

# MGA-Net:融合多尺度与通道移位的超声图像分割网络

吴雨芯,邵党国,马磊

(昆明理工大学信息工程与自动化学院,昆明 650500)

E-mail:wu\_yx2025@163.com

**摘要:**针对医学超声图像分割中存在的斑点噪声干扰、病灶边界模糊及多尺度特征捕获不足等问题,本文提出了一种融合多尺度特征聚合和多组通道移位策略的轻量化网络 MGA-Net. 该网络通过引入多组通道移位下采样模块(MgCSD)和组聚合桥接模块(GAB),在降低模型复杂度的同时提升分割精度. MgCSD 模块采用通道空间移位策略与尺度不变卷积,通过分组循环移位操作建立中心区域细节与边缘轮廓的关联性,结合全局加权增强机制保留多尺度纹理特征,相较传统下采样方法减少 30%~40% 参数量;GAB 模块通过深度可分离卷积对齐编码器高级语义特征与解码器低级细节特征,利用掩膜引导的分组多尺度空洞卷积实现跨层级特征高效融合,有效缓解语义鸿沟问题. 实验表明,在 BUSI 乳腺超声数据集上,MGA-Net 的 Dice 系数达到 93.4%,交并比 IoU 88.9%,平均对称表面距离 ASSD 1.02 像素,参数量仅 23.8M,较 U-Net 提升 6.2% 分割精度的同时降低 23% 计算复杂度. 可视化结果显示,本文的方法在保留病灶细微结构、抑制伪影干扰方面具有显著优势,为临床医学影像的计算机辅助诊断提供了高效可靠的技术支持.

**关键词:**医学图像分割;多尺度特征;空洞卷积;卷积神经网络

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2026)04-1008-08

## MGA-Net: an Ultrasound Image Segmentation Network Incorporating Multiscale and Channel Shift

WU Yuxin, SHAO Dangguo, MA Lei

(Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

**Abstract:** Aiming at the problems of speckle noise interference, blurred lesion boundaries and insufficient multi-scale feature capture in medical ultrasound image segmentation, this paper proposes a lightweight network, MGA-Net, which integrates multi-scale feature aggregation and multi-group channel shift strategy. This network reduces the model complexity while improving the segmentation accuracy through the introduction of the multi-group channel shift downsampling module (MgCSD) and the group aggregation bridging module (GAB). The MgCSD module adopts the channel space shift strategy with scale-invariant convolution to establish the correlation between center region details and edge contours through group cyclic shift operation. The MgCSD module employs a channel space shifting strategy with scale-invariant convolution to establish the correlation between the center region details and the edge contours through group cyclic shifting operation, which is combined with a global weighting enhancement mechanism to retain the multi-scale texture features, and reduces the number of covariates by 30%~40% compared to the traditional downsampling method; the GAB module aligns the high-level semantic features of the encoder and the low-level detail features of the decoder through a depth-separable convolution. GAB module aligns encoder high-level semantic features with decoder low-level detail features through depth-separable convolution, and utilizes mask-guided grouped multi-scale cavity convolution to achieve efficient cross-layer feature fusion, effectively alleviating the semantic divide problem. Experiments show that on the BUSI breast ultrasound dataset, MGA-Net achieves a Dice coefficient of 93.4%, an intersection-to-merge ratio IoU of 88.9%, an average symmetric surface distance ASSD of 1.02 pixels, and a parameter count of only 23.8M, which improves segmentation accuracy by 6.2% while reducing computational complexity by 23% compared to U-Net. The visualization results show that this method has significant advantages in preserving the fine structure of the lesion and suppressing the interference of artifacts, which provides efficient and reliable technical support for computer-aided diagnosis of clinical medical images.

**Keywords:** medical image segmentation; multiscale features; null convolution; convolutional neural networks

## 0 引言

医学超声图像分割作为计算机辅助诊断的核心技术,在

新生儿脑发育评估<sup>[1]</sup>、甲状腺结节良恶性鉴别、心脏动态分析<sup>[2]</sup>等临床场景中具有重要应用价值. 与传统成像模式相比,超声图像具有实时、无创、低成本等优势,因而在临床检查

收稿日期:2025-03-25 收修改稿日期:2025-04-21 基金项目:国家自然科学基金项目(62266025)资助. 作者简介:吴雨芯,女,1999年生,硕士研究生,CCF会员,研究方向为图像处理、深度学习、人工智能;邵党国(通信作者),男,1979年生,博士,副教授,CCF会员,研究方向为计算机视觉、自然语言处理、医学图像;马磊,男,1978年生,硕士,高级工程师,研究方向为医疗信息技术、智慧医疗、计算机视觉.

中被广泛应用<sup>[3]</sup>。然而,由于超声成像系统受成像技术机制的限制,超声图像中难免会出现斑点噪声,会严重影响图像质量,从而对临床诊断和对超声图像的分割任务等都产生一定的影响<sup>[4]</sup>。另外,医学成像技术的快速发展使得对医学图像的分析成为医生诊断疾病的重要依据,但由于诊断过程费时费力且非常依赖医生的专业知识,大大增加了医生的工作量<sup>[5]</sup>。因此通过计算机对医学图像自动分割可以显著减轻医生的压力。由于超声图像目标区域与背景区域对比度低、分割边界模糊且病灶的尺度差异较大,对医学超声图像的分割是一项具有挑战性的任务。医学图像分割的目的是将目标区域从图像中准确定位并分割出来,不仅有助于量化病变的大小和形状,而且对于后续的疾病诊断、治疗规划和疗效评估具有重要意义<sup>[6]</sup>。

