

基于深度学习的轻量级实时图像分割方法研究

李建锋^{1,2}, 熊明强^{1,2}, 陈园琼^{1,2,3}, 王宗达^{1,2}, 向涛^{1,2}, 孙培玮^{1,2}

(1. 吉首大学通信与电子工程学院, 湖南 吉首 416000;

2. 吉首大学计算机科学与工程学院, 湖南 吉首 416000;

3. 中南大学计算机学院, 湖南 长沙 410083)

摘要: 针对深度学习在各领域应用中因模型复杂度提升而引发的计算与存储负担, 尤其在图像分割任务中面临的算法复杂性、实时响应不足及高内存占用问题, 提出了一种轻量级且高效的分割网络架构——多尺度叠加融合网络 (MSFNet)。MSFNet 设计了一个双分支多尺度边界融合模块, 该模块通过融合不同尺度的特征信息与边界细节, 有效提升了图像分割精度, 同时显著减少了模型参数量。实验结果表明, MSFNet 在 3 个公开数据集上表现优异, 其模型参数量仅为 0.6×10^6 , 在 RTX 3070 GPU 上处理大小为 800×800 像素的图像仅需 12 ms, 显著提升了分割任务的执行效率和资源利用率。因此, 该模型特别适合应用于资源有限的边缘设备或移动设备中, 为实时图像分割应用提供了有力的技术支撑。

关键词: 图像分割; 轻量级实时网络; 双分支多尺度边界融合模块

中图分类号: TN92

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2025026

Research on lightweight real-time image segmentation methods based on deep learning

LI Jianfeng^{1,2}, XIONG Mingqiang^{1,2}, CHEN Yuanqiong^{1,2,3}, WANG Zongda^{1,2}, XIANG Tao^{1,2}, SUN Peiwei^{1,2}

1. School of Communication and Electronic Engineering, Jishou University, Jishou 416000, China

2. School of Computer Science and Engineering, Jishou University, Jishou 416000, China

3. School of Computer, Central South University, Changsha 410083, China

Abstract: In response to the computational and storage burdens caused by the increasing model complexity in deep learning applications, especially in image segmentation tasks where algorithmic complexity, insufficient real-time responsiveness, and high memory usage were prevalent, a lightweight and efficient segmentation network architecture——multi-scale superposition fusion network (MSFNet) was proposed. MSFNet featured a dual-branch multi-scale boundary fusion module, which effectively enhanced segmentation accuracy by integrating feature information and boundary details from different scales. At the same time, it significantly reduced the model parameter count. Experimental results show that MSFNet outperforms other models on three public datasets, with a model size of only 0.6×10^6 parameters. On the RTX 3070 GPU, it processes 800×800 pixels images in just 12 ms, significantly improving the execution efficiency and resource utilization of segmentation tasks. Therefore, this model is particularly well-suited for deployment on resource-constrained edge or mobile devices, providing a favorable technical foundation for real-time image segmentation applications.

Keywords: image segmentation, lightweight real-time network, dual-branch multi-scale boundary fusion module

收稿日期: 2024-12-23; 修回日期: 2025-02-07

通信作者: 陈园琼, yqchen@vip.163.com

基金项目: 国家自然科学基金资助项目 (No.61962023)

Foundation Item: The National Natural Science Foundation of China (No.61962023)

0 引言

图像分割作为计算机视觉中的基础任务之一,长期以来在多个领域中扮演着至关重要的角色。尤其是在医学影像分析^[1]、自动驾驶、安防监控、遥感图像分析等领域中,图像分割技术不仅能够提供关键的图像内容信息,还能够为后续的图像理解和决策提供重要支持。例如,在医学影像分析中,精确的图像分割技术有助于医生识别病变区域,从而为疾病的早期诊断和治疗提供依据^[2]。在自动驾驶中,实时准确的图像分割能够帮助车辆识别周围环境,保障行车安全^[3]。在遥感图像分析中,图像分割能够提取出地理信息,用于环境监测、土地利用^[4]等应用。

随着深度学习技术的快速发展,基于卷积神经网络(CNN, convolutional neural network)的图像分割算法在这些领域中得到了广泛应用,并且显著提高了分割精度和计算效率。然而,这些深度学习技术的复杂性也使计算资源和存储资源需求大幅增加。尤其是在嵌入式设备、物联网终端、移动设备等资源受限的环境中,深度网络的高计算要求往往难以得到满足。因此,如何在保证分割精度的同时优化计算效率,成为当前图像分割技术发展的关键问题。针对这一挑战,研究者们提出了多种优化方法,包括设计轻量级网络架构、减少模型参数、采用高效的卷积操作以及利用硬件加速等手段。这些优化方法不仅使得图像分割能够在资源受限的设备上高效运行,同时也使得实时性得以保障,从而进一步推动了实时图像分割技术的普及,为各类应用提供了可靠的解决方案。

本文方法在性能、效率与资源利用率之间实现了良好平衡,为实时语义分割和小型设备应用提供了新的解决方案。本文的主要贡献可总结如下。

1) 提出了一种轻量级实时图像分割网络模型——多尺度叠加融合网络(MSFNet, multi-scale superposition fusion network)。该模型在优化参数规模和推理速度的同时,显著提升了图像分割性能,综合效果卓越。

2) MSFNet使用少量标准卷积、非对称卷积、深度可分离(DS, depthwise separable)卷积和残差深度可分离(RDS, residual depthwise separable)卷积进行特征提取,并通过补丁合并操作实现下采样。相比传统方法,这种设计显著提高了模型的推

理速度,减少了图像中的信息损失,确保分割结果更加精准。

3) 设计了一个双分支多尺度边界融合(DBMBF, dual-branch multi-scale boundary fusion)模块,用于图像特征提取。该模块包括边界特征提取辅助(BFEA, boundary feature extraction assistance)分支和多尺度特征信息融合(MFIF, multi-scale feature information fusion)分支。通过融合2个分支的特征图,显著提升了模型的分割精度和鲁棒性。

4) MSFNet对边缘设备、移动设备以及资源受限的应用场景进行了优化。模型参数量仅为 0.6×10^6 ,在RTX 3070 GPU上处理大小为800像素×800像素的图像仅需12 ms,为移动设备上的实时图像分割提供了高效且实用的解决方案。

1 相关工作

1.1 图像分割

随着深度学习技术的快速发展,基于深度神经网络的图像分割方法已经成为主流,与传统方法(如阈值分割^[5]、区域生长^[6]、边缘检测^[7]等)相比,深度学习技术具有显著优势。传统机器学习方法通常依赖于人工设计的特征,如颜色、纹理、灰度值等,且对复杂背景、噪声和不规则形态的分割效果较差。例如,阈值分割方法通过设定固定阈值来区分不同区域,但在复杂场景下容易失效。区域生长方法依赖种子点的选择,并且在图像质量较差或目标不明显时容易产生错误。边缘检测方法容易受到噪声影响,导致边缘模糊。深度学习技术能够在大规模数据集上自动学习到更为精细和复杂的特征,从而显著提高分割精度和鲁棒性。而基于深度神经网络的图像分割方法,特别是卷积神经网络及其变种,能够在大规模数据集上自动学习到更加丰富和复杂的特征,克服了传统方法的局限性。深度学习技术不仅能够从大量数据中提取出高级语义信息,还能在图像分割过程中实现自适应的特征学习,使分割结果更加精确和鲁棒。尤其在医学图像、卫星遥感图像以及自动驾驶等高要求领域,深度学习技术显著提高了分割精度,推动了图像分割技术的发展。

尽管现有的深度学习图像分割模型在精度上取得了显著突破,但这些模型的高性能通常伴随着计算复杂度和存储开销的显著增加。特别是在处理高

分辨率图像时，这一问题尤其突出。高分辨率图像包含丰富的细节信息，要求分割模型具备更强的特征提取能力和更精细的分割精度。然而，为了满足这些需求，许多高性能的分割模型（如 AlexNet^[8]、SegNet^[9]、DeepLab^[10-13]等）通常依赖深度卷积神经网络来提取多层次和复杂的特征。这些深度卷积神经网络尽管能够捕获丰富的上下文信息和细节特征，却往往具有庞大的参数量和极高的计算复杂度。

