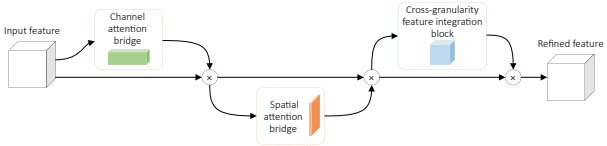
OEE | 光电工程 Opto-Electronic Engineering

Article

2026 年 1 月, 第 53 卷, 第 1 期

膨胀门控与自适应桥接的轻量化皮肤病变分割网络



杨本臣，苏焕哲*，金海波，曲海成

辽宁工程技术大学软件学院，辽宁 葫芦岛 125105

摘要：针对皮肤病变分割任务中因边界模糊、形态多变、毛发遮挡等复杂特征干扰导致病灶区域定位困难，以及现有方法参数量大、计算复杂度高的问题，本文提出一种膨胀门控与自适应桥接的轻量化皮肤病变分割网络。首先，设计膨胀门控感知增强模块，结合差异化膨胀策略、门控注意机制与渐进感知结构，强化病变区域的边界表征与上下文感知。其次，构建自适应特征桥接模块，利用通道-空间双注意机制与跨粒度协同模块实现跨层级通道校准，提升对病灶尺度变化与毛发干扰的鲁棒性。然后，设计轻量初始结构引导模块，融合标准卷积与大核深度可分卷积构建长程依赖，提供高质量特征基元。最后，引入高效通道注意力，利用自适应卷积核缓解上采样过程中的信息损失。在 ISIC2016、ISIC2017 与 ISIC2018 数据集上进行验证，平均交并比 (mean intersection over union, mIoU) 分别达到 85.78%、79.81%、80.65%，Dice 相似系数 (dice similarly coefficient, DSC) 分别达到 91.90%、88.73%、89.37%，且参数量仅为 0.30 M，计算复杂度为 0.22 GFLOPs。同当前主流方法相比，本文方法在具备更优分割精度的同时实现了模型的轻量化设计，具有良好的实际部署价值与临床应用前景。

关键词：图像处理；皮肤病变分割；U-Net；轻量化；注意力机制

The English title and abstract appear at the end of the paper.

DOI: 10.12086/oee.2026.250121 | CSTR: 32245.14.oee.2026.250121 | 中图分类号：TP391 | 文献标识码：A

杨本臣，苏焕哲，金海波，等. 膨胀门控与自适应桥接的轻量化皮肤病变分割网络 [J]. 光电工程，2026，**53**(1): 250121

Yang B C, Su H Z, Jin H B et al. A lightweight skin lesion segmentation network based on dilated gating and adaptive bridging[J]. *Opto-Electron Eng*, 2026, **53**(1): 250121

1 引　言

皮肤癌作为全球范围内高发病率的恶性肿瘤类型，每年累计数百万临床病例[1]。其中，黑色素瘤作为最具侵袭性的皮肤源性癌变，黑色素瘤的攻克已成为全球公共卫生领域的重大挑战[2]。其发病率持续上升，若未实现早期准确诊断，患者的 5 年生存率将大幅下降至 15% 以下[3]，早期发现和干预对于提高患者生存率至关重要。然而，黑色素瘤病灶因常呈现边界模糊、形状不规则、易受毛发遮挡、对光照变化敏感、以及与周围皮肤对比度低等高度异质性的视觉特征，使其与正常组织的区分极具挑战性[4]。相比于结构清晰、形态规则的 CT 或 MRI 影像，皮肤病变图像更具非结构性和局部扰动性，对模型的特征提取与边界刻画能力提出了更高要求，成为皮肤病变分割任务中的关键科学难题。

在早期诊断中，皮肤科医生依靠视觉检查并识别皮肤病变及潜在的癌症迹象[5]。然而，皮肤镜图像的判读过程不仅耗时、烦琐，而且诊断结果易受到医生主观经

收稿日期：2025-04-24；修回日期：2025-07-31；录用日期：2025-07-31；网络出版日期：2026-01-25

基金项目：国家自然科学基金资助项目 (62173171)；辽宁省教育厅基本科研项目 (JYTMS20230804)；辽宁省高等学校基本科研项目 (LJKMZ20220699)

*通信作者：苏焕哲，2865775393@qq.com。

版权所有©2026 中国科学院光电技术研究所

验的影响[6]。为提升病灶精准识别与分析效能，计算机辅助诊断 (computer aided diagnosis, CAD) 技术[7-8] 应运而生。早期方法多采用基于灰度阈值的分割算法、边缘梯度检测及传统监督学习方法[9]，其依赖人工设计特征难以有效提取深层病理特征，且抗干扰能力较弱，致使分割精度受限。

随着深度学习在医学图像分割中的迅速发展，卷积神经网络 (convolutional neural networks, CNN)[10] 在皮肤病变分割任务中展现出卓越的性能。Ronneberger 等[11] 提出的 U-Net 结构，已成为皮肤病变分割领域的奠基性工作。为增强模型对目标区域的感知能力，Oktay 等[12] 提出的 Attention U-Net，整合了注意力门控机制，提高了模型特征选择的准确性。但单一的注意力机制存在感知维度有限，难以捕捉多层次病灶信息的问题。因此，Sun 等[13] 提出的 MSCA-Net，通过结合四种注意力机制，捕获皮肤病变图像中多尺度上下文信息，提升了模型对皮肤病变区域的定位能力。尽管上述基于 CNN 的皮肤病变分割方法在捕捉细粒度特征方面表现优越，然而受限于卷积操作固有的局部感受野，其在应对皮肤病图像中复杂病灶区域的结构表征与边缘刻画时仍存在局部依赖性强，全局感知能力不足等挑战。

为增强对皮肤病变区域的全局建模能力，研究者将 ViT[14] 的自注意力机制引入 U 型网络，构建融合局部感知与长程建模优势的混合皮肤病变分割网络。Zhang 等[15] 提出的 TransFuse，将 Transformer 和 CNN 以并行方式结合，能够高效地捕捉全局依赖关系和低级空间细节，提高了皮肤病变分割的准确性。赵宏和王枭[16] 提出了一种基于图像降噪和 Transformer 的双分支混合模型，利用双分支特征提取模块，有效提升了病灶边界的分割效果。此外，梁礼明等[17] 构建了一种自适应特征融合级联 Transformer 算法，有效剔除病变图像背景噪声，提高了模型分割性能。然而，上述方法普遍存在参数规模庞大、计算开销较高的问题，在实际部署中存在一定局限性。因此，众多轻量化皮肤病变分割网络[18-19] 涌现出来。Yin 等[20] 设计的 AMSUnet，重构了编码器结构，引入空洞卷积与多尺度卷积操作，提升了模型对皮肤病变区域的特征提取能力与计算效率。李大湘等[21] 提出的 UH-Net 通过构造多头门控位置交叉自注意力编码器，在提高网络捕获皮肤病变区域的形状和边界信息能力的同时，有效压缩了参数规模。尽管上述方法在参数压缩和计算效率上取得进展，但多数轻量化模型因结构设计相对简约、建模能力有限，难以应对病灶模糊、形态不规则及图像对比度低等复杂皮肤病变情况，模型不能精确捕捉边缘细节特征，导致分割精度下降，进而引发误判和漏检风险。

为解决以上问题，本文设计一种膨胀门控与自适应桥接的轻量化皮肤病变分割网络，以实现高精度、低计算开销的黑色素瘤分割任务。

2 膨胀门控与自适应桥接的轻量化皮肤病变分割网络

2.1 网络总体结构

本文网络整体采用 6 阶段 U 型网络架构，以一张尺寸为 256×256 的皮肤病变图像作为输入，通过各阶段共同作用，最终生成一张大小为 256×256 的分割预测图。网络由四个核心模块组成，利用轻量初始结构引导模块 (lightweight initial structure-guided module, LISG) 增强浅层特征的结构性与判别性表达，为深层特征提取提供更稳定且具指导性的基础支持；使用膨胀门控感知增强模块 (dilated gated perceptual enhancement module, DGPE) 实现全局信息和局部信息动态耦合优化，提升感知范围与表征能力；通过高效通道注意力 (efficient channel attention, ECA) 进行跨通道多粒度语义信息的动态选择与校准，强化特征通道间的有效交互；使用自适应特征桥接模块 (adaptive feature bridge module, AFB) 引入多维注意机制与多尺度建模策略，自适应调节跨层信息融合过程中的关键语义流，抑制冗余干扰。整体网络结构如图 1 所示。

2.2 膨胀门控感知增强模块

皮肤病变图像分割属于密集预测任务，全局与局部信息的协同捕获对提升模型性能具有决定性作用。因此，本文提出 DGPE，首先通过分裂膨胀单元 (split dilated unit, SD)，利用差异化膨胀率的深度可分离卷积并行提取局部与全局特征。再利用门控注意单元 (gated attention unit, GA) 抑制冗余背景信息。最后经过渐进感知增强单元 (progressive perceptual enhancement unit, PPE) 增强多尺度上下文感知能力，从而有效提升边缘建模与形态结构保真度。DGPE 结构如图 2 所示。

1) 分裂膨胀单元。首先采用通道维度特征解耦策略，通过四等分操作 ($Channel_4$) 将输入特征图分解为四组子特征 (x_1, x_2, x_3, x_4)，并对四组子特征采用差异化膨胀率

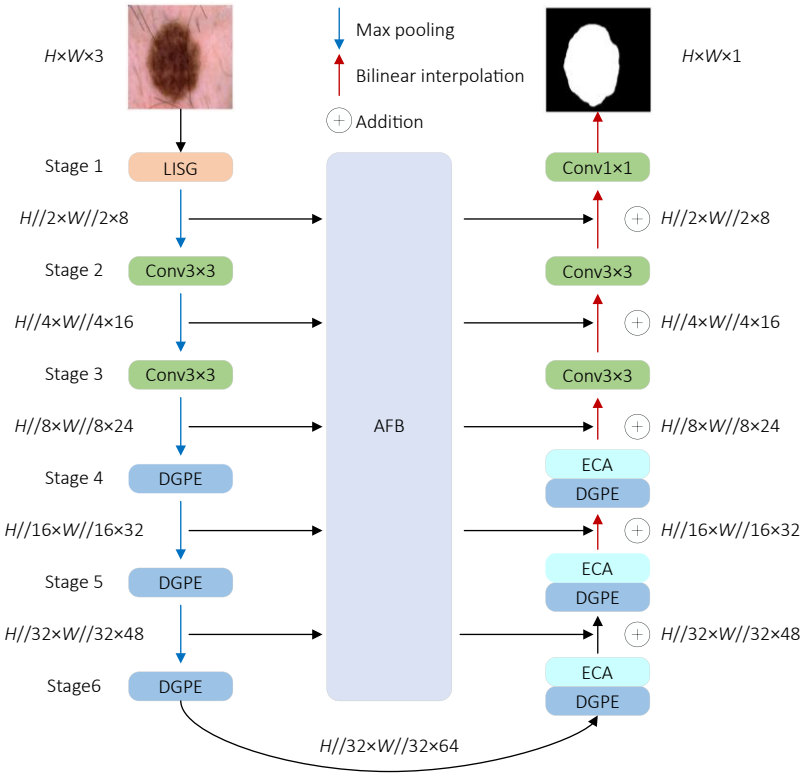


图 1 膨胀门控与自适应桥接的轻量化皮肤病变分割网络

Fig. 1 Lightweight skin lesion segmentation network based on dilated gating and adaptive bridging

