西北大学学报(自然科学版)
2026 年 10 月,第 56 卷第 5 期,Oct. ,2026,Vol. 56,No. 5
Journal of Northwest University (Natural Science Edition)

·深度学习视觉处理与分析·

基于边缘特征增强的 U-Net 视网膜血管分割

王　怡,程　旭,王　嫄

天津科技大学 人工智能学院,天津　300457

摘要　针对视网膜血管分割过程中,微小血管结构复杂、边界模糊导致血管分割不完全或微小血管断裂等问题,提出一种基于边缘特征增强的 U 形网络 EFE U-Net(edge feature enhancement U-Net)。通过引入边缘特征增强模块代替 U-Net 模型中的二维卷积,在增加 U-Net 感受野的同时,增强视网膜血管图像中的边缘信息,边缘特征增强模块由边缘信息提取模块、边缘信息融合模块、通道和空间注意力模块构成。在 DRIVE、CHASE_DB1 和 STARE 数据集上进行实验,EFE U-Net 灵敏度为 82.64%、84.41% 和 84.45%,F1 分数为 83.09%、84.49% 和 82.84%,2 个指标相较于基准模型算法均有所提升。实验结果表明 EFE U-Net 能够有效对视网膜血管进行分割。

关键词　视网膜血管分割;边缘特征增强;U 形网络

中图分类号:TP391.4　　**DOI**:10.16152/j.cnki.xdxbzr.2026-05-013

U-Net retinal vessel segmentation based on edge feature enhancement

Wang Yi, Cheng Xu, Wang Yuan

College of Artificial Intelligence, Tianjin University of Science and Technology, Tianjin 300457, China

Abstract　In the process of retinal vessel segmentation, issues such as incomplete vessel segmentation or the presence of micro-vessel fractures arise due to the complex structure of small vessels and blurred boundaries. To address these challenges, we propose an edge feature enhancement U-Net (EFE U-Net) based on edge feature enhancement. By introducing an edge feature enhancement module to replace the 2D convolution in the U-Net model, the receptive field of U-Net is expanded while enhancing the edge information in retinal vessel images. The edge feature enhancement module consists of an edge information extraction module, an edge information fusion module, and a channel and spatial attention module. Experiments on the DRIVE, CHASE_DB1, and STARE datasets demonstrate that EFE U-Net achieves sensitivities of 82.64%, 84.41% and 84.45%, the F1 scores are 83.09%, 84.49% and 82.84%. The two metrics show improvemens over the baseline model. The experimental results indicate that EFE U-Net can effectively segment retinal vessels.

Keywords　retinal vessel segmentation; edge feature enhancement; U-shaped network

对视网膜血管进行有效分割涉及从视网膜图　　像中识别和提取血管结构,这对早期诊断和检测

收稿日期:2026-01-21　　修回日期:2026-03-13

基金项目:西藏自治区自然科学基金(Z202401ZR0008);天津市科委技术创新引导专项(20YDTPJC00560)。

第一作者:王怡,男,副教授,硕士生导师,从事深度学习、智能图像处理研究,wy@tust.edu.cn。

多种眼科疾病,如糖尿病视网膜病变、青光眼和高血压性视网膜病变等具有重要意义。传统的医学图像分割方法依赖于医生的主观判断,不仅耗时而且易受个体经验和技术水平的限制,因此利用自动化分割技术至关重要。利用自动化的视网膜血管分割算法,可以有效提高诊断的准确率和效率。为临床医生提供定量的血管参数,从而辅助诊断和治疗决策[1-2]。

近年来,深度学习技术在医学图像分割领域获得了显著进展。其中,全卷积神经网络(fully convolutional network, FCN)[3]作为语义分割领域的开创性工作,具有高精度、高速度、自动化处理的优势,为医学图像分割提供了强有力的支持。U-Net[4]是 Ronneberger 等通过对 FCN 进行改进而提出的一种 U 型结构医学图像分割模型,U-Net 由编码器和解码器两部分及跳跃连接组成,跳跃连接有效融合了多尺度空间特征信息,减少了血管细节上的丢失。由 Zhuang 等提出的 LadderNet[5]通过将 2 个 U-Net 网络相连,构建多路径融合机制,并在残差块中实现 2 个卷积层共享权重,增强模型对复杂结构的理解。Oktay 等提出了 Attention U-Net[6],通过增加 Attention Gate 模块增强相关区域的权重抑制不相关区域。

