

# MRAU-Net: 基于改进 U-Net 和注意力机制的视网膜血管分割<sup>①</sup>



谢佳锋<sup>1</sup>, 廖光忠<sup>2</sup>

<sup>1</sup>(华中科技大学 计算机科学与技术学院, 武汉 430065)

<sup>2</sup>(华中科技大学 智能信息处理与实时工业系统湖北省重点实验室, 武汉 430065)

通信作者: 谢佳锋, E-mail: [x\\_feng0626@163.com](mailto:x_feng0626@163.com)

**摘要:** 由于眼底图像背景复杂、毛细血管细小且模糊以及噪声干扰等原因, 传统视网膜血管分割算法容易出现识别不准确和断连等问题. 针对这些问题, 提出一种基于改进 U-Net 和注意力机制的视网膜血管分割算法 (MRAU-Net). 为解决特征提取不充分问题, 设计了多尺度残差卷积模块 (MSRCB) 来代替 U-Net 传统的卷积块; 为了减少信息丢失和噪声干扰, 在瓶颈层嵌入双维注意力优化模块 (DAOM); 为了减少在编解码过程中造成的信息丢失, 构建了一种新的多尺度密集空洞卷积块 (MDCB), 并与传统的跳跃连接相结合. 在 DRIVE 和 CHASE\_DB1 两个公开数据集上进行了实验, *F1-score* 分别为 82.92%、83.75%, *AUC* 分别为 98.87%、98.96%, 灵敏度分别为 84.50%、83.82%, 准确率分别为 97.11%、97.63%. 实验结果表明 MRAU-Net 较现有优秀算法拥有更优异的性能表现.

**关键词:** 视网膜血管; 图像分割; 深度学习; U-Net; 注意力机制; 多尺度; 密集空洞卷积

引用格式: 谢佳锋, 廖光忠. MRAU-Net: 基于改进 U-Net 和注意力机制的视网膜血管分割. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/9868.html>

## MRAU-Net: Retinal Blood Vessel Segmentation Based on Improved U-Net and Attention Mechanism

XIE Jia-Feng<sup>1</sup>, LIAO Guang-Zhong<sup>2</sup>

<sup>1</sup>(School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430065, China)

<sup>2</sup>(Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan University of Science and Technology, Wuhan 430065, China)

**Abstract:** Due to the complex background of fundus images, thin and blurred capillaries, and noise interference, traditional retinal vessel segmentation algorithms often experience issues of inaccurate recognition and disconnections. To address these problems, a retinal blood vessel segmentation algorithm based on improved U-Net and attention mechanism (MRAU-Net) is proposed. To resolve the issue of insufficient feature extraction, a multi-scale residual convolution block (MSRCB) is designed to replace the traditional convolution blocks of U-Net. To reduce information loss and noise interference, a dual-dimensional attention optimization module (DAOM) is embedded in the bottleneck layer. To further mitigate information loss during the encoding-decoding process, a new multi-scale dense convolution block (MDCB) is constructed and combined with traditional skip connections. Experiments conducted on two public datasets of DRIVE and CHASE\_DB1 yield *F1-scores* of 82.92% and 83.75%, *AUCs* of 98.87% and 98.96%, sensitivities of 84.50% and 83.82%, and accuracies of 97.11% and 97.63%, respectively. These results show that MRAU-Net outperforms existing outstanding algorithms.

<sup>①</sup> 基金项目: 武汉市重点研发计划 (2022012202015070)

收稿时间: 2024-11-01; 修改时间: 2024-11-28; 采用时间: 2024-12-25; csa 在线出版时间: 2025-03-24

**Key words:** retinal vessel; image segmentation; deep learning; U-Net; attention mechanism; multi-scale; dense dilated convolution

视网膜血管的构造和形状对于眼科疾病的诊断与治疗具有极其重要的意义。例如, 糖尿病视网膜病变、青光眼和高血压性视网膜病变等眼科疾病均会呈现特定的血管异常特征。因此, 自动化的视网膜血管分割技术不仅能够协助医生快速而准确地检测这些病变, 还可以在大规模筛查中减少人力负担。然而, 视网膜图像中血管结构的复杂多变性以及病变区域对正常血管的干扰, 使得实现高精度血管分割的任务依然充满挑战。

近年来, 随着深度学习技术的兴起, 基于卷积神经网络 (CNN)<sup>[1]</sup> 的视网膜血管分割技术取得了显著进展。这些技术通过分析大量标注数据, 能够细致地捕捉血管的细微结构。其中, Ronneberger 等<sup>[2]</sup> 提出的 U-Net 模型因其出色的对称结构和跳跃连接的简单结构等有点而备受推崇。该模型通过跳跃连接有效地整合了不同分辨率的特征图, 实现了图像特征的融合。Jin 等<sup>[3]</sup> 提出的 DUNet 网络模型, 集成了可变形卷积网络, 以适应血管的不同尺寸和形状, 从而提高分割的准确性。Azad 等<sup>[4]</sup> 提出的 BCDU-Net 网络模型, 通过结合双向卷积 LSTM (BConvLSTM) 和密集连接的卷积网络来加强特征传播和鼓励特征重用, 提高了其模型的性能。Guo 等<sup>[5]</sup> 提出轻量化网络 SA-UNet, 通过在 U-Net 中引入注意力机制, 提高了模型的表示能力。Zhuang 等<sup>[6]</sup> 提出的 LadderNet 是一种基于多分支 U-Net 的卷积神经网络, 通过增加编码器-解码器分支的数量并在不同分支间引入跳跃连接, 克服了视网膜血管分割中信息流路径受限的问题。Gu 等<sup>[7]</sup> 提出的 CE-Net 通过密集空洞卷积 (DAC) 块和残差多核池化 (RMP) 块组成上下文提取器模块, 充分捕获更多高级信息, 很大程度解决了原始 U-Net 血管断连现象。