1.2 实时图像分割

实时图像分割旨在快速、准确地对静态动态图像或视频中的每个像素进行分类。为实现低时延和高效处理，通常采用轻量化网络结构、轻量化卷积和高效的上采样下采样技术，优化计算量和内存消耗，确保分割精度与速度的平衡。通过这些技术，实时图像分割能够在资源受限的环境中提供快速响应和精准的分割结果，满足实时应用的需求。近年来，基于实时分割的算法一个接一个地出现。以下是一些基于卷积神经网络的优秀算法。文献[14]采用了深度可分离卷积和低计算成本的特征提取模块，减少了模型的计算量和内存消耗。网络采用逐像素分类和层次化结构，将图像分割任务拆解为多个模块，结合下采样和上采样操作，有效融合了空间和语义信息，从而在保证较高分割精度的同时，实现了低延迟和快速响应。文献[15]提出了一种基于图像合成和特征对齐（ISFA, image synthesis and feature alignment）的无监督域适应方法，用于对眼底图像上的视盘和杯进行分割。利用基于生成对抗网络（GAN, generative adversarial network）的图像合成机制以及视盘和视杯的边界信息生成类目标查询图像，作为源域图像和目标域图像之间的中间潜在空间缓解域移位问题，在提升分割精度的同时降低模型参数量。文献[16]提出了 PocketNet 范式，通过限制卷积神经网络中通道数量的增长来减小深度

学习模型的大小。将模型中的通道数量限制在小而同一的大小从而大大降低模型参数量。文献[17]引入了一种基于卷积神经网络和多层感知机（MLP, multilayer perceptron）的轻量级医学图像分割网络（LcmUNet）。在 CNN 部分，提出了一个轻量级的卷积模块，它结合了非对称卷积、深度可分离卷积和注意力机制，以减少网络参数的数量，同时保持强大的特征提取能力。在 MLP 部分，提出了一个轻量级的多层感知器模块，该模块有助于增强上下文信息，同时关注本地信息，并在保持高推理速度的同时提高分割准确性。该网络还涵盖了编码器和解码器之间各个级别的跳过连接。其网络在广泛的实验中准确地获得了实时分割结果。文献[18]提出了一种新的卷积层，即通道分割循环卷积（CSR-Conv），其中拆分输出通道以生成长度为 T 的数据序列，作为具有共享权重的循环层的输入。因此，通过简单用 CSR-Conv 层替换一些线性卷积层来构建轻量级卷积网络。文献[19]引入了一种称为 ResX 块的新型主干网络，它利用矩形和膨胀卷积来捕获多尺度特征，并且所提 ResX 块在准确性和鲁棒性方面均优于主流特征提取块。其研究证实了准确的多尺度特征提取对于提高 CNN 在医学图像分割中性能的重要性。此外，还验证了矩形和膨胀卷积在医学图像分割中捕获多尺度特征的潜力，并且该模块在降低模型参数的同时提升了模型性能。

受上述实时图像分割网络研究的启发，本文设计了一种轻量化架构，以确保组件既能独立运行又能协同工作，从而提升网络整体性能。该架构采用以下轻量化技术优化分割过程：1) 使用少量标准卷积与补丁合并操作，提取图像中的空间细节信息；2) 为了降低参数数量和计算成本，使用 DS 卷积代替普通卷积，这样可使参数和计算复杂度降低至标准卷积的 $\frac{1}{9} \sim \frac{1}{8}$ （当 $D_k=3$ 时）。参数和计算成本结果如式(1)所示。

$$\begin{cases} C_{fc} = \frac{D_k \times D_k \times M \times N \times H_p \times W_p}{D_k \times D_k \times M \times N \times W_p + M \times N \times H_p \times W} = \frac{D_k^2 \times N}{D_k^2 + N} \\ C_{pc} = \frac{D_k \times D_k \times M \times N}{D_k \times D_k \times M + M \times N} = \frac{D_k^2 \times N}{D_k^2 + N} \end{cases} \quad (1)$$

其中，分子表示标准卷积，分母表示深度可分离卷积， D_k 表示卷积核大小， W_p 和 H_p 分别表示输入图

像的宽度和高度， M 表示输入通道数， N 表示输出通道数， C_{fc} 表示浮点运算成本比， C_{pc} 表示参数成

本比。

非对称卷积 (AC, asymmetric convolution) 与普通卷积操作相比, 采用 AC 可以节省大约 $\frac{k}{2}$ 的参数数量, 如式(2)所示。

$$R_p = \frac{k \times 1 + 1 \times k}{k \times k} \quad (2)$$

其中, 分子表示非对称卷积, 分母表示标准卷积, k 表示卷积核的大小。

残差深度可分离卷积以减少计算复杂度; 非对称金字塔结构 (APS, asymmetric pyramid structure) [20] 用来提取空间细节信息和上下文信息; 多尺度边界信息融合 (MBIF, multi-scale boundary information fusion) [21-23] 模块以便于模型精准提取分割图像的边界信息。

2 方法

2.1 网络结构详解

本文介绍了模型的3个主要组成部分: 空间信息特征获取模块 (模块A)、全局上下文信息特征获取模块 (模块B) 以及分类器模块 (模块C), 其网络结构如图1所示。

模块A负责提取图像中的空间细节特征, 包含2个关键部分: 下采样模块和DBMBF模块, 图1中显示了DBMBF模块。在下采样过程中, 采用补丁合并操作以最大限度保留眼底图像的原始信息。每次补丁合并后, 添加一个RDS卷积层, 增强对空间细节信息的提取能力。RDS卷积通过残差结构将输出特征与原始输入直接连接, 从而有效提高空间交互信息的提取效率。DBMBF模块的引入显著提升了模型在多尺度特征和边界空间细节信息提取方面的能力。通过融合来自不同尺度的特征, DBMBF模块增强了模型对关键区域的敏感性, 从而提升图像分割性能。如图1所示, 模块A展现了下采样模块与DBMBF模块的组合设计, 进一步增强了模型在不同尺寸特征图上的空间特征提取能力, 为后续模块的全局信息整合和分类提供了更加丰富的细节信息支撑。

模块B负责提取全局上下文信息特征, 由3个高效的逆残差瓶颈块和一个DBMBF模块组成。逆残差瓶颈块采用来自MobileNetV2[24]的2种逆残差瓶颈结构和一个压缩和激励 (SE, squeeze-and-excitation) [9] 注意力模块。在减小模型参数的同时,

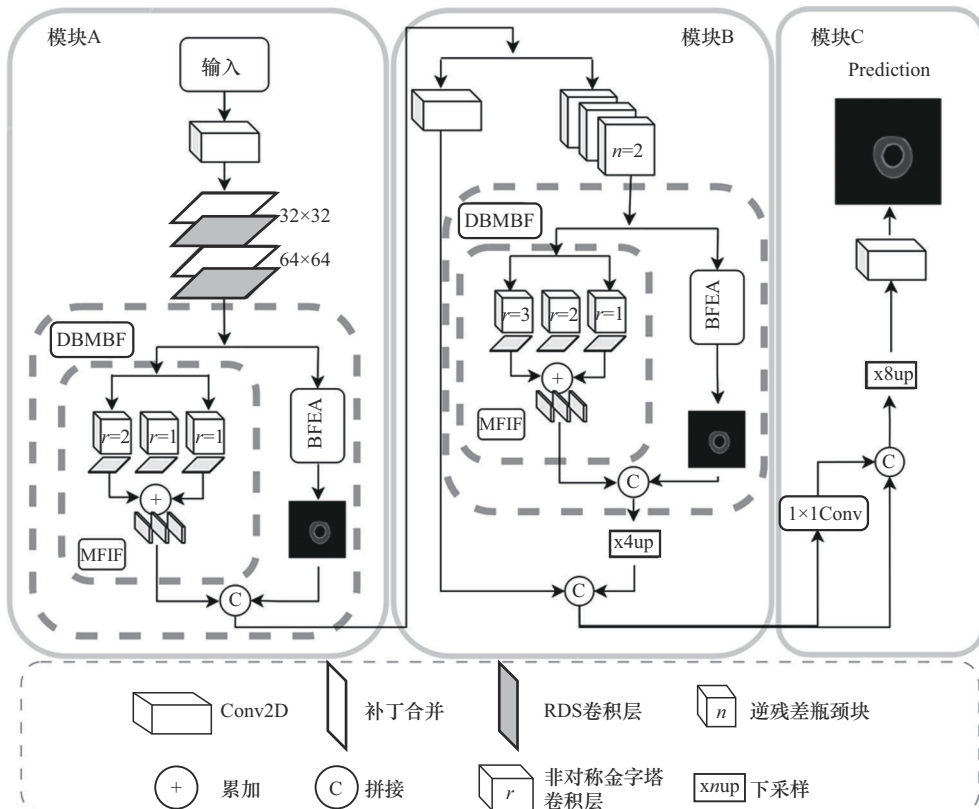


图1 MSFNet网络结构

SE 注意力模块通过自适应调整通道特征的权重，增强了网络对关键信息的关注，抑制无关信息的干扰。这种组合弥补了传统方法在特征选择和鲁棒性方面的不足，显著提高了模型在复杂任务中的性能表现。每个逆残差瓶颈块的输出特征通道数量逐渐增大（32→64→128）。在前 2 个块中，第一个逆残差瓶颈块的步长设置为 2，以调整输入特征图的大小，后续的步长设置为 1，以保持较简洁的操作。模块 A 生成的空间细节特征图被映射至 3 个逆残差瓶颈块，经过处理得到一个深度为 128、尺寸为 16×16 的上下文特征图。接下来，将该上下文特征图输入 DBMBF 模块，通过不同尺度的卷积块强化上下文信息与边界细节信息的提取和融合，最终生成一个深度为 129、尺寸为 16×16 的边界增强上下文特征图。为了降低参数数量和计算成本，使用 DS 卷积代替普通卷积，这样可使参数和计算复杂度降低至标准卷积的 $\frac{1}{9} \sim \frac{1}{8}$ （当 $D_k=3$ 时）。