深度学习中卷积神经网络是其中一种特殊且重要的框架,卷积神经网络凭借它具有的局部感知能力和权值共享功能在计算机视觉领域取得了许多成果。近年来,基于深度学习的分割方法取得显著进展。2015年 Shelhamer 等人<sup>[7]</sup>提出的全卷积神经网络(FCN)首次实现端到端像素级预测,通过反卷积层恢复空间分辨率,在自然图像语义分割任务中表现出了强大的性能,但其跳跃连接的简单叠加难以捕捉医学图像的复杂拓扑结构,后来, Ronneberger 等人<sup>[8]</sup>提出的具有编码-解码结构和跳跃连接的 U-Net 网络通过对称编码器-解码器架构与跨层特征融合,在电子显微镜图像分割中取得突破性进展,为医学图像分割带来了巨大的进步,大大的提升了医学图像的分割精度。随后, U-Net ++<sup>[9]</sup> 通过嵌套密集连接优化特征传递路径,但参数量增加导致计算效率下降,后来的轻量化改进方法虽然在一定程度上降低了计算复杂度,却牺牲了轮廓细节的捕捉能力。另外,医学超声图像与自然图像不同,医学超声图像的质量相对较低,目标往往不明显,并且不同结构之间的边界比较模糊,这些问题导致网络在进行超声图像分割时会出现一些错误分割或者对小目标区域的遗漏,并且, U-Net 网络还具有计算复杂度高的问题,后来的一些轻量化方法虽降低资源消耗,却牺牲了超声图像中关键的轮廓信息。 Attention U-Net<sup>[10]</sup> 通过门控单元抑制无关区域特征响应,用于解决医学图像中目标区域与背景对比度低、边界模糊等问题<sup>[11]</sup>,注意力机制能聚焦于肿瘤区域,抑制周围肝实质的干扰,但在超声图像中容易受到噪声干扰,注意力模块可能放大噪声区域的权重,导致分割结果出现异常凸起<sup>[12]</sup>,而且注意力模块增加了额外的参数和计算量,推理速度难以满足术中实时需求; Swin-UNet<sup>[13]</sup> 采用纯 Transformer 架构建立全局依赖, Transformer 的自注意力机制捕捉器官间的空间关系,避免了卷积神经网络中局部感受野的局限性,但由于缺乏局部归纳偏置,在超声图像中,对于斑点噪声和微小钙化灶的局部细节敏感性不足,导致边缘分割粗糙。 DPMUNet<sup>[14]</sup> 通过双池化自注意力模块的协同作用,双池化策略能抑制斑点噪声的影响,结合自注意力模块的全局上下文建模,提高了模型在低信噪比区域的鲁棒性,但需要大量数据学习有效权重,若实验仅基于小规模数据集,可能泛化能力不足,尤其是对于罕见斑块类型(如溃疡性斑块)上表现不佳。最近, Ali 等<sup>[15]</sup> 提出了一种基于自适应进化算法的深度学习架构搜索框架,通过进化优化策略自动设计或调整网络结构,以提升分割精度并

适应超声图像的特性,但其高昂的计算成本、泛化性不足及可解释性缺陷制约了临床实用性。这些方法在复杂解剖结构分割中仍存在 3 个主要缺陷:1) 浅层网络对细节特征捕获不足导致边缘模糊<sup>[16]</sup>; 2) 跨尺度特征交互效率低下导致不同层次特征之间的语义差异较大,使得它们在融合时难以有效结合<sup>[17]</sup>; 3) 单一模态训练限制模型泛化能力<sup>[18]</sup>。

基于以上背景与医学超声图像数据的特性,本文提出了一种用于医学超声图像分割的多尺度特征聚合网络 MGA-Net (Multi-scale Global Aggregation Network), 通过动态通道分组与移位操作,在降低特征图分辨率的同时保留高频细节信息。在公开数据集 BUSI 上进行实验验证。

## 1 相关技术

### 1.1 多组通道移位下采样模块

传统的 U-Net 架构有大量的卷积层和池化层,浅层卷积层能够有效地获取局部边缘特征,但有可能丢失轮廓结构信息,更深层次的卷积层虽然能够通过扩展感受野并捕捉轮廓特征的上下文关系,也有可能丢失像边缘线这样的细节信息。因此在重复的卷积池化操作过程中会丢失大量的细节信息,为了弥补这一损失,本文引入了一种基于多组通道空间移位加权的下采样方法 (Multi-group Channel Shift Downsampling, MgCSD), 用来平衡全局与局部特征的提取,通过通道移位策略和特征融合机制,在减少信息损失的同时增强轮廓特征的表达力。

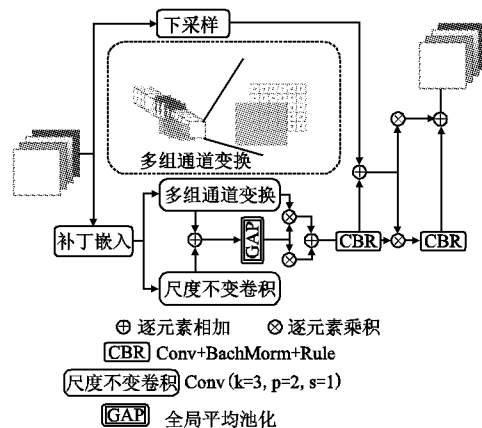


图 1 多组通道移位下采样模块

Fig. 1 Multi-group channel shift downsampling module

移位操作如图 1 所示,其中特征根据通道维度被分成组。第 1 组保持原始位置,保留中心区域的细节特征,其余的根据循环向上和循环向右移位,在宽度和高度上移位,保留区域中元素的相对位置。移位后,非特征区域用零填充。移位后的特征用  $X_{shift}$  表示。移位操作将边缘区域与中心特征联系起来,但导致移位区域的详细信息丢失。因此,使用尺度不变卷积来获得细节信息,用  $X_{supply}$  表示。公式如式(1)所示:

$$X_{supply} = Conv_{3 \times 3}(X_{shift}) \quad (1)$$

受通道自注意力模块的启发,该模块通过利用全局平均池化的结果,在通道维度上增强了  $X_{shift}$  和  $X_{supply}$  的性能,并将其融合以获得全局增强权重  $X_L$ :

$$X_L = \text{GAP}(X_{\text{shift}} \oplus X_{\text{supply}}) \otimes (X_{\text{shift}} \oplus X_{\text{supply}}) \quad (2)$$

其中 $\otimes$ 和 $\oplus$ 是逐元素操作. 下采样过程保留完整的局部特征, 记作 $X_R$ . 将 $X_L$ 和 $X_R$ 相加的结果映射得到一个交互关系矩阵 $\text{Omixer}$ . 对于 $X_R$ ,  $\text{Omixer}$ 保留了右侧分支的空间特征与全局特征. 权重与文献[19]中的空间自注意力模块相同, 融合过程表示如式(3)所示:

$$X_L = \text{GAP}(X_{\text{shift}} \oplus X_{\text{supply}}) \otimes (X_{\text{shift}} \oplus X_{\text{supply}}) \quad (3)$$

其中,  $\text{Omixer}, X_L, X_R \in \mathbf{R}^{c \times h \times w}$ .

