

## CAMU-Net: 基于 Attention U-Net 的视网膜血管分割改进模型

唐云飞<sup>1,2</sup>, 但志平<sup>1,2</sup>, 洪郑天<sup>1,2</sup>, 陈永麟<sup>1</sup>, 程沛霖<sup>3</sup>, 成果<sup>4</sup>, 刘芳婷<sup>5</sup>

1. 三峡大学计算机与信息学院水电工程智能视觉监测湖北省重点实验室, 湖北 宜昌 443002; 2. 三峡大学计算机与信息学院智慧医疗宜昌市重点实验室, 湖北 宜昌 430002; 3. 华中科技大学同济医学院附属同济医院, 湖北 武汉 430030; 4. 清华大学深圳国际研究生院, 广东 深圳 518055; 5. 红安县人民医院, 湖北 黄冈 438400

**【摘要】**提出一种改进的 U-Net 模型 (CAMU-Net), 以达到精准分割视网膜血管的目的。CAMU-Net 模型通过添加残差增强模块来提取区域特征中的重要信息, 增强模型对区域特征的了解; 通过添加特征细化模块来促进特征的提取, 提高新模型的全局特征收集能力; 通过添加通道注意力机制模块来捕捉图像特征, 精确分割结果; 通过引入多尺度特征融合结构来提升模型感知目标边界等细节的能力。在 DRIVE 数据集上进行消融实验, 得出各模块的实际效果, 验证各模块对于本模型视网膜血管分割各方面提升的作用; 在 DRIVE 和 STARE 数据集上和其他主流网络模型进行对比分析, 结果表明 CAMU-Net 模型优于其他模型。

**【关键词】**视网膜血管; 图像分割; 深度学习; CAMU-Net; 注意力机制

**【中图分类号】**R318

**【文献标志码】**A

**【文章编号】**1005-202X(2024)08-0960-09

## CAMU-Net: an improved model for retinal vessel segmentation based on Attention U-Net

TANG Yunfei<sup>1,2</sup>, DAN Zhiping<sup>1,2</sup>, HONG Zhengtian<sup>1,2</sup>, CHEN Yonglin<sup>1</sup>, CHENG Peilin<sup>3</sup>, CHENG Guo<sup>4</sup>, LIU Fangting<sup>5</sup>

1. Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China; 2. Yichang Key Laboratory of Intelligent Medicine, College of Computer and Information Technology, China Three Gorges University, Yichang 430002, China; 3. Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China; 4. Tsinghua Shenzhen International Graduate School, Shenzhen 518055, China; 5. Hong'an County People's Hospital, Huanggang 438400, China

**Abstract:** An improved U-Net model (channel attention module U-Net, CAMU-Net) is proposed to achieve precise segmentation of retinal vessels. CAMU-Net model enhances its understanding of regional features by employing residual enhancement convolution to extract important information from the regions, improves the global feature acquisition capability by introducing feature refinement module to promote feature extraction, realizes precise segmentation by adding channel attention module to capture image features accurately, and enhances its capability to perceive target boundaries and details through a multi-scale feature fusion structure. The ablation study on the DRIVE dataset validates the role of each module in retinal vessel segmentation. The comparison with other mainstream network models on DRIVE and STARE datasets verify that CAMU-Net model is superior to other models.

**Keywords:** retinal vessel; image segmentation; deep learning; CAMU-Net; attention mechanism

### 前言

作为眼底图像中可以检测到的主要解剖结构, 视网膜血管的形态信息能反映出很多病症, 如糖尿

病、白血病和高血压都容易引起视网膜的特殊病变<sup>[1-2]</sup>。视网膜血管的形态结构对各类眼科和心血管疾病的评判具有重要作用, 医生能通过患者眼底图像的改变分析视网膜血管的形态变化, 从而进行诊断, 以便后续的治疗<sup>[3-4]</sup>。由于眼底图像自身存在噪声, 血管背景像素比例失衡, 图像中的视网膜血管分布密集且毫无规律, 具有高曲率和多尺度的特点, 极难获取形态信息, 通过医疗人员手动分割视网膜血管不仅需要花费极长的时间, 还难以形成统一的标准, 误差极大<sup>[5-6]</sup>。随着计算机技术研究不断深入, 国

**【收稿日期】**2024-04-25

**【基金项目】**国家自然科学基金(U1401252)

**【作者简介】**唐云飞, 硕士研究生, 研究方向: 智慧医疗, E-mail: 1357196465@qq.com

**【通信作者】**但志平, 博士, 教授, 研究方向: 计算机视觉、自然语言处理, E-mail: zp\_dan@ctgu.edu.cn

内外学者就视网膜血管分割提出很多的方法<sup>[7]</sup>。与传统方法相比,全卷积网络(Fully Convolutional Network, FCN)通过像素级的分类在语义级别上解决图像的分割问题,使得模型由于使用像素块导致重复计算和存储的情况大幅减少<sup>[8]</sup>;但是,FCN相对于其他模型在训练方面更为繁琐,训练结果不尽人意,而且没有关注到像素之间的关系,使得眼底图像缺乏空间一致性<sup>[9]</sup>。U-Net模型具有两个特点,一是捕捉上下文信息的收缩路径,二是允许精确定位的对称扩展路径,允许网络向更高层分辨率传播上下文信息,在医学图像分割领域具有显著优势<sup>[10]</sup>。U-Net模型在进行分割时,有时会出现一些边界模糊或分割错误的情况,在复杂的图像中准确率不高。注意力机制能将模型赋予人类注意力的特性,使得模型对任务的关键信息进行锁定并加强关注,且兼容性很强。将注意力机制与U-Net模型进行合理的结合既可以满足模型的高性能,也可以将模型有效地应用在医学图像分割中<sup>[11-12]</sup>。Oktay等<sup>[13]</sup>在U-Net模型中添加注意力机制,使改进后的模型可在对显著性区域的关注以及对无关背景区域的抑制方面达到比较好的效果。由于视网膜血管分布十分密集且毫无规律,Attention U-Net很难对细小的血管进行分割处理,也难以解决血管边界的不确定性问题<sup>[14]</sup>。