在获得边界增强上下文特征图后，为加强模块间的相关性，特征图经过 4 次上采样，尺寸从 16×16 增至 64×64。与此同时，使用 1×1 卷积调整模块 A 生成的空间信息特征图的维度（65→129），使其与上下文特征图的维度一致，最终实现两者的融合，得到一个同时包含空间细节信息与上下文信息的全局上下文细节特征图。图 1 中已标出模块 B 的位置，展示了该模块在全局信息提取中的关键作用。

模块 C 是分类器模块，包含一个 1×1 辅助卷积

和一个标准卷积，操作步长均设置为 1。模块 B 生成的最终融合特征图将作为输入传递至模块 C。在模块 C 中，采用辅助分支机制，将模块 B 生成的特征图传送至辅助卷积层。通过该辅助卷积层获取的边界特征信息与原始特征图进行融合，以提高分割边界的精确度。接着，使用标准卷积对特征图像素进行分类，生成尺寸为 64×64 的图像。为了进一步提升分割结果的分辨率，最后对该图像进行 8 倍上采样（从 64×64 增至 512×512）操作，以获得与输入图像相同尺寸的分割图。

2.2 补丁合并操作

图像分割网络通常采用平均池化或最大池化操作进行下采样，但这些操作容易导致有效特征信息的丢失，从而影响模型性能。为了解决这一问题，本文在下采样阶段采用了补丁合并操作，以减少有效特征信息的损失。此方法通过缩小特征图尺寸的同时调整其深度，灵感来源于 Swin transformer^[25]，如图 2 所示。

与传统的池化方法（每次在行列方向上减半分辨率）不同，补丁合并操作在行列方向上每隔 2 个位置选取一个元素，并将这些元素组合成 2×2 的补丁小块。然后，将补丁中相同位置的元素合并，生成 4 个特征图。接着，4 个特征图被连接在一起，并通过归一化层进行标准化。随后，利用全连接层将特征图的深度从 C 线性转换为 $\frac{C}{2}$ 。经过这一操作，特征图的深度被缩小为原来的一半，其高度和宽度则减半。

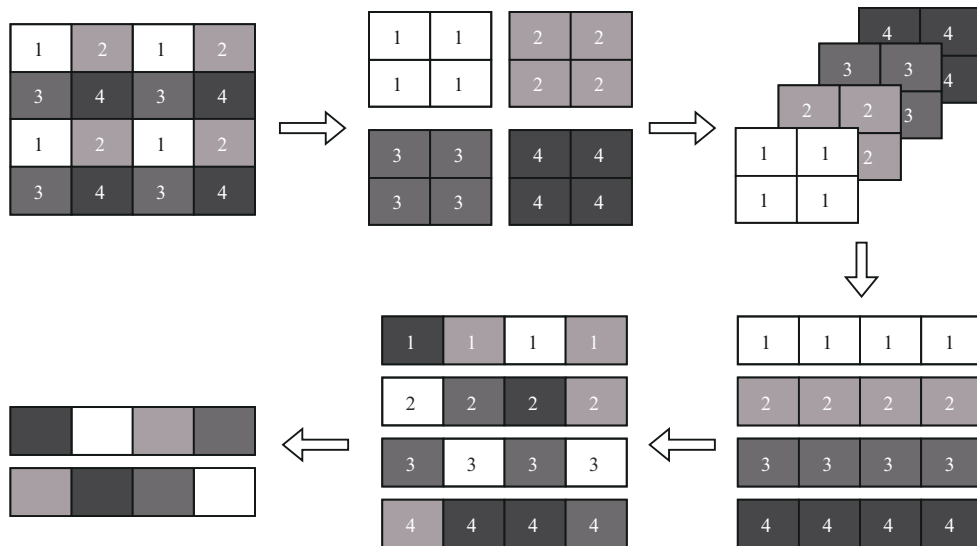
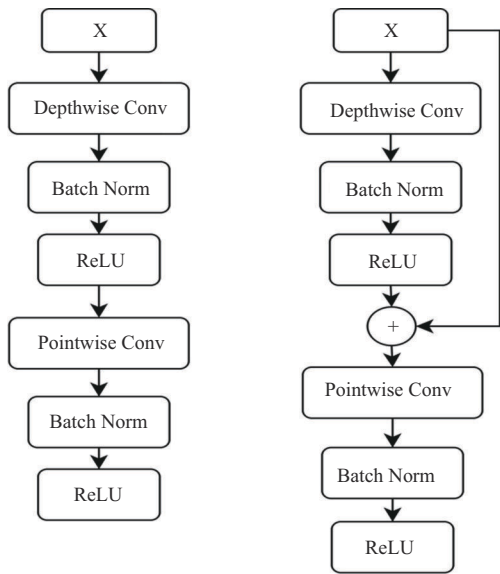


图 2 补丁合并操作

这种补丁合并操作的突出优势在于其能够在保持图像分辨率的同时,有效融合空间信息。这使得网络能够更加精确地捕捉图像中的局部特征,且不丢失关键信息,尤其是在进行图像分割任务时,能够更好地利用像素级别的细节。相比传统的池化方法,补丁合并操作不仅能够减少有效特征信息的损失,还能提高网络在处理复杂图像和高维数据时的表达能力和性能,最终提升了分割精度和模型鲁棒性。通过这一操作,网络能够更好地利用空间信息的关联性,增强对图像关键区域的关注,从而在图像分割任务中获得更高的精度和更强的泛化能力。这种方法的引入不仅优化了下采样过程,还提升了整个网络架构的表现,使模型在复杂和多样化的应用场景下展现出更强的适应性。

2.3 残差深度可分离卷积

尽管深度可分离卷积(如图3(a)所示)在参数和计算复杂度方面相较于标准卷积具有显著优势,但由于其在空间和通道维度上独立计算,容易导致空间交互信息的丢失。为了克服这一问题,本文引入了残差深度可分离卷积(如图3(b)所示)。



(a) 深度可分离卷积 (b) 残差深度可分离卷积

图3 卷积优化结构对比

残差深度可分离卷积与深度可分离卷积的计算过程基本相同,但在深度卷积、归一化和激活函数处理之后,并不直接执行逐点卷积。相反,残差操作将深度卷积后的特征图与原始的残差信号进行连接。这一操作显著减少了深度卷积带来的空间交互

信息的损失,提升了训练过程的稳定性,降低了模型退化的风险。与普通卷积和深度可分离卷积相比,残差深度可分离卷积不仅在降低参数数量的同时提高了特征提取能力,而且增强了网络的整体性能。此外,残差深度可分离卷积还提高了模型训练过程的稳定性。由于残差连接的引入,它提供了一条快捷的路径,允许信息流可以在网络中更容易地传递,减少了训练过程中的梯度衰减问题,特别是在较深的网络中。通过这种方式,模型能够更容易地训练出深层次的特征,同时避免了因层数过多而导致的训练困难。与传统的普通卷积和深度可分离卷积相比,残差深度可分离卷积不仅在降低模型参数数量方面发挥了重要作用,还显著提高了特征提取能力。通过减少计算开销和参数数量,残差深度可分离卷积在保持模型轻量化的同时,提升了网络对复杂特征的捕捉能力。

2.4 双分支多尺度边界融合模块

DBMBF 模块旨在增强对图像中的边界信息和细节信息的提取能力。该模块包含 BFEA 分支和 MFIF 分支。BFEA 分支采用标准 1×1 卷积和最大池化,以步长为 1 提取图像中的分割边界信息。在 MFIF 分支中,通过连接 3 个单尺度特征提取子分支的输出特征图作为总输出特征图,从而减少卷积过程中的特征损失。每个单尺度特征提取子分支由 3 个大小分别为 3、5 和 7 的 AC 组成的特征金字塔结构构成,卷积的扩张系数参考了文献[26]。在此结构中,每个卷积核包含 $k \times 1$ 和 $1 \times k$ 的卷积操作。与普通卷积操作相比,采用 AC 可以节省大约 $\frac{2}{k}$ 的参数数量。该分支通过整合并融合各子分支的特征图,生成包含多尺度细节信息的特征融合图,从而增强了边界信息的提取能力并减少了特征损失。