MgCSD模块代替和增强了传统的下采样模块, 能够充当编码阶段的直接替代物. 但因为全局加权分支是以特征而不是原始图像为目标的, 所以在将原始图像映射到特征空间时, 保留第1层的简单卷积. 尽管MgCSD模块在下采样过程中增加了操作的复杂性, 但它通过减少特征图的尺寸降低了后续处理阶段的计算负担. 与传统卷积下采样相比, MgCSD在相同输入尺寸下参数数量减少30%~40%.

## 1.2 组聚合桥接模块

在U-Net及其各类变体的网络系列中, 位于编码器与解码器之间的跳跃连接, 会把来自编码器的浅层、低级别且细粒度的特征映射, 与解码器的深层、高级别以及粗粒度的特征映射相互结合. 实验证明, 跳跃连接能够有效地还原目标的细节, 即便在复杂的背景环境下, 也能够生成精细的分割掩码. 在医学图像分割领域, 相较于于自然图像, 对分割精度要求更高. 因为在许多场景中, 自然图像并不一定需要精细的分割掩码, 然而在医学图像中, 需要分割的结果尽可能地精细. 所以, 设计出更为有效的跳跃连接, 以解决由于低分辨率与高分辨率之间语义差距过大, 而引发的特征映射模糊问题, 进而有效地分割出医学图像中的目标细节是很有必要的.

组聚合桥接模块(Group Aggregation Bridge, GAB)模块, 用于融合高级特征、低级特征和掩膜, 通过通道和宽高对齐, 使用不同空洞率的空洞卷积进行特征融合. 有效解决了传统U-Net跳跃连接在复杂医学图像中的局限性. 用GAB模块作为编码器与解码器之间的桥接模块, 可以融合来自深层网络的高级语义特征、来自浅层网络的低级细节特征和来自解码器的多尺度输出的中间掩膜预测.

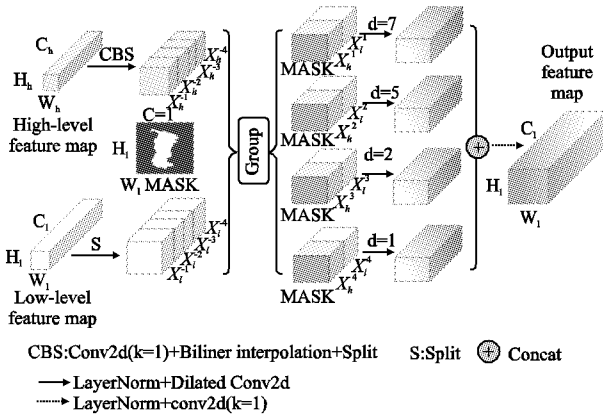


图2 GAB模块结构

Fig. 2 GAB module structure

基本结构如图2所示. 核心结构分为3个处理阶段. 首先使用深度可分离卷积(DW)和双线性插值来调整高级特征的

大小, 以匹配低级特征的大小, 深度可分离卷积的参数数量远小于标准卷积. 然后, 将两个特征图沿着通道维度划分为4组, 并将低级特征中的一组与高级特征中的一组连接起来, 以获得4组融合特征. 对于每组融合特征, 将掩码与之连接. 接下来, 应用不同空洞率的空洞卷积(核大小为3, 空洞率分别为{1, 2, 5, 7})到不同的组上, 以提取不同尺度的信息. 最后, 将这4组沿着通道维度连接起来, 用核大小为1的普通卷积, 实现不同尺度特征之间的交互. GAB模块通过掩膜引导的分组多尺度融合与深度监督优化, 在极低参数数量下实现了跨层级特征的高效整合. 设计结构充分考虑了医学图像中病灶形态多样性和边界模糊性的特点, 为轻量化模型的多尺度处理提供了创新思路.

## 2 基于多尺度特征聚合的分割模型

MGA-Net的模型框架在经典U-Net架构基础上进行了多维度改进如图3所示. 其核心改进体现在特征提取与融合机制的重构, 通过引入多组通道移位下采样模块(MgCSD)和组聚合桥接模块(GAB), 构建了具有自适应多尺度感知能力的深度网络体系. 模型以基于编码器和解码器组成的U形架构U-Net为基础架构, 在编码器部分, 网络采用六级渐进式特征提取结构: 前三级沿用传统 $3 \times 3$ 卷积核捕获局部纹理信息, 后三级则替换为MgCSD模块. 该模块通过通道分组移位操作建立跨区域特征关联, 并辅以尺度不变卷积补充细节信息, 显著提升了全局特征表征能力. 这种分阶段设计既保留了浅层网络的细节捕获优势, 又通过深层网络的高阶抽象实现了特征表达的升维, 同时有效控制了模型参数数量.

与U-Net中简单的跳跃连接不同, 本文提出的模型在每一级编码器和解码器之间都连接了一个GAB模块. 该模块融合编码器的多级特征与解码器生成的动态掩膜信息, 通过分组空洞卷积策略提取多尺度上下文特征. 具体而言, GAB将低级细节特征、高级语义特征及中间掩膜预测进行多模态融合, 借助门控机制实现特征选择与增强, 形成具有空间自适应的特征传递路径, 利用深层网络生成的不同尺度的掩膜预测, 以此作为GAB模块的输入之一. 在每个阶段生成的掩膜用于指导后续的特征提取, 最终通过最大池化和加法操作整合特征. 模型通过6个阶段逐步提取和融合特征, 生成多个中间掩膜(Mask1至Mask5), 各阶段生成的中间掩膜不仅作为监督信号优化网络训练, 更通过反馈机制引导后续特征提取过程. 最终通过多尺度特征聚合技术, 将最大池化获得的全局信息与逐级细化特征进行融合, 实现分割结果的精准输出.