残差连接通常用于缓解深度神经网络中的梯度消失问题, 并促进信息的传递<sup>[8]</sup>。它让网络中的每个层可以直接学习输入与输出之间的残差, 从而使网络可以更容易地捕捉到复杂的特征。Diakogiannis 等<sup>[9]</sup> 提出的 ResUNet 通过引入残差连接, 使得梯度能够通过较短的路径直接传播, 这减少了训练过程中的梯度消失问题, 并提高了模型的分割精度。

基于以上研究, 本文对 U-Net 网络模型进行了改

进和优化, 提出了一种基于改进 U-Net 和注意力机制的视网膜血管分割算法 (multi-scale residual attention U-Net, MRAU-Net) 用于眼底视网膜血管分割。本文的贡献主要有以下几点。

一是设计了多尺度残差卷积模块 (multi-scale residual convolution block, MSRCB) 并替代 U-Net 编码器中原有的卷积块, 提高模型在不同尺度上特征信息的提取能力。

二是在瓶颈层嵌入双维注意力优化模块 (dual-dimensional attention optimization module, DAOM) 使得网络可以同时关注空间维度和通道维度的特征, 减少信息丢失和噪声干扰。

三是在跳跃连接处构建了一种新的多尺度密集空洞卷积块 (multi-scale dense dilated convolution block, MDCB), 以减少在编解码过程造成的信息损失, 更准确地分割细小血管。

## 1 网络模型

### 1.1 网络总体架构

本文提出的视网膜血管分割网络结构如图 1 所示。该网络参考了 U-Net 的典型结构, 并进行了针对性地改进与优化。首先, 采用本文提出的多尺度残差卷积模块来代替 U-Net 原有的卷积块, 以有效地提取多个尺度的特征信息, 进而解决 U-Net 在处理不同尺度特征时缺乏灵活性的问题。其次, 瓶颈层部分嵌入了双维注意力优化模块, 旨在空间和通道两个维度优化网络提取的特征信息, 这有效缓解了 U-Net 模型中存在的梯度消失问题, 从而提高了深层次特征提取的精度。最后, 为了降低 U-Net 在跳跃连接过程中的信息损失并提升特征融合的效果, 本文在跳跃连接处引入了多尺度密集空洞卷积模块, 以增强特征融合的能力, 弥补了 U-Net 模型在细小血管分割时信息流失和特征融合不足的缺点。

### 1.2 MSRCB

在视网膜血管分割领域中最经典的 U-Net 网络模型因其结构的简洁性和易于实现而受到青睐, 但随着分割任务的复杂度的不断提高, 对于模型的要求也

越来越高, 这导致传统的卷积模块在提取复杂眼底图像特征信息的方面的能力已经难以满足当下的需求. 因此本文设计了多尺度残差卷积模块 (MSRCB) 来代替 U-Net 中原有的卷积块, 旨在通过多尺度特征提取来增强网络的性能, 其基本结构如图 2 所示. 其核心结构包括卷积层、批量归一化 (batch normalization, BN)

层、修正线性单元 (rectified linear unit, ReLU) 激活层以及残差连接. 该模块的设计思想是通过连续应用不同尺寸 (1×1、3×3、5×5) 的卷积核并结合 BN 和 ReLU 对输入的眼底图像数据进行多尺度的特征提取. 并通过残差连接将输入特征信息直接添加到卷积层的输出, 实现特征信息的跳跃传递, 促进深层特征融合.