($d=1,2,5,7$) 的深度可分离卷积W_1、W_2、W_5和W_7，并行提取到局部细粒度特征 (低膨胀率：$d=1,2$) 和全局上下文信息 (高膨胀率：$d=5,7$)。最后，经通道维度重组操作 ($Concat$)，使特征图恢复到原始大小，再通过标准卷积操作 (W)，实现全局和局部特征之间的协同交互，得到分解重组后的特征X'。具体计算过程如下：

$$x_1, x_2, x_3, x_4 = Channel_4(X)\,, \tag{1}$$

$$x_1', x_2', x_3', x_4' = W_1(x_1), W_2(x_2), W_5(x_3), W_7(x_4)\,, \tag{2}$$

$$X' = W(Concat(x_1', x_2', x_3', x_4'))\,. \tag{3}$$

2) 门控注意单元。将输入特征X'分别输入到三个并行的处理分支，主路径通过标准卷积操作 (W) 保持特征完整性，在辅助路径中，两个深度可分离卷积支路分别执行空间注意力建模与特征变换。

其中注意力建模支路通过 Sigmoid 激活函数生成空间注意力掩膜Att，该掩膜通过动态权重分配机制有效抑制非关键区域对全局上下文的干扰。此过程采用空间注意力机制原理，计算如式 (4) 所示：

$$Att = S(DW(X'))\,, \tag{4}$$

式中：DW代表深度可分离卷积操作；S是用于激活的 Sigmoid 函数。

同时，另一支路对原始特征X'进行深度可分离卷积编码，将其与注意力掩膜Att进行逐元素相乘，以实现特征重校准，再经深度可分离卷积进行二次特征融合，形成空间注意力强化的中间特征。最终通过残差连接将中间特征与主路径的标准卷积特征进行自适应融合，在保留原始特征流形结构的同时，显著提升了关键区域的信息传递效率，输出优化后的特征表示Out_1。该过程用公式表示如下：

$$Out_1 = DW(DW(X') \otimes Att) + W(X')\,, \tag{5}$$

式中：⊗表示逐元素相乘。

3) 渐进感知增强单元。为了扩大感知场和更好地聚集局部信息，本文使用了一个核大小为 3×3 的深度可分离卷积，并在其中引入一个空洞卷积[22]，通过三阶段渐进式处理，增强模型多尺度特征提取的能力。

考虑到空洞卷积会丢失一些特征信息，在使用空洞卷积之前，首先将输入Out_1通过 3×3 深度可分离卷积处理，以减少特征信息的损失，再经过具有 3×3 卷积核的空洞深度可分离卷积聚集局部信息，生成第一阶段输出Out_2。此过程表示如下：

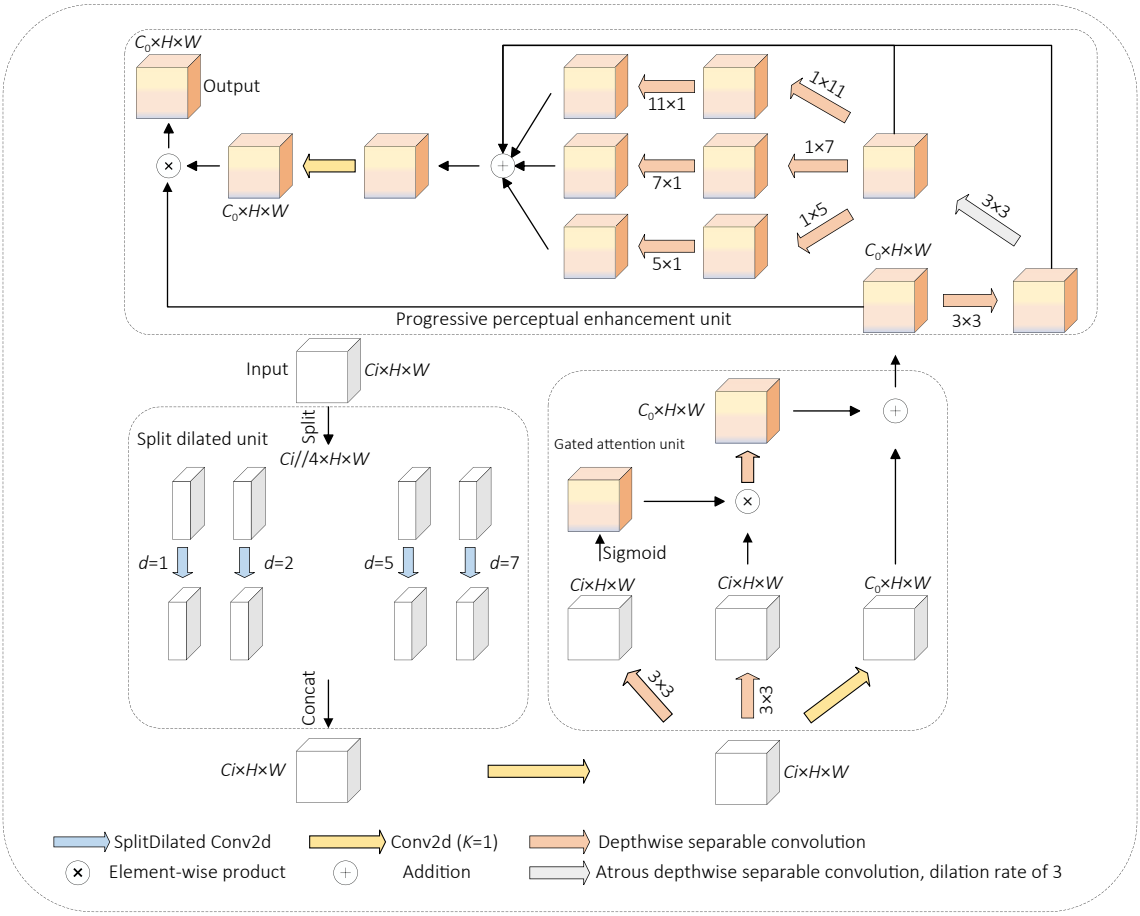


图 2 膨胀门控感知增强模块

Fig. 2 Dilated gated perceptual enhancement module

$$Out_2 = ADW\text{-}Conv_{3\times3}(DW\text{-}Conv_{3\times3}(Out_1)), \quad (6)$$

式中：$ADW\text{-}Conv_{3\times3}$表示核大小为3×3的空洞深度可分离卷积。

对于第二阶段，使用并行多分支深度可分离卷积捕获多尺度上下文信息。最后，在第三阶段中使用 1×1 卷积用来实现跨通道信息聚合。将渐进感知增强单元的最初输入Out_1与其逐元素相乘，得到膨胀门控感知增强模块的最终输出。上述过程可表示为

$$Out = Out_1 \otimes Conv_{1\times1}\left(\sum_{i=0}^{3} Scale_i(Out_2)\right), \quad (7)$$

式中：$Scale_i$表示第 i 个分支，其中$i \in \{0,1,2,3\}$。

2.3 自适应特征桥接模块

当前皮肤病变分割研究主要集中于编码器下采样模块的改进，而对特征融合机制的优化尚未充分解决。其次，各类 U 型模型通常采用简单的线性相加操作，难以实现跨层次特征的有效融合。

为解决以上问题，本文构建自适应特征桥接模块 (AFB)，利用通道注意力桥块 (channel attention bridge block, CAB) 融合多层级特征并引入动态通道权重分配机制，实现跨尺度语义信息的有效整合与细粒度特征增强。再使用空间注意力桥块 (spatial attention bridge block, SAB) 通过双模态空间特征交互与大感受野卷积以增强空间上下文建模能力，提升模型对关键区域的定位与响应精度。最后通过跨粒度特征协同模块 (cross-granularity feature integration block, CGFI) 利用多尺度深度可分离卷积与通道混洗机制实现跨粒度特征融合，强化语义协同与多尺度形态感知能力，有效提升模型对病灶区域的解析与表征精度。AFB 结构如图 3 所示。

1) 通道注意力桥块。如图 4 所示，通道注意力桥块通过整合编码器不同层级的特征输入，构建跨尺度的特征融合机制。首先，对于具有异构空间尺寸的多层级特征，先进行全局平均池化，将其压缩为统一的 1×1 细粒

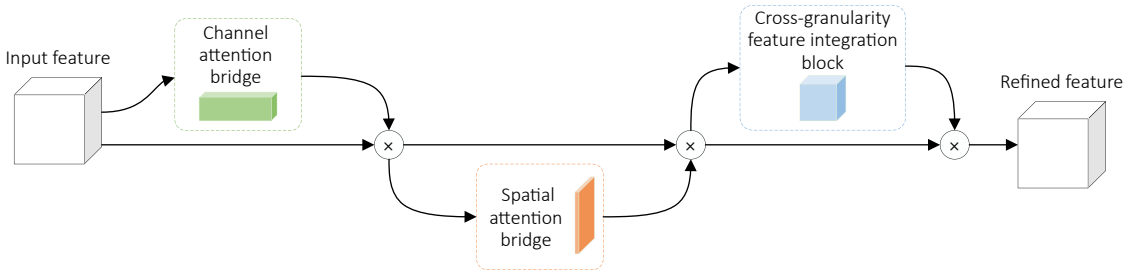


图 3 自适应特征桥接模块

Fig. 3 Adaptive feature bridge module

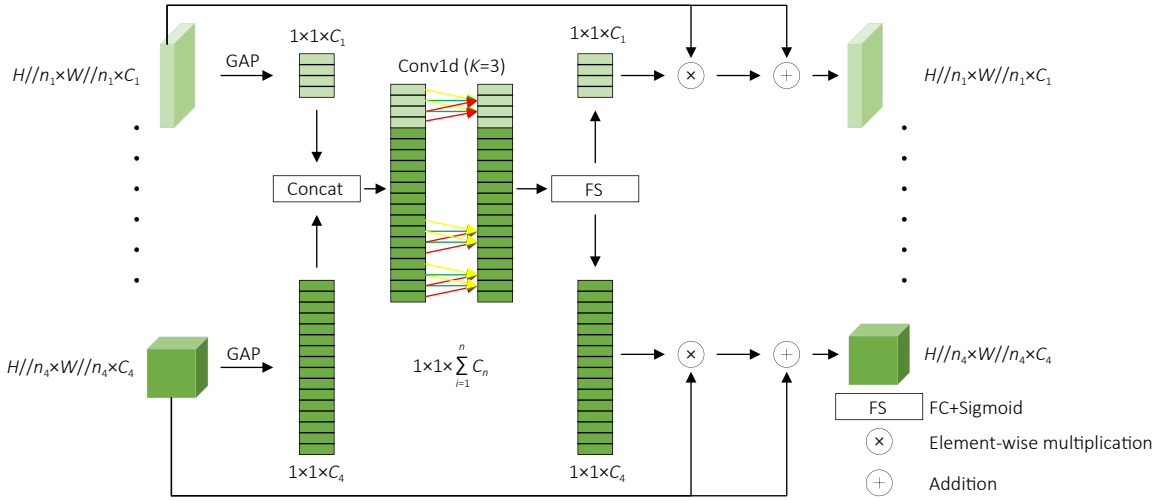


图 4 通道注意力桥块

Fig. 4 Channel attention bridge block

度特征表示。其次，将细粒度特征沿通道维度进行拼接形成多尺度特征描述符。此过程可表示为

$$x'_n = GAP(x_n)\,, \tag{8}$$

$$X = Concat(x'_1, x'_2, ..., x'_n)\,, \tag{9}$$

式中：x'_n表示从编码器的不同级别获得的特征映射；GAP代表全局平均池化；$Concat$表示将$x'_1, x'_2, \cdots, x'_n$连接在一起的操作。