这种架构实现了3个维度的性能提升: MgCSD模块增强了跨区域特征关联能力, GAB模块建立了动态特征选择机制, 而渐进式掩膜生成策略则强化了模型的自适应性. 三者协同作用使网络能够充分挖掘图像的多层次特征, 在保持U-Net架构简洁性的同时显著提升了复杂场景下的分割鲁棒性.

## 3 实验验证

### 3.1 数据集和预处理

BUSI(Breast Ultrasound Image)是一个包含乳腺超声图

像的分类和分割数据集. 数据集包括了 2018 年收集的乳腺超声波图像, 涵盖了 25 ~ 75 岁的 600 名女性患者. 数据集由 780 张图像组成, 包含 133 张正常图像, 表现为光滑、均匀的界限,

回声为低回声或中等回声; 良性肿瘤图像 437 张, 表现为边界清晰的高回声结节, 可能有晕圈; 恶性肿瘤图像 210 张, 表现为边界不清晰的高回声结节, 常伴有后方阴影和边缘不规则.

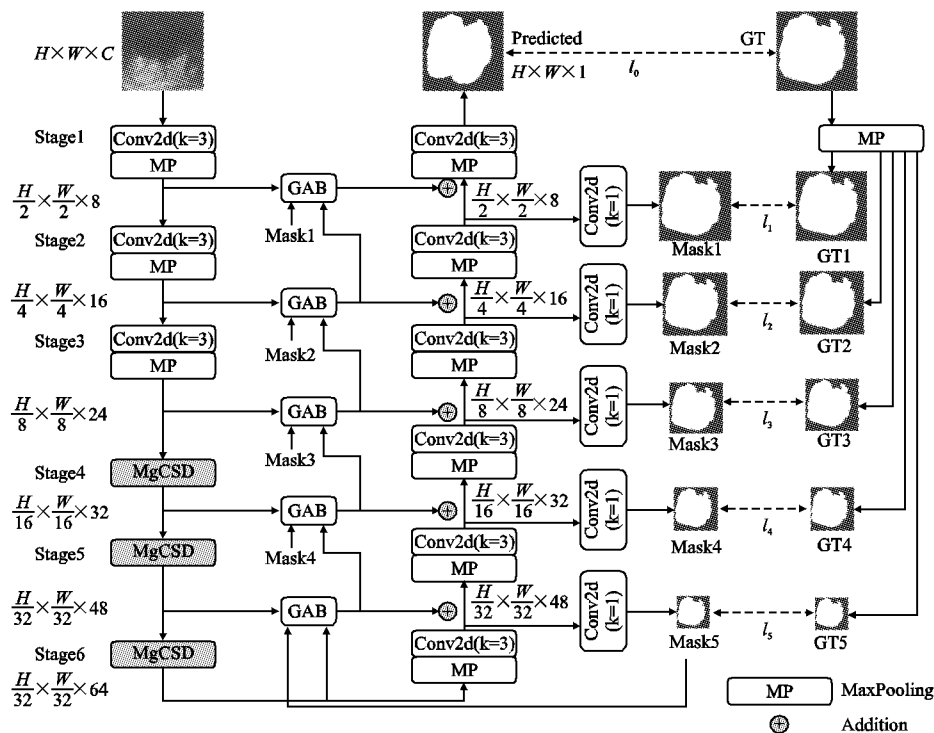


图 3 MGA-Net 的架构

Fig. 3 Architecture of MGA-Net

每张图像的平均大小为  $500 \times 500$  像素. 而在良性和恶性乳腺超声图像中, 还包含了对应胸部瘤的详细分割标注, 为深入研究和精准诊断提供了关键信息. 本实验中去除了 133 张正常的没有结节的超声图片, 将剩余的 647 张图像按 80%, 10%, 10% 随机划分为训练集, 验证集和测试集. 图像尺寸统一调整为  $256 \times 256$  像素, 以方便训练和测试. 通过将每个像素值除以 255 把像素值从 0 ~ 255 的原始范围缩放到 0 ~ 1 的新范围来进行数据标准化. 由于良恶性类别不平衡, 还采用水平翻转、放大裁剪、旋转等方法增强数据集. 将所有图像转换为灰度图, 以减少计算资源消耗.

### 3.2 损失函数

医学超声图像分割可以看作是像素级别的二元分类任务. 研究中选择骰子损失 (Dice loss) 作为损失函数. 骰子损失是两个样本之间重叠区域的度量指标, 适用于二值图像分割任务, 可以在一定程度上缓解正负样本之间的数量不平衡问题. 其具体公式如式 (4) 所示:

$$Dice = \frac{2 \sum_{i=1}^n x_i \cdot y_i + \varepsilon}{\sum_{i=1}^n x_i + \sum_{i=1}^n y_i + \varepsilon} \quad (4)$$

$$\Gamma_{Dice} = -\ln(Dice) \quad (5)$$

$x_i \in [0, 1]$ ,  $y_i \in [0, 1]$  分别表示模型分割区域和 Ground Truth 分割区域的像素值.  $\varepsilon$  为一个很小的值, 取  $10^{-5}$ , 目的是避免分母为 0,  $n = H \times W$  表示像素数,  $\Gamma_{Dice}$  表示骰子损失,  $\ln$  表示以无理数  $e$  为底的对数.

### 3.3 评价指标

4 个指标被用于评估模型的性能, 分别为骰子得分 (Dice)、交并比 (IoU)、平均对称表面距离 (ASSD) 和准确率 (Precision).

$$Dice = \frac{2 \sum_{i=1}^n x_i \cdot y_i + \varepsilon}{\sum_{i=1}^n x_i + \sum_{i=1}^n y_i + \varepsilon} \quad (6)$$

$$IoU = \frac{\sum_{i=1}^n x_i \cdot y_i + \varepsilon}{\sum_{i=1}^n x_i + \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot y_i + \varepsilon} \quad (7)$$

$x_i \in [0, 1]$ ,  $y_i \in [0, 1]$  分别表示模型分割区域和 Ground Truth 分割区域的像素值.  $\varepsilon$  为一个很小的值, 取  $10^{-5}$ , 目的是避免分母为 0,  $n = H \times W$  表示像素数.