此外,在图 2 的 DBMBF 模块中,可以看到标注的 r 值。由于模块 A 和模块 B 中均设置了相同的 DBMBF 模块,为了在模块 A 中提取更多原始图像的边界信息和细节信息,将其扩张系数 r 设置为 (1,1,2)。为了在模块 B 中获得更精确的分割精度,将其扩张系数 r 设置为 (1,2,3)。

MFIF 分支的计算复杂度不仅低,而且在提取分割图像边界信息方面非常有效。在后续的消融实验中验证了这一点,实验结果显示了该模块在分割精度上的优势。DBMBF 模块结构如图 4 所示。

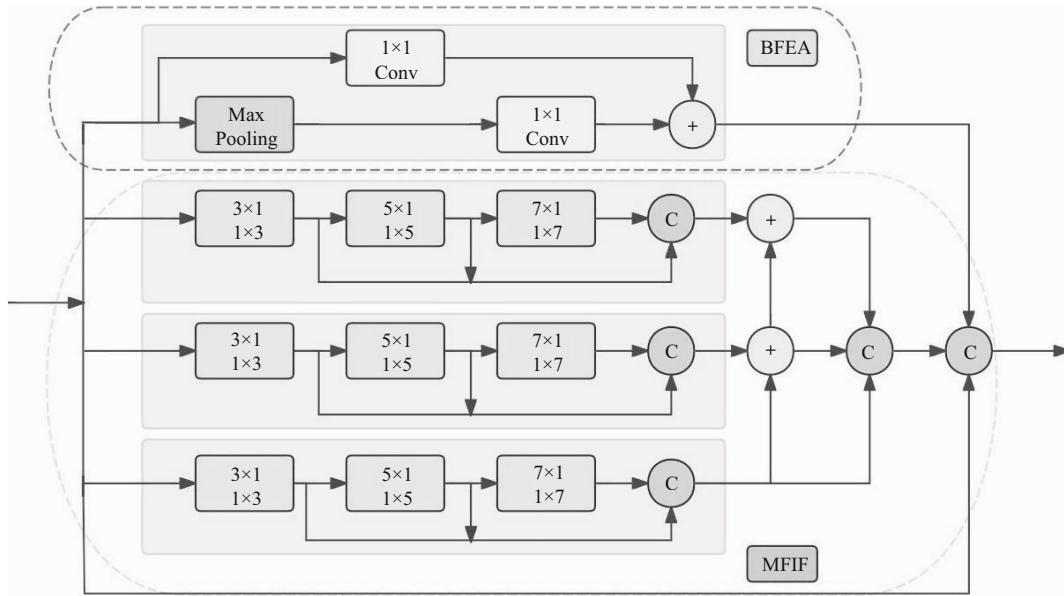


图4 DBMBF 模块结构

3 实验与结果

3.1 数据集与预处理

本节使用 3 个公开的青光眼眼底图像数据集来评估分割网络的性能。青光眼是一种严重的眼科疾病，会对负责视觉传递的视神经造成不可逆且逐渐加重的损害，最终导致视力丧失。一旦确诊，该疾病就无法被完全治愈。垂直杯盘比 (VCDR, vertical cup-to-disc ratio) 是青光眼眼底图像中最具代表性的指标之一。通常情况下，当 VCDR 超过 0.65^[1] 时，患者被视为确诊青光眼。因此，快速、准确地分割视杯 (OC, optic cup) 和视盘 (OD, optic disc) 以精确计算 VCDR 值，对于青光眼的诊断和管理具有重要意义。

第一个数据集是 Drishti-GS^[27]，包含 101 张眼底图像。第二个数据集是 RIM-ONE-r3^[28]，包含 159 张眼底图像。第三个数据集是 REFUGE^[29]，包含 1 200 张眼底图像，但由于设备和数据集标注原因在本文中仅使用了训练部分。各数据集具体信息如表 1 所示。

表 1 数据集具体信息

数据集	训练集图像/张	测试集图像/张	总数/张
Drishti-GS	50	51	101
RIM-ONE-r3	99	60	159
REFUGE	320	80	400

由于医学图像的高复杂性，需要提前进行预处理操作。裁剪医学图像可以初步划分目标区域，同时减少分割网络的工作量并最小化分割错误。鉴于所选公开医学数据集的规模较小，本文只能通过使用不同的数据增强方法来增加数据多样性。本文使用的数据增强方法包括弹性变换、对比度调整、翻转、随机缩放、随机擦除和添加噪声。同时，还使用了形态学操作，如腐蚀、中值滤波和孔填充，以优化输出图像，使医学图像的分割边界更加平滑和自然，从而提高模型的处理效率和准确性。其原理如式(3)所示。

$$X \ominus Y = \{ t | (Y) t \subseteq X t \} \quad (3)$$

其中， X 表示一幅二值图像， Y 表示结构元素， t 表示平移单位， \ominus 表示 X 通过 Y 进行腐蚀操作。

$$B(x,y) = \text{med} A(x - m,y - n), (m,n) \in w \quad (4)$$

其中， $A(x,y)$ 表示原始图像， $B(x,y)$ 表示处理后的图像， w 表示一个二维模板， med 为中位数函数。

$$Y_n = (Y_n - 1 \oplus C) \cap I^c, k = 1, 2, 3, \dots \quad (5)$$

其中， Y_n 表示第 k 次扩展的结果， C 表示结构元素， n 表示迭代次数， I^c 表示原始图像的补集。

3.2 训练参数

为了确保实验的公平性，本文中的所有实验均在相同硬件设备上训练，使用单个 NVIDIA RTX 3070 GPU。本文使用较大的批量大小 (设置为 12) 以充分利用 GPU 内存，其中学习率的衰减

策略如式(6)所示。

$$lr = (\text{base} - lr) \left(1 - \frac{\text{iter}}{\text{max} - \text{iter}} \right)^{\text{power}} \quad (6)$$

其中, 动量power设置为0.9, base-lr为 1×10^{-5} 。

同时, 本文使用Adam优化器进行网络模型训练, 动量设置为0.9。由于数据集大小的差异, 在训练较小的数据集(Drishti-GS和RIM-ONE-r3)时, 模型经过了800个周期才达到了损失函数的收敛。而对于较大的REFUGE数据集, 损失函数在训练了600个周期后便已收敛。

3.3 评价指标

本文中的所有网络分析和比较均基于一系列广泛使用的语义分割评估指标: Dice相似系数(DICE)、灵敏度(Sensitivity)、Jaccard系数和VCDR, 其定义分别为

$$\text{DICE} = \frac{2N_{\text{TP}}}{2N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}} \quad (7)$$

$$\text{Jaccard} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}} \quad (8)$$

$$\text{Sensitivity} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (9)$$

$$\text{VCDR} = \frac{\text{OC}}{\text{OD}} \quad (10)$$

其中, N_{TP} 、 N_{FP} 和 N_{FN} 分别表示正确识别的正样本

数、错误识别的正样本数和正确识别的负样本数。本文使用 δ 来估计 VCDR_p 和 VCDR_g 之间的绝对偏差, 其中下标 p 和 g 分别表示预测值和真实值。

$$\delta = \text{abs}(\text{VCDR}_p - \text{VCDR}_g) \quad (11)$$

3.4 实验结果

本文从定量和定性2个角度对实验结果进行了比较和分析。表2和表3将本文方法与一些经典方法和新方法在Drishti-GS和RIM-ONE-r3数据集上的结果进行了比较。部分方法无法公开获取, 仅能从原始论文中获取已知结果。对于公开可用的方法, 本文在相同的实验条件下进行了实验, 并在REFUGE数据集上将本文方法与其进行了比较, 并取得了优异的实验结果, 如表4所示。

在青光眼视杯分割中, 视杯附近的区域容易受到血管和黄斑等因素的影响, 这通常会导致分割效果较差。然而, 实验结果表明, 本文方法在实验结果中展现了显著的性能优势, 与现有主流方法相比, 本文方法在多个关键指标上性能表现突出。在视盘分割任务中, 本文方法的分割精度(如Dice相似系数、Jaccard系数)处于先进水平, 敏感度表现尤为优异。同时, 与大多数方法相比, 本文方法在眼底误差控制(δ)方面展现了更低的数值, 提升幅度为10%~20%。在视杯分割任务中, 本文方法在多个指标(包括Dice相似系数和灵敏度)