$$Att_n = S(FC_n(Conv_{1D}(X)))\,, \tag{10}$$

$$Out_n = x_n + x_n \otimes Att_n\,, \tag{11}$$

式中：$Conv_{1D}$表示一维卷积；FC_n表示第 n 层的全连接层。

2) 空间注意力桥块。如图 5 所示，空间注意力桥块通过构建双模态空间特征交互机制实现全局上下文建模。将通道注意桥块输出的特征Out_n作为输入，采用双通道池化策略，即分别执行最大值池化与均值池化操作，沿通道维度提取异构特征统计量，生成具有空间位置敏感性的特征描述符。再使用一个扩张率为 3，核大小为 7×7 的共享卷积建立大感受野的局部上下文关联，并结合 Sigmoid 激活函数，用来增强特征之间的局部上下文关系和计算注意力权重。最终通过哈达玛乘积实现特征图的自适应校准，并引入残差连接保留原始特征分布，得到输出特征 Y。这一过程增强了模型识别和响应相关空间特征的能力，能够更好地整合各阶段的特征。用公式表示如下：

$$Y = S(LC([Ch_{\max}(Out_n), Ch_{\text{avg}}(Out_n)])) \otimes Out_n\,, \tag{12}$$

式中：$Ch_{\max}$表示最大池化操作；Ch_{avg}表示平均池化操作；LC表示一个扩张率为 3，核大小为7×7的大核卷积。

3) 跨粒度特征协同模块。如图 6 所示，该模块通过异构卷积核并行计算机制实现多粒度特征表征。首先使用逐点卷积PWC_1扩展通道数量，再通过批归一化层和 ReLU6[23] 激活层进行非线性变换。其次，使用多尺度深度可分卷积 (MSDC)，如图 7 所示，并行提取多尺度空间特征，有效捕获跨分辨率的上下文信息。针对深度卷

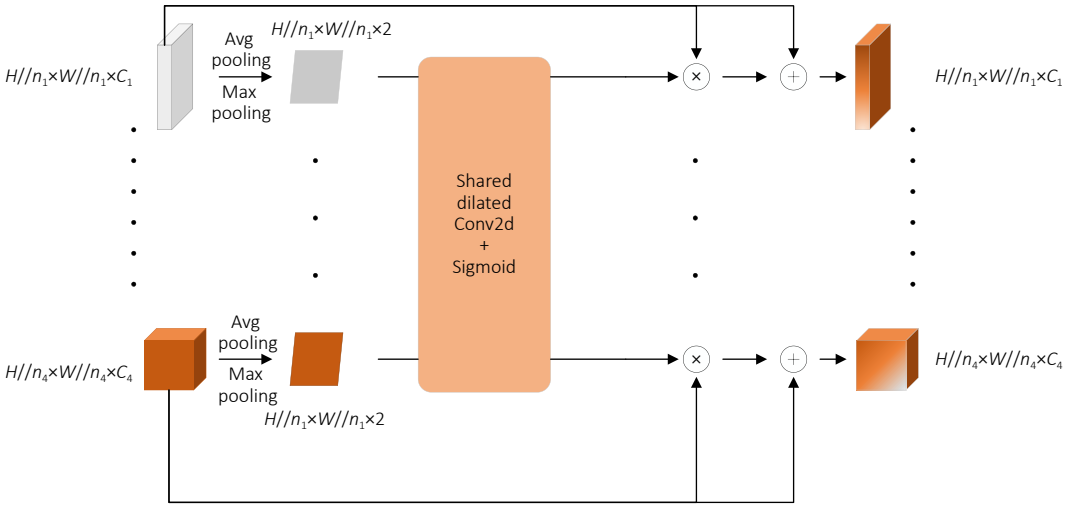


图 5 空间注意力桥块

Fig. 5 Spatial attention bridge block

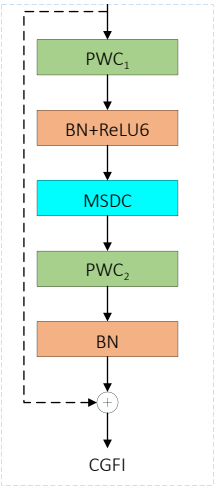


图 6 跨粒度特征协同模块

Fig. 6 Cross-granularity feature integration block

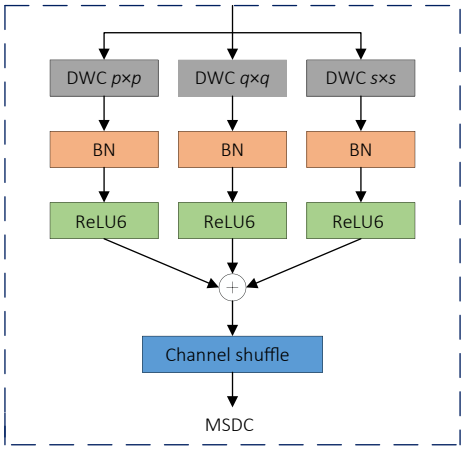


图 7 多尺度深度可分卷积

Fig. 7 Multi-scale depth-wise convolution

积固有的通道间交互缺失问题，引入通道混洗机制促进跨通道信息流动，增强特征映射的语义协同表达能力。最后通过另一个逐点卷积PWC_2和批归一化操作将特征维度恢复至初始尺度，生成自适应特征桥接模块的最终输出。这种分层特征处理策略在抑制冗余背景干扰的同时，显著提升了模型对病灶区域多尺度形态特征的解析能力。此过程用公式表示为

$$Out = BN(PWC_2(CR(MSDC(R_6(BN(PWC_1(Y))))))) . \quad (13)$$

不同内核大小 (K) 的并行 MSDC 表示为

$$D_k(x) = R_6(BN(DWC_k(x))) , \quad (14)$$

$$MSDC(x) = \sum\nolimits_{k\in K} D_k(x) , \quad (15)$$

式中：CR 表示通道混洗操作；DWC_k是内核大小为k的深度卷积；x 为输入特征；BN和R_6分别为批归一化和 ReLU6 激活函数。

2.4 高效通道注意力

针对单通道注意力机制在全局上下文建模中面临的特征交互不足与参数冗余双重挑战。本文通过使用高效通道注意力 (efficient channel attention, ECA)，在特征维度保留约束下，通过解耦局部特征强化与全局依赖建模过程，构建参数效率优化框架，在抑制冗余参数增长的同时，实现通道间多粒度语义信息的自适应校准，ECA 结构如图 8 所示。

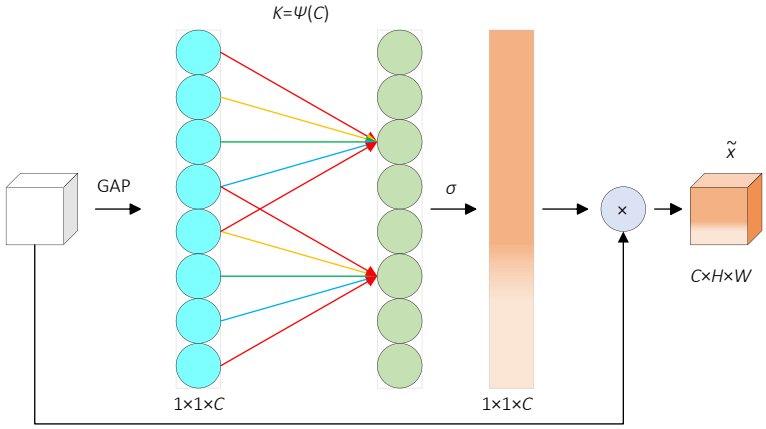


图 8 高效通道注意力

Fig. 8 Efficient channel attention

ECA 首先对输入特征图 X 进行全局平均池化，构建特征全局统计描述符。然后采用基于自适应卷积核策略的核大小为 K 的快速一维卷积操作生成通道权重系数，其中 K 是根据通道维度 C 映射自适应确定的。最后，该权重经 Sigmoid 激活后与原始特征进行逐通道调制，生成最终输出 $\tilde{X}$。该机制通过动态跨通道交互建模，在解码阶段有效维持特征映射的维度完整性，补偿上采样过程中的空间信息耗散，同时采用参数高效的计算拓扑结构，显著降低特征校准过程中的计算复杂度，实现模型精度与效率的协同优化。

此过程可用以下公式表示为

$$X' = GAP(X)\,, \tag{16}$$

$$\tilde{X} = X \otimes S(G_{K=\Psi(C)}(X'))\,, \tag{17}$$

式中：G 表示规则卷积操作，$G_{K=\Psi(C)}$ 表示由通道维度 C 与 K 之间的映射函数 $\Psi(C)$ 决定核大小为 K 的卷积，表达式为

$$C = \varphi(K)\,. \tag{18}$$

它们之间最简单的映射关系是

$$\varphi(K) = 2K - 1\,. \tag{19}$$

然而，这种简单的线性关系有限。具体来说，通道大小通常被设置为 2 的幂次，因此将线性函数扩展为非线性函数，即

$$C = \varphi(K) = 2^{2K-1}\,. \tag{20}$$

给定通道维数 C，内核大小 K 可自适应确定，表达式为

$$K = \Psi(C) = \left|\frac{\log_2 C}{2} + \frac{1}{2}\right|(odd)\,, \tag{21}$$

式中：$|*|(odd)$ 为 $*$ 最接近的奇数。

2.5 轻量初始结构引导模块

在目前的皮肤病图像分割网络中，浅层特征提取模块存在的感受野受限与计算冗余问题。本文对第一个下采样阶段进行结构重构，提出轻量初始结构引导模块 (lightweight initial structure-guided block, LISG)，如图 9 所示。该模块在不增加模型参数量的同时增强浅层特征的判别性表达，为后续深层网络提供鲁棒的特征基础。

