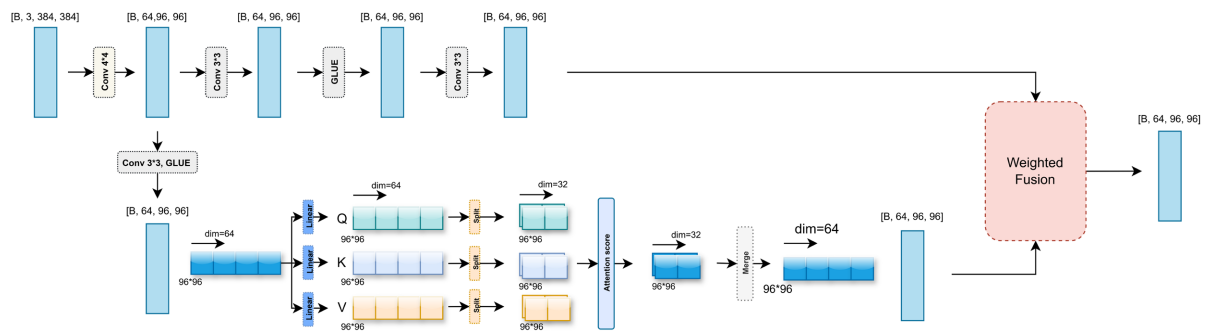


Figure 2. Dual-branch structure (taking the first-layer encoder as an example)

图 2. 双分支结构(以第一层编码器为例)

Table 2. Hyperparameters of transformer branch

表 2. Transformer 分支的超参数设置

阶段	通道数	注意力头数量	层数
编码器 1	64	2	2
编码器 2	128	4	2
瓶颈层	256	4	2
解码器 1	128	4	-
解码器 2	64	2	-

## 4.2. 跨层特征融合机制(Weighted Fusion)

为有效集成局部细节特征  $F_{\text{local}}$  与全局语义特征  $F_{\text{global}}$ , 本网络在两个分支的对应层级之间设计了一个加权融合模块(Weighted Fusion)。该模块通过学习一组自适应的通道权重向量  $\alpha \in R^{1 \times C \times 1 \times 1}$  和  $\beta \in R^{1 \times C \times 1 \times 1}$  (其中  $\alpha + \beta = 1$ ), 对两个分支的特征进行通道级的加权融合:

$$\mathcal{F}_{\text{total}} = \alpha \cdot \mathcal{F}_{\text{local}} + \beta \cdot \mathcal{F}_{\text{global}} \quad (1)$$

这里  $\odot$  表示逐通道乘法。权重  $\alpha$  和  $\beta$  由基于全局上下文的门控机制(fusion gate)根据输入特征动态生成, 该门控首先对拼接后的特征进行全局平均池化, 将空间信息压缩为通道描述向量; 随后通过  $1 \times 1$  卷积实现通道间信息交互, 并通过 Sigmoid 函数生成归一化权重系数, 使网络能依据输入内容自适应地决定对局部与全局信息的信赖程度。最终输出的融合特征  $F_{\text{fused}}$  兼具细节分辨力与语义一致性, 为下游的分

割或检测头提供了更鲁棒、更全面的特征表示。

在编码器前三个层级，融合后的特征分别经过下采样进入下一层。

### 4.3. 梯度空间注意力门控(GSA-Gate)

在编码器的前两层，我们引入了梯度空间注意力门控模块，其结构如图 3 所示。该模块首先通过 Sobel 算子计算输入特征的梯度幅值，将梯度特征与原始特征拼接后，送入轻量级门控网络(Gating Network)生成空间注意力权重图。该权重图能够显式增强边缘区域的响应，抑制均匀组织区域的激活，为解码器提供精确的边界引导信息。