表2 不同方法在Drishti-GS数据集上的定量结果

方法	OD			OC			δ
	DICE	Jaccard	Sensitivity	DICE	Jaccard	Sensitivity	
文献[30]	0.967 8	0.938 6	0.971 1	0.861 8	0.773 0	0.882 2	0.092
文献[31]	0.963 8	0.930 1	0.948 8	0.879 3	0.784 6	0.876 5	—
文献[32]	0.948 7	0.904 9	0.919 8	0.829 0	0.725 4	0.823 9	0.081
文献[33]	0.955 9	0.917 7	0.936 7	0.863 5	0.773 0	0.866 9	0.060
文献[34]	0.967 5	0.937 8	0.971 0	0.864 5	0.777 7	0.886 4	0.064
文献[35]	0.949 0	0.904 2	0.926 8	0.828 2	0.711 3	0.741 3	—
文献[36]	0.964 2	0.932 3	0.975 9	0.881 8	0.800 6	0.881 9	0.076
文献[37]	0.962 0	0.885 0	—	—	0.815 0	—	—
文献[38]	0.861 4	0.838 5	—	0.808 7	0.763 3	—	—
文献[39]	0.943 0	0.893 0	—	0.889 0	0.801 0	—	—
文献[40]	0.960 6	—	—	0.888 1	—	—	—
本文方法	0.964 9	0.933 2	0.991 3	0.889 7	0.814 0	0.933 6	0.049

表3 不同方法在 RIM-ONE-r3 数据集上的定量结果

方法	OD			OC			δ
	DICE	Jaccard	Sensitivity	DICE	Jaccard	Sensitivity	
文献[30]	0.952 6	0.911 4	0.948 1	0.834 8	0.730 0	0.814 6	0.059
文献[31]	0.940 1	0.887 0	0.923 6	0.839 7	0.723 7	0.813 3	—
文献[32]	0.927 3	0.887 8	0.905 5	0.743 4	0.635 0	0.681 2	0.104
文献[33]	0.954 3	0.914 4	0.931 8	0.801 6	0.689 8	0.741 7	0.078
文献[34]	0.958 2	0.921 9	0.957 1	0.833 0	0.733 4	0.841 4	0.077
文献[35]	0.903 6	0.828 9	0.873 7	0.690 3	0.556 7	0.905 2	—
文献[36]	0.952 7	0.911 5	0.950 2	0.843 5	0.742 4	0.835 2	0.059
文献[37]	0.956 0	0.880 0	—	0.854 0	0.762 0	—	—
文献[38]	0.858 2	0.838 5	—	0.733 3	0.663 3	—	—
文献[39]	0.910 0	0.838 0	—	0.649 0	0.770 0	—	—
文献[40]	0.943 8	—	—	0.856 0	—	—	—
本文方法	0.957 9	0.922 2	0.963 7	0.863 0	0.771 2	0.890 2	0.055

表4 不同方法在 REFUGE 数据集上的定量结果

方法	OD			OC			δ
	DICE	Jaccard	Sensitivity	DICE	Jaccard	Sensitivity	
文献[32]	0.945 7	0.898 0	0.913 7	0.860 8	0.762 1	0.817 3	0.068
文献[33]	0.953 0	0.910 8	0.932 5	0.864 5	0.767 4	0.853 4	0.063
文献[34]	0.960 3	0.925 2	0.961 9	0.877 9	0.791 5	0.902 9	0.059
文献[41]	0.962 7	0.920 8	0.969 9	0.877 9	0.788 4	0.900 1	0.056
文献[42]	0.951 0	—	—	0.869 0	—	—	—
本文方法	0.961 7	0.926 8	0.961 5	0.882 3	0.795 4	0.924 3	0.052

上均优于主流模型，如文献[30,32,37]。尤其是在高难度的视杯边界分割任务中，本文方法表现出更强的细节捕捉能力，误差减少幅度接近30%~50%，展现了显著的精确性和鲁棒性。

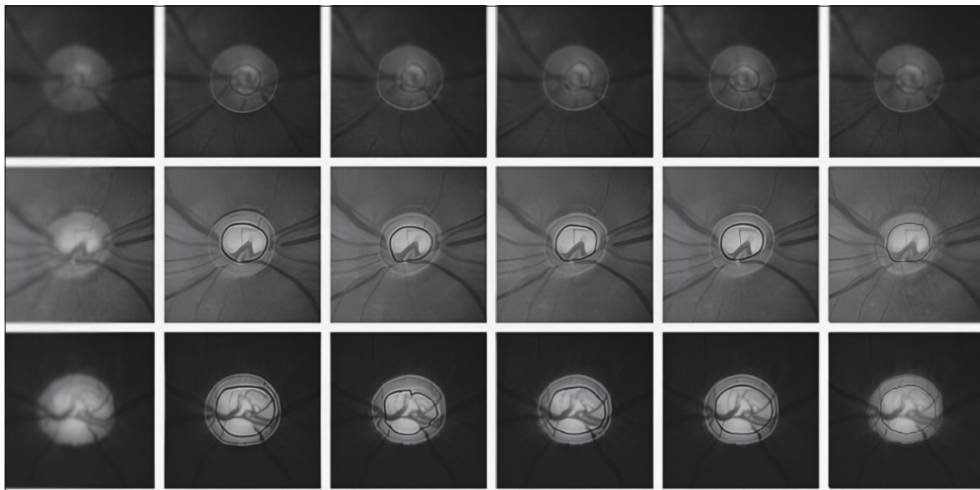
总体而言，本文方法在多个数据集上的实验结果验证了其在高分辨率图像分割任务中的竞争力。无论是分割精度还是误差控制，均优于或接近最先进的对比方法，表明其在复杂分割任务中具有较高的应用潜力和实际价值。与文献[6,32-33]方法进行比较的定性实验结果如图5~图7所示，图中外围大圈表示OD边界，内围小圈表示OC边界。

3.5 模型性能分析

本文通过 ROC (receiver operating characteris-

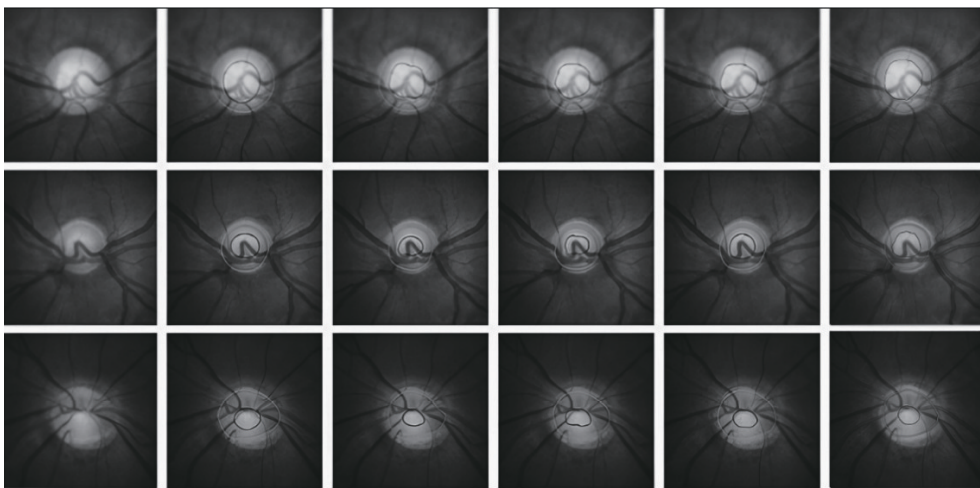
tic curve) 概率曲线^[43]和曲线下面积 (AUC, area under the curve) 验证了 VCDR 估计值与青光眼病例在相应数据集中的相关性。AUC 反映了模型区分类别的能力，AUC 值越高，表明算法性能越优秀。通过分析 ROC 概率曲线，本文评估了 MSFNet 的分割性能。图8展示了3个公开数据集的 ROC 概率曲线及对应的 AUC 值。

为深入探讨 AUC 与平均 VCDR 误差 δ 的相关性，图9展示了3个公开数据集的 AUC 与 δ 之间的双折线图。尽管 AUC 与 δ 并不总是完全相关，但数据表明，AUC 仍具有一定的指导意义。在青光眼研究领域，平均 VCDR 误差 δ 依然是青光眼诊断中的重要工具。