轻量初始结构引导模块采用多粒度特征引导策略，

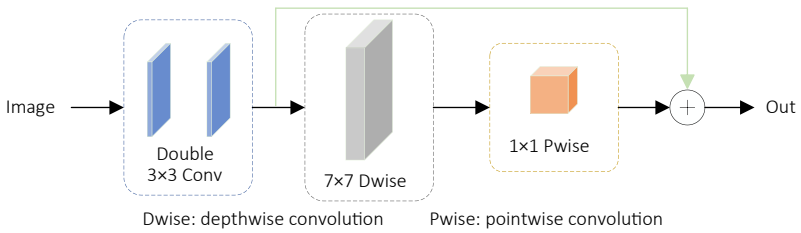


图 9 轻量初始结构引导模块

Fig. 9 Lightweight initial structure-guided module

由两层级联的标准 3×3 卷积单元与深度可分离卷积单元构成复合计算拓扑。其中，标准卷积单元通过局部特征强化机制提取细粒度纹理信息，而深度可分离卷积单元将传统卷积解耦为空间域与通道域独立计算。7×7 大核深度卷积以极低的参数代价构建跨像素长程依赖关系，有效捕获病灶区域的空间分布模式。1×1 逐点卷积则通过通道维度线性变换实现跨通道语义信息融合，形成多尺度特征协同表达。这种解耦设计不仅显著降低计算复杂度，更通过大感受野特征提取有效抑制浅层语义信息在逐层传递过程中的特征耗散现象。

针对深度可分离卷积在低维特征空间易引发的表征能力退化问题，该模块在下采样关键路径引入经典 U-Net 双 3×3 卷积层级作为特征增强单元。该结构通过堆叠小尺寸卷积核渐进式扩展感受野，结合残差连接构建梯度稳定通路，形成特征提取与梯度传播的双向优化机制。此外，模块内部采用动态特征校准机制，通过 ReLU 非线性激活与批归一化操作的协同作用，实现特征分布的自适应规整，进一步提升模型对皮肤病变图像复杂噪声环境的鲁棒性。轻量初始结构引导模块的设计在保证计算效率的同时，为后续深层网络提供富含多尺度语义信息的特征基元，奠定精准分割的底层基础。

3 实验结果与分析

3.1 数据集

本文所使用的数据集均为皮肤镜图像领域广泛认可的公开基准数据集，包括 ISIC2016[24]、ISIC2017[25]、ISIC2018[26] 和 PH^2[27] 数据集。这四个数据集分别由国际皮肤成像协作组织 (international skin imaging collaboration, ISIC) 和葡萄牙波尔图大学医院皮肤科 (dermatology service of hospital Pedro Hispano, Matosinhos, Portugal) 收集并公开。在实验中，ISIC2016、ISIC2017 和 ISIC2018 数据集均按照 7:3 的比例随机划分为训练集和测试集。PH^2 数据集的 200 张皮肤病变图像都用来做泛化性实验，以验证模型的泛化能力。各数据集的详细情况如下：

1) ISIC2016 数据集

其包含 1279 张带有分割掩码标签的皮肤镜图像，训练集包含 900 张皮肤镜图像，测试集包含 379 张皮肤镜图像。分为良性 (非黑色素瘤) 病变和恶性黑色素瘤两种类别。

2) ISIC2017 数据集

其包含 2150 张皮肤镜图像，训练集 1500 张，测试集 650 张。分为黑色素瘤、痣和脂溢性角化症三种类别，其中黑色素瘤代表恶性病变，痣和脂溢性角化病均属于良性病变。

3) ISIC2018 数据集

其包含 2694 张皮肤镜图像，训练集 1886 张，测试集 808 张。分为黑色素瘤、痣、基底细胞癌、光化性角化病、良性角化病、皮肤纤维瘤和血管病变，共 7 种类别。明确的恶性类别是黑色素瘤和基底细胞癌，而光化性角化病是重要的癌前病变，有发展为鳞状细胞癌的风险。其余类别均为良性。

4) PH^2 数据集

其包含 200 张皮肤病变图像，分为黑色素瘤和痣两种类别。在实验中，将 PH^2 数据集的所有图像都用来做交叉验证实验，以验证模型的泛化能力。

3.2 图像预处理

皮肤病变图像在采集过程中存在受毛发遮挡、光照不均以及对比度低等影响，导致病变区域模糊。为提升皮肤病变区域的分割精度，本文对图像进行预处理操作，步骤如下：

1) 利用形态学操作对毛发遮挡的皮肤病变图像进行有效处理。采用形态学腐蚀操作收缩目标边界，缩小病变区域以消除毛发遮挡。利用膨胀操作可以增大病变区域，填补空洞。针对毛发密集的图像，使用以上两个形态学操作，可缓解毛发对分割精度的影响。

2) 利用一系列数据扩充操作，增加样本数量。由于公开数据集中样本数量有限，为避免模型训练时过拟合，采用多种数据增强方法，包括垂直和水平翻转、灰度变换、弹性变换、随机通道扰动、高斯模糊等方法以增加训练样本数量。此外，将所有图像统一调整为 256×256 分辨率。图像预处理操作可视化如图 10 所示。

3.3 实验环境及评价指标

本实验在 Ubuntu 18.04 操作系统的计算平台上进行，硬件配置包括 NVIDIA GeForce RTX 3090 GPU (24 GB 显存) 和 80 GB 内存。实验基于 Python 3.8，采用 PyTorch 1.8.1 深度学习框架，并运行于 CUDA 11.1 环境。使用 AdamW 作为优化器，批量大小为 8，学习率设置为 0.001，训练周期为 300 轮。损失函数为 BceDice 损失，可用式 (22) 表示。

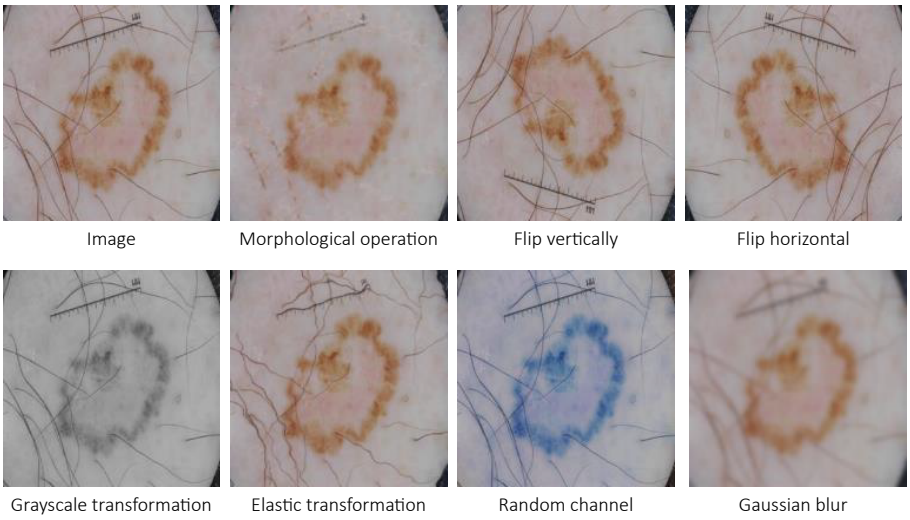


图 10 图像预处理操作可视化

Fig. 10 Visualization of image preprocessing operations

$$\begin{cases} L_{\mathrm{Bce}} = -\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(y_i\log(p_i)+(1-y_i)\log(1-p_i)) \\ L_{\mathrm{Dice}} = 1-\dfrac{2|X\cap Y|}{|X|+|Y|} \\ L_{\mathrm{BceDice}} = \alpha_1 L_{\mathrm{Bce}}+\alpha_2 L_{\mathrm{Dice}} \end{cases}, \tag{22}$$

式中：N为样本总数；y_i为真实标签；p_i为预测值；$|X|$和$|Y|$分别代表真实值和预测值；α_1和α_2为两个损失函数的权值，两个权重的值都设置在 0 和 1 之间，以 0.1 为步长进行多次对比实验，最终确定两个权重的值都为 1。

本文采用 5 个常用的评估指标，包括平均交并比 (mean intersection over union, mIoU)、Dice 相似系数 (dice similarly coefficient, DSC)、准确率 (accuracy, Acc)、灵敏度 (sensitivity, Sen)、特异性 (specificity, Spe) 五种指标。除此之外，参数个数用 Params 表示，单位为百万 (million, M)。计算复杂度是根据浮点运算符 (GFLOPs) 的个数来计算的。模型的 Params 和 GFLOPs 是在 256×256 输入尺寸下测量的。各指标如式 (23) 所示。

$$\begin{cases} mIoU = \dfrac{TP}{TP+FP+FN} \\ DSC = \dfrac{2TP}{2TP+FP+FN} \\ Acc = \dfrac{TP+TN}{TP+TN+FP+FN} \\ Sen = \dfrac{TP}{TP+FN} \\ Spe = \dfrac{TN}{TN+FP} \end{cases}, \tag{23}$$

式中：TP为真阳性 (模型预测和实际标签都为正类)；FP为假阳性 (标签为负类而模型将其预测为正类)；TN为真阴性 (模型预测和标签都为负类)；FN为假阴性 (标签为正类而模型将其预测为负类)。

3.4 对比实验

为验证本文方法的有效性，在 ISIC2016 数据集、ISIC2017 数据集和 ISIC2018 数据集上进行对比实验。所选取的对比模型包括 U-Net[10]、U-Net++[28]、Attention U-Net[11] 和 SANet[29] 经典分割模型，还包括近年来用于医学图像分割任务的将 Transformer 与 CNN 相结合的混合模型 TransFuse[15]、UTNetV2[30]、ATTENTION SWINU-NET[31] 和 MHorUNet[32]，以及近几年的轻量级医学图像分割模型 UNeXt-S[33]、MobileViTv2[34] 和 MSS-UNet[35]。

在这些对比模型中，TransFuse 以并行的方式将 Transformer 和 CNN 结合，有效兼顾了全局上下文建模能力与低层次空间细节的保留。MHorUNet 引入具备高阶空间交互的递归门控卷积，并在跳跃连接路径中使用空间与通道注意力机制，构建增强型 U 型结构，提高融合多级和多尺度信息的能力。UNeXt-S 则提出一种基于卷积多层感知器 (MLP) 的医学图像分割框架，通过在潜在空间中使用标记化的 MLP 模块，有效压缩模型参数规模与计算开销，实现轻量化。经过实验，ISIC2016、ISIC2017 和 ISIC2018 数据集实验结果分别如表 1、表 2 和表 3 所示。结果表明，本文所提方法在不同数据分布与病灶复杂度条件下均展现出良好且稳定的分割性能。

以 ISIC2017 数据集为例，本文方法在 mIoU 和

表 1 在 ISIC2016 数据集上的对比实验结果

Table 1 Comparative experiment results on the ISIC2016 dataset

Model	Params/M	GFLOPs	mIoU/%	DSC/%	Acc/%	Spe/%	Sen/%
U-Net	7.77	13.78	83.38	89.68	95.19	96.39	90.12
U-Net++	9.16	34.90	83.75	89.92	95.63	95.84	91.09
Attention U-Net	8.73	16.74	84.19	90.07	95.72	95.92	90.98
TransFuse	26.27	11.53	85.27	91.40	95.81	96.94	91.34
UTNetV2	12.80	15.50	84.03	90.79	95.64	97.03	88.85
UNeXt-S	**0.30**	**0.10**	84.65	91.35	95.55	96.76	90.93
ATTENTION SWINU-NET	46.91	14.18	84.59	90.96	95.51	96.59	91.06
MSS-UNet	0.33	15.98	85.12	90.89	95.53	96.63	91.68
MHorUNet	9.58	2.25	85.39	91.13	95.41	96.17	92.21
Ours	**0.30**	0.22	**85.78**	**91.90**	**95.92**	**97.15**	**92.54**