在 GSA-Gate 中，门控网络(Gating Network)采用轻量级卷积结构实现，其具体形式为两层卷积网络。首先，将原始特征与归一化梯度幅值在通道维度进行拼接，输入至第一层  $3 \times 3$  卷积(通道数由  $2C$  降至  $C/r$ ，其中  $r$  为压缩比，本文设为 4)，并依次经过 Batch Normalization 与 ReLU 激活函数。随后，通过第二层  $3 \times 3$  卷积将特征映射恢复至  $C$  维，并通过 Sigmoid 函数生成最终的空间注意力权重图。该设计在保证表达能力的同时，有效控制了参数规模与计算开销。

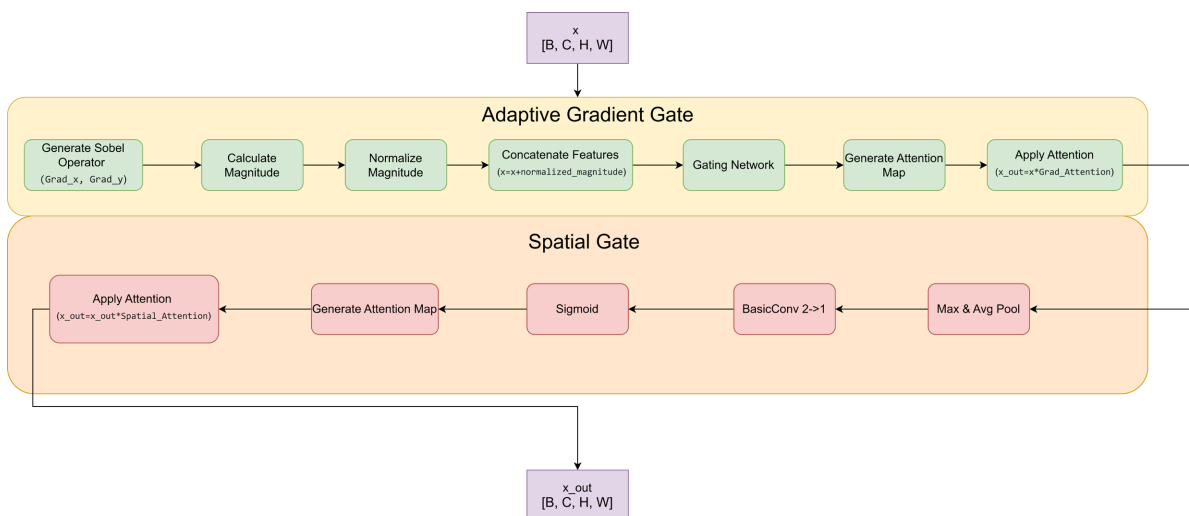


Figure 3. Workflow of GSA-Gate

图 3. GSA-Gate 的工作流程

### 4.4. 解码器设计

解码器同样采用双分支结构，通过转置卷积(UpConv  $2 \times 2$ )逐步恢复空间分辨率。在每层上采样后，同样使用 GSA-Gate 模块对特征进行边缘增强，并与对应层级的编码器特征进行跳跃连接(Skip Connection)，确保空间细节的完整保留。解码器最后一层将两个分支的特征进行最终融合，通过  $1 \times 1$  卷积输出分割概率图，通过 4 倍线性上采样恢复图像尺度。

整个网络采用端到端方式训练，通过结合 CNN 的局部感知能力和 Transformer 的全局建模优势，在有限标注数据下实现精确的乳腺病灶分割。

## 5. 实验结果及其分析

### 5.1. 训练策略

为保证实验结果的稳健性与统计可靠性，本研究在 Dataset1 与 Dataset2 上均采用五折交叉验证进行

模型训练与评估。在每一折实验中，数据集首先随机划分出 20% 作为独立测试集，用于模型最终性能评估；剩余 90% 数据再进一步划分为 80% 训练集与 20% 验证集，其中验证集用于模型参数调优与早停策略的判定。该划分方式保证了测试集在整个训练过程中完全不可见，从而避免信息泄露并提升评估的公平性。所有模型均在相同的训练策略下进行公平比较。训练过程中采用 AdamW 优化器进行参数更新，其兼顾自适应学习率与权重衰减，有利于缓解过拟合并提升收敛稳定性。初始学习率设置为  $1 \times 10^{-4}$ ，并在验证集性能不再提升时采用学习率衰减策略进行动态调整，以促进模型收敛至更优解。在每一折训练中，模型以验证集上的损失与 mIoU 作为主要监控指标，用于模型选择与早停，以防止过拟合并提高泛化能力。