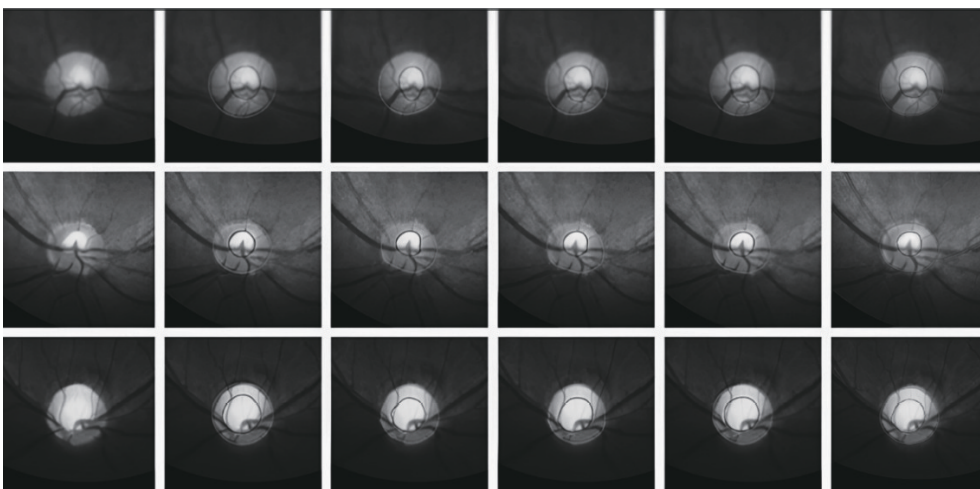
(a) 原始图像 (b) 手工真实图像 (c) 文献[32]方法 (d) 文献[33]方法 (e) 文献[6]方法 (f) 本文方法

图5 Drishti-GS 数据集中的视盘和视杯分割结果



(a) 原始图像 (b) 手工真实图像 (c) 文献[32]方法 (d) 文献[33]方法 (e) 文献[6]方法 (f) 本文方法

图6 RIM-ONE-r3 数据集中的视盘和视杯分割结果



(a) 原始图像 (b) 手工真实图像 (c) 文献[32]方法 (d) 文献[33]方法 (e) 文献[6]方法 (f) 本文方法

图7 REFUGE 数据集中的视盘和视杯分割结果

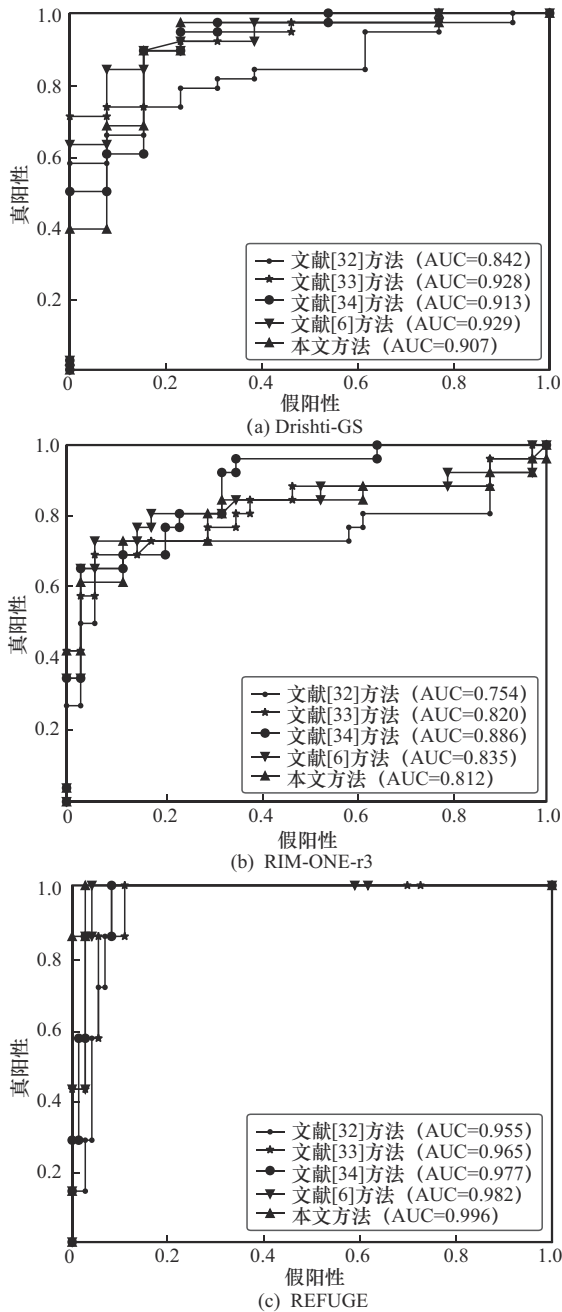


图8 3个公开数据集上的ROC概率曲线和AUC评分

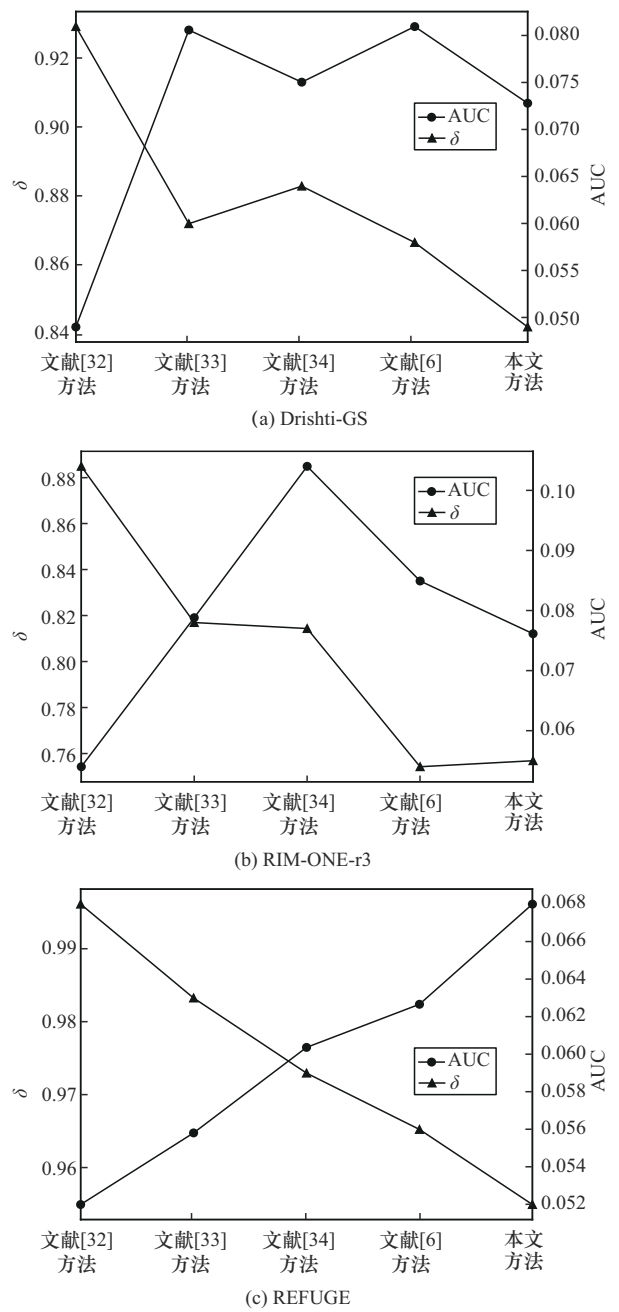


图9 3个公开数据集上AUC评分与平均VCDR误差 δ 的相关性

本文在 NVIDIA RTX 3070 GPU 设备上进行了实验，输入图像大小为 800 像素×800 像素。为了确保实验的严谨性，表 5 中的每种公开算法都在相同的实验环境下进行了测试。为展示本文模型的轻量化与高效性，本文将模型大小、计算复杂度、推理时间和内存使用情况等参数与一些先进方法进行了对比。实验结果表明，本文方法在大多数评价参数上优于其他方法，并且在视杯和视盘分割任务中保持了极强的竞争力。

此外，从表 5 可以看出，本文所提网络模型相比其他网络模型在模型大小上有显著优势，数值甚至小于 1×10^6 ，极其适用于计算资源有限的环境。在推理时间方面与文献[34]方法相当，仅为其他方法推理时间的一半左右，能够显著提升实时性要求较高的任务。模型在每秒浮点运算次数 (FLOPS) 参数上具有极强的竞争力，达到了最小的 FLOPS 值，仅为文献[14]方法的 66%。这些优势共同表明，本文所提网络模型在资源受限的应用场景中，具有较强的竞争力和实际应用潜力。