表 2 在 ISIC2017 数据集上的对比实验结果

Table 2 Comparative experiment results on the ISIC2017 dataset

Model	Params/M	GFLOPs	mIoU/%	DSC/%	Acc/%	Spe/%	Sen/%
U-Net	7.77	13.78	76.98	86.99	95.65	97.43	86.82
U-Net++	9.16	34.90	77.18	87.12	95.82	96.83	87.09
Attention U-Net	8.73	16.74	77.89	87.57	95.89	96.90	86.96
TransFuse	26.27	11.53	79.21	88.40	96.17	97.98	87.14
UTNetV2	12.80	15.50	77.35	87.23	95.84	98.05	84.85
UNeXt-S	**0.30**	**0.10**	78.26	87.80	95.95	97.74	86.91
ATTENTION SWINU-NET	46.91	14.18	78.55	88.59	95.91	97.69	84.92
MSS-UNet	0.33	15.98	78.66	88.05	95.93	97.57	87.78
MHorUNet	9.58	2.25	79.37	87.99	95.74	97.15	88.86
Ours	**0.30**	0.22	**79.81**	**88.73**	**96.27**	**98.23**	**88.94**

表 3 在 ISIC2018 数据集上的对比实验结果

Table 3 Comparative experiment results on the ISIC2018 dataset

Model	Params/M	GFLOPs	mIoU/%	DSC/%	Acc/%	Spe/%	Sen/%
U-Net	7.77	13.78	77.86	87.55	94.05	96.69	85.86
U-Net++	9.16	34.90	78.31	87.83	94.02	95.75	88.65
Attention U-Net	8.73	16.74	78.43	87.91	94.13	96.23	87.60
SANet	23.90	5.96	79.52	88.59	94.39	95.97	89.46
UTNetV2	12.80	15.50	78.97	88.25	94.32	96.48	87.60
MobileViTv2	1.87	0.70	79.88	88.81	94.40	95.94	89.52
UNeXt-S	**0.30**	**0.10**	79.09	88.33	94.39	96.62	87.15
ATTENTION SWINU-NET	46.91	14.18	79.92	88.59	94.31	96.53	87.29
MSS-UNet	0.33	15.98	80.10	88.67	94.42	96.60	87.46
MHorUNet	9.58	2.25	80.16	88.95	94.49	96.58	88.87
Ours	**0.30**	0.22	**80.65**	**89.37**	**94.75**	**96.73**	**89.74**

DSC 指标上分别达到 79.81% 和 88.73%，相比于经典 U-Net 模型，分别实现了 2.83% 与 1.74% 的性能提升。同时，模型参数量显著减少为 3.86%，计算复杂度也降低为 1.60%，充分体现出其在保证精度的同时具备卓越的计算效率。从实验结果可知，TransFuse 和 MHorUNet 的 mIoU 指标分别为 79.21% 与 79.37%，DSC 指标分别为 88.40% 和 87.99%，与本文其他对比方法相比表现较好，但这与其具有巨大的计算量是分不开的。与 TransFuse 相比，本文方法在参数量和计算复杂度方面分别降低为 1.41% 和 1.91%，显著减轻了硬件负担。本文方法与同为轻量化模型的 MSS-UNet 相比，参数量减少了 0.03 M，相差不大，但计算复杂度降低为 1.38%，MSS-UNet 即使利用双空间移位 MLP 模块和外部注意力在参数压缩方面取得一定成效，但计算复杂度仍远高于本文方法，说明其结构在计算效率方面尚存优化空间，且高计算负荷可能导致过拟合风险，从而影响模型的性能。相比之下，UNeXt-S 虽具有较低的计算复杂度，仅为 0.10 GFLOPs，但其 mIoU 和 DSC 指标分别仅为 78.26% 与 87.80%，在分割精度方面仍存在明显差距，且其他指标性能也不如本文算法。本文方法在该对比下分别提升 1.55% 的 mIoU 与 0.93% 的 DSC，体现出了优异的性能与效率协同的全面优势。

通过对比实验可得，本文方法在参数量、计算复杂度和分割精度上取得了先进的平衡。具体而言，通过与传统 CNN 医学分割模型相比，本文方法显著降低了模型冗余，通过结构设计优化实现了更高效的特征提取与融合。相较于基于 Transformer 的混合模型，本文方法在实现精度提高的同时，大幅减少了计算资源的消耗，显著缓解了部署难题。而面对同类轻量化模型，本文方法不仅在效率方面排在前列，且在分割精度上也保持了更高水平。

此外，为进一步验证本文方法的分割性能，本文在 ISIC2017 与 ISIC2018 数据集上开展了可视化对比实验，结果分别如图 11 与图 12 所示。所选图像涵盖病变边缘模糊、毛发遮挡以及病灶区域形态不规则等挑战场景，旨在全面评估各方法的分割性能。

从可视化结果分析可见，尽管现有分割方法均能大致勾勒出皮肤病变区域的轮廓，但仍存在一定差异性表现。U-Net 由于结构设计较为基础，在面对形态复杂或边界模糊的病变图像时表现不佳，存在明显的欠分割或误分割现象。U-Net++模型通过引入嵌套密集跳跃连接在一定程度上提升了特征融合能力，分割结果较 U-Net 有所改善，但在边缘细节刻画方面仍显不足。Attention U-Net 通过引入注意力机制增强了对关键区域的响应，但其注意力机制设计相对单一，难以兼顾通道与空间维度的特征分布差异，导致其在病变定位及边界刻画方面精度受限，尤其在存在毛发遮挡等干扰因素时表现不稳定。从 ISIC2017 可视化结果可见，MSS-UNet 和 MHorUNet 在病灶边界模糊、区域形态不规则的样本中呈现出分割轮廓不清、边界不完整等问题，说明其在高

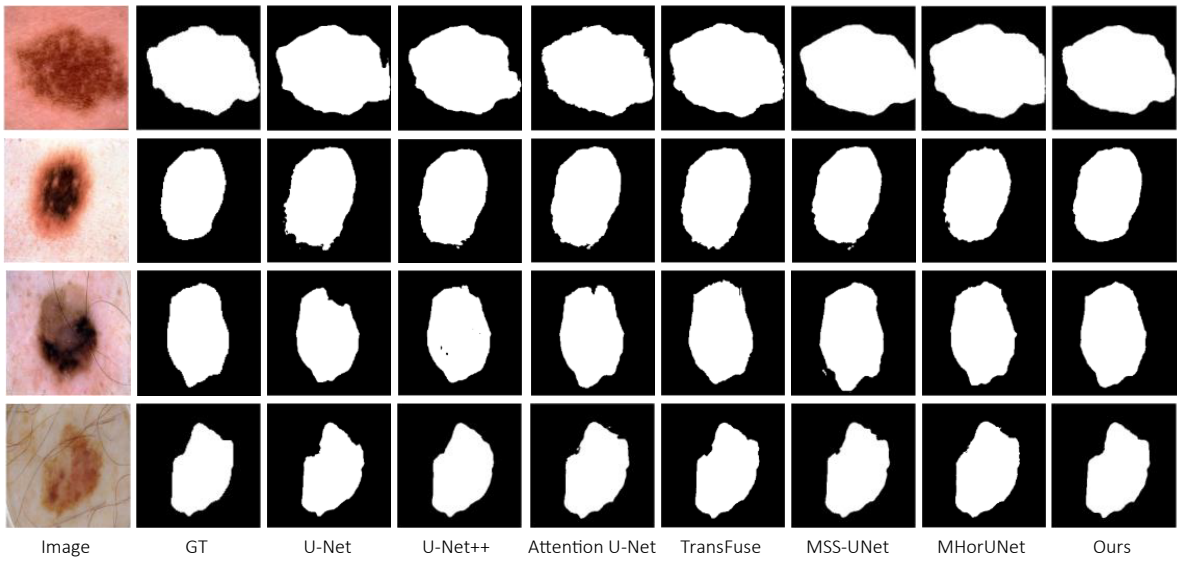


图 11 不同方法在 ISIC2017 数据集上的可视化结果

Fig. 11 Visualized results of different methods on the ISIC2017 dataset

阶语义建模或跨尺度特征融合方面尚有改进空间。而在ISIC2018数据集的对比中，MobileViTv2与UNeXt-S作为轻量化网络，尽管具备较低的计算代价，但受限于网络结构设计的简化，其在处理病变边界扩展显著、形态复杂、图像对比度低等场景时表现不理想，具体表现为病变区域轮廓模糊、边缘细节缺失以及背景误识别等问题，暴露出其在高精度分割任务中易受表征能力限制的不足。相对比而言，本文方法在能够更准确地定位病灶区域，边缘轮廓更加清晰完整，充分验证了所提出模块在增强特征表达、提升结构辨识能力等方面的有效性。

为了更充分验证本文方法在复杂皮肤病变图像分割任务中的有效性，本文从测试集中选取若干具有代表性的复杂样本进行对比实验。所选取的复杂图像涵盖多种挑战性情况：包含多个相互粘连病变区域、受光照影响的小型病变、表面反光、受气泡及血液等噪声干扰，以及存在毛发遮挡且病变区域模糊的图像。通过与近年提出的分割网络进行可视化对比，进一步验证了本文方法的分割性能。复杂皮肤病变分割可视化结果如图13所示。从结果可以看出，轻量化网络UNeXt-S和MSS-UNet对复杂皮肤病变分割的能力较差。在面对毛发遮挡、气泡和血液等噪声干扰时，其分割效果明显下降，尤其在处理多病灶粘连的复杂图像时，出现严重的欠分割现象。MHorUNet虽然能大致分割出病变区域，但对病变区域边缘细节的提取不够充分，难以准确勾勒出病变区域的真实轮廓。相比之下，本文提出的方法在应对表面反光、气泡、血液和毛发遮挡等噪声干扰的复杂图像时，通过所构建的膨胀门控感知增强模块、自适应特征桥接模块和其他模块的协同作用，在有效提取皮肤病变关键特征信息的同时，也显著提升了模型对多阶段、多尺度特征信息的融合能力，实现了对复杂皮肤病变图像的准确分割。

综上所述，现有方法在病灶区域的整体感知与边界细化方面表现欠佳，根本原因在于病变相关特征提取不充分、特征融合能力受限，以及信息传递过程中的语义丢失现象。为此，本文设计轻量初始结构引导模块，使用深度可分离卷积与残差连接机制，在有效控制模型复杂度的前提下增强浅层特征的结构表达能力，提升远程空间依赖建模能力，最大程度保留图像原始细节信息。构建膨胀门控感知增强模块，通过融合分裂膨胀单元、门控注意单元和渐进感知增强单元，在保持模型轻量化特性的同时，显著提升多尺度病变特征的表征能力，增强对病灶区域的精准定位。进一步地，本文在跳跃连接通路中提出自适应特征桥接模块，联合空间注意力、通道注意力与跨粒度特征协同机制，促进跨阶段、多尺度特征的高效融合，有效缓解语义传递过程中的信息退化问题。最后，通过集成高效通道注意力机制，不引入额外参数和计算负担，在不降低特征维度的情况下有效捕获跨通道交互信息。最终，本文提出的分割方法在实现模型高度轻量化的同时，有效提升了皮肤病变区域的分割精度。