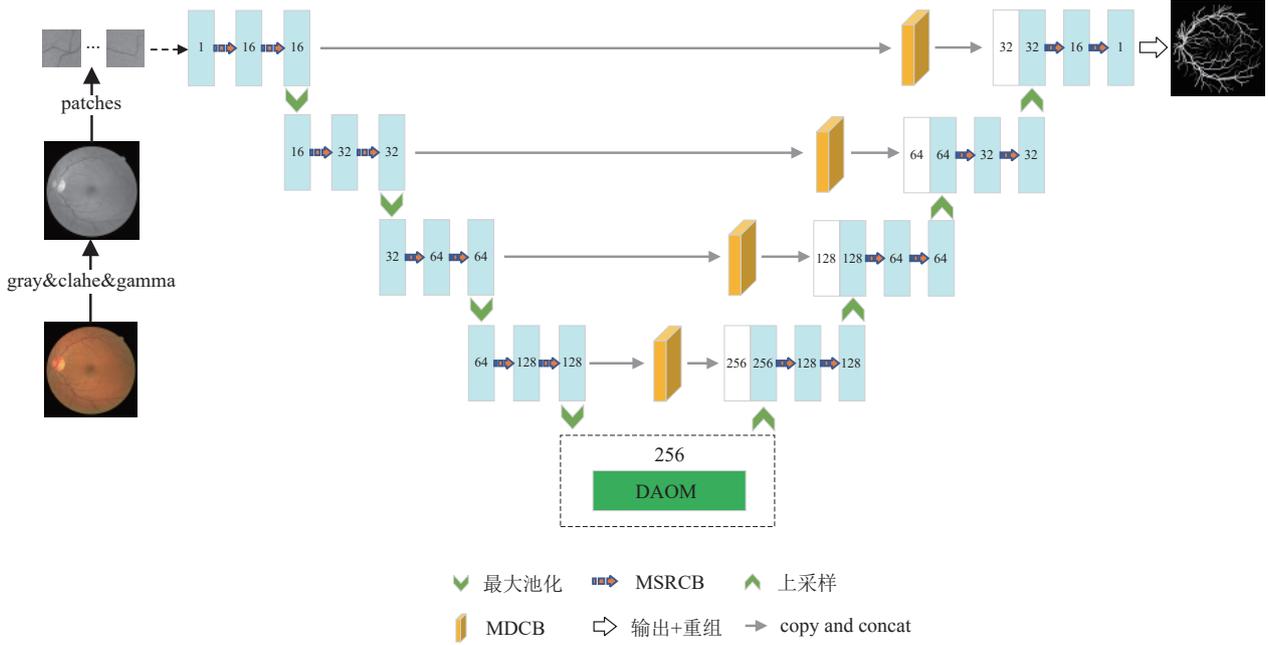


图 1 MRAU-Net

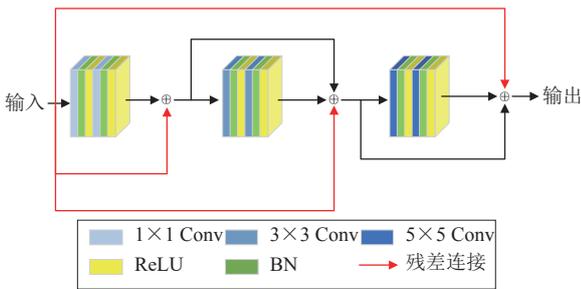


图 2 多尺度残差卷积模块

首先输入数据经过 1×1 卷积来改变通道数, 实现特征图的线性变换, 同时减少参数数量, 提高计算效率. 其次经过 3×3 卷积捕捉局部特征, 主要包括捕捉图像中的边缘和纹理信息. 最后经过 5×5 卷积捕捉更大范围的特征, 有助于提取更复杂的模式和结构信息. 在每个尺度的卷积后结合 BN 以减少内部协变量偏移, 加速训练过程, 同时提高模型的泛化能力, 以及结合 ReLU 激活函数引入非线性, 使得模型能够学习更复杂的特征.

假如  $I_{in}$  作为 MSRCB 模块的输入, 那么第 1 尺度

$S^{1 \times 1}$  的信息提取如式 (1) 所示. 其中  $f^{1 \times 1}$  是内核大小为 1×1 的深度可分离卷积,  $R$  是线性单元 (ReLU) 激活函数,  $\beta_n$  是批量归一化操作.

$$S^{1 \times 1} = R(\beta_n(f^{1 \times 1}(R(\beta_n(f^{1 \times 1}(I_{in})))))) + I_{in} \quad (1)$$

第 2 尺度  $S^{3 \times 3}$  提取的信息如式 (2) 所示. 其中  $f^{3 \times 3}$  是内核大小为 3×3 的深度可分离卷积.

$$S^{3 \times 3} = R(\beta_n(f^{3 \times 3}(R(\beta_n(f^{3 \times 3}(S^{1 \times 1})))))) + I_{in} \quad (2)$$

最后, 第 3 尺度  $I_{out}$  中提取的信息如式 (3) 所示. 其中  $f^{5 \times 5}$  是内核大小为 5×5 的深度可分离卷积.

$$I_{out} = R(\beta_n(f^{5 \times 5}(R(\beta_n(f^{5 \times 5}(S^{3 \times 3})))))) + I_{in} \quad (3)$$

### 1.3 DAOM

在视网膜血管图像分割任务中, 网络模型往往能够轻易地捕捉到粗大血管的特征, 但对细微血管的识别往往难以做到准确地识别. 即便在网络的较浅层次能够提取到细微血管的某些特征, 但这些特征也可能

在向深层次传递的过程中被噪声影响或者直接被丢失. 尤其在使用 U-Net 架构时, 噪声有可能在下采样的过程中逐渐扩散至网络的瓶颈层而影响模型最终的分割精度. 受到文献[10,11]的启发, 本文在 U-Net 的瓶颈层中嵌入了 DAOM 模块, 以缓解噪声干扰并提升细微特征的捕捉能力, 如图 3 的网络结构所示. 该模块的引入旨在增强网络对微小血管特征的表示力, 同时抑制噪声的不利影响, 以达到更优的分割效果.

该模块在特征提取时从空间和通道两个维度进行操作. 在空间维度上, 该模块专注于像素空间内垂直和水平方向的注意力权重编码, 采用双路径池化方法捕获更精细的空间关系建模. 首先对眼底图像中任意像素点

的空间位置关系进行编码, 将输入特征图  $F \in \mathbb{R}^{H \times W}$  送入空间维度优化块, 其中包含了两个并行路径. 一条路径执行垂直池化  $P_w$  操作, 目的是识别特征图中列像素间的相互关系, 输出为  $T_w \in \mathbb{R}^W$ ; 另一条路径执行水平池化  $P_h$  操作, 旨在识别特征图中行像素间的相互关系, 输出为  $T_h \in \mathbb{R}^H$ . 之后通过矩阵乘法将  $T_w$  与  $T_h$  相结合, 构建出粗空间注意力图  $A_s \in \mathbb{R}^{H \times W}$ , 使网络能够实时计算出整个特征图像的像素空间位置关系. 再将  $A_s$  通过 Sigmoid 函数归一化, 将每个像素空间位置权重调整到  $(0, 1)$  区间内, 从而得出精细化的注意力图  $A \in \mathbb{R}^{H \times W \times C}$ . 最终, 将优化后的注意力图  $A$  与原始输入特征图进行逐元素乘法操作, 得到经过空间注意力加权的特征图  $F_s = A \otimes F$ .