## 5.2. 损失函数

针对乳腺超声图像分割中前景 - 背景类别严重不平衡的问题，本研究采用二值交叉熵损失(Binary Cross-Entropy, BCE)与 Dice 损失的加权组合作为最终优化目标：

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{BCE}} + \beta \cdot \mathcal{L}_{\text{Dice}} \quad (2)$$

其中，BCE 损失用于保证像素级分类的稳定性，而 Dice 损失能够直接优化区域重叠度，对小目标与边界区域更为敏感，两者的结合可以在提升整体分割精度的同时，更好地保持病灶区域的完整性与轮廓一致性。在本研究中，BCE 损失与 Dice 损失的权重系数分别设置为  $\alpha = 0.5$  和  $\beta = 0.5$ ，以在像素级分类精度与区域重叠度之间取得平衡。

## 5.3. 对比实验结果

表 3 展示了各方法在 Dataset1 与 Dataset2 上的实验结果，GMG-TCNet (Ours) 在两个数据集上均取得了最优性能。以 Dataset1 为例：GMG-TCNet 的 mDice 达到 85.84%，较 U-Net (82.17%) 提升约 3.67 个百分点，较 Swin-UNet (79.47%) 提升约 6.37 个百分点；在边界敏感的 mIoU 指标上，GMG-TCNet 达到 77.18%，显著优于 U-Net (72.74%)、PSPNet (73.59%) 等方法，体现出其在边缘分割任务中的精准性优势。

在 Dataset2 中，GMG-TCNet 的性能提升更为明显：其 mDice 达到 95.47%，较 U-Net (92.32%) 提升约 3.15 个百分点，较 Swin-UNet (90.77%) 提升约 4.7 个百分点；Recall 指标达到 93.11%，同样领先于所有对比方法，进一步验证了 GMG-TCNet 在不同数据集上的泛化能力与分割精度优势。

**Table 3.** Test results of different methods on Dataset1 and Dataset2

**表 3.** 不同方法在数据集 Dataset1 和 Dataset2 上的测试结果

Datasets	Methods	mDice (%)	mIoU (%)	Precision (%)	Recall (%)	Specificity (%)	Accuracy (%)	Auc_pr (%)
Dataset1	U-Net	82.17	72.74	77.54	58.59	<b>98.61</b>	95.58	72.14
	PSPNet	82.85	73.53	76.11	61.80	98.39	95.62	74.52
	DeepLabV3+	84.32	75.29	<b>78.98</b>	64.18	98.60	95.99	77.12
	ViT	77.15	67.07	62.56	53.59	97.32	94.01	60.49
	Swin-UNet	79.47	69.52	64.75	59.65	97.30	94.45	66.96
	<b>GMG-TCNet (Ours)</b>	<b>85.84</b>	<b>77.18</b>	75.70	<b>72.09</b>	98.09	<b>96.12</b>	<b>80.36</b>

续表

	U-Net	92.32	86.26	86.73	86.39	98.10	96.63	93.72
	PSPNet	93.41	88.32	89.02	90.32	98.25	97.67	94.23
	DeepLabV3+	95.30	<b>91.55</b>	<b>93.18</b>	91.00	<b>99.04</b>	<b>98.03</b>	<b>97.74</b>
Dataset2	ViT	88.44	80.39	82.01	77.57	97.55	95.04	88.23
	Swin-UNet	90.77	82.32	85.67	79.40	97.62	96.12	90.13
	<b>GMG-TCNet (Ours)</b>	<b>95.47</b>	90.84	89.25	<b>93.11</b>	98.35	97.81	96.82