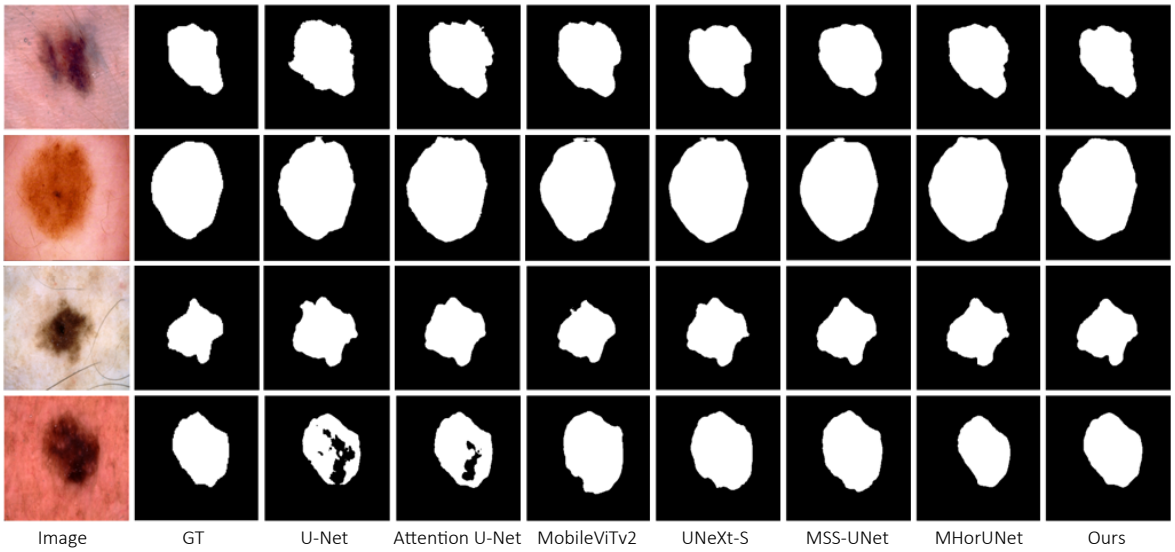


图 12 不同方法在ISIC2018数据集上的可视化结果

Fig. 12 Visualized results of different methods on the ISIC2018 dataset

此外，为全面评估所提模型在不同数据域下的泛化性能，本文在 ISIC2018 与 PH^2 数据集上开展了交叉数据集验证实验。通过在 ISIC2018 数据集上对模型进行训练，并将其直接迁移应用于 PH^2 数据集进行测试，以检验模型在未知数据分布下的适应性与稳健性。相关实验结果如表 4 所示，与最具代表性方法 MHorUNet 相比，本文模型在 mIoU、DSC 与 Acc 三项核心指标上分别提升了 0.46%、0.38% 和 0.39%，充分验证了所提方法在跨数据集场景下的良好表现。该性能提升主要得益于所提模块能够增强特征提取和特征融合的能力，并关注有用信息，从而显著提升分割模型的泛化能力与鲁棒性。

3.5 消融实验

U-Net[11] 作为经典的 5 阶段 U 型结构模型。其各阶段的特征通道数依次为{32,64,128,256,512}，每一阶段均采用两层标准卷积操作。为进一步优化模型结构，本文将 5 阶段的 U 型架构扩展为 6 阶段，并将各阶段的卷积操作简化为单层，以控制模型复杂度。表 5 展示了在

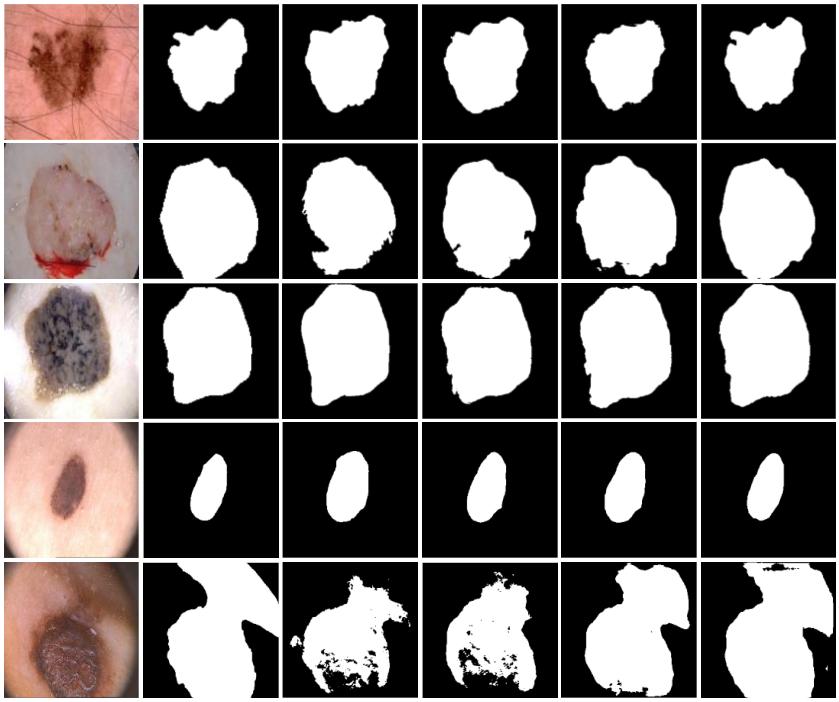


图 13 复杂皮肤病变分割可视化结果

Fig. 13 Visualized results of complex skin lesion segmentation

表 4 对 ISIC2018 和 PH^2 数据集的交叉数据集评估结果

Table 4 Cross-dataset evaluation results of the ISIC 2018 and PH^2 datasets

Model	Params/M	GFLOPs	mIoU/%	DSC/%	Acc/%	Spe/%	Sen/%
U-Net	7.77	13.78	83.09	89.48	94.09	92.29	93.82
U-Net++	9.16	34.90	83.52	89.67	94.49	91.87	94.24
Attention U-Net	8.73	16.74	84.04	89.95	94.68	91.90	94.15
TransFuse	26.27	11.53	85.16	91.26	94.71	92.79	94.56
UTNetV2	12.80	15.50	83.97	90.54	94.59	93.04	92.73
UNeXt-S	**0.30**	**0.10**	84.25	91.68	94.42	92.19	94.09
MHorUNet	9.58	2.25	85.29	91.95	94.47	92.08	95.35
Ours	**0.30**	0.22	**85.75**	**92.33**	**94.86**	**94.34**	**95.94**

ISIC2018 数据集上，不同通道设置下的 U-Net 的性能表现。综合模型精度、参数规模和其他性能指标等因素，最终选取通道数设置为{8,16,24,32,48,64}的 6 阶段 U-Net 作为本文方法的主干网络 (backbone)，并在该结构基础上针对 ISIC2018 数据集开展系列消融实验，以验证各模块的有效性与性能贡献。消融实验结果如表 6 所示，消融实验可视化如图 14 所示。

1) Backbone + LISG

与 Backbone 相比，通过在编码器下采样的第一层引入轻量初始结构引导模块 (LISG) 代替常规卷积操作，不仅增强了远距离空间依赖信息的建模能力，还显著提升了低层语义特征的提取质量，有效缓解了浅层信息的

表 5　不同通道设置下的 U-Net 性能

Table 5　U-Net performance with different channel settings

Channel setting	Params/M	GFLOPs	mIoU/%	DSC/%	Acc/%	Spe/%	Sen/%
{32,64,128,256,512}	7.77	13.78	77.68	87.55	94.05	96.69	85.86
{16,32,64,128,160,256}	1.30	0.42	78.59	88.01	94.15	96.07	88.16
{8,16,32,64,128,160}	0.57	0.14	78.50	87.96	94.24	96.75	86.42
{8,16,32,48,64,96}	0.21	0.10	78.44	87.92	94.11	96.08	87.98
{8,16,24,32,48,64}	0.11	0.07	78.22	87.77	94.09	96.32	87.15
{32,64,128,256,512}	7.77	13.78	77.68	87.55	94.05	96.69	85.86
{16,32,64,128,160,256}	1.30	0.42	78.59	88.01	94.15	96.07	88.16

表 6　消融实验结果

Table 6　Ablation experiment results

Method	Params/M	GFLOPs	mIoU/%	DSC/%	Acc/%	Spe/%	Sen/%
Backbone	0.11	0.07	78.22	87.77	94.09	96.32	87.15
Backbone + LISG	0.11	0.17	78.89	88.17	94.25	96.16	88.31
Backbone + DGPE	0.25	0.08	79.02	88.23	94.39	95.85	88.09
Backbone + AFB	0.15	0.11	78.96	88.19	94.35	96.93	87.43
Backbone + ECA	0.11	0.07	78.80	88.12	94.20	97.53	87.32
Ours	0.30	0.22	80.65	89.37	94.75	96.73	89.74

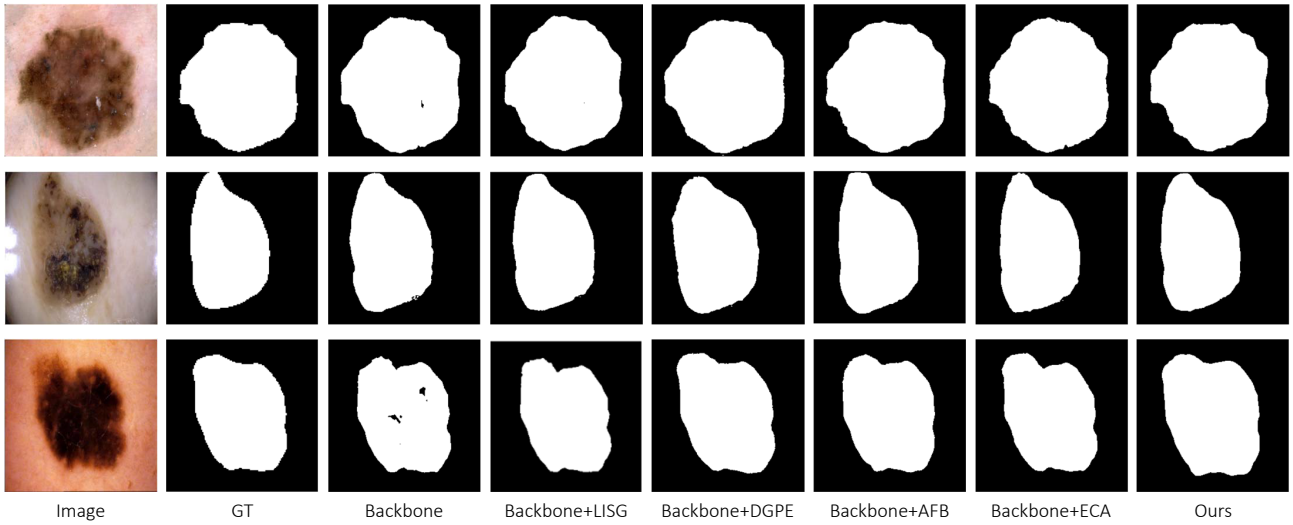


图 14　消融实验可视化

Fig. 14　Visualization of ablation experiment

损失，保留了更加完整的原始特征表示。由表 6 可见，引入 LISG 模块后，在不增加模型参数量的前提下，mIoU 与 DSC 指标分别提升了 0.67% 和 0.40%。尽管 Spe 指标略微下降了 0.16%，但这也从侧面反映出模型对前景区域的感知能力得以增强，而非过度关注背景区域。结合图 14 的消融实验可视化结果，LISG 模块的引入提高了病变区域的定位精度，进一步验证了其在浅层引导中的有效性。