残差连接是一种能解决梯度消失和信息传递问题的技术,常被用于深度神经网络中<sup>[15-16]</sup>,它可将引入跨层的直接连接作用于不同层的神经网络中,达到在跳跃连接的协助下将网络输出传递到后续层的效果,以便模型能保存更多的原始信息和梯度。Diakogiannis等<sup>[17]</sup>通过加入残差连接,使梯度可以直接流经较短的路径,减少训练过程中的梯度消失问题,更好地捕捉图像的细节和上下文信息,提高分割的准确性;但ResU-Net会造成上下文信息的丢失,影响血管的拓扑结构。

Transformer能通过自注意力机制来进行学习,具有自我感知上下文的能力,相较循环神经网络(Recurrent Neural Network, RNN)和卷积神经网络(Convolutional Neural Network, CNN)具有高并行性和高计算效率的特点。Chen等<sup>[18]</sup>在U-Net模型的基础上引入混合编码器,将CNN和Transformer结合起来,使得TransU-Net模型能更好地处理长距离依赖性,捕获图像中的语义信息,提高模型的代表能力和泛化性能<sup>[19-21]</sup>;但是TransU-Net在特征提取和上采样阶段卷积中的感受野太小。

针对上述模型在视网膜血管分割方面存在的一些问题,本研究根据视网膜血管边界模糊、分布散乱、具有大量细小血管和对比度低等特点<sup>[22-23]</sup>,提出

一种高效的全卷积神经网络架构(Channel Attention Module U-Net, CAMU-Net),以达到显著提升模型分割精度的目的<sup>[24]</sup>。本模型受U-Net结构和残差式连接的启发,在U-Net结构的基础上,使用改进的残差增强模块(Residual Enhancement Convolution Module, REConv)替代原始编码器模块来进行主干特征提取,通过增加通道和空间关注来提取区域特征中的必要信息,从而增强对区域特征的了解和动态场景感知的效果。此外,本模型使用特征细化模块(Feature Refinement Module, FRM)来提高模型的上下文特征信息提取效率,提高分割的准确性<sup>[25-26]</sup>;使用通道注意力机制模块(Channel Attention Module, CAM)来提高图像分割任务的精准度并引入多尺度特征融合来加强对全局信息的关注。

## 1 相关工作

### 1.1 网络总体架构

本研究提出的CAMU-Net结构见图1,该网络基于编码器-解码器-跳跃连接的U-Net结构,在编码器部分使用改进的REConv进行主干特征提取。待处理的图像被送入编码器中,通过REConv模块进行优化区域特征提取、动态场景信息表示和特征信息分配,利用残差式结构传递关键特征,增加非线性和复杂特征层,借助池化和卷积操作实现特征增强和信息流组合。为解决解码器中普通的卷积容易丢失上下文特征信息,影响分割准确性的缺点,利用FRM将3×3卷积的结果和初始值与批量归一化(Batch Normalization, BN)层结果进行通道融合,通过构建通道间相互依赖,选择高级特征信息并优化排序,以增强关键特征,利用特征细化块和多级连结结构增强语义表达与全局信息收集能力。为提高模型的分割精度,提出CAM,在特征处理中结合可形变卷积、GELU激活函数及BN操作,通过坐标信息理解空间结构,实现注意力权重计算和Softmax归一化,提高模型通用性,特别在处理具有明确空间关系的数据中发挥优势。为解决模型全局信息关注不足的问题,引入多尺度特征融合模块,通过编码器层级输出融合增强空间信息,在不同尺度下利用信息提高分割模型对细节感知能力,促进图像内容理解与准确性提升。本研究通过CAMU-Net完成眼底视网膜血管分割任务并在精确标注的眼底视网膜血管数据集下进行实验,取得很好的分割效果。