#### 5.4. 消融实验结果

为了验证 GSA-Gate 模块的有效性,本研究在 Dataset1 与 Dataset2 上针对 GMG-TCNet 进行了消融实验,结果如表 4 所示。实验数据表明:在仅采用基础双分支结构(Base)时,两数据集的 mDice 指标分别为 84.88%和 94.49%;引入梯度门控(G-Gate)后,性能小幅提升至 85.08%与 94.67%;进一步集成空间注意力门控(SAM-Gate)后,指标持续优化至 85.51%和 95.01%;而当应用完整的 GSA-Gate 模块(Ours)时,模型性能达到最优,在两数据集上分别实现了 85.84%与 95.47%的最高 mDice。上述结果充分证明了 GSA-Gate 各组件间具有显著的协同效应,能够有效且逐步地提升模型的分割精度。

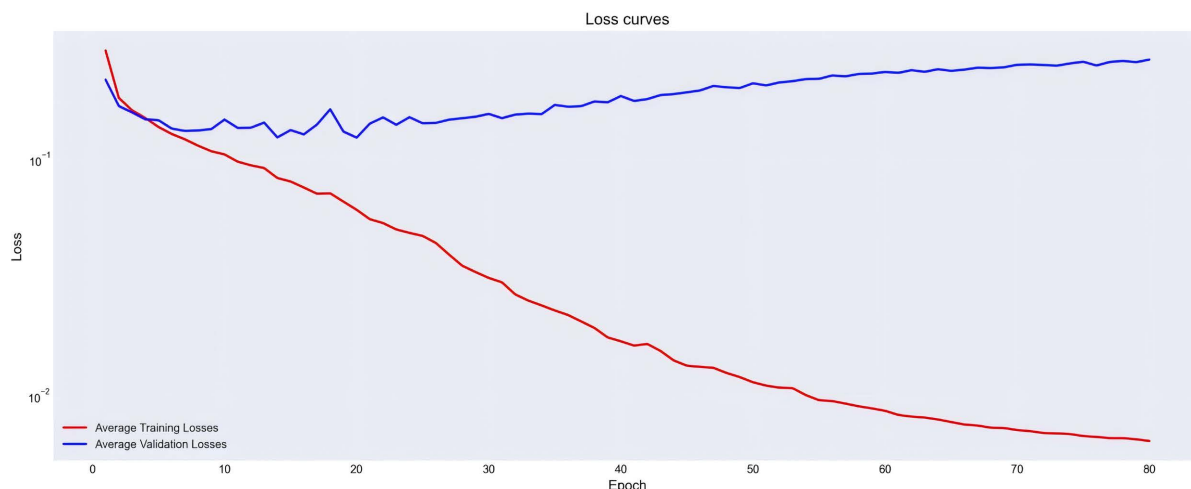
**Table 4.** Test results of GMG-TCNet variants on datasets Dataset1 and Dataset2

**表 4.** GMG-TCNet 的变体在数据集 Dataset1 和 Dataset2 上的测试结果

Datasets	Variants	mDice (%)	mIou (%)	Precision (%)	Recall (%)	Specificity (%)	Accuracy (%)	Auc_pr (%)
Dataset1	Base	84.88	75.93	73.01	71.32	97.81	95.80	79.39
	Base G-Gate	85.08	76.21	75.82	69.21	98.19	95.99	79.74
	Base + SAM-Gate	85.51	76.80	<b>76.68</b>	70.29	<b>98.23</b>	96.11	79.83
	<b>Base + GSA-Gate (Ours)</b>	<b>85.84</b>	<b>77.18</b>	75.70	<b>72.09</b>	98.09	<b>96.12</b>	<b>80.36</b>
Dataset2	Base	94.49	89.85	88.55	92.34	98.28	97.54	96.33
	Base + G-Gate	94.67	90.04	<b>89.49</b>	92.60	98.30	97.60	96.53
	Base + SAM-Gate	95.01	90.50	89.20	92.92	<b>98.44</b>	97.71	96.64
	<b>Base + GSA-Gate (Ours)</b>	<b>95.47</b>	<b>90.84</b>	89.25	<b>93.11</b>	98.35	<b>97.81</b>	<b>96.82</b>