2) Backbone + DGPE

在 Backbone 的第 4 至第 6 阶段设计膨胀门控感知增强模块 (DGPE) 以替代原有的常规卷积块，旨在增强模型对全局语义结构与局部细节信息的协同建模能力。由表 6 可知，引入 DGPE 模块后，mIoU 与 DSC 指标分别提升了 0.80% 和 0.46%。从可视化实验结果来看，模型在病变区域的整体定位与边缘轮廓刻画方面均表现出更高的精度，显著提升了分割性能。上述结果进一步验证了 DGPE 模块的有效性，一方面增强了模型对异常区域空间范围的感知与表达能力。另一方面提升了对病变边缘细节的辨析能力，使模型在分割皮肤病变与正常组织之间的边界时更加准确。

为验证膨胀门控感知增强模块 (DGPE) 中不同膨胀率组合的深度可分离卷积对于分割精度的影响，在相同实验条件下，进行了 7 组具有代表性的实验来分析不同膨胀率组合对算法性能的影响，实验结果如表 7 所示。

从表 7 可知，膨胀率组合为{1,2,5,7}时，模型的性能最好，该设计具有良好的局部和全局尺度分布，低膨胀率 (1,2) 增强其对细节的感知，高膨胀率 (5,7) 提升结构性区域建模的能力，多尺度感受野合理覆盖，融合效果良好，各项指标得到明显提升。膨胀率组合为{2,3,5,7}时，模型也取得了较好的分割性能，但缺乏对小感受野的覆盖，在处理小目标和边缘细节方面稍差一些。而膨胀率组合为{1,1,2,2}和{1,4,7,10}时，模型分割性能较差，前者是因为过度集中于小尺度，多分支感受野高度重叠，特征表达冗余，大区域病灶的建模能力较弱。后者虽然膨胀率跨度大，但缺乏连续性，易造成特征分布不均，融合困难的问题。膨胀率组合为{1,2,3,4}和{1,2,3,6}时，模型具备一定的多尺度覆盖能力，但缺乏更大感受野，限制了对大病变区域的整体建模能力。

综上所述，本文选用膨胀率组合为{1,2,5,7}的膨胀门控感知增强模块，利用低膨胀率和高膨胀率的协同组合，使模型在多尺度建模能力方面得到很大提高，进而提高模型的分割性能。

3) Backbone + AFB

在 Backbone 的跳跃连接部分构建自适应特征桥接模块 (AFB)，旨在解决传统跳跃连接机制中信息传递粗糙、语义融合不足的问题。AFB 通过融合通道注意力、空间注意力与跨粒度特征协同模块，实现跨层级、多尺度特征的动态调整与自适应增强，有效提升了浅层与深层语义信息之间的协调表达能力。如表 6 所示，引入 AFB 模块后，在仅增加极少量参数与计算开销的前提下，mIoU 和 DSC 指标分别提升了 0.74% 和 0.42%，显著增强了模型的分割精度。可视化消融实验结果进一步表明，AFB 能够提升模型对边缘区域的感知与刻画能力，使得病变轮廓更加平滑与清晰。

4) Backbone + ECA

在 Backbone 上采样阶段引入高效通道注意力 (ECA)，能够在不引入额外参数量和计算复杂度的前提下，有效建模跨通道的交互关系，无需降低特征维度，即可精准挖掘关键语义特征，从而增强模型对重要区域的响应能力。如表 6 所示，引入 ECA 模块后，mIoU 和 DSC 分别提升了 0.58% 和 0.35%，验证了该模块在分割性能上的显著增益。进一步从图 14 的可视化结果可见，

表 7　不同膨胀率组合的算法性能

Table 7　Algorithm performance for different expansion rate combinations

Dilation rate setting	mIoU/%	DSC/%	Acc/%	Spe/%	Sen/%
{1,1,2,2}	78.23	87.67	93.75	95.60	87.13
{1,2,3,4}	78.39	87.89	94.15	95.69	87.20
{1,2,3,6}	78.46	87.93	94.21	95.75	87.24
{1,2,5,7}	**79.02**	**88.23**	**94.39**	**95.85**	**88.09**
{2,3,5,7}	78.85	88.16	94.29	95.74	87.82
{1,3,6,9}	78.33	87.85	94.08	95.67	87.18
{1,4,7,10}	78.28	87.77	93.99	95.64	87.15

ECA 在保持结构轻量化的同时，有效缓解了上采样阶段可能导致的位置信息丢失问题，使模型在边缘细节定位方面更加精确。

5) 所提模型

最后，在四个模块协同作用下，模型的整体性能得到显著提升。相比于 Backbone、mIoU 与 DSC 指标分别提升了 2.43% 和 1.60%，体现出所提方法在分割精度上的优势。如图 14 所示，各模块共同作用使模型在病变区域的识别与边缘定位方面均表现出更强的鲁棒性与准确性，分割效果达到最优。综合实验结果表明，本文提出的皮肤病变分割网络在显著提升模型特征提取能力与多尺度特征融合效果的同时，仅引入少量额外参数与计算复杂度，在模型分割性能与复杂度之间实现了良好的平衡。

4 结　论

为解决当前皮肤病变图像分割方法普遍存在分割精度不足与模型参数冗余的问题，本文提出一种基于膨胀门控与自适应桥接的轻量化皮肤病变分割网络。设计膨胀门控感知增强模块 (DGPE)，结合分裂膨胀单元、门控注意单元和渐进感知增强单元实现多尺度特征的充分挖掘与全局-局部信息的协同建模，有效缓解特征细节的丢失问题。构建自适应特征桥接模块 (AFB)，利用通道注意力、空间注意力和跨粒度特征协同机制，增强跨层特征间的有效交互与融合，提升语义信息表达的精准性。在网络初始阶段提出轻量初始结构引导模块 (LISG)，以提升浅层特征的空间表达能力与低级语义感知力，为深层表征提供更加稳健的引导信息。最后，在上采样阶段嵌入高效通道注意力 (ECA)，无需增加模型复杂度，即可捕获关键通道间的关联特征，强化边缘区域的定位能力。本文在四个公开数据集上进行了系统评估，实验结果表明，所提方法在确保高分割精度的同时，显著减少了参数量与计算开销，整体性能优于现有主流对比模型，并通过消融实验进一步验证了各模块的有效性与协同优势。本研究为临床环境中皮肤病变图像的分析提供了更可靠的解决方案。在未来的工作中，将尝试融合更多特征增强策略与结构优化方法，进一步提高模型泛化性，使模型适用于其他疾病研究。

作者贡献

杨本臣：论文方向的确定以及论文的审核与指导；苏焕哲：论文方法的提出，论文的构思、撰写与修改，实验的设计和数据整理；金海波、曲海成：对论文研究方向进行国内外调查、收集与分析，获取资助。

利益冲突

所有作者声明无利益冲突

参考文献

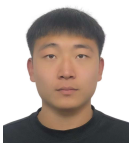
1. Sarwar N, Irshad A, Naith Q H, et al. Skin lesion segmentation using deep learning algorithm with ant colony optimization[J]. *BMC Med Inform Decis Mak*, 2024, **24**(1): 265.
2. Siegel R L, Miller K D, Wagle N S, et al. Cancer statistics, 2023[J]. *CA A Cancer J Clin*, 2023, **73**(1): 17−48.
3. Wang X H, Jiang X D, Ding H H, et al. Bi-directional dermoscopic feature learning and multi-scale consistent decision fusion for skin lesion segmentation[J]. *IEEE Trans Image Process*, 2020, **29**: 3039−3051.
4. Wu H S, Chen S H, Chen G L, et al. FAT-Net: feature adaptive transformers for automated skin lesion segmentation[J]. *Med Image Anal*, 2022, **76**: 102327.
5. Algarin Y A, McCullum C, Patel V A. 33256 Skin cancer screening practices among dermatologists: a survey study[J]. *J Am Acad Dermatol*, 2022, **87**(3): AB203.
6. Yin W J, Huang J H, Ji Y F. Skin tumor classification method based on improved dense convolutional network[J]. *Comput Eng*, 2023, **49**(7): 288−294.
殷文君, 黄建华, 纪元法. 基于改进密集卷积网络的皮肤肿瘤分类方法[J]. 计算机工程, 2023, **49**(7): 288−294.
7. Iqbal S, Khan T M, Naqvi S S, et al. LDMRes-Net: a lightweight neural network for efficient medical image segmentation on IoT and edge devices[J]. *IEEE J Biomed Health Inform*, 2024, **28**(7): 3860−3871.
8. Khan T M, Naqvi S S, Meijering E. ESDMR-Net: a lightweight network with expand-squeeze and dual multiscale residual connections for medical image segmentation[J]. *Eng Appl Artif Intell*, 2024, **133**: 107995.
9. Jiang Y, Cao S, Tao S, et al. Skin lesion segmentation based on multi-scale attention convolutional neural network[J]. *IEEE Access*, 2020, **8**: 122811−122825
10. Anwar S M, Majid M, Qayyum A, et al. Medical image analysis using convolutional neural networks: a review[J]. *J Med Syst*, 2018, **42**(11): 226.
11. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//*18th International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 2015: 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
12. Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: learning where to look for the pancreas[Z]. arXiv: 1804.03999, 2018. https://doi.org/10.48550/arXiv.1804.03999.
13. Sun Y H, Dai D W, Zhang Q N, et al. MSCA-Net: multi-scale contextual attention network for skin lesion segmentation[J]. *Pattern Recogn*, 2023, **139**: 109524.
14. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale[C]// *9th International Conference on Learning Representations*, 2021.