其中骰子得分所表示的是模型预测结果的像素集合与 Ground Truth 的像素集合之间的像素重叠数, 它可以充分的衡量分割结果与真实标签之间的样本重叠程度, 它对细致微小的特征或边界信息特别敏感.

IoU 指标又叫做 Jaccard 指数, 用于评估模型的分割质量. 它表示的含义是模型分割区域的像素集合与 Ground Truth 的像素集合之间的交集与并集的比值, 所以被称为交并比.

$$ASSD = \frac{1}{|S_x| + |S_y|} \times \left( \sum_{x \in S_x} d(x, S_y) + \sum_{y \in S_y} d(y, S_x) \right) \quad (8)$$

$S_x$  和  $S_y$  分别表示模型的分割边界点集合和 Ground Truth 的分割边界点集合,  $d(n, S_x) = \min_{m \in S_x} (\|n - m\|)$  表示  $n$  与  $S_x$  上的所有点的最小欧氏距离。

平均对称曲面距离 (Average Symmetric Surface Distance, ASSD) 被用于评估模型的分割结果与 Ground Truth 之间的几何形状差距。ASSD 通过计算模型预测结果以及 Ground Truth 之间所有点最短距离的平均值来评估两者之间的相近程度。

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

其中,  $TP$  和  $FP$  分别表示正确分割的像素和被错误的标记为病灶的背景像素。准确率表示模型分割结果的所有正样本中的正确分割数所占的比例, 即模型真正预测正确的病灶区域占真实病灶区域的比例。

### 3.4 实验设置

本研究基于 PyTorch 深度学习框架搭建完整的实验体系, 采用模块化设计思想构建端到端的训练流程。在优化策略方面, 选用具有动态学习率调节能力的自适应矩估计优化器 (Adam), 参数设置为  $\beta_1 = 0.9, \beta_2 = 0.999$ , 通过指数衰减机制平衡历史梯度的一阶矩和二阶矩估计, 有效缓解参数震荡问

题。针对图像分割任务中常见的类不平衡现象, 采用骰子损失 (Dice Loss) 作为目标函数, 该损失通过计算预测掩膜与真实标签的交并比, 有效缓解边缘区域误分割问题, 其平滑常数设置为  $1 \times 10^{-5}$  以避免除零错误。训练参数初始学习率设置为  $1 \times 10^{-4}$ , 权重衰减设置为  $1 \times 10^{-8}$ , 数据批次规模 (Batch Size) 设置为 16, 在保证训练稳定性的同时充分利用 GPU 显存资源。学习率调度采用阶梯式衰减策略, 每完成 256 次迭代即按 0.5 倍率进行动态调整。实验硬件平台搭载 NVIDIA GeForce RTX 3090 GPU, 配合 CUDA 11.0 并行计算架构与 cuDNN 8.0.4 加速库, 实现混合精度训练与张量核心运算的深度融合。实验设计采用五重交叉验证方案, 最终性能指标取五次独立实验的算术平均值。

### 3.5 消融实验

为了确定最适合本实验的 MgCSD 模块分组数, 本研究在 Magnetic Tile Dataset 数据集上进行了消融实验。在 U-Net (Baseline) 模型中引入不同分组数的 MsCSD 模块, 分组数的选择参考了多种先进网络, 根据不同的组合得到了 4 种结果如表 1 所示, 表中数值的前半部分代表数据的均值, 后半部分则代表数据与均值之间的差。

表 1 MgCSD 模块不同分组数对模型性能的影响

Table 1 Effect of different number of subgroups of MgCSD module on model performance

Methods	Dice (%)	IoU (%)	Precession (%)	ASSD (pix)	Parameters (M)
Baseline + MgCSD (2)	89.6 ± 0.25	81.7 ± 0.21	86.2 ± 0.05	1.94 ± 0.02	29.5
Baseline + MgCSD (4)	<b>91.7 ± 0.27</b>	<b>85 ± 0.31</b>	<b>88.9 ± 0.54</b>	<b>1.3 ± 0.05</b>	23.7
Baseline + MgCSD (6)	91.3 ± 0.22	84.9 ± 0.36	88.5 ± 0.26	1.45 ± 0.01	<b>23.0</b>
Baseline + MgCSD (8)	90.9 ± 0.36	84.1 ± 0.25	87.9 ± 0.50	1.68 ± 0.02	25.9

在消融实验中, 所有的实验对象都在相同的实验环境以及数据增强下运行。值得注意的是, Dice、IoU、Precision 指标越高表示分割结果越好, 而 ASSD 和 Parameters 指标则与其相反。从表 1 中可以看出, 当分组数为 4 时评价指标 Dice、IoU、Precision 和 ASSD 取得了最好的成绩, 而分组数为 6 时评价指标 Parameters 取得最好的结果, 综合考虑, 选择第 2 组参数, 分组数为 4。

此外, 本文一共提出了两个模块 MgCSD 模块和 GAB 模

块, 为了证明各模块的有效性, 进行了消融实验来比较不同模块的性能。本次实验选取了 U-Net 网络作为 Baseline, 通过在 Baseline 上分别或同时添加 MgCSD 模块和 GAB 模块, 一共获得了 4 种方法, 分别为 Baseline 方法、Baseline + MgCSD 方法、Baseline + GAB 方法、Baseline + MgCSD + GAB 方法, 其中, MgCSD 模块的参数选取上面的消融实验中的第 2 组参数。一共获得了 4 组消融实验用于分割 BUSI 数据集, 其性能评价指标如表 2 所示。消融实验中所有方法的实验环境和数

表 2 不同模块对模型性能的影响

Table 2 Effect of different modules on model performance

Methods	Dice (%)	IoU (%)	Precession (%)	ASSD (pix)	Parameters (M)
Baseline (U-Net)	87.2 ± 0.57	79.1 ± 0.42	83.5 ± 0.52	2.81 ± 0.34	31.0
Baseline + MgCSD	91.7 ± 0.27	85.0 ± 0.31	88.9 ± 0.54	1.3 ± 0.05	<b>23.7</b>
Baseline + GAB	91.8 ± 0.47	86.5 ± 0.20	89.0 ± 0.22	1.25 ± 0.04	24.1
Baseline + MgCSD + GAB (Ours)	<b>93.4 ± 0.03</b>	<b>88.9 ± 0.21</b>	<b>91.0 ± 0.14</b>	<b>1.02 ± 0.01</b>	23.8