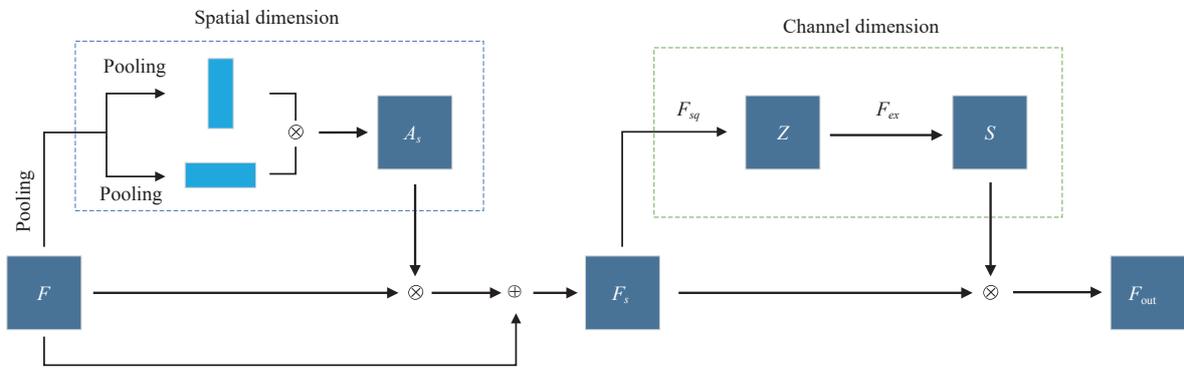


图 3 双维注意力优化模块

在通道维度上, DAOM 通过结合通道注意力机制<sup>[12]</sup>可以在原有的基础上增强通道间的特征表示能力, 更好地捕捉细微血管的分割. 因此接着对  $F_s$  进行压缩 (squeeze) 操作  $F_{sq}$  中的全局平均池化得到通道维度的全局特征  $Z \in \mathbb{R}^C$ .

$$Z = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_s(i, j, c) \quad (4)$$

其次, 执行激励 (excitation) 操作  $F_{ex}$ , 通过两个全连接层和 Sigmoid 激活函数来计算通道注意力权重  $S$ . 其中,  $W_1$  是降维矩阵,  $W_2$  是升维矩阵.

$$S = \sigma(W_2 \cdot R(W_1 \cdot Z)) \quad (5)$$

最后, 将经过空间注意力加权的特征图  $F_s$  根据通道注意力权重  $S$  进行加权处理, 从而得到最终经过空间和通道双维度注意力优化后的输出特征图  $F_{out}$ .

$$F_{out}(i, j, c) = S \cdot F_s(i, j, c) \quad (6)$$

#### 1.4 MDCB

由于传统 U-Net 的跳跃连接不充分, 导致低级和

高级特征难以充分融合, 并且容易丢失空间信息和缺乏对全局上下文信息的感知. 针对以上问题, 本文设计了 MDCB, 其基本结构如图 4 所示. 该模块首先对于编码器所输入的特征图  $F \in \mathbb{R}^{H \times W \times C}$  进行  $3 \times 3$  卷积再使用批量归一化 BN 和修正线性单元 ReLU 和来优化特征图, 接着对该特征图分别进行空洞率 (Rate) 为 1、2、4 的空洞卷积操作, 扩大模型感受野, 增强模型对长距离依赖的感知能力, 再将经过多尺度空洞卷积后的特征图进行加权求和得到  $F_m$ , 最后通过  $1 \times 1$  卷积降维, 改变特征图的深度, 而不改变其空间维度, 进而提升模型的表征能力. 此外, 通过引入密集连接机制, 以实现特征信息的高效复用和连续传递. 在密集连接中每层输入通道都与先前各层输出进行连接, 作为下一层的输入. 这样不仅增强了特征的融合能力, 还有助于缓解梯度消失问题, 从而提高了模型的整体性能. 最后将输入特征直接传递到更深层, 进一步增强了特征的融合与传递效率.

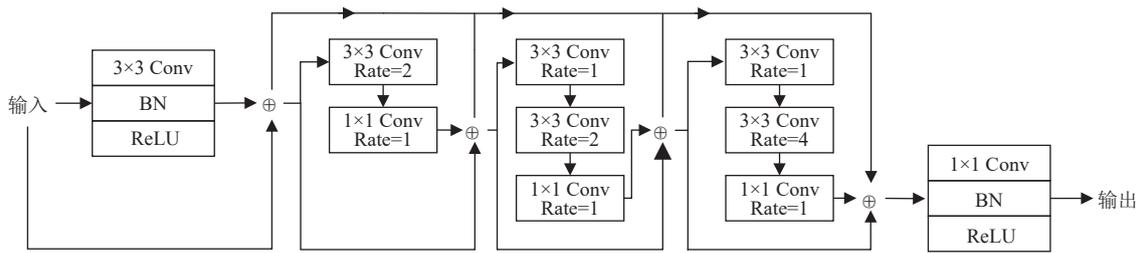


图4 多尺度密集空洞卷积块

## 2 实验结果与分析

### 2.1 实验设计

本文实验环境具体配置如下: 核心处理器为 Intel Core i7-12700H 2.30 GHz, 系统内存为 16 GB RAM, 显卡为 NVIDIA 3060, 操作系统为 Windows 10 操作系统. 深度学习框架为 PyTorch, CUDA 版本为 12.3, 优化器使用带有默认参数的 Adam, 初始学习率为 0.0005, 批量大小为 2, 并且每训练 20 次将学习效率衰减为原来的一半. 本文在 CHASE\_DB1 和 DRIVE 这两个具有代表性的公开数据集上进行了实验, 以验证本文方法的有效性, 并与一些表现优异的经典算法进行比较.