### 1.2 REConv

图像中不同区域对分割任务的贡献并不同等重要,所以本研究提出REConv(图2),使用池化后的特征进行进一步卷积,让网络专注于颜色较差的处理

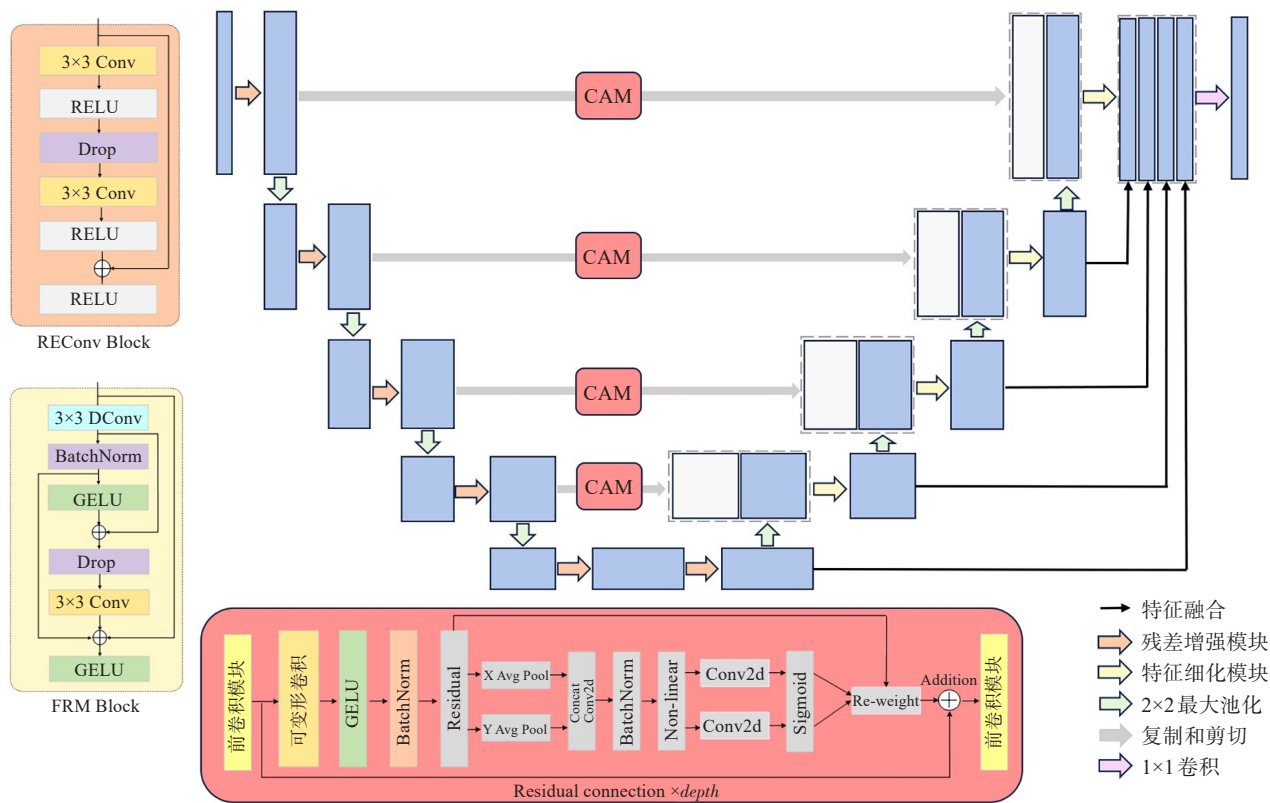


图1 CAMU-Net网络结构  
Figure 1 CAMU-Net architecture

区域和关键部分,增加通道和空间关注,可以从区域特征中提取必要信息,更好地理解空间区域特征,将动态场景信息表示为视觉显著性,并结合通道注意力机制进行详细的特征信息分配和处理,将图像纹理解释为空间域的上下文信息。REConv的贡献是双重的:(1)为待分割的图像提供更准确的特征提取能力,使更多的空间信息作用于图像上;(2)使用一种残差式的结构来抑制当前信息量较少的特征,并将更多有价值的特征传递到网络的下一层,这对图像分割非常有价值。一个简单的3×3卷积被用来增加模型的非线性并生成更复杂的特征层。在特征相关性的作用下,REConv允许来自不同池化层的信息流进行多级特征的融合,从而提供重新校准局部信息的注意掩码。

1.3 FRM

相互依赖是通道层的基本特征之一,本研究模型通过构建通道间的相互依赖性,选择更高级的特征信息,并利用其来抑制无效的特征。在特征细化块的辅助下,本研究提出FRM(图3),利用堆叠卷积核和级联结构从不同的接收域提取特征,以促进特征提取;同时,重新衡量相关属性的重要性,并通过排序优化进一步强化关键特征。作为CAMU-Net的基本单元,FRM有助于增强模型的特征提取能力,而

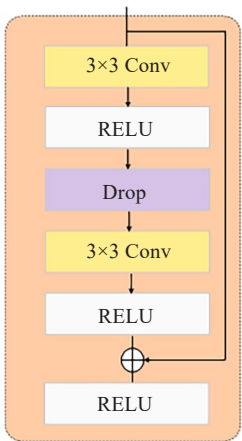


图2 残差增强模块  
Figure 2 Residual enhancement convolution module

高级表示则确保了对输入信息的更好理解。首先添加可形变卷积以提取信息;然后将得到的结果送入BN层,利用BN层的归一化操作来实现非线性的较强的表达能力,同时进一步加快网络收敛<sup>[9]</sup>,将可形变卷积的信息和BN层的信息进行融合,以便后续的分割操作;最后将通道融合的信息利用GELU激活函数来解决梯度消失的问题。FRM采用多级连结构的形式增强语义表达能力,纠正注意机制的建模方向,提高全局信息收集的能力。



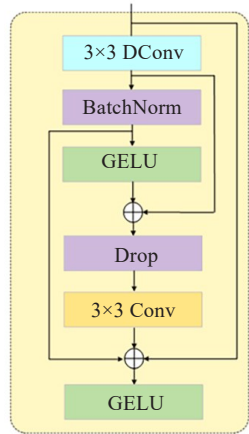


图3 特征细化模块

Figure 3 Feature refinement module

1.4 CAM

为进一步提高图像分割任务的精准度和效果,本研究提出 CAM(图 4),旨在优化模型对输入数据的空间结构的理解和利用,从而更准确地捕捉图像中的重要特征,并提升分割结果的质量。在该模块中,跳跃连接上每个层级的特征首先经过可变形卷积、GELU 激活函数以及 BN 操作,接着这些特征会与刚刚输出的特征进行连接。经过 BN 后,特征会根据输入数据的每个位置生成对应的坐标信息,这一步骤有助于模型更好地理解不同位置之间的关系。利用整合后的特征,计算得到注意力权重,进行线性变换后再计算注意力分数,最终通过 softmax 归一化得到最终的注意力权重。

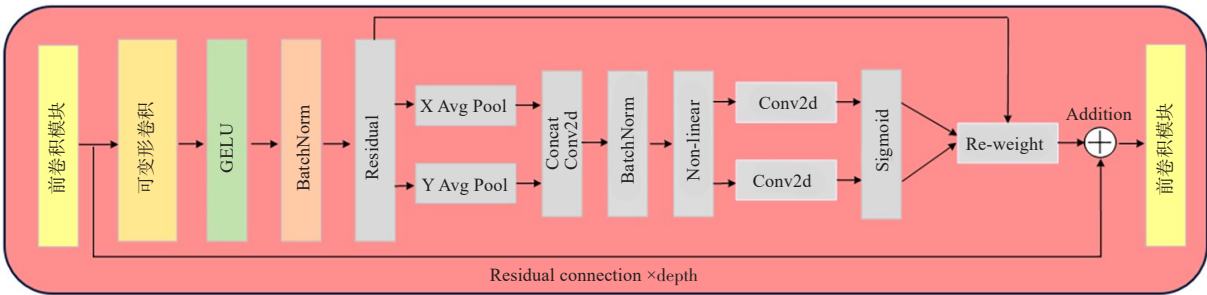


图4 通道注意力机制模块

Figure 4 Channel attention module

CAM 的引入具有诸多优势:(1)CAM 能有效地捕捉到输入数据的空间结构,这在处理具有明确空间关系的数据(如图像)时显得尤为重要。通过综合考虑坐标信息,该机制使模型能更全面地理解不同位置之间的相关性,从而促进精细的特征提取和分割效果的提升。(2)CAM 有助于模型更好地泛化到不同大小的输入数据,因为它并不依赖于输入的绝对位置信息,而是更注重相对位置信息的利用,这使得模型更具通用性,可以灵活地适应不同尺寸和形状的输入数据,从而提高模型的泛化能力和适用范围。