感受野对于视网膜血管分割具有重要意义,其大小直接影响模型对于血管的识别能力。Jin 等人将可变型卷积引入 U-Net,提出了 DU-Net 模型[7],在解码器中引入了多尺度特征融合机制,利用可变型卷积自适应调整感受野大小,从而适应血管的形状和尺度。Gu 等提出的 CE-Net 模型[8]将密集空洞卷积和残差多核池化模块引入 U-Net 解码器,扩大感受野,在不增加参数量的情况下捕捉更多全局信息,同时,引入空间注意力机制进一步优化特征选择能力,这些模块能够捕获更高层次的特征并保留空间信息,在保留血管特征的同时抑制无关特征,有效提取血管的精细特征信息。Zhang 等提出的 ET-Net 模型[9],结合边缘感知和 Transformer 机制,在捕捉目标区域的边缘信息同时,使用 Transformer 的全局建模能力提升分割精度。Pan 等提出的 MSFA-Net 模型[10],基于多光谱滤光阵列从原图中直接提取纹理特征,通过多尺度特征聚合机制提升分割精度。张瑞等提出一种融合多尺度特征和反向注意力机制的深度网络[11],进而提升改善 U-Net 感受野受限的问题,提升了网络对不同大小目标的分割能力。

尽管自动化视网膜血管分割技术取得了一定的进展,但仍然面临着以下问题:由于视网膜血管结构复杂且眼底图像的对比度低,现有算法在处理细小血管时,仍存在边缘细节丢失和血管断裂的问题。

针对上述问题,本文提出一种基于 U-Net 模型的改进网络模型 EFE U-Net(edge feature enhancement U-Net),该模型使用边缘特征增强(edge feature enhancement module, EFEM)模块代替传统 U-Net 中的二维卷积。边缘特征增强模块由边缘特征提取模块、边缘特征融合模块和通道空间注意力模块构成。通过使用边缘特征提取模块可以有效改善 U-Net 感受野受限的问题并且提取眼底图像中包含的血管边缘信息,通过边缘特征融合模块可以有效增强图像中包含的血管边缘信息,使用通道空间注意力模块增强模型对局部细节信息和全局信息的提取,提升模型的准确性和鲁棒性。

1 本文方法

本文提出的 EFE U-Net 网络结构如图 1 所示,网络架构由编码器、解码器和跳跃连接组成。为了有效提取多尺度边缘特征,本文提出了边缘特征增强模块(EFEM),并将其融合到网络编码器和解码器中,从而改善 U-Net 网络感受野受限和无法充分提取边缘特征信息的问题,提高了模型对多尺度信息提取和利用能力,增强了视网膜血管分割模型的精准性和鲁棒性。

1.1 边缘特征增强模块(EFEM)

为了提升模型对视网膜图像中的边缘特征信息和上下文信息的获取能力,改善 U-Net 网络在卷积过程中可能出现的边缘特征信息丢失问题,如图 1 所示,本文在传统 U-Net 模型中引入边缘特征增强模块,实现卷积过程中增强视网膜血管的边缘特征。

边缘特征增强模块主要由 3 个部分组成:边缘特征信息提取模块、边缘特征融合模块(edge feature fusion module, EFFM)和通道空间注意力模块(channel space attention module, CSAM)。边缘特征信息提取模块的主要职能是提取水平和垂直方向的边缘特征信息。边缘特征融合模块的作用是将获取到的垂直边缘特征信息和水平边缘特

征信息与 3 ×3 卷积核处理后的特征图进行融合,以增强特征图中包含的水平和垂直边缘特征信息。通道空间注意力模块则通过整合通道特征和空间特征,并对它们进行校准,从而增强模型对特征图中包含的通道信息和空间信息进行有效关注,减少不重要的信息干扰。