据增强方式均相同, 以此来保证实验的公正性。此外, 表中数值的前半部分代表数据的均值, 后半部分则代表数据与均值之间的差距。

由表 2 可以看出, Baseline 的参数量为 31.0M, 与之相比 MGA-Net 降低了 7.2M。通过单独添加 MgCSD 和 GAB 模块, 绝大部分评价指标均值都有所提升, 而同时添加两个模块, 4 个指标均值均取得了最优异的成绩。其中, Dice 均值提升了 6.2%, IoU 均值提升了 9.8%, Precision 均值提升了 7.5%。而 ASSD 均值在 Baseline 的基础上下降了 1.79 个点。此外, 与单

独添加 MgCSD 模块相比, 同时添加两个模块后模型的参数量仅上升了约 0.1M, 通过消融实验可以看出, 相较于 U-Net, MGA-Net 的各性能指标均有所提升。因此, MGA-Net 是一个高效且轻量的网络。

图 4 对消融实验进行了可视化展示。其中, 图 4(a) 表示原图, 图 4(b) 表示原图对应的基准真值, 其他则为对比方法的分割图片。由图 4(d) 可以看出, 对于不同尺度大小的病灶区域, Baseline + MgCSD 方法充分考虑到了相邻特征之间的特征差异较小, 通过通道移位策略和特征融合机制, 使得网络

获取到的特征信息更为丰富全面,提高了网络对不同尺度的特征信息的理解能力,进而提高了分割的准确率.而 Baseline 方法图 4(c) 在定位感兴趣区域上略有不足,在处理小尺度的

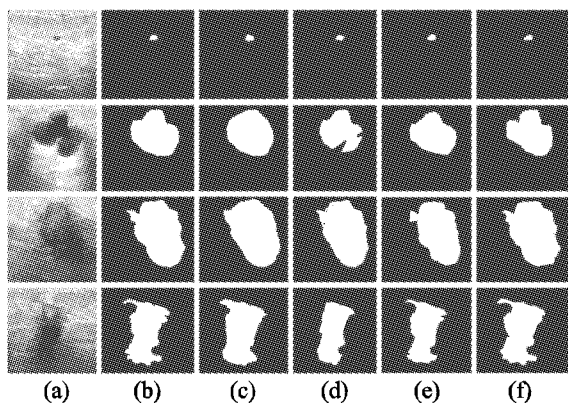


图 4 (a) 原图 (b) Ground Truth (c) U-Net  
(d) Baseline + MgCSD (e) Baseline + GAB (f) Ours

Fig. 4 (a) Original (b) Ground Truth (c) U-Net  
(d) Baseline + MgCSD (e) Baseline + GAB (f) Ours

目标区域时定位与 Ground Truth 有所偏差.对于分割边界复杂的目标区域, U-Net 通过在跳跃连接部分添加 GAB 模块,将边缘、纹理等低级特征丰富的编码器特征图以及高级语义

表 3 不同网络在 BUSI 数据集上的分割性能比较

Table 3 Comparison of segmentation performance of different networks on BUSI dataset

Methods	Dice (%)	IoU (%)	Precession (%)	ASSD (pix)	Parameters (M)
U-Net <sup>[8]</sup>	87.2 ± 0.57	79.1 ± 0.42	83.5 ± 0.52	2.81 ± 0.34	31.0
U-Net + <sup>[9]</sup>	89.4 ± 0.45	82.3 ± 0.51	85.6 ± 0.48	2.15 ± 0.27	32.5
CResU-Net <sup>[19]</sup>	88.7 ± 0.60	80.9 ± 0.55	84.8 ± 0.62	2.40 ± 0.30	33.0
UltraLightUNet <sup>[20]</sup>	90.1 ± 0.50	83.5 ± 0.47	86.2 ± 0.53	1.98 ± 0.25	24.3
TransUNet <sup>[21]</sup>	91.5 ± 0.35	85.2 ± 0.39	87.8 ± 0.41	1.75 ± 0.20	35.5
Ours	<b>93.4 ± 0.03</b>	<b>88.9 ± 0.21</b>	<b>91.0 ± 0.14</b>	<b>1.02 ± 0.01</b>	<b>23.8</b>

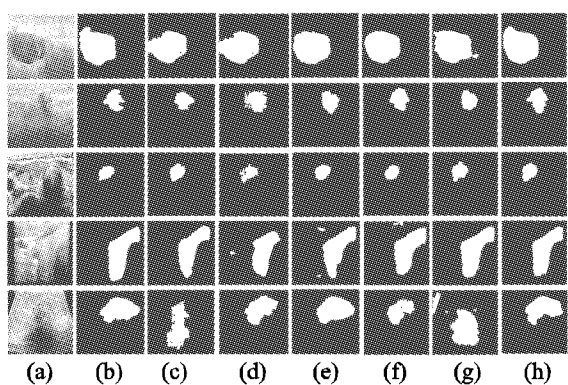


图 5 (a) 原图 (b) Ground Truth (c) U-Net (d) U-Net + +  
(e) CResU-Net (f) UltraLightUNet (g) TransUNet (h) Ours

Fig. 5 (a) Original (b) Ground Truth (c) U-Net (d) U-Net + +  
(e) CResU-Net (f) UltraLightUNet (g) TransUNet (h) Ours

方法的分割图片.其中, U-Net + +, 首先在 U-Net 的基础上引入了密集跳跃连接和嵌套结构,减少了编码器和解码器之间的语义差距,提了网络对局部信息的提取能力,使得网络获取的病灶区域更加集中,如图 5(d); CResU-Net 通过 ResNet 的残差跳跃连接保留低级特征,减少下采样导致的信息丢失,同

信息充足的解码器特征图融合,从而细化边缘信息,使得网络在捕获边缘特征方面更有优势.相反, Baseline 方法在定位边界信息时则不够准确,最终导致分割边界不清晰且将过多的背景区域标记为病变区域.整体看来,图 4(f) Ours 的分割结果更为精确.