1.5 多尺度特征融合

由于 U-Net 模型没能很好地关注全局信息,本研究提出一种多尺度特征融合模块,通过将编码器每一层的输出进行融合,可以在一定程度上改善空间信息缺乏的问题,此外,该模块可以同时利用不同尺度下的信息,从而更好地捕捉图像中的空间信息,这有助于提高分割模型对目标边界等细节的感知能力。不同尺度的特征往往包含不同层次的语义信息,多尺度特征融合可以使模型更好地理解图像内容,提高分割的准确性和鲁棒性。

2 实验数据与评价标准

2.1 数据集

本研究采用两种公开的用于研究视网膜病变的视网膜血管图像数据集,分别是 DRIVE (Digital Retinal Images for Vessel Extraction) 和 STARE (Structured Analysis of the Retina)。DRIVE 数据集图像来自荷兰的糖尿病视网膜病变筛查计划,该数据集包含 40 名 25~90 岁的糖尿病受试者,其中 7 张图像有轻度早期糖尿病视网膜病变痕迹呈现,33 张图像没有任何糖尿病视网膜病变痕迹显现;图像分辨率均为 584×565,为划分出图像的重点关注区域,每一张视网膜图像均附带一个视场角(Field of View, FOV)掩膜,掩膜是半径 270 像素的圆形。本研究将 DRIVE 数据集中的图像划分为测试集和训练集,划分比例为 1:1。STARE 数据集的图像由美国佛蒙特大学的 Michael Goldbaum 博士创建,旨在帮助开发和研究血管分割算法,数据集总共包含 20 张视网膜图像,每个图像都配备了两名专家的手工分割结果作为金标准,其中 10 张有病变,如黄斑变性、高血压性视网膜病变、糖尿病性视网膜病变等,另外 10 张没有病变,图像分辨率均为 700×605,以 PPM 格式进行

保存。两个眼底视网膜血管图像数据集信息如图5所示。

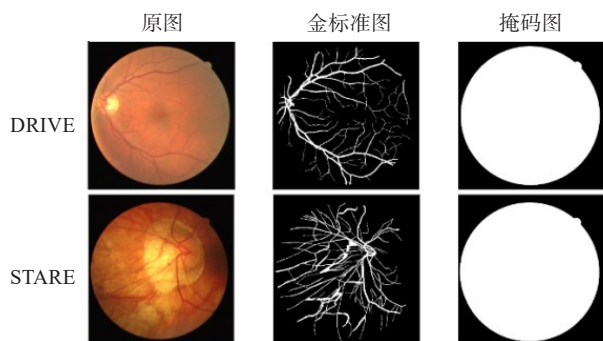


图5 眼底视网膜血管图像数据集  
Figure 5 Fundus retinal vascular image datasets

## 2.2 图像预处理

因为数据集存在图像光照不均和视网膜血管与图像背景对比度较低的情况,所以本研究在实验前对DRIVE数据集所训练的视网膜血管图像进行图像预处理操作,以获取更详细的视网膜血管特征,提高视网膜血管分割准确率。首先,将初始的RGB眼底图像通过加权法的灰度化处理转化为灰度图像,提

高图像的对比度;然后,通过对对比度的调控来限制灰度图像的自适应直方图均衡化,从而改善图像的细节;最后,针对图像光照不均匀或者有虚影干扰的部分,使用自适应伽马矫正来进行改善(图6)。

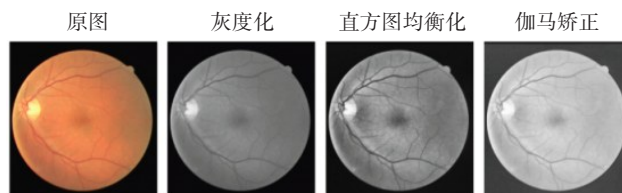


图6 DRIVE数据集预处理图像  
Figure 6 Preprocessed images from DRIVE dataset

因为数据集图像数量过少会导致模型出现过拟合现象使得模型的性能大打折扣,而DRIVE数据集中的视网膜图像只有40张,因此本研究决定通过重叠随机采样分块的方式,从DRIVE训练集图像中提取图像块尺寸为 $64 \times 64$ 像素的200 000个局部图像块,同时在金标准图像中提取与训练集中相对应的200 000个 $64 \times 64$ 像素局部图像块,以达到增大实验训练集图像的目的。图7分别表示训练集局部图像块和金标准局部图像块。

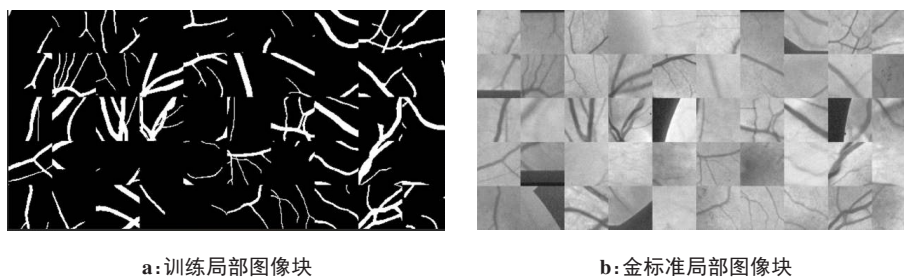


图7 局部图像块  
Figure 7 Partial image block

## 2.3 损失函数

### (1) 交叉熵损失函数

对于正样本来说,损失越小则表示输出值越大;对于负样本来说,损失越小则表示输出值也越小。交叉熵的定义为:

$$H = -y \log y' - (1 - y) \log (1 - y') = \begin{cases} -y \log y', & y = 1 \\ -(1 - y) \log (1 - y'), & y = 0 \end{cases} \quad (1)$$