### 2.2 数据集

在视网膜血管分割研究中, DRIVE 和 CHASE\_DB1<sup>[13]</sup> 是两个常用且最具有代表性的公开数据集. DRIVE 数据集是基于荷兰糖尿病视网膜病变筛查项目的成果, 整合了 400 名参与筛查的糖尿病患者的医疗影像数据. 在这些数据中, 研究者随机抽取了 40 幅图像, 形成了包含 20 幅训练图像和 20 幅测试图像的 DRIVE 数据集, 这些图像的分辨率统一为 565×584 像素. 而 CHASE\_DB1 数据集则源自英格兰针对儿童进行的一项健康研究, 包含了儿童视网膜的扫描图像. 该数据集共收录了 28 幅彩色视网膜图像, 其中前 20 幅用于训练, 后 8 幅则用于测试. 这些图像的分辨率达到了 999×960 像素, 高于 DRIVE 数据集的分辨率.

### 2.3 数据预处理

由于公开的原始彩色眼底图像存在光照不均、对比度低以及噪声与伪影等问题, 因此为了提高算法的血管分割效果, 本文首先对这些图像做预处理, 主要包括: 灰度转换、对比度受限的自适应直方图均衡化 (contrast limited adaptive histogram equalization, CLAHE)<sup>[14]</sup>和伽马校正<sup>[15]</sup>. 并且由于公开数据集的图像数量较少, 因此为了让网络拥有更好地训练效果, 本文对使用的眼底图像数据集通过一些处理进行扩充, 主

要包括随机旋转、平移、水平和垂直翻转以及切片, 预处理后的眼底图像如图 5 所示, 部分切片图以及对应该真实标签切片图如图 6 所示.

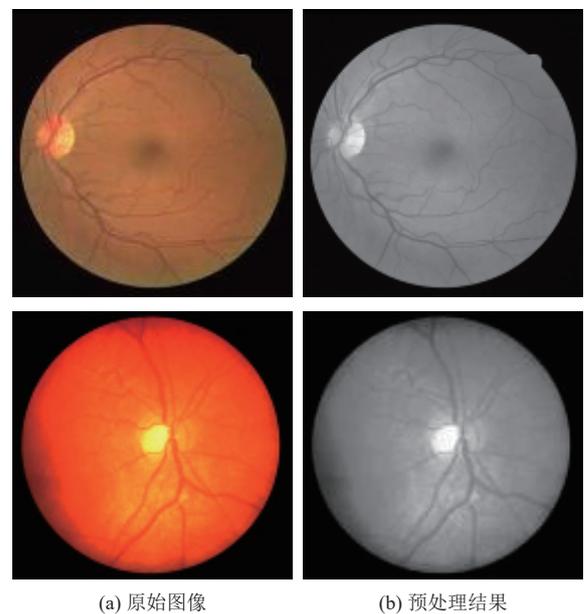


图5 预处理结果示例

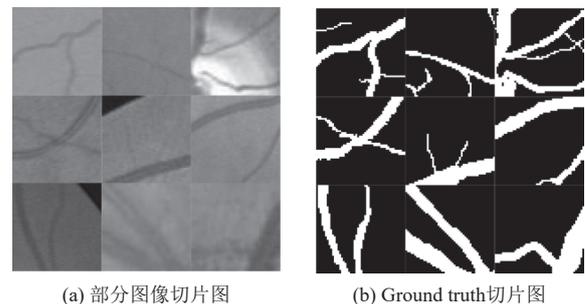


图6 部分图像切片示例

### 2.4 评价指标

本实验算法性能的评价指标为: 准确率 (accuracy, Acc)、灵敏度 (sensitivity, Se)、特异性 (specificity, Sp)、F1 值 (F1-score, F1) 以及曲线下面积 (area under curve, AUC), 以上指标计算公式如下:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$Se = \frac{TP}{TP+FN} \quad (8)$$

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

$$F1 = \frac{2TP}{2TP+FP+FN} \quad (10)$$

其中,  $TP$  表示真阳性;  $FP$  表示假阳性;  $TN$  表示真阴性;  $FN$  假阴性.  $Acc$  表示正确分类的像素占总像素的比例.  $Se$  表示正确分类血管像素的概率.  $Sp$  代表正确分类背景像素的概率.  $F1$  表示模型的预测结果与真实分割结果的相似程度.

## 2.5 损失函数

在视网膜血管分割任务中, 我们使用了二元交叉熵损失函数与 Dice 损失函数的结合来优化模型. 选择这两种损失函数的原因在于它们各自具有独特的优点, 交叉熵损失确保了整体分割的稳定性, 而 Dice 损失则促进了在血管区域的精确分割.