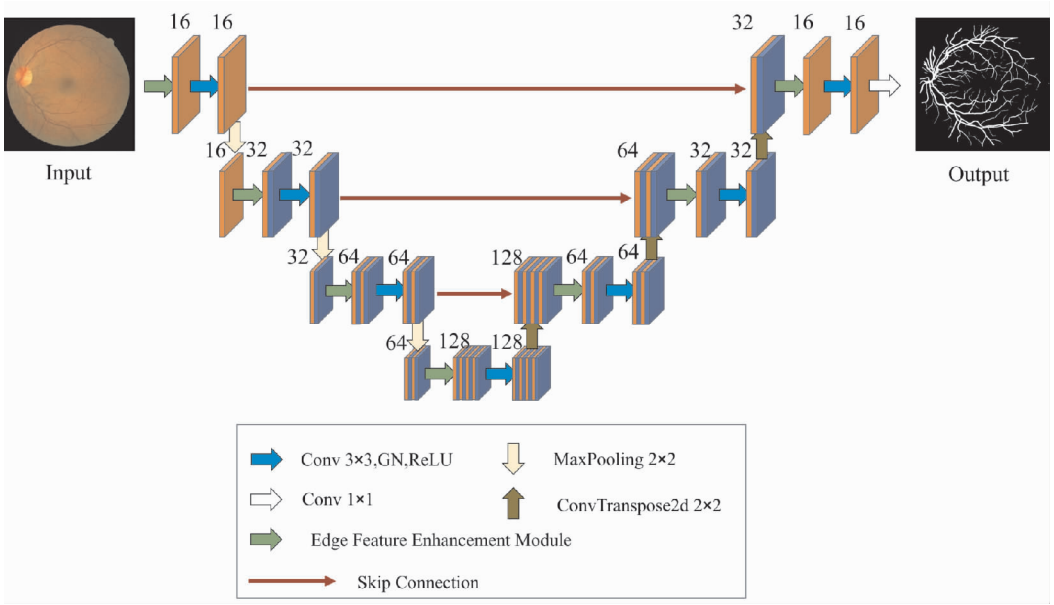


图 1　EFE U-Net 网络模型结构

Fig. 1　Structure of EFE U-Net network

在模型的训练阶段,输入特征图会依次通过不同配置的二维卷积层进行处理。如图 2 所示,输入会先经过 3 个不同卷积核大小的卷积操作,这些卷积操作分别由卷积核大小为 3 ×3、3 ×1、1 ×3 的卷积和批量归一化以及 ReLU 激活函数组成,旨在从输入图像中提取多尺度的特征。经过这一系列操作后,生成 3 组特征图,分别记为 $\boldsymbol{F}_1$、$\boldsymbol{F}_2$ 和 $\boldsymbol{F}_3$。公式为:

$$\boldsymbol{F}_1 = \mathrm{ReLU}(\mathrm{BN}(\mathrm{Con}_{3\times3}(\boldsymbol{F}_{\mathrm{input}}))) \tag{1}$$

$$\boldsymbol{F}_2 = \mathrm{ReLU}(\mathrm{BN}(\mathrm{Con}_{1\times3}(\boldsymbol{F}_{\mathrm{input}}))) \tag{2}$$

$$\boldsymbol{F}_3 = \mathrm{ReLU}(\mathrm{BN}(\mathrm{Con}_{3\times1}(\boldsymbol{F}_{\mathrm{input}}))) \tag{3}$$

式中:$\boldsymbol{F}_{\mathrm{input}}$ 表示输入特征图;$\mathrm{Con}_{3\times3}$ 表示卷积核大小为 3 ×3 的卷积;$\mathrm{Con}_{1\times3}$ 表示卷积核大小为 3 ×1 的卷积;$\mathrm{CON}_{1\times3}$ 表示卷积核大小为 1 ×3 的卷积;BN 表示批归一化。