其中, $y'$ 是模型的输出,输出值为0~1。

### (2) Dice损失函数

Dice损失函数是由Dice系数而得名的,Dice系

数是一种用于评估两个样本相似性的度量函数,其值越大意味着这两个样本越相似,其范围为 $[0, 1]$ 。

$$\text{Dice} = 2 \frac{|V_{\text{seg}} \cap V_{\text{gt}}|}{|V_{\text{seg}}| + |V_{\text{gt}}|} = \frac{2\text{TP}}{\text{FP} + 2\text{TP} + \text{FN}} \quad (2)$$

其中,TP表示预测为正样本且标签为正样本的像素数量,FP表示预测为正样本但标签为负样本的像素数量,FN表示预测为负样本但标签为正样本的像素数量。Dice损失函数在感兴趣部分仅占据扫描非常小的区域,从而使学习过程陷入损失函数的局部最小值。所以,要加大前景区域的权重:

$$\text{Loss}_{\text{Dice}} = 1 - \text{Dice} \quad (3)$$

最后,给出 CAMU-Net 的总体优化目标,即上述损失函数的线性加权融合如下:

$$L_{total} = \lambda_1 Loss_{CE} + \lambda_2 Loss_{Dice}$$

(4)

其中,  $\lambda_1$  和  $\lambda_2$  是权重参数,  $\lambda_1$  和  $\lambda_2$  的值都在  $[0, 1]$  内。当  $\lambda_1=0$  时,只有交叉熵损失是相关的。随着  $\lambda_1$  的增大,交叉熵损失逐渐发挥主导作用。相同的,当  $\lambda_2$  逐渐增大时, Dice 损失在分割任务中占据主导作用。本研究经过大量实验证明,采用上述损失约束模型训练学习能取得最好的分割效果。

2.4 评价指标

本实验使用 64 位 Windows 10 系统,训练显卡使用英伟达的 4060ti 显卡,训练框架为 PyTorch,实验将 DRIVE 和 STARE 数据集的验证集和训练集比例随机划分为 1:9,模型的性能通过本文构建的损失  $L_{total}$  来约束,其中  $\lambda_1$  和  $\lambda_2$  分别为 0.4 和 0.6;并通过 Adam 优化器的辅助来进行共同优化,模型设置的初始学习率为 0.000 1, batch size 设定为 64, epoch 设定为 200。训练终止的标准设定为 20 轮训练后模型的损失仍未有下降,与此同时,保存每次训练过程中最优异的各项评价指标和模型。为验证模型的分割性能,本研究采用灵敏度(Sensitivity)、特异性(Specificity)、准确率(Accuracy)、F1 值和曲线下面积(AUC)作为评价指标。

灵敏度是衡量分类模型在正确识别正样本(血管像素)方面的能力,表示将被正确分类为正样本的血管像素与真实血管像素的比例:

$$Sensitivity = \frac{TP}{TP + FN}$$

(5)

特异性是衡量分类模型在正确识别负样本(非血管像素)方面的能力,表示将被正确分类为负样本的非血管像素与真实非血管像素的比例:

$$Specificity = \frac{TN}{TN + FP}$$

(6)

准确率是衡量分类模型总体性能的指标,表示分类模型正确分类样本的比例:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

(7)

F1 值是综合考虑模型的准确率和召回率的度量指标,其中,分类模型中实际为正样本的样本在预测为正样本的样本中所占比例代表准确率,模型正确找到正样本的比例代表召回率:

$$F1-score = \frac{2TP}{2TP + FP + FN}$$

(8)

AUC 的作用是衡量血管分割结果综合性能, AUC 越大,代表算法的分割性能越优异。

3 实验结果与分析

3.1 消融实验

为验证本研究基于 U-Net 并在编码器采用 REConv 替换普通卷积、解码器添加 FRM、跳跃连接添加 CAM 以及引入多尺度特征融合结构对模型分割结果优化的有效性,设计一系列实验,验证添加不同种类模块后的分割结果(表 1)。在尺寸大小为 512×512 的图像上进行实验。U-Net 的准确率、灵敏度和 F1 值明显低于其他模型,说明该模型对血管提取能力有待提高,主要原因是 U-Net 结构一般比较关注局部特征,对全局信息的理解不够充分,这可能导致图像的一些全局性特征被忽略,如果输入图像中存在噪声,可能导致模型难以准确识别目标边界和类别,从而影响模型的分割效果。在基础的 M1 模型上加入 REConv 能显著提升模型的灵敏度和 F1 值;在 M2 基础上加入 FRM 能显著提升模型的 AUC 值;在 M3 的基础上加入 CAM 能显著提升模型的准确率和特异性;在 M4 的基础上加入多尺度特征融合结构能显著提升模型的灵敏度。

表 1 DRIVE 数据集上的消融研究  
Table 1 Ablation studies on DRIVE dataset

方法	准确率/%	特异性/%	灵敏度/%	F1 值/%	AUC
M1	95.94	97.94	78.62	80.03	0.979 0
M2	96.11	98.05	80.22	81.43	0.970 2
M3	96.21	98.03	80.97	82.51	0.981 8
M4	96.52	98.61	81.33	82.64	0.983 0
M5	96.76	98.66	81.60	82.79	0.983 2

M1 表示 U-Net 基线;M2 表示在 U-Net 中加入 REConv;M3 表示在 M2 中加入 FRM;M4 表示在 M3 中加入 CAM;M5 表示在 M4 中加入多尺度特征融合结构,即本研究提出的 CAMU-Net

M1 相较于其他模型的 F1 值和灵敏度较低,说明模型对于特征信息的提取和空间预测的血管信息能力不足,而且容易受到噪声的干扰,而通过改进模型后,模型的各项数据都得到大幅度提高,特别是 F1 值和灵敏度。添加模块后,模型通过多阶段的重复高分辨率与低分辨率信息交换和信息融合来逐步改进现有分辨率特征图像,使得低分辨率图像优化生成高分辨率特征图像,提高空间预测血管信息的准确率,大大提高模型全局获取信息的能力和分割精度。CAMU-Net 可以更有效地增强模型对特征获取的能力,提高模型的分割精度。