表5 不同模型参数性能对比

方法	内存占用($\times 10^6$)	参数($\times 10^6$)	乘加操作数($\times 10^9$)	浮点运算次数($\times 10^9$)	推理时间/ms
文献[14]	1.1	377	4.2	2.1	11
文献[15]	7.0	8 022	200.6	98	—
文献[32]	4.3	1 857	196.4	98.3	27
文献[33]	13.1	1 182	217	95.8	25
文献[34]	3.2	773	9.8	4.9	12
文献[44]	11.7	652	62.6	32.8	16
文献[45]	10.2	4 692	381.7	173.6	33
文献[46]	46.3	23 012.8	924	458.6	—
文献[47]	109.3	—	—	78.0	—
文献[41]	5.8	1 576	129	64.6	22
文献[48]	—	172.72	148.05	—	65
文献[42]	5.3	495.4	48.8	24.4	24
文献[49]	85	—	—	85.9	—
本文方法	0.6	267	2.8	1.4	12

4 消融实验

本文所提网络结构的核心模块包括MFIF模块和BFEA模块。为了验证这2个模块对网络分割性能的贡献,本文在RIM-ONE-r3数据集上进行了一系列消融实验,并以移除这2个模块后的网络结构作为基准网络(Baseline)。实验结果通过表6和图10进行了详细的定量和定性分析,旨在阐明MFIF和BFEA模块在提升网络分割性能方面的显著作用。

视杯分割面临的主要挑战之一是存在诸多干扰因素,如血管、黄斑等,这些因素显著增加了分割难度。为了有效分析本文所提模块的优势,通过收集并展示各模块中视杯的消融熵图,进一步揭示了在视杯分割任务中如何抑制这些干扰因素。图11展示了不同模型在视杯分割任务中的表

现。其中,Baseline+BFEA虽然能够突出分割图像的边界结构信息,但由于其简单的结构和单一的操作,未能有效抑制边界噪声,分割精度受限。相比之下,Baseline+MFIF能够有效降低分割边缘预测图的熵值,但仍存在一定的边界噪声。这表明MFIF模块在有效提升边界提取的精度方面具有一定优势。

当MFIF和BFEA模块结合使用时,结果显示这2个模块不仅能有效提取图像中的边界信息,还能显著抑制分割边界周围的干扰信息,从而大幅提升模型的分割性能。通过对比消融实验中的不同配置,本文进一步验证了MFIF和BFEA模块在图像分割中的有效性,并且证实了它们在处理图像信息时的优势。图10和图11通过选取2组图片从多个方向深入探讨了本文方法在图像分割任务中的应用,为其有效性和优势提供了充分的理论支持。

表6 消融实验对比

方法	OD			OC			δ
	DICE	Jaccard	Sensitivity	DICE	Jaccard	Sensitivity	
Baseline	0.924 5	0.866 2	0.990 8	0.822 9	0.716 4	0.914 0	0.070
Baseline+BFEA	0.929 5	0.874 6	0.989 2	0.829 4	0.726 6	0.920 4	0.062
Baseline+MFIF	0.957 4	0.919 8	0.974 4	0.848 1	0.749 6	0.881 0	0.059
本文方法	0.957 9	0.922 2	0.963 7	0.863 0	0.771 2	0.890 2	0.055

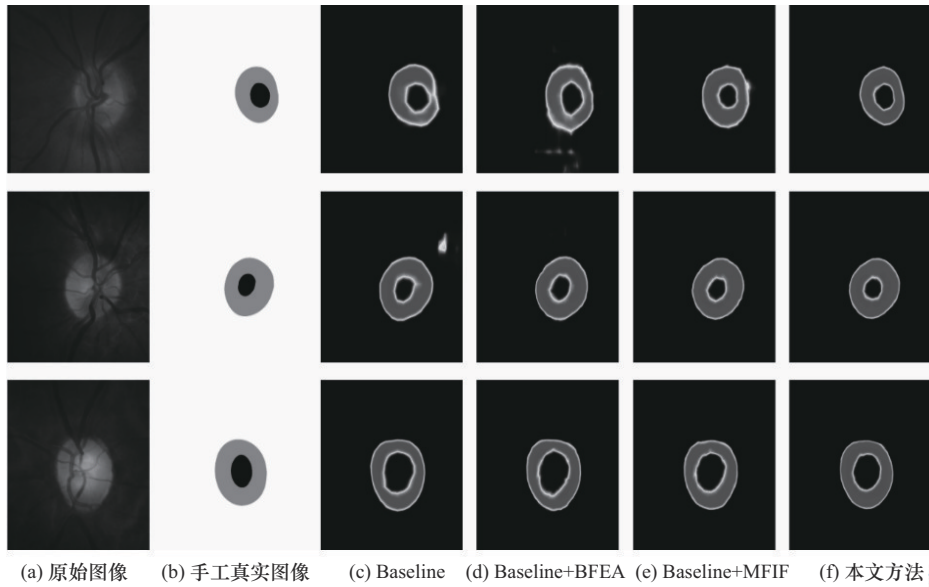


图 10 RIM-ONE-r3 数据集的消融实验中 OD 和 OC 的熵

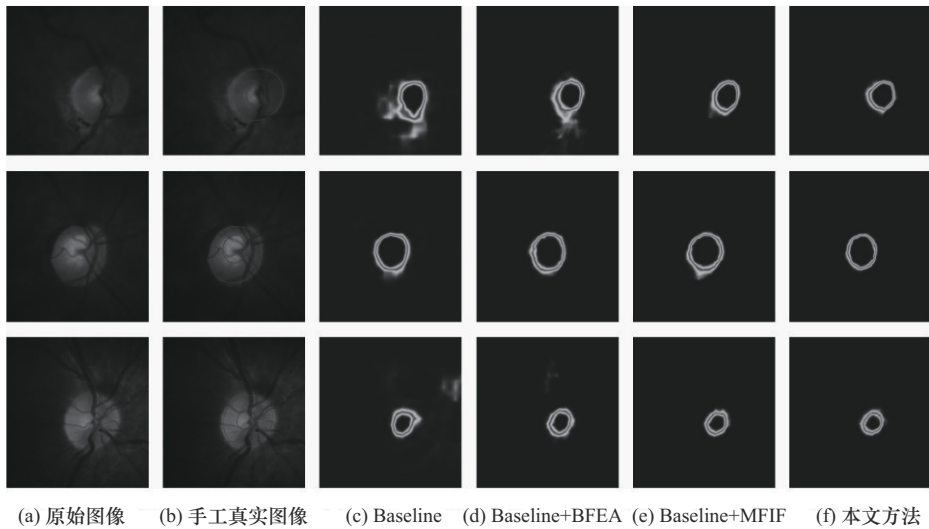


图 11 RIM-ONE-r3 数据集的消融实验中 OC 的熵

5 结束语

本文提出了一种新的轻量级实时图像分割网络 MSFNet。该网络通过引入补丁合并操作替代传统池化方法，避免了有效信息的丢失，并设计了一个 DBMBF 模块。DBMBF 模块融合了 MFIF 分支与 BFEA 分支，从而显著增强了模型对边界信息和语义信息的捕捉能力，进而提升了分割精度。实验结果表明，MSFNet 在 3 个公开数据集上的分割性能表现优异，诊断准确度分别为 90.69%、81.22% 和 99.61%，充分展示了其在图像分割任务中的强大能力。在消融实验中进一步的分析表明，

MFIF 模块通过促进多尺度特征间的交互与融合，显著改善了模型对复杂图像中细节信息的捕捉能力。BFEA 模块则通过突出边界结构，使模型能更加精准地聚焦于目标区域。综上所述，MSFNet 结合 MFIF 和 BFEA 模块，在性能上超越了传统方法，展现了更强的稳定性与适应性，验证了这些模块在图像分割中的重要价值。同时，该网络在模型参数量、计算复杂度和内存使用方面均优于大多数现有算法，充分体现了其轻量化和高效性的优势。尽管如此，图像分割在精度上仍存在挑战，未来研究将进一步致力于提高检测精度，解决这一技术难题。

参考文献:

- [1] AKRAM M U, TARIQ A, KHALID S, et al. Glaucoma detection using novel optic disc localization, hybrid feature set and classification techniques[J]. *Australasian Physical & Engineering Sciences in Medicine*, 2015, 38(4): 643-655.
- [2] SHEN D G, WU G R, SUK H I. Deep learning in medical image analysis[J]. *Annual Review of Biomedical Engineering*, 2017, 19: 221-248.
- [3] ZHU W, WU Y, WU Z, et al. Deep learning in autonomous driving: a review[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(6): 3426-3440.
- [4] LI X, LI S, ZHANG Y, et al. A review on image segmentation in remote sensing[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 178: 1-20.
- [5] MUNEEER F, IQBAL J, ZHANG L, et al. Thresholding-based image segmentation techniques: a review[J]. *Journal of Imaging*, 2019, 5(2): 25.
- [6] HUANG Y, ZHANG C, LIU S, et al. A review of region growing techniques for image segmentation[J]. *Multimedia Tools and Applications*, 2019, 78(10): 14409-14426.
- [7] XIE L, WANG J, YANG X, et al. A survey on edge detection algorithms and their applications in image segmentation[J]. *International Journal of Advanced Computer Science and Applications*, 2018, 9(6): 445-452.
- [8] YUAN Z W, ZHANG J. Feature extraction and image retrieval based on AlexNet[C]//*Proceedings of the Eighth International Conference on Digital Image Processing (ICDIP 2016)*. SPIE, 2016, 10033: 65-69.
- [9] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. *arXiv Preprint*, arXiv: 1412.7062, 2014.
- [11] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [12] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. *arXiv Preprint*, arXiv: 1706.05587, 2017.
- [13] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//*Computer Vision-ECCV 2018*. Berlin: Springer, 2018: 833-851.
- [14] POUDEL R P K, LIWICKI S, CIPOLLA R. Fast-SCNN: fast semantic segmentation network[J]. *arXiv Preprint*, arXiv: 1902.04502, 2019.
- [15] LEI H J, LIU W X, XIE H, et al. Unsupervised domain adaptation based image synthesis and feature alignment for joint optic disc and cup segmentation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(1): 90-102.
- [16] CELAYA A, ACTOR J A, MUTHUSIVARAJAN R, et al. PocketNet: a smaller neural network for medical image analysis[J]. *IEEE Transactions on Medical Imaging*, 2023, 42(4): 1172-1184.
- [17] ZHANG S, NIU Y M. LcmUNet: a lightweight network combining CNN and MLP for real-time medical image segmentation[J]. *Bioengineering*, 2023, 10(6): 712.
- [18] WU G J, ZHANG X, ZHANG Z M, et al. Learning lightweight neural networks via channel-split recurrent convolution[C]//*Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. Piscataway: IEEE Press, 2023: 3847-3857.
- [19] QI S M, LEE Z Q, LIU J L, et al. ResX: feature extraction block for medical image segmentation[J]. *IEEE Access*, 2024, 12: 28775-28783.
- [20] ZANG H Y, FOO S Y, BERNADIN S, et al. Facial emotion recognition using asymmetric pyramidal networks with gradient centralization[J]. *IEEE Access*, 2021, 9: 64487-64498.
- [21] LIU X W, YANG L, CHEN J G, et al. Region-to-boundary deep learning model with multi-scale feature fusion for medical image segmentation[J]. *Biomedical Signal Processing and Control*, 2022, 71: 103165.
- [22] FAN T L, WANG G L, LI Y, et al. A multi-scale information fusion level set for breast tumor segmentation[J]. *Journal of Medical Imaging and Health Informatics*, 2021, 11(8): 2124-2132.
- [23] AKRAM N, ADNAN S, ASIF M, et al. Exploiting the multiscale information fusion capabilities for aiding the leukemia diagnosis through white blood cells segmentation[J]. *IEEE Access*, 2022, 10: 48747-48760.
- [24] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//*Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE Press, 2018: 4510-4520.
- [25] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//*Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Piscataway: IEEE Press, 2021: 9992-10002.
- [26] DING X H, GUO Y C, DING G G, et al. ACNet: strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C]//*Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Piscataway: IEEE Press, 2019: 1911-1920.
- [27] SIVASWAMY J, KRISHNADAS S R, JOSHI G D, et al. Drishti-GS: retinal image dataset for optic nerve head (ONH) segmentation[C]//*Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. Piscataway: IEEE Press, 2014: 53-56.
- [28] FUMERO F, ALAYON S, SANCHEZ J L, et al. RIM-ONE: an open retinal image database for optic nerve evaluation[C]//*Proceedings of the 2011 24th International Symposium on Computer-Based Medical Systems (CBMS)*. Piscataway: IEEE Press, 2011: 1-6.
- [29] ORLANDO J I, FU H Z, BREDA J B, et al. REFUGE challenge: a unified framework for evaluating automated methods for glaucoma assessment from fundus photographs[J]. *Medical Image Analysis*, 2020, 59: 101570.
- [30] FU H Z, CHENG J, XU Y W, et al. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(7): 1597-1605.
- [31] ZHAO X Y, WANG S Y, ZHAO J, et al. Application of an attention U-Net incorporating transfer learning for optic disc and cup segmentation[J]. *Signal, Image and Video Processing*, 2021, 15(5): 913-921.
- [32] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//*Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015*. Berlin: Springer, 2015: 234-241.
- [33] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//*Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE Press, 2015: 3431-3440.
- [34] HOWARD A, SANDLER M, CHEN B, et al. Searching for mobileNetV3[C]//*Proceedings of the 2019 IEEE/CVF International Confer-*

ence on Computer Vision (ICCV). Piscataway: IEEE Press, 2019: 1314-1324.

- [35] AL-BANDER B, WILLIAMS B M, AL-NUAIMY W, et al. Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis[J]. Symmetry, 2018, 10(4): 87.
- [36] GU Z W, CHENG J, FU H Z, et al. CE-Net: context encoder network for 2D medical image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281-2292.
- [37] WANG J K, LI X, CHENG Y Z. Towards an extended efficientNet-based U-Net framework for joint optic disc and cup segmentation in the fundus image[J]. Biomedical Signal Processing and Control, 2023, 85: 104906.
- [38] JIANG L C, TANG X Y, YOU S, et al. BEAC-Net: boundary-enhanced adaptive context network for optic disk and optic cup segmentation[J]. Applied Sciences, 2023, 13(18): 10244.
- [39] TADISETTY S, CHODAVARAPU R, JIN R M, et al. Identifying the edges of the optic cup and the optic disc in glaucoma patients by segmentation[J]. Sensors, 2023, 23(10): 4668.
- [40] VARSHAN V Y S, KATTAMURI H G, SAHAYAM S, et al. Integrating edge information into ground truth for the segmentation of the optic disc and cup from fundus images[J]. arXiv Preprint, arXiv: 2408.05052, 2024.
- [41] LUO L, XUE D Y, PAN F, et al. Joint optic disc and optic cup segmentation based on boundary prior and adversarial learning[J]. International Journal of Computer Assisted Radiology and Surgery, 2021, 16(6): 905-914.
- [42] CHEN Y Q, LIU Z J, MENG Y J, et al. Lightweight optic disc and optic cup segmentation based on MobileNetv3 convolutional neural network[J]. Biomimetics, 2024, 9(10): 637.
- [43] FAWCETT T. An introduction to ROC analysis[J]. Pattern Recognition Letters, 2006, 27(8): 861-874.
- [44] WANG L B, LI R, ZHANG C, et al. UNetFormer: a UNet-like transformer for efficient semantic segmentation of remote sensing urban scene imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 190: 196-214.
- [45] LOU A G, GUAN S Y, LOEW M. DC-UNet: rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation[C]//Proceedings of the Medical Imaging 2021: Image Processing. SPIE, 2021, 11596: 758-768.
- [46] LIU B Y, PAN D R, SONG H. Joint optic disc and cup segmentation based on densely connected depthwise separable convolution deep network[J]. BMC Medical Imaging, 2021, 21(1): 14.
- [47] ZHENG Y Y, ZHANG X T, XU X Y, et al. Deep level set method for optic disc and cup segmentation on fundus images[J]. Biomedical Optics Express, 2021, 12(11): 6969-6983.
- [48] WU J D, FANG H H, SHANG F X, et al. SeATrans: learning segmentation-assisted diagnosis model via transformer[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2022: 677-687.
- [49] ZEDAN M J M, ABDANI S R, LEE J, et al. RMHA-Net: robust optic disc and optic cup segmentation based on residual multiscale feature extraction with hybrid attention networks[J]. IEEE Access, 2025, 13: 7715-7735.

[作者简介]



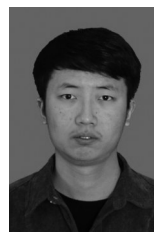
李建锋 (1979-), 男, 湖南张家界人, 博士, 吉首大学教授、硕士生导师, 主要研究方向为人工智能、数字图像处理、计算机视觉、医学图像处理。



熊明强 (2000-), 男, 湖南益阳人, 吉首大学硕士生, 主要研究方向为人工智能、医学图像处理。



陈园琼 (1985-), 女, 湖南张家界人, 吉首大学讲师, 主要研究方向为深度学习、图像处理、计算机视觉。



王宗达 (1998-), 男, 江苏连云港人, 吉首大学硕士生, 主要研究方向为人工智能、医学图像处理。



王涛 (2000-), 男, 湖南吉首人, 吉首大学硕士生, 主要研究方向为人工智能、深度学习。



孙培玮 (2001-), 男, 湖南吉首人, 吉首大学硕士生, 主要研究方向为人工智能、深度学习。