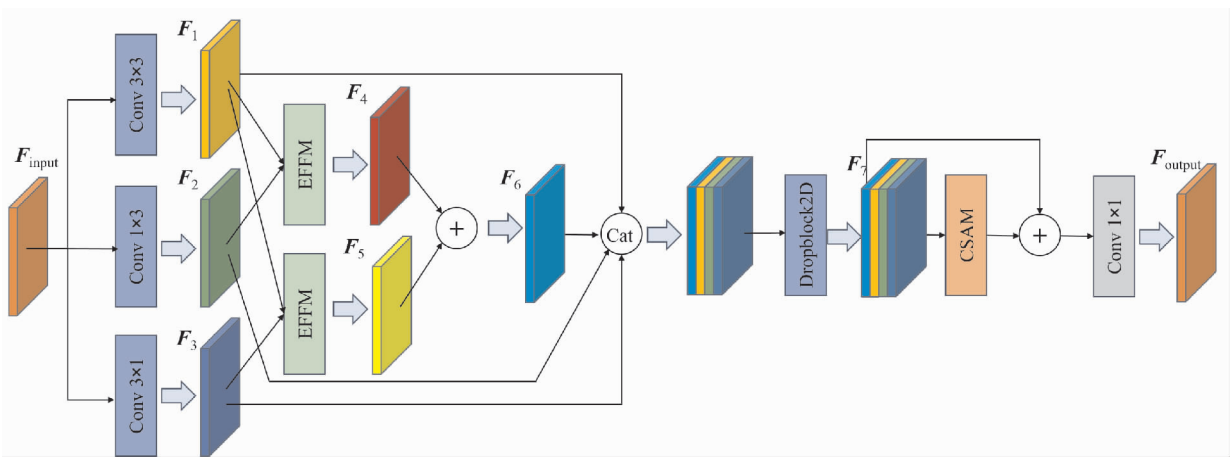


图 2　边缘特征增强模块

Fig. 2　Edge feature enhancement module

接下来,特征图 $\boldsymbol{F}_1$ 和 $\boldsymbol{F}_2$ 通过边缘特征融合模块进行特征融合,产生特征图 $\boldsymbol{F}_4$。同理,将 $\boldsymbol{F}_1$ 和 $\boldsymbol{F}_3$ 使用边缘特征融合模块进行融合,产生特征图 $\boldsymbol{F}_5$。随后,将特征图 $\boldsymbol{F}_4$ 和 $\boldsymbol{F}_5$ 相加,生成特征图 $\boldsymbol{F}_6$,

完成边缘特征的增强处理。随后,将特征图 $\boldsymbol{F}_1$、$\boldsymbol{F}_2$、$\boldsymbol{F}_3$ 以及 $\boldsymbol{F}_6$ 进行拼接操作,进入 Dropblock 模块进行正则化。Dropblock 模块[12] 的作用是减少模型过拟合的风险。之后,使用通过通道空间注意力模块,将通道和空间特征合并和校准,然后,与 Dropblock 模块产生的特征图 $\boldsymbol{F}_7$ 建立残差连接,最后,通过一个 1 × 1 的卷积进行降维,最终,输出特征图 $\boldsymbol{F}_{output}$。上述操作的公式为:

$$\boldsymbol{F}_4 = \mathrm{EFFM}(\boldsymbol{F}_1, \boldsymbol{F}_2) \tag{4}$$

$$\boldsymbol{F}_5 = \mathrm{EFFM}(\boldsymbol{F}_1, \boldsymbol{F}_3) \tag{5}$$

$$\boldsymbol{F}_6 = \boldsymbol{F}_4 + \boldsymbol{F}_5 \tag{6}$$

$$\boldsymbol{F}_7 = \mathrm{Dropblock}(\mathrm{Concat}(\boldsymbol{F}_1, \boldsymbol{F}_2, \boldsymbol{F}_3, \boldsymbol{F}_6)) \tag{7}$$

$$\boldsymbol{F}_{output} = \mathrm{Con}_{1\times1}(\boldsymbol{F}_7 + \mathrm{CSAM}(\boldsymbol{F}_7)) \tag{8}$$

式中:EFFM 为边缘特征融合模块;Concat 为张量拼接操作;CSAM 为通道空间注意力模块;$\mathrm{Con}_{1\times1}$ 为卷积核大小为 1 的卷积。

1.2 边缘特征融合模块(EFFM)

为克服传统的简单加法或者乘法的特征融合方式无法有效获取特征图中的关键边缘特征信息,本文提出的边缘特征融合模块(EFFM)受到时间特征融合模块(TFAM)[13] 的启发,通过融合 2 个不同的特征图进而增强边缘特征信息。它结合多尺度特征提取以及通道和空间注意力机制,通过使用多尺度卷积捕捉不同尺度的特征信息,然后通过注意力机制动态调整特征的重要性,保证关键特征在融合过程中能得到更高的权重。如图 3 所示,EFFM 通过将输入特征图 $\boldsymbol{F}_{input2}$ 和 $\boldsymbol{F}_{input3}$ 分别送入多尺度特征提取模块,得到 2 个多尺度特征图 $\boldsymbol{F}_8$ 和 $\boldsymbol{F}_9$。多尺度特征提取模块由 3 个不同尺寸的卷积核(3 × 3、5 × 5、7 × 7)组成的二维卷积和 ReLU 激活函数构成,输入数据会分别经过这 3 个卷积层的处理,每个卷积核分别提取不同尺度的特征。这些特征图随后通过 ReLU 激活函数进行非线性变换,以增强特征的表达能力。最后,将这 3 个不同尺度的特征图相加,得到新的输出特征图 $\boldsymbol{F}_8$ 和 $\boldsymbol{F}_9$。通过这一系列过程,有效地整合了不同尺度的特征信息,克服了单一尺度的卷积核可能无法充分捕捉特征图中的多尺度信息,从而限制了模型在处理不同尺度和复杂特征时的表现问题。