3.2 与其他主流方法的比较

为证明本研究提出的 CAMU-Net 在眼底血管图像分割方面的优势,将本文方法与现有的图像分割



网络 (SCUNet、U-Net、Attention U-Net、LinkNet、FCN、DenseU-Net) 进行对比。为更直观地体现 CAMU-Net 模型视网膜血管分割性能,还对比了不同算法在 DRIVE 和 STARE 数据集上的 ROC 曲线走势

(图 8)。CAMU-Net 的 ROC 曲线在两个数据集中相比其他算法都取得了很好的效果,表明本文算法相比其他视网膜血管分割算法的综合性能更强,能有效地识别细小血管和消除噪音干扰。

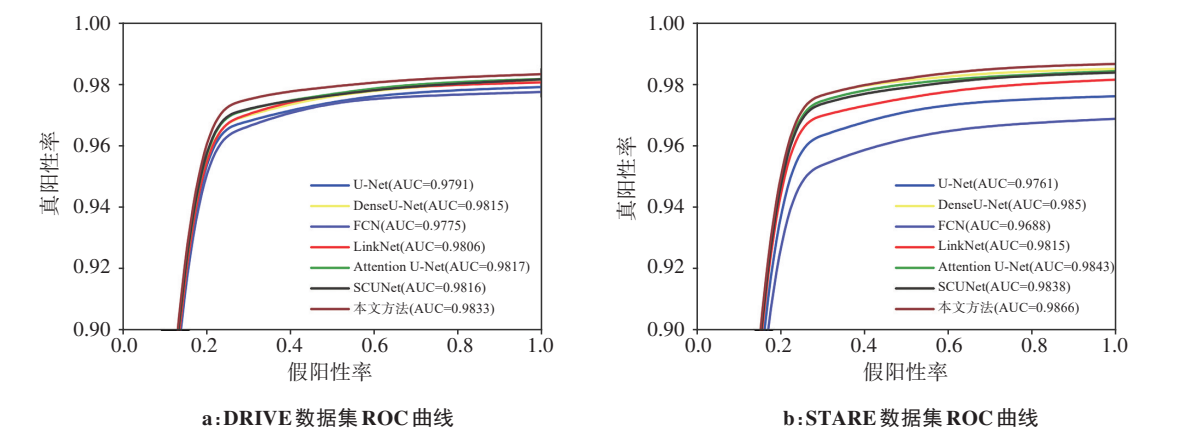


图8 不同算法的 ROC 曲线  
Figure 8 ROC curves of different algorithms

本研究提出的 CAMU-Net 模型相比较其他 6 种模型在 DRIVE 和 STARE 数据集上有更好的分割效果(表 2)。DRIVE 眼底血管图像分割结果见图 9,从左到右依次为 DRIVE 数据、标签图、Attention U-Net、DenseU-Net、FCN、LinkNet、SCUNet、U-Net、本文方法。U-Net 和 FCN 等模型的各项数值均较低,需要进一步提升;Attention U-Net 通过添加注意力机制,使得模型受到噪音的干扰变少,对血管细节分割能力增强,模型分割的轮廓清晰,可以看出眼底血管位置的某些细节特征,但其边缘模糊不够精细,说明模型在信息提取过程中有不少信息丢失;LinkNet 分割效

果较差,分割出来的结果只能大体反映眼底血管区域的位置,并且分割得到的边缘粗糙;DenseUnet 在准确率和灵敏度方面得到比较大的提升,但容易受到噪音影响;尽管 SCUNet 在某些特定的图像分割任务中表现出色,但其架构可能需要适应眼底血管图像分割任务才能达到比较好的分割效果。本研究提出的 CAMU-Net 在准确率、特异性、灵敏度、F1 值和 AUC 都有显著的提升,说明 CAMU-Net 不仅可以提取充分的特征信息,还可以有效地减小噪音干扰,准确地反映眼底血管区域的特征信息,达到较强的分割水平。

表 2 DRIVE 和 STARE 数据集上的分割结果  
Table 2 Segmentation results on DRIVE and STARE datasets

模型	DRIVE					STARE				
	准确率/%	特异性/%	灵敏度/%	F1 值/%	AUC	准确率/%	特异性/%	灵敏度/%	F1 值/%	AUC
SCUNet <sup>[27]</sup>	96.39	98.48	81.34	82.45	0.9816	96.25	97.91	81.80	81.85	0.9838
U-Net <sup>[28]</sup>	95.94	97.94	78.62	80.03	0.9790	95.89	97.75	79.76	80.04	0.9760
Attention U-Net <sup>[29]</sup>	96.10	97.93	80.16	80.95	0.9817	96.37	98.08	78.89	81.78	0.9842
LinkNet <sup>[30]</sup>	95.76	98.06	79.10	82.79	0.9805	96.18	97.31	77.75	80.81	0.9814
FCN <sup>[31]</sup>	96.09	97.40	76.05	80.09	0.9775	95.59	96.99	79.73	75.91	0.9688
DenseU-Net <sup>[32]</sup>	96.22	97.93	81.33	81.64	0.9815	96.42	98.01	80.60	83.01	0.9850
本文方法	96.76	98.66	81.60	82.79	0.9832	96.55	98.19	82.28	83.14	0.9866