### 3.6 对比实验

为了验证 MGA-Net 的性能,本文在 BUSI 数据集上进行了对比实验.实验选取了 5 个优秀的网络作为竞争对手,且其中的许多网络均是为了医学图像分割而设计的,为了保证实验结果的可靠性,所有的竞争对手均在同样的实验环境和数据增强下运行.此外,表中数据的构成均与消融实验中的相同.表 3 定量地列出了 MGA-Net 与这 5 种对比方法在 BUSI 数据集上的性能指标的比较情况.

由表 3 可以看出, MGA-Net 无论是在参数量,还是在其它评价指标均值上均表现最好.其中, Dice 均值为 93.4%, 比 U-Net 高了 6.2%. 综合以上各项指标,本文提出的方法在所有评估指标上都表现出色,不仅在分割准确性和边界精度上有所提升,而且在模型轻量化方面也有显著优势.这表明本文的方法在保持高精度的同时,具有较高的效率和较轻的模型大小,是一种有效的分割方法.

图 5 为在 BUSI 数据集上进行可视化比较其中,图 5(a) 表示原图,图 5(b) 表示原图对应的基准真值,其他则为对比

时引入“Co-Block”增强多尺度特征融合,通过增加上下文感知来增强模型对语义信息的理解能力,获得的分割结果与 Ground Truth 较为接近,如图 5(e) 所示; Ultra-lightUNet 属于轻量化的改进,参数较少但性能有所降低;而 TransUNet 结合了 Transformer 和 U-Net 的优势,它在 U-Net 网络的基础上添加自注意力机制来提高模型对关键信息的辨别能力,使得模型再识别感兴趣区域的边界信息上比 U-Net 更有优势,获得的分割图片边界信息更加清晰,如图 5(g) 所示.与其它网络相比,本文所提出的网络的分割结果则更加细致,如图 5(h) 所示,感兴趣区域的边界与 Ground Truth 更加靠拢.对于目标区域尺度过大或过小的图像如图 5 的第 1 行和第 3 行所示,与 U-Net 相比, MGA-Net 逐步融合了相近两层解码器的输出特征图,并利用空间注意力机制对融合信息的像素权重进行调整,使得网络能够准确捕捉感兴趣区域的所在位置,进而提高了模型的分割精度.

## 4 讨论

针对消融实验以及对比实验的结果展开分析与探讨,图 6(a) 和图 6(b) 对消融实验与对比实验的结果进行了可视化

展示.从消融实验的结果能够看出,MGA-Net 在分割性能以及参数量方面相较于 U-Net 均实现了一定程度的提升.除此之外,通过对比实验可以看出,MGA-Net 具备良好的泛化性,在 BUSI 数据集上取得了优异的成绩,分割结果与 Ground Truth 较为接近.具体而言,与典型的分割网络 U-Net 相比较,MGA-Net 通过多组通道移位下采样模块对全局与局部特征的提取进行了平衡,借助通道移位策略以及特征融合机制,在减少信息损失的同时增强了轮廓特征的表达能力,使得网络

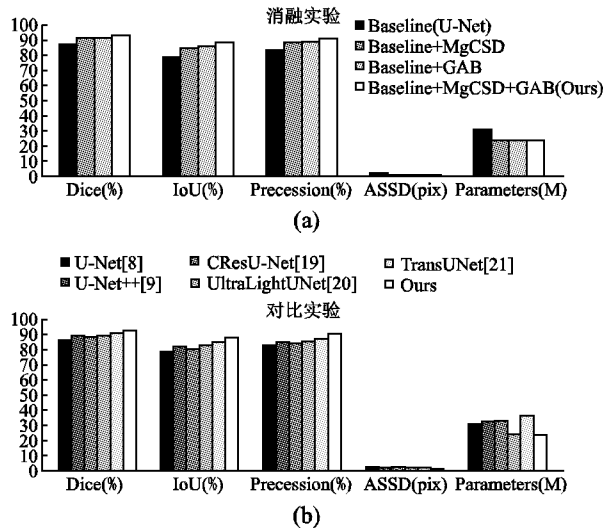


图6 消融实验与对比实验的结果可视化展示  
Fig. 6 Visualization of the results of the ablation experiment and the comparison experiment

区域的边界更为敏感,分割的边界与 Ground Truth 更加贴合,最终在乳腺超声数据集上获得了比 U-Net 更为好的分割结果,并且参数量减少了 23%. 与基于 U-Net 的其他改进网络相比,MGA-Net 的分割结果也更为出色. 相较于轻量级网络 UltraLightUNet 来说,MGA-Net 的参数量更小,然而分割结果却优于 UltraLightUNet. 本文提出的网络 MGA-Net 实现了更为良好的分割性能,并且参数量并未出现明显的增加,这充分地表明,MGA-Net 在分割准确率以及复杂度上均呈现出较好的表现. 因此,MGA-Net 是一个高效且轻量的分割网络. 然而,本文所提出的方法依然存在着一一些问题与不足之处. 首先,通过 MGA-Net 所得到的分割结果仍然存在着一一些过分割与欠分割的区域,对于精度要求极高的医学图像领域而言,这是有所欠缺的,因为哪怕是一丁点的失误都有可能对后续医生对于疾病的诊断和治疗产生影响,所以持续提升模型的分割精度依然是今后研究的主要任务. 其次,本文目前所处理的数据集属于全注释的数据集,而在现实中对数据集进行标注需要耗费大量的时间和精力,因此后续的研究将会从这一方面着手,去应对弱注释或半注释的分割任务. 与此同时,为了对模型的泛化性和鲁棒性进行验证,今后需要在其他类型的医学图像数据集上对该网络的性能进行验证,例如 CT 图像、MR 图像等. 最后,MGA-Net 仅仅能够将病灶区域分割出来,尚无法识别其病变类型,所以,为了更加便捷地辅助医生对疾病进行诊断,后续需要对分割结果进行分类,并添加相应的文字描述.