二元交叉熵损失函数是针对二分类问题的标准损失函数, 特别适合于像素级别的二值分割任务. 在视网膜血管分割任务中, 每个像素都被分类为前景或背景, 因此二元交叉熵损失函数可以有效度量每个像素的预测值与真实值之间的差异. 其定义为:

$$L_{bce} = - \sum_{i=1}^N [y_i \log(x_i) + (1 - y_i) \log(1 - x_i)] \quad (11)$$

其中,  $x$  表示模型的预测值,  $y$  表示真实标签.

Dice 损失函数特别适合于语义分割任务, 尤其是处理不平衡数据时. 由于视网膜血管在图像中占据的比例较小, 单独使用交叉熵损失可能导致对背景的偏重. 而 Dice 损失通过直接优化 Dice 相似系数, 使得模型能够更加准确地捕捉到血管的细节. 其定义为:

$$L_{dice} = 1 - \frac{2TP}{2TP+FP+FN} \quad (12)$$

为了同时利用交叉熵损失和 Dice 损失的优点, 我们将二者结合, 形成一个综合损失函数:

$$Loss = \alpha L_{bce} + \beta L_{dice} \quad (13)$$

其中,  $\alpha$  和  $\beta$  为权重参数, 值都在  $[0, 1]$  范围内, 用于控制两部分损失的相对重要性. 本研究经过大量实验证明, 采用上述混合损失函数来训练网络, 可以使网络达到最优分割效果.

## 2.6 消融实验

为了评估 MRAU-Net 中各模块的有效性, 本文分别在 DRIVE 和 CHASE\_DB1 这两个数据集上进行了消融实验. 在表 1 中, M1 表示原始 U-Net, M2 表示编解码器加入 MSRCB 的 U-Net, M3 表示瓶颈层嵌入 DAOM 的 U-Net, M4 表示跳跃连接部分加入 MDCB 的 U-Net, M5 表示编解码器加入 MSRCB 的同时在瓶颈层嵌入 DAOM 的 U-Net, M6 表示本文提出的 MRAU-Net 算法.

表 1 DRIVE 和 CHASE\_DB1 数据集上的消融实验结果

数据集	算法	Acc	Se	Sp	F1	AUC
DRIVE	M1	0.9559	0.8102	0.9733	0.8137	0.9775
	M2	0.9580	0.8158	0.9721	0.8189	0.9789
	M3	0.9641	0.8243	0.9832	0.8246	0.9810
	M4	0.9666	0.8331	0.9808	0.8249	0.9825
	M5	0.9682	0.8397	0.9836	0.8273	0.9862
	M6	<b>0.9711</b>	<b>0.8450</b>	<b>0.9851</b>	<b>0.8292</b>	<b>0.9887</b>
CHASE_DB1	M1	0.9563	0.8097	0.9679	0.8123	0.9783
	M2	0.9577	0.8139	0.9725	0.8140	0.9786
	M3	0.9586	0.8197	0.9745	0.8238	0.9789
	M4	0.9599	0.8195	0.9756	0.8245	0.9793
	M5	0.9681	0.8335	0.9771	0.8281	0.9813
	M6	<b>0.9763</b>	<b>0.8382</b>	<b>0.9798</b>	<b>0.8375</b>	<b>0.9896</b>

通过表 1 可以发现, M2 相较于 M1, 可以看出 MSRCB 可以从多个尺度提取输入的特征信息, 并且结合残差连接可以有效促进模型收敛, 从而使得模型在整体有略微提升. M3 相较于 M1, 可以看出 DAOM 模块在提取空间维度的基础上结合通道注意力机制可以充分提取空间和通道两个维度的特征信息, 以降低噪声干扰, 提升模型的分割精度, 因此使得模型在  $Sp$  和  $F1$  值都有较大的提升. M4、M5 相比于 M1, 在各个指标上都有一定程度的提升. 最后本文算法 M6 与基础的 U-Net 相比, 对于指标  $Acc$ 、 $Se$ 、 $Sp$ 、 $F1$  和 AUC 在 DRIVE 数据集上分别提升了 1.52%、3.48%、1.18%、1.55%、1.12%. 在 CHASE\_DB1 数据集上分别提升了 2.00%、2.85%、1.19%、2.52%、1.13%. 说明 MRAU-Net 可以更好地识别细小血管, 对视网膜血管的整体识别精度更高.

## 2.7 不同分割算法对比实验

为验证本文算法在视网膜血管分割任务中的有效性, 将本文方法与现有的一些经典血管分割算法分别在 DRIVE 和 CHASE\_DB1 两个数据集上进行比较实验, 表 2、表 3 为各网络的性能指标, 图 7、图 8 为各网络的分割结果. 为了便于观察到各网络分割结果的差异, 对图中部分区域进行了标注.

表2 在 DRIVE 数据集上的指标对比

算法	Acc	Se	Sp	F1	AUC
U-Net <sup>[2]</sup>	0.9559	0.8102	0.9733	0.8137	0.9775
AttU-Net <sup>[16]</sup>	0.9610	0.8016	0.9793	0.8095	0.9817
IterNet <sup>[17]</sup>	0.9573	0.7735	0.9838	0.8205	0.9816
LadderNet <sup>[6]</sup>	0.9561	0.7856	0.9810	0.8202	0.9793
DenseNet <sup>[18]</sup>	0.9622	0.8133	0.9793	0.8164	0.9815
SA-UNet <sup>[5]</sup>	0.9698	0.8212	0.9840	0.8263	0.9864
MC-UNet <sup>[19]</sup>	0.9678	0.8100	<b>0.9879</b>	0.8149	0.9828
本文方法	<b>0.9711</b>	<b>0.8450</b>	0.9851	<b>0.8292</b>	<b>0.9887</b>