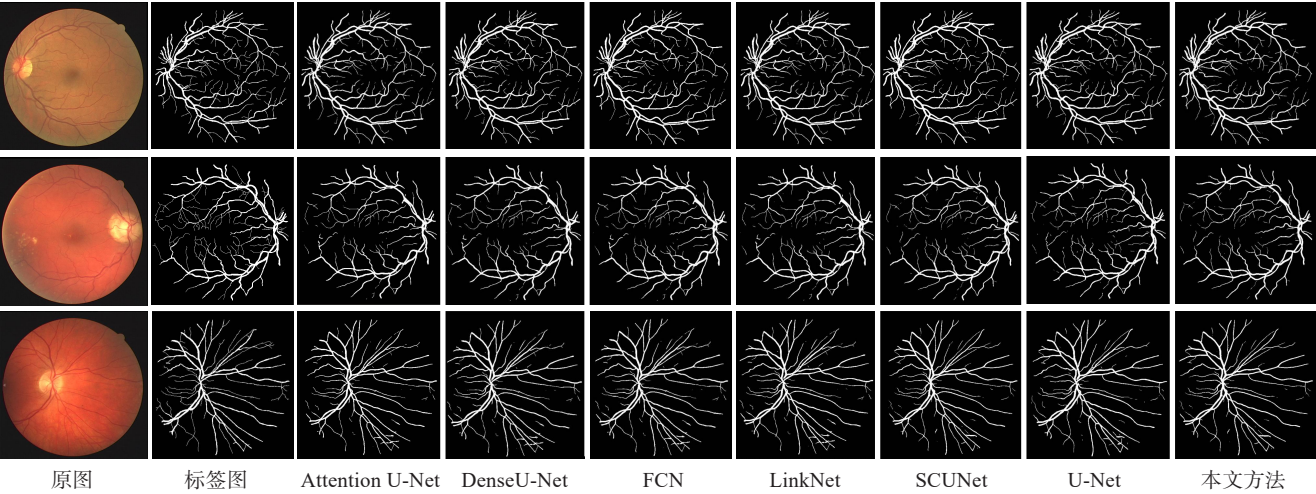


图9 DRIVE数据集上不同网络模型的分割结果

Figure 9 Segmentation results of different network models on DRIVE dataset

为更好地验证本文构建的网络在眼底血管图像分割任务上的优势,本研究在 STARE 数据上进行对比实验,进而说明本文模型的优势(图 10)。在 STARE 眼底血管图像分割任务中,CAMU-Net 在 AUC 值上取得了最高分(0.986 6),表现优于其他所有对比的网络,说明本文方法在眼底血管分割任务中具有最强的区分能力和准确性。DenseU-Net 和 Attention U-Net 也表现出色,分别取得了 0.985 0 和 0.984 2 的 AUC 值。SCUNet 和 LinkNet 分别获得了 0.983 8 和 0.981 4 的 AUC 值,显示出它们在此任务中

的有效性。相比之下,U-Net 和 FCN 的 AUC 值稍低,分别为 0.976 0 和 0.968 8,这表明虽然它们是经典的分割网络,但在处理复杂的眼底血管图像时,其性能略逊于其他改进和更复杂的网络架构。本文方法在 STARE 眼底血管图像分割任务中表现最佳,其在处理眼底血管图像分割任务上具有显著优势,DenseU-Net 和 Attention U-Net 紧随其后,而 SCUNet 和 LinkNet 表现良好,但 U-Net 和 FCN 的表现相对较弱。

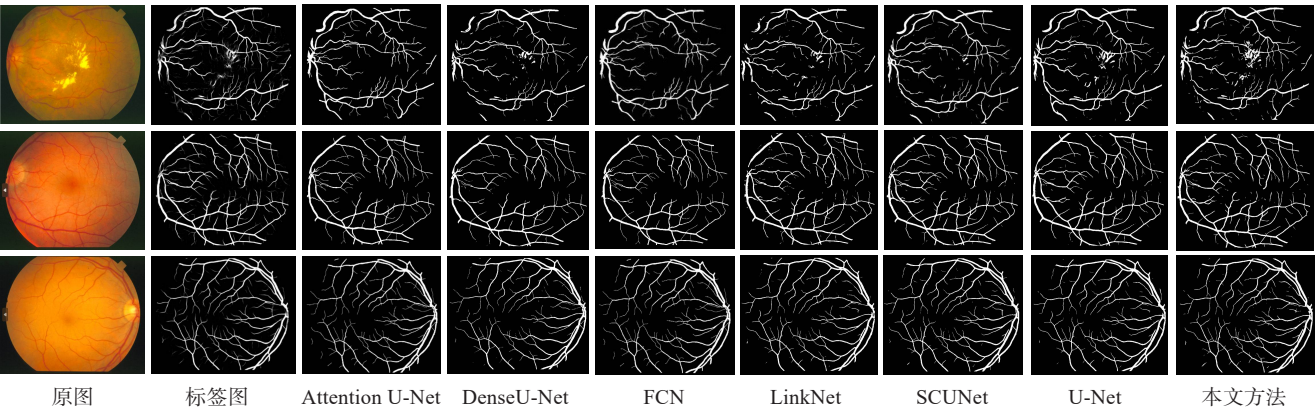


图10 STARE数据集上不同网络模型的分割结果

Figure 10 Segmentation results of different network models on STARE dataset

**4 结 论**

考虑到视网膜血管分布不均、细小复杂等特点,本研究通过添加多种模块和引入多尺度特征融合结构来改进模型的识别能力。在编码器部分,通过

REConv 的通道和空间关注,模型可以将重点放在颜色缺陷部分和重要信息部分,以达到更好的提取和分配特征信息的目的。模块的残差式结构允许不同池化层的信息流组合成残留特征,达到抑制低价值



特征,凸显高价值特征的作用。在解码器部分,通过FRM的可形变卷积进行信息提取,利用BN层来加快网络收敛,添加通道注意力机制来提高语义属性,并使用GELU激活函数来解决梯度消失等问题,提高模型的性能;通过CAM来有效捕捉图像中的重要特征,使得模型更具有通用性,提高模型的泛化能力。通过引入多尺度特征融合结构,解决模型缺乏对全局信息的充分关注的问题,增强模型对不同层次语义信息的理解,提高分割结果的精确性和鲁棒性。最后在实验部分进行消融实验和对比试验来验证优化模型的性能,通过实验发现CAMU-Net能在视网膜血管分割方面取得比较好的效果,提高血管分割的精度。