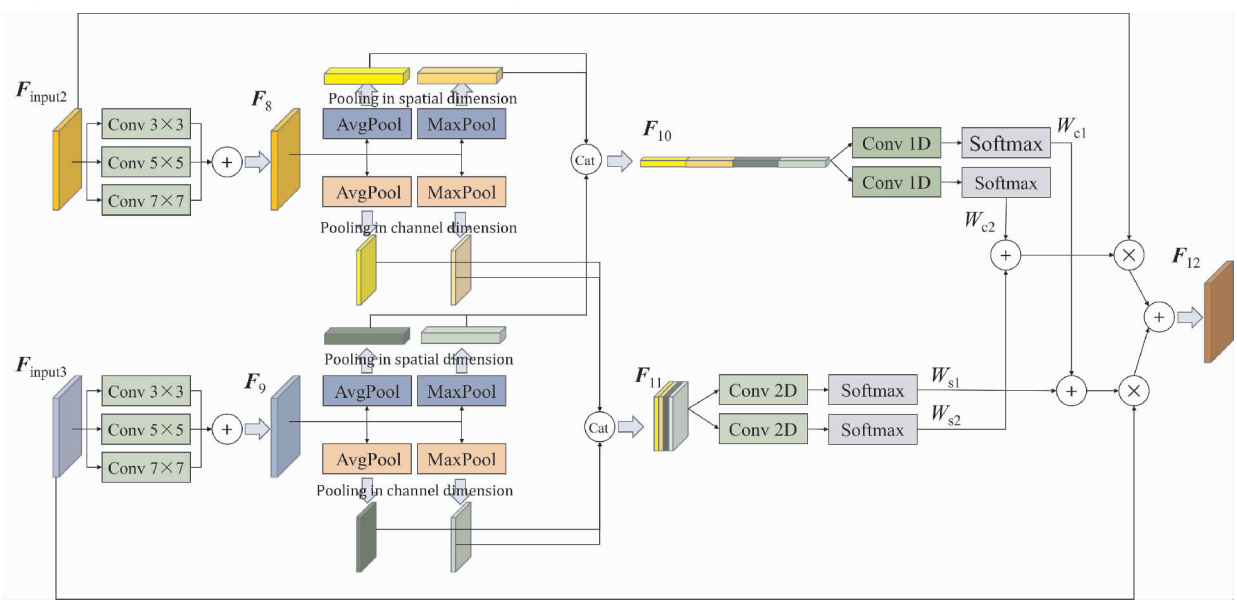


图 3 边缘特征融合模块

Fig. 3 Edge feature fusion module

这种多尺度处理方法使模型能够捕捉到图像中的局部细节和全局上下文信息,从而更有效地识别和融合。通过这种方式,EFFM 提高了模型对图像中复杂结构和模式的理解能力,尤其在处理具有多种尺度变化的特征时。其公式为:

$$\boldsymbol{F}_8 = \mathrm{ReLU}(\mathrm{Con}_{3\times3}(\boldsymbol{F}_{input2})) + \mathrm{ReLU}(\mathrm{Con}_{5\times5}(\boldsymbol{F}_{input2})) + \mathrm{ReLU}(\mathrm{Con}_{7\times7}(\boldsymbol{F}_{input2})) \tag{9}$$

$$\boldsymbol{F}_9 = \mathrm{ReLU}(\mathrm{Con}_{3\times3}(\boldsymbol{F}_{input3})) + \mathrm{ReLU}(\mathrm{Con}_{5\times5}(\boldsymbol{F}_{input3})) + \mathrm{ReLU}(\mathrm{Con}_{7\times7}(\boldsymbol{F}_{input3})) \tag{10}$$

式中,$\boldsymbol{F}_{input2}$ 和 $\boldsymbol{F}_{input3}$ 代表 2 个不同的输入数据。

随后,模型通过通道注意力和空间注意力确定特征中重要的部分,在通道特征处理过程中,特征图 $\boldsymbol{F}_8$ 和 $\boldsymbol{F}_9$ 分别应用全局最大池化和全局平均池化,提取全局最大信息和全局平均信息。然后将 2 个特征图产生的 4 个池化结果拼接在一起,将聚合后的通道特征图 $\boldsymbol{F}_{10}$ 通过一维卷积和 Softmax 激活函数处理,以生成通道注意力权重 W_{c1} 和 W_{c2}。其公式为:

$$\boldsymbol{F}_{10} = \mathrm{Concat}(\mathrm{GAvg}(\boldsymbol{F}_8), \mathrm{GMax}(\boldsymbol{F}_8), \mathrm{GAvg}(\boldsymbol{F}_9), \mathrm{GMax}(\boldsymbol{F}_9)) \tag{11}$$

$$W_{c1}, W_{c2} = \mathrm{Softmax}(\mathrm{Con}_{1d}(\boldsymbol{F}_{10})) \tag{12}$$