表3 在 CHASE\_DB1 数据集上的指标对比

算法	Acc	Se	Sp	F1	AUC
U-Net <sup>[2]</sup>	0.9563	0.8097	0.9679	0.8123	0.9783
AttU-Net <sup>[16]</sup>	0.9694	0.8047	0.9818	0.8092	0.9813
IterNet <sup>[17]</sup>	0.9655	0.7970	0.9823	0.8073	0.9851
LadderNet <sup>[6]</sup>	0.9656	0.7978	0.9818	0.8031	0.9839
DenseNet <sup>[18]</sup>	0.9539	0.8155	0.9711	0.7959	0.9766
SA-UNet <sup>[5]</sup>	0.9755	0.8151	<b>0.9835</b>	0.8153	0.9778
MC-UNet <sup>[19]</sup>	0.9714	0.8366	0.9829	0.7741	0.9818
本文方法	<b>0.9763</b>	<b>0.8382</b>	0.9798	<b>0.8375</b>	<b>0.9896</b>

由表2、表3可见,在DRIVE数据集上,虽然MC-UNet在特异性上提升了0.28%,但其余指标均低于MRAU-Net.这表明MC-UNet在处理背景噪声和非血管区域时表现有所优化,但在精确捕捉血管特征方面仍存在不足.相较之下,MRAU-Net在准确率、灵敏度、F1值和AUC关键指标上表现优异,表明其能够精准识别和分割细小血管,同时有效降低背景像素被误分类为血管的概率,并减少了血管断裂现象.另外,MRAU-Net的灵敏度较SA-UNet提高了2.38%,在准确率上也略优于SA-UNet,这进一步说明MRAU-Net在提升前景识别能力、减少假阴性的同时,成功避免了更多背景像素被误分类为血管(假阳性),拥有更好的综合分割性能.在CHASE\_DB1数据集上,MRAU-Net虽在特异性上略低于SA-UNet,但在准确率、灵敏度、F1值和AUC上均取得了最优的效果,尤其是F1值相较于SA-UNet高出了2.22%,说明MRAU-Net能够更精确地提取细小血管并减少血管的漏判和误判,具有较高的精准率和鲁棒性,综合分割性能更优.