#### 5.5. 模型收敛性与拟合状态分析

图 4 所展示的损失曲线(Loss curves)直观反映了 GMG-TCNet 模型在训练过程中的学习状态。



**Figure 4.** Training loss curve of GMG-TCNet on Dataset1  
**图 4.** GMG-TCNet 在 Dataset1 上的训练损失曲线图

由图 4 所示结果，可以得出如下结论：

#### 1. 模型具备良好的收敛效率

训练过程中，训练损失(红色曲线)与验证损失(蓝色曲线)在训练前期均呈现显著下降趋势，表明模型具备较快的收敛速度，能够迅速学习到数据中的有效特征。这一现象反映出模型初始化及优化策略的有效性，为后续训练提供了良好的起点。

#### 2. 模型拟合程度充分且训练稳定

随着训练轮次的增加，训练损失持续下降并逐渐趋于平缓，说明模型对训练数据的表征能力逐步增强，拟合过程稳定推进。验证损失在经历初期下降后维持在相对稳定的低水平区间，未出现显著波动，表明模型在训练过程中未出现明显的欠拟合现象，具备充分的学习能力。

#### 3. 模型具有较强的泛化能力

值得注意的是，验证损失并未随着训练损失的进一步降低而出现回升趋势，训练损失与验证损失之间始终保持合理的差距，未呈现明显的“发散”现象。这证明模型在提升对训练数据拟合能力的同时，并未牺牲其泛化性能，有效规避了过拟合风险，显示出良好的正则化效果与结构稳定性。

综上所述，本模型在训练过程中表现出理想的动态特征：收敛迅速、拟合充分且泛化性能稳定。训练曲线符合监督学习中模型训练的预期规律，反映出模型结构设计、损失函数选择以及超参数设置的合理性。这一稳健的训练状态为模型在后续任务中的可靠性能提供了理论支持与实践保障。

## 6. 结论

本研究针对乳腺超声图像分割中边界模糊、形态多变的问题，提出了一种基于梯度引导注意力机制的 Transformer-CNN 混合网络 GMG-TCNet。通过引入可学习的梯度门控与空间注意力模块，网络能够显式增强边缘特征，并实现全局与局部信息的有效融合。在多个公开数据集上的实验表明，GMG-TCNet 在分割精度、边界一致性等方面均优于现有主流方法。未来研究可进一步优化计算效率，并将其拓展至更多医学影像分析任务中。

## 基金项目

云南大学大学生创新创业训练项目资助(校级, 20251238)。

## 参考文献

- [1] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., *et al.*, Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Springer International Publishing, 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [2] Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention Is All You Need. <http://arxiv.org/abs/1706.03762>
- [3] Chen, Y., Li, J., Xiao, H., *et al.* (2018) Dual Path Networks. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 4470-4478.
- [4] 张恩琿, 林帅, 陈金令, 等. PFTransCNN: 基于 CNN-Transformer 双分支融合的病理图像分割[J]. 微电子学与计算机, 2026, 43(3): 88-97.
- [5] Wu, T., Tang, S., Zhang, R., Cao, J. and Zhang, Y. (2021) CGNet: A Light-Weight Context Guided Network for Semantic Segmentation. *IEEE Transactions on Image Processing*, **30**, 1169-1179. <https://doi.org/10.1109/tip.2020.3042065>
- [6] 邓酩, 徐锦凡, 肖洪祥, 等. 改进 TransUNet 的高效通道注意力医学图像分割网络[J]. 计算机应用, 2025, 45(12): 4037-4044.
- [7] Chen, J.N., Lu, Y.Y., Yu, Q.H., *et al.* (2021) TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. <https://doi.org/10.48550/arXiv.2102.04306>
- [8] Cao, H., Wang, Y.Y., Chen, J., *et al.* (2021) Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. <https://arxiv.org/abs/2105.05537>