15. Zhang Y D, Liu H Y, Hu Q. TransFuse: fusing transformers and CNNs for medical image segmentation[C]//*24th International Conference on Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, 2021: 14–24. https://doi.org/10.1007/978-3-030-87193-2_2.
16. Zhao H, Wang X. Study on lesion segmentation of melanoma images based on Swin-Transformer[J]. *Comput Eng*, 2024, **50**(8): 249−258.
赵宏, 王枭. 基于 Swin-Transformer 的黑色素瘤图像病灶分割研究[J]. 计算机工程, 2024, **50**(8): 249−258.
17. Liang L M, Lu B H, Long P W, et al. Adaptive feature fusion cascade Transformer retinal vessel segmentation algorithm[J]. *Opto-Electron Eng*, 2023, **50**(10): 230161.
梁礼明, 卢宝贺, 龙鹏威, 等. 自适应特征融合级联 Transformer 视网膜血管分割算法[J]. 光电工程, 2023, **50**(10): 230161.
18. Ruan J C, Xiang S C, Xie M Y, et al. MALUNet: a multi-attention and light-weight UNet for skin lesion segmentation[C]//*2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2022: 1150–1156. https://doi.org/10.1109/BIBM55620.2022.9995040.
19. Hu B B, Zhou P, Yu H F, et al. LeaNet: lightweight U-shaped architecture for high-performance skin cancer image segmentation[J]. *Comput Biol Med*, 2024, **169**: 107919.
20. Yin Y C, Han Z M, Jian M W, et al. AMSUnet: a neural network using atrous multi-scale convolution for medical image segmentation[J]. *Comput Biol Med*, 2023, **162**: 107120.
21. Li D X, Yang F J, Liu Y, et al. Skin lesion segmentation network with cross-attention coding[J]. *Opt Precis Eng*, 2024, **32**(4): 609−621.
李大湘, 杨福杰, 刘颖, 等. 融入交叉注意力编码的皮肤病变分割网络[J]. 光学 精密工程, 2024, **32**(4): 609−621.
22. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[C]//*4th International Conference on Learning Representations*, 2016.
23. Krizhevsky A, Hinton G. Convolutional deep belief networks on CIFAR-10[EB/OL].[2025-04-24]. https://www.cs.toronto.edu/~kriz/conv-cifar10-aug2010.pdf
24. Gutman D, Codella N C F, Celebi E, et al. Skin lesion analysis toward melanoma detection: a challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)[Z]. arXiv: 1605.01397, 2016. https://doi.org/10.48550/arXiv.1605.01397.
25. Berseth M. ISIC 2017-skin lesion analysis towards melanoma detection[Z]. arXiv: 1703.00523, 2017. https://doi.org/10.48550/arXiv.1703.00523.
26. Codella N, Rotemberg V, Tschandl P, et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC)[Z]. arXiv: 1902.03368, 2019. https://doi.org/10.48550/arXiv.1902.03368.
27. Mendonça T, Ferreira P M, Marques J S, et al. PH^2-A dermoscopic image database for research and benchmarking[C]//*2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013: 5437–5440. https://doi.org/10.1109/EMBC.2013.6610779.
28. Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation[C]//*4th International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018: 3–11. https://doi.org/10.1007/978-3-030-00889-5_1.
29. Wei J, Hu Y W, Zhang R M, et al. Shallow attention network for polyp segmentation[C]//*24th International Conference on Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, 2021: 699–708. https://doi.org/10.1007/978-3-030-87193-2_66.
30. Gao Y H, Zhou M, Liu D, et al. A multi - scale transformer for medical image segmentation: architectures, model efficiency, and benchmarks[Z]. arXiv: 2203.00131, 2022. https://doi.org/10.48550/arXiv.2203.00131.
31. Aghdam E K, Azad R, Zarvani M, et al. Attention Swin U-Net: cross-contextual attention mechanism for skin lesion segmentation[C]//*2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, 2023: 1–5. https://doi.org/10.1109/ISBI53787.2023.10230337.
32. Wu R K, Liang P C, Huang X, et al. MHorUNet: high-order spatial interaction UNet for skin lesion segmentation[J]. *Biomed Signal Process Control*, 2024, **88**: 105517.
33. Valanarasu J M J, Patel V M. UNeXt: MLP-based rapid medical image segmentation network[C]//*25th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2022: 23–33. https://doi.org/10.1007/978-3-031-16443-9_3.
34. Mehta S, Rastegari M. Separable self-attention for mobile vision transformers[Z]. arXiv: 2206.02680, 2022. https://doi.org/10.48550/arXiv.2206.02680.
35. Zhu W H, Tian J Y, Chen M Z, et al. MSS-UNet: a Multi-Spatial-Shift MLP-based UNet for skin lesion segmentation[J]. *Comput Biol Med*, 2024, **168**: 107719.

作者简介



杨本臣 (1975-)，男，博士，副教授，硕士生导师，主要从事深度学习、图像处理、智能数据处理方面的研究。
E-mail：yangbenchen@lntu.edu.cn



【通信作者】苏焕哲 (2002-)，男，硕士研究生，主要从事医学图像分割方面的研究。
E-mail：2865775393@qq.com

A lightweight skin lesion segmentation network based on dilated gating and adaptive bridging

Yang Benchen, Su Huanzhe*, Jin Haibo, Qu Haicheng

School of Software, Liaoning University of Engineering and Technology, Huludao, Liaoning 125105, China

Abstract

Objective Skin lesions are often accompanied by complex features such as blurred boundaries, variable shapes, and hair occlusion, which makes it difficult to accurately locate and finely characterize the lesion area in the segmentation task. At the same time, existing methods often have problems such as large number of parameters and high computational complexity, making them difficult to deploy on resource-constrained terminal devices. To this end, this paper proposes a novel lightweight skin lesion segmentation network based on dilated gating and adaptive bridging mechanisms.

Methods The proposed method is designed to strike a balance between efficiency and segmentation performance, addressing the need for high-precision medical image analysis under resource-constrained environments.Specifically, a Dilated Gated Perceptual Enhancement (DGPE) module is introduced to improve both local detail modeling and global contextual understanding. This module employs a channel decoupling strategy combined with parallel depthwise separable convolutions of varying dilation rates to extract features at multiple scales. By integrating a gating attention mechanism and a progressive perception structure, the model effectively suppresses redundant background information and enhances edge structure preservation. To further bridge the semantic gap between encoder and decoder features, an Adaptive Feature Bridging (AFB) module is constructed. This module incorporates a dual attention mechanism—including Channel Attention Bridge (CAB) and Spatial Attention Bridge (SAB)—to perform cross-level feature calibration and establish global receptive field associations. In addition, a Cross-Granularity Feature Integration (CGFI) block is introduced to merge multi-resolution semantic features, enhancing the model's capability to handle complex lesion structures with varying shapes and sizes. Furthermore, a Lightweight Initial Structure-Guided (LISG) module is developed to enhance shallow feature extraction. It integrates standard convolutions with large-kernel depthwise separable convolutions, forming long-range spatial dependencies in the early stages of the network. This is further supported by residual connections and cascaded small-kernel convolutions, ensuring effective gradient flow and feature stability across network layers. Lastly, an Efficient Channel Attention (ECA) mechanism is incorporated to model channel-wise dependencies using lightweight, adaptive one-dimensional convolution, effectively mitigating information loss during upsampling and preserving feature consistency throughout the decoding process.

Results and Discussions Taking the ISIC2017 dataset as an example, our proposed method achieves mIoU and DSC metrics of 79.81% and 88.73%, respectively, representing performance improvements of 2.83% and 1.74% compared to the classic U-Net model. Simultaneously, the number of model parameters is significantly reduced by 25 times, and the computational complexity is reduced by 62 times, fully demonstrating its excellent computational efficiency while maintaining accuracy. Experimental results show that TransFuse and MHorUNet achieve mIoU of 79.21% and 79.37%, and DSC metrics of 88.40% and 87.99%, respectively, performing better than other methods compared in our paper. However, this is inseparable from their enormous computational burden. Compared to TransFuse, our proposed method reduces the number of parameters and computational complexity by 87 times and 51 times, respectively, significantly alleviating the hardware burden. Compared to MSS-UNet, another lightweight model, our proposed method reduces the number of parameters by 0.03M, a negligible difference, but achieves a 72-fold reduction in computational complexity. Even with the advancements in parameter compression achieved through dual-space shift MLP modules and external attention, MSS-UNet's computational complexity remains significantly higher than our proposed method, indicating room for improvement in computational efficiency. Furthermore, its high computational load may lead to overfitting, impacting model performance. In contrast, while UNeXt-S boasts a lower computational complexity of only 0.10 GFLOPs, its mIoU and DSC metrics are only 78.26% and 87.80%, respectively, showing a significant gap in segmentation accuracy. Its performance in other metrics is also inferior to our proposed algorithm. Our proposed method, in comparison, improves mIoU by 1.55% and DSC by 0.93%, demonstrating a superior balance between performance and efficiency.Comparative

Foundation item: National Natural Science Foundation of China Fund (62173171), Basic Scientific Research Project of Liaoning Provincial Department of Education Fund (JYTMS20230804), and Basic Scientific Research Project of Liaoning Provincial Universities Fund (LJKMZ20220699)

*Correspondence: Su H Z, E-mail: 2865775393@qq.com

experiments show that our proposed method achieves an advanced balance between parameter count, computational complexity, and segmentation accuracy. Specifically, compared to traditional CNN-based medical segmentation models, our proposed method significantly reduces model redundancy and achieves more efficient feature extraction and fusion through structural design optimization. Compared to Transformer-based hybrid models, our method improves accuracy while significantly reducing computational resource consumption, thus alleviating deployment challenges. Furthermore, compared to similar lightweight models, our method not only ranks among the best in efficiency but also maintains a higher level of segmentation accuracy.

Conclusions To address the common problems of insufficient segmentation accuracy and redundant model parameters in current skin lesion image segmentation methods, this paper proposes a lightweight skin lesion segmentation network based on dilatational gating and adaptive bridging. A dilatational gating perception enhancement module (DGPE) is designed, combining split dilatation units, gated attention units, and progressive perception enhancement units to fully mine multi-scale features and co-model global-local information, effectively mitigating the loss of feature details. An adaptive feature bridging module (AFB) is constructed, utilizing channel attention, spatial attention, and cross-granularity feature collaboration mechanisms to enhance effective interaction and fusion between cross-layer features, improving the accuracy of semantic information representation. A lightweight initial structure guidance module (LISG) is proposed in the initial stage of the network to improve the spatial representation ability and low-level semantic perception of shallow features, providing more robust guidance information for deep representation. Finally, efficient channel attention (ECA) is embedded in the upsampling stage, capturing the correlation features between key channels without increasing model complexity, and strengthening the localization ability of edge regions. This paper conducts a systematic evaluation on four publicly available datasets. Extensive experiments conducted on the ISIC2016, ISIC2017, and ISIC2018 datasets demonstrate the superior performance of the proposed model. It achieves mIoU scores of 85.78%, 79.81%, and 80.65%, and DSC scores of 91.90%, 88.73%, and 89.37%, respectively. With only 0.30 M parameters and 0.22 GFLOPs, the model significantly outperforms existing methods in terms of accuracy and efficiency, showcasing strong potential for real-world deployment in mobile medical applications and intelligent diagnostic systems, especially under constrained computing environments and limited storage conditions. Experimental results show that the proposed method significantly reduces the number of parameters and computational overhead while ensuring high segmentation accuracy. Its overall performance outperforms existing mainstream comparative models. Ablation experiments further validate the effectiveness and synergistic advantages of each module. This study provides a more reliable solution for the analysis of skin lesion images in clinical settings. Future work will explore integrating more feature enhancement strategies and structural optimization methods to further improve the model's generalization ability and make it applicable to research on other diseases.

Keywords: image processing; skin lesion segmentation; U-Net; lightweight; attention mechanism

Yang B C, Su H Z, Jin H B et al. A lightweight skin lesion segmentation network based on dilated gating and adaptive bridging[J]. *Opto-Electron Eng*, 2026, **53**(1): 250121; DOI: 10.12086/oee.2026.250121

Opto-Electronic Journals Group
www.oejournal.org

扫描二维码，获取 PDF 全文