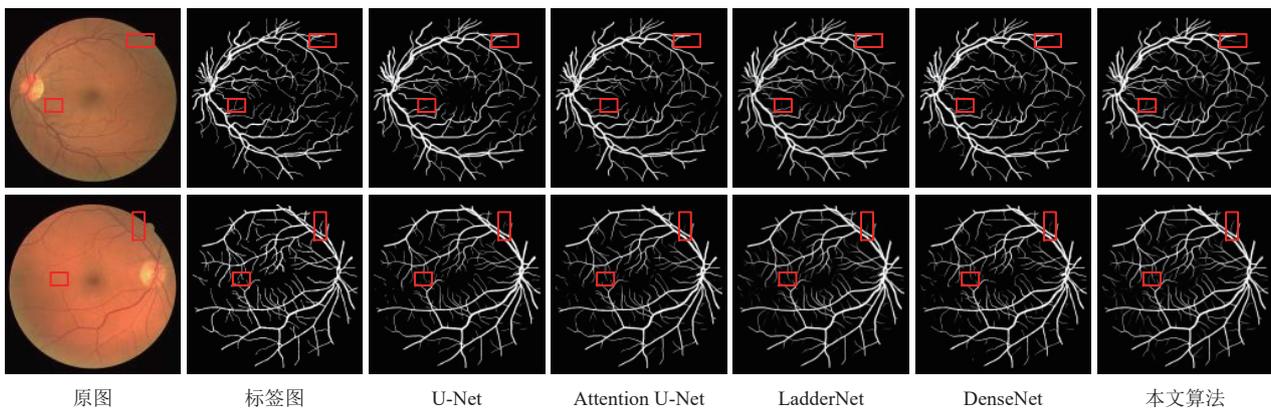


图7 在DRIVE数据集上不同算法的分割结果

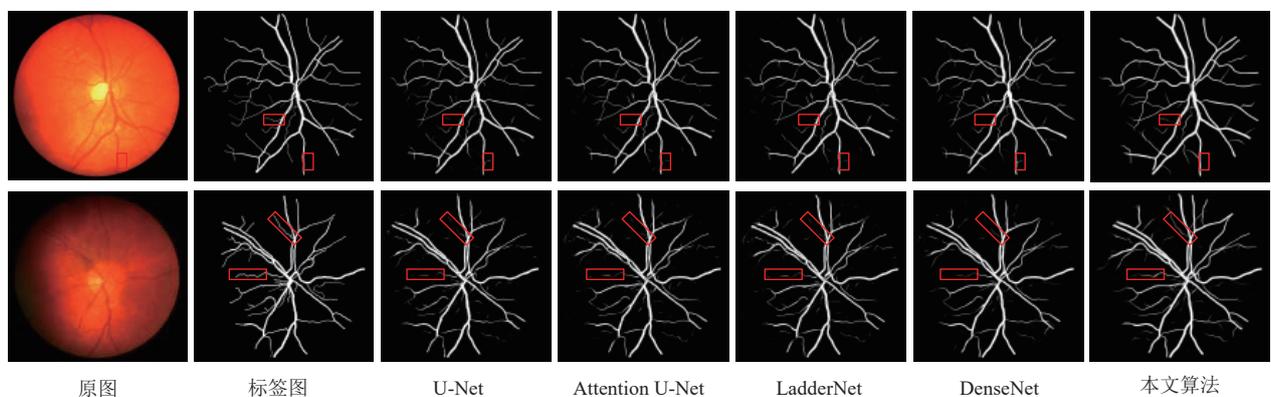


图8 在CHASE\_DB1数据集上不同算法的分割结果

通过对图7和图8中分割结果的比较可以看出,本文所提出的算法在DRIVE和CHASE\_DB1这两个数据集上,解决了其他算法在捕捉微小血管结构时常出现的模糊和遗漏问题,能够更有效地识别细小血管并在清晰度上表现更加突出.另外针对其他算法中血管断连问题,本文算法展现了更优异的细节提取能力,有效改善了其他算法对血管不连续、识别不准确的不足.

### 3 结论与展望

本文提出一种基于改进U-Net和注意力机制的视网膜血管分割算法(MRAU-Net)的视网膜血管分割方法,在该模型中,首先使用多尺度残差卷积模块(MSRCB)代替传统的卷积块,通过应用连续不同大小的卷积核并结合残差连接提取多尺度的特征信息;在瓶颈层嵌入双维注意力优化模块(DAOM),通过空间和通道两个维度对血管信息进行深入挖掘,减少了信息丢失和噪声干扰;最后在跳跃连接处构建了一种新的多尺度密集空洞卷积块(MDCB)以减少编解码过程中造成的信息损失.实验结果表明,相较于其他网络算法,本文算法能够生成更加平滑连续的血管边界,并且在细小血管的识别上性能更优.但是算法仍然存在一定的误分割问题,未来我们将继续优化算法以提高其分割性能.

### 参考文献

- 1 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84–90. [doi: [10.1145/3065386](https://doi.org/10.1145/3065386)]
- 2 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. *Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention*. Munich: Springer, 2015. 234–241.
- 3 Jin QG, Meng ZP, Pham TD, *et al.* DUNet: A deformable network for retinal vessel segmentation. *Knowledge-based Systems*, 2019, 178: 149–162. [doi: [10.1016/j.knosys.2019.04.025](https://doi.org/10.1016/j.knosys.2019.04.025)]
- 4 Azad R, Asadi-Aghbolaghi M, Fathy M, *et al.* Bi-directional ConvLSTM U-Net with densely connected convolutions. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop*. Seoul: IEEE, 2019. 406–415.
- 5 Guo CL, Szemenyei M, Yi YG, *et al.* SA-UNet: Spatial attention U-Net for retinal vessel segmentation. *Proceedings of the 25th International Conference on Pattern Recognition (ICPR)*. Milan: IEEE, 2021. 1236–1242.
- 6 Zhuang JT. LadderNet: Multi-path networks based on U-Net for medical image segmentation. *arXiv:1810.07810*, 2018.
- 7 Gu ZW, Cheng J, Fu HZ, *et al.* CE-Net: Context encoder network for 2D medical image segmentation. *IEEE Transactions on Medical Imaging*, 2019, 38(10): 2281–2292. [doi: [10.1109/TMI.2019.2903562](https://doi.org/10.1109/TMI.2019.2903562)]
- 8 Ma WA, Yu S, Ma K, *et al.* Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification. *Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention*. Shenzhen: Springer, 2019. 769–778.
- 9 Diakogiannis FI, Waldner F, Caccetta P, *et al.* ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 162: 94–114. [doi: [10.1016/j.isprsjprs.2020.01.013](https://doi.org/10.1016/j.isprsjprs.2020.01.013)]
- 10 Li KQ, Qi XQ, Luo YW, *et al.* Accurate retinal vessel segmentation in color fundus images via fully attention-based networks. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(6): 2071–2081. [doi: [10.1109/JBHI.2020.3028180](https://doi.org/10.1109/JBHI.2020.3028180)]
- 11 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018. 3–19.
- 12 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 7132–7141.
- 13 Siddique N, Paheding S, Elkin CP, *et al.* U-Net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, 2021, 9: 82031–82057. [doi: [10.1109/ACCESS.2021.3086020](https://doi.org/10.1109/ACCESS.2021.3086020)]
- 14 路皓翔, 刘振丙, 郭棚跃, 等. 多尺度卷积结合自适应双区间均衡化的图像增强. *光子学报*, 2020, 49(10): 1010002.
- 15 马鑫, 喻春雨, 陈刚, 等. 分区自适应伽马校正的非均匀光照图像增强. *激光与光电子学进展*, 2024, 61(14): 1437006.
- 16 Oktay O, Schlemper J, Le Folgoc L, *et al.* Attention U-Net: Learning where to look for the pancreas. *arXiv:1804.03999*, 2018.
- 17 Li LZ, Verma M, Nakashima Y, *et al.* IterNet: Retinal image segmentation utilizing structural redundancy in vessel networks. *Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision*. Snowmass: IEEE, 2020. 3645–3654.
- 18 Aldoj N, Biavati F, Michallek F, *et al.* Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-Net. *Scientific Reports*, 2020, 10(1): 14315. [doi: [10.1038/s41598-020-71080-0](https://doi.org/10.1038/s41598-020-71080-0)]
- 19 Zhang T, Li J, Zhao Y, *et al.* MC-UNet multi-module concatenation based on U-shape network for retinal blood vessels segmentation. *arXiv:2204.03213*, 2022.

(校对责编: 张重毅)