## 5 总结

本文针对超声图像分割中存在的病灶边界模糊、模型复杂度、小目标易漏分割等问题,提出了一种融合多尺度特征聚合和多组通道移位策略的轻量化网络 MGA-Net. 通过引入多组通道移位下采样模块 (MgCSD) 与组聚合桥接模块 (GAB),在 U-Net 架构基础上实现特征提取与跨层融合的双重创新. MgCSD 模块利用通道空间移位策略和尺度不变卷积,通过分组循环移位操作建立中心区域细节与边缘轮廓的关联性,结合全局加权增强机制保留多尺度纹理特征,与 U-Net 相比卷积下采样减少了 30% ~ 40% 参数量;GAB 模块通过将编码器高级语义特征、解码器低级细节特征及中间掩膜预测进行通道对齐,采用分组多尺度空洞卷积实现跨层级特征融合,有效缓解了深层次特征语义不匹配问题. 实验基于 BUSI 数据集验证,MGA-Net 达到 93.4% 的 Dice 系数,参数量仅 23.8M,且在病灶边缘分割任务中敏感度提升 7.5%.

## References:

- [1] CHI M X, AN H, JIN X, et al. N-shaped neural network for brain tumor segmentation [J]. *Journal of Chinese Computer Systems*, 2025, 46(2): 365-372.
- [2] Yadav K V, Singhai J. Segmentation of lungs from chest X-ray images based on deep atrous attention UNet (DAA-UNet) model[J]. *Medical & Biological Engineering & Computing*, 2025, (prepublish): 1-13, doi:10.1007/s11517-025-03344-8.
- [3] WANG H P, DING W P, HUANG J S, et al. FTransCNN: fusing transformer and CNN based on fuzzy fusion for uncertain medical image segmentation [J]. *Journal of Chinese Computer Systems*, 2024, 45(6): 1426-1435.
- [4] Gongping C, Lei L, Jianxun Z, et al. Rethinking the unpretentious U-net for medical ultrasound image segmentation[J]. *Pattern Recognition*, 2023, 142: 109728, doi:10.1016/j.patcog.2023.109728.
- [5] LI G P, ZHAI Y J, ZHANG X L, et al. Application of deep learning in breast tumor pathology image analysis [J]. *Beijing Biomedical Engineering*, 2025, 44(1): 81-89.
- [6] ZHU Y T, TIAN F, DONG B L. Ultrasound image segmentation method for thyroid nodules based on HA-UNet++ [J]. *Computer and Modernization*, 2025, (3): 93-98 + 105. doi:10.3969/j.issn.1006-2475.2025.03.014.
- [7] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [8] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [J]. *CoRR*, 2015, abs/1505.04597.
- [9] Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. UNet++: redesigning skip connections to exploit multiscale features in image segmentation [J]. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 1856-1867.
- [10] Huang L, Lin Y, Cao P, et al. Automated detection and segmentation of pleural effusion on ultrasound images using an attention U-net [J]. *Journal of Applied Clinical Medical Physics*, 2024, 25(1), doi:10.1002/acm2.14231.

- [11] Chen G, Li L, Dai Y, et al. AAU-net: an adaptive attention U-net for breast lesions segmentation in ultrasound images [J]. IEEE Transactions on Medical Imaging, 2023, 42(5): 1289-1300.
- [12] Zhao J Q, Zhang M, Zhou Z H, et al. CFFormer: a cross-fusion transformer framework for the semantic segmentation of multi-source remote sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 1-17, doi: 10. 1109/TGRS. 2024. 3507274.
- [13] Cao H, Wang Y, Chen J, et al. Swin-Unet: unet-like pure transformer for medical image segmentation [C]//European Conference on Computer Vision, Cham: Springer Nature Switzerland, 2022: 205-218.
- [14] Wang X, An H, Zhang J, et al. DPSMUNet: a new network based on a dual-pooling self-attention module for carotid artery plaque segmentation in ultrasound images [J]. Supercomput, 2025, 81, 267, doi: 10. 1007/s11227-024-06770-x.
- [15] Zifan A, Zhao K, Lee M, et al. Adaptive evolutionary optimization of deep learning architectures for focused liver ultrasound image segmentation [J]. Diagnostics, 2025, 15(2): 117, doi: 10. 3390/diagnostics15020117.
- [16] Kareem D N A, Fiaz M, Novershtern N, et al. Medical image segmentation using directional window attention [C]//IEEE International Symposium on Biomedical Imaging (ISBI), 2024: 1-5.
- [17] Khan R F, Lee B D, Lee M S. Transformers in medical image segmentation: a narrative review [J]. Quantitative Imaging in Medicine and Surgery, 2023, 13(12): 8747-8767.
- [18] Che Z, Zhang Z, Wu Y, et al. Disentangle and then fuse: a cross-modal network for synthesizing gadolinium-enhanced brain MR images [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2025, doi: 10. 1109/TCSVT. 2025. 3528981.
- [19] Pham T M, Do N, Pham H T T, et al. CResU-net: a method for landslide mapping using deep learning [J]. Machine Learning-Science and Technology, 2024, 5(3): 035008, doi: 10. 1088/2632-2153/ad5f17.
- [20] Isensee F, Rokuss M, Krämer L, et al. NnInteractive: redefining 3D promptable segmentation [J]. arXiv, 2025, doi: 10. 48550/arXiv. 2503. 08373.
- [21] Chen J N, Mei J R, Li X H, et al. TransUNet: rethinking the U-Net architecture design for medical image segmentation through the lens of transformers [J]. Medical Image Analysis, 2024, 103280, doi: 10. 1016/j. media. 2024. 103280.

#### 附中文参考文献:

- [1] 迟孟贤, 安虹, 金旭, 等. 用于脑肿瘤分割的N形神经网络 [J]. 小型微型计算机系统, 2025, 46(2): 365-372.
- [3] 王海鹏, 丁卫平, 黄嘉爽, 等. FTransCNN: 基于模糊融合的 Transformer-CNN 不确定性医学图像分割模型 [J]. 小型微型计算机系统, 2024, 45(6): 1426-1435.
- [5] 李冠鹏, 翟羽佳, 张晓丽, 等. 深度学习在乳腺肿瘤病理图像分析中的应用 [J]. 北京生物医学工程, 2025, 44(1): 81-89.
- [6] 朱永田, 田飞, 董宝良. 基于 HA-UNet++ 的甲状腺结节超声图像分割方法 [J]. 计算机与现代化, 2025, (3): 93-98 + 105. doi: 10. 3969/j. issn. 1006-2475. 2025. 03. 014.