## 【参考文献】

- [1] Wu YC, Xia Y, Song Y, et al. NFN+: a novel network followed network for retinal vessel segmentation[J]. Neural Netw, 2020, 126: 153-162.
- [2] Huang HM, Lin LF, Tong RF, et al. UNet 3+: a full-scale connected UNet for medical image segmentation[C]//ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ, USA: IEEE, 2020: 1055-1059.
- [3] Li X, Jiang YC, Li ML, et al. Lightweight attention convolutional neural network for retinal vessel image segmentation[J]. IEEE Trans Industr Inform, 2021, 17(3): 1958-1967.
- [4] Cheung CY, Xu DJ, Cheng CY, et al. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre[J]. Nat Biomed Eng, 2021, 5(6): 498-508.
- [5] Sun L, Shao W, Wang ML, et al. High-order feature learning for multi-atlas based label fusion: application to brain segmentation with MRI[J]. IEEE Trans Image Process, 2019, 29: 2702-2713.
- [6] Hesamian MH, Jia WJ, He XJ, et al. Deep learning techniques for medical image segmentation: achievements and challenges[J]. J Digit Imaging, 2019, 32(4): 582-596.
- [7] Sun JD, Peng YJ, Guo YF, et al. Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN[J]. Neurocomputing, 2021, 423: 34-45.
- [8] 田丰, 李莹, 王静. 基于多尺度小波变换融合的视网膜血管分割[J]. 光学学报, 2021, 41(4): 76-86.  
Tian F, Li Y, Wang J. Retinal blood vessel segmentation based on multi-scale wavelet transform fusion[J]. Acta Optica Sinica, 2021, 41(4): 76-86.
- [9] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [10] Gu R, Wang GT, Song T, et al. CA-net: comprehensive attention convolutional neural networks for explainable medical image segmentation[J]. IEEE Trans Med Imaging, 2021, 40(2): 699-711.
- [11] Cao B, Zhao JW, Lü ZH, et al. Multiobjective evolution of fuzzy rough neural network via distributed parallelism for stock prediction[J]. IEEE Trans Fuzzy Syst, 2020, 28(5): 939-952.
- [12] Chen SY, Zou YN, Liu PX. IBA-U-net: attentive BConvLSTM U-net with redesigned inception for medical image segmentation[J]. Comput Biol Med, 2021, 135: 104551.
- [13] Oktay O, Schlemper J, Le Folgoc L, et al. Attention U-net: learning where to look for the pancreas[EB/OL]. (2018-05-20)[2024-01-17]. <https://arxiv.org/abs/1804.03999>.
- [14] Sarhan A, Rokne J, Alhajj R, et al. Transfer learning through weighted loss function and group normalization for vessel segmentation from retinal images[C]//2020 25th International Conference on Pattern Recognition (ICPR). Piscataway, NJ, USA: IEEE, 2021: 9211-9218.
- [15] Ma WA, Yu S, Ma K, et al. Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Cham: Springer International Publishing, 2019: 769-778.
- [16] Hu JF, Wang H, Gao SB, et al. S-UNet: a bridge-style U-net framework with a saliency mechanism for retinal vessel segmentation[J]. IEEE Access, 2019, 7: 174167-174177.
- [17] Diakogiannis FI, Waldner F, Caccetta P, et al. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data[J]. ISPRS J Photogramm Remote Sens, 2020, 162: 94-114.
- [18] Chen LC, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Trans Pattern Anal Mach Intell, 2018, 40(4): 834-848.
- [19] Chen JN, Lu YY, Yu QH, et al. TransUNet: transformers make strong encoders for medical image segmentation[EB/OL]. (2021-02-08)[2024-01-13]. <https://arxiv.org/abs/2102.04306>.
- [20] Wang B, Wang SP, Qiu S, et al. CSU-net: a context spatial U-net for accurate blood vessel segmentation in fundus images[J]. IEEE J Biomed Health Inform, 2021, 25(4): 1128-1138.
- [21] Zhu XZ, Hu H, Lin S, et al. Deformable ConvNets V2: more deformable, better results[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 9300-9308.
- [22] Li XT, You AS, Zhu Z, et al. Semantic flow for fast and accurate scene parsing[C]//Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 775-793.
- [23] Zhao YT, Rada L, Chen K, et al. Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images[J]. IEEE Trans Med Imaging, 2015, 34(9): 1797-1807.
- [24] 朱承璋, 崔锦恺, 邹北骥, 等. 基于多特征融合和随机森林的视网膜血管分割[J]. 计算机辅助设计与图形学学报, 2017, 29(4): 584-592.  
Zhu CZ, Cui JK, Zou BJ, et al. Retinal vessel segmentation based on multiple feature fusion and random forest[J]. Journal of Computer-Aided Design and Computer Graphics, 2017, 29(4): 584-592.
- [25] 徐光柱, 胡松, 陈莎, 等. U-net与Dense-net相结合的视网膜血管提取[J]. 中国图象图形学报, 2019, 24(9): 1569-1580.  
Xu GZ, Hu S, Chen S, et al. Retinal blood vessel extraction by combining U-net and Dense-net[J]. Journal of Image and Graphics, 2019, 24(9): 1569-1580.
- [26] Zhang K, Li YW, Liang JY, et al. Practical blind image denoising via Swin-Conv-UNet and data synthesis[J]. Mach Intell Res, 2023, 20(6): 822-836.
- [27] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
- [28] 梁礼明, 黄朝林, 石霏, 等. 融合形状先验的水平集眼底图像血管分割[J]. 计算机学报, 2018, 41(7): 1678-1692.  
Liang LM, Huang ZL, Shi F, et al. Retinal vessel segmentation using level set combined with shape priori[J]. Chinese Journal of Computers, 2018, 41(7): 1678-1692.
- [29] Chaurasia A, Culurciello E. LinkNet: exploiting encoder representations for efficient semantic segmentation[C]//2017 IEEE Visual Communications and Image Processing (VCIP). Piscataway, NJ, USA: IEEE, 2017: 1-4.
- [30] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(4): 640-651.
- [31] Kande GB, Savithri TS, Subbaiah PV. Retinal vessel segmentation using histogram matching[C]//APCCAS 2008-2008 IEEE Asia Pacific Conference on Circuits and Systems. Piscataway, NJ, USA: IEEE, 2008: 129-132.
- [32] Zhou YP, Chang HY, Lu XL, et al. DenseUNet: improved image classification method using standard convolution and dense transposed convolution[J]. Knowl Based Syst, 2022, 254: 109658.

(编辑:谭斯允)