式中:$\boldsymbol{F}_{10}$ 为合并后的通道特征图;GAvg 和 GMax 表示全局平均池化和全局最大池化。

类似地,在空间注意力分支中,对 2 个输入特征图 $\boldsymbol{F}_8$ 和 $\boldsymbol{F}_9$ 分别应用通道维度上的最大池化和空间平均池化,进而提取空间信息。之后将产生的 4 个空间特征结果拼接在一起,将聚合后的空间特征图 $\boldsymbol{F}_{11}$ 通过二维卷积和 Softmax 激活函数处理,生成空间注意力权重 W_{s1} 和 W_{s2}。其公式为:

$$\boldsymbol{F}_{11} = \mathrm{Concat}(\mathrm{CAvg}(\boldsymbol{F}_8), \mathrm{CMax}(\boldsymbol{F}_8), \mathrm{CAvg}(\boldsymbol{F}_9), \mathrm{CMax}(\boldsymbol{F}_9)) \tag{13}$$

$$W_{s1}, W_{s2} = \mathrm{Softmax}(\mathrm{Con}_{2d}(\boldsymbol{F}_{11})) \tag{14}$$

式中:$\boldsymbol{F}_{11}$ 表示合并后的空间特征图;CAvg 和 CMax 表示通道维度的平均池化和通道维度的最大池化。

最后,将空间注意力权重和通道注意力权重应用于输入特征图 $\boldsymbol{F}_{\mathrm{input2}}$ 和 $\boldsymbol{F}_{\mathrm{input3}}$,进行特征融合,并输出最终的特征图 $\boldsymbol{F}_{12}$。

$$\boldsymbol{F}_{12} = (W_{c1} + W_{s1}) \odot \boldsymbol{F}_{\mathrm{input2}} + (W_{c2} + W_{s2}) \odot \boldsymbol{F}_{\mathrm{input3}} \tag{15}$$

式中,⊙ 表示逐元素乘法。

1.3　通道空间注意力模块(CSAM)

为了使网络在关注局部细节的同时兼顾全局信息,通道空间注意力模块的应用使网络输出更为全面,进而提升模型的泛化能力。该模块结合了通道注意力(SE)[14]和空间注意力机制[15]。其中,通道模块借助全局平均池化及 1×1 卷积操作生成通道权重,空间模块则通过 1×1 卷积与 Sigmoid 函数生成空间权重。最终,模块通过取两者逐元素的最大值,动态协调特征图的通道与空间信息,从而强化特征图的表达能力。

通道空间注意力模块的结构如图 4 所示。通道注意力机制:首先,输入特征图 $\boldsymbol{F}_{\mathrm{input4}}$ 经过全局平均池化,将特征图的每个通道压缩为一个标量,提取全局信息;然后,通过一个 1×1 卷积层进行降维,接着应用 ReLU 激活函数,生成通道注意力权重;最终,将这些通道注意力权重与原始输入特征图 $\boldsymbol{F}_{\mathrm{input4}}$ 相乘,实现特征重校准,并得到特征图 $\boldsymbol{F}_{13}$。空间注意力机制:输入特征图 $\boldsymbol{F}_{\mathrm{input4}}$ 经过 1×1 卷积层,生成空间权重;然后,通过 Sigmoid 函数归一化;最后,将空间注意力权重应用到输入特征图 $\boldsymbol{F}_{\mathrm{input4}}$ 上,以实现空间级别的特征重校准,并得到特征图 $\boldsymbol{F}_{14}$。通道注意力和空间注意力校准后的特征图 $\boldsymbol{F}_{13}$ 和 $\boldsymbol{F}_{14}$ 通过逐元素最大值操作(Max)融合,以结合通道和空间的注意力信息,得到最终的特征图 $\boldsymbol{F}_{15}$。

$$\boldsymbol{F}_{13} = \boldsymbol{F}_{\mathrm{input4}} \odot \mathrm{ReLU}(\mathrm{Con}_{1\times1}(\mathrm{GAve}(\boldsymbol{F}_{\mathrm{input4}}))) \tag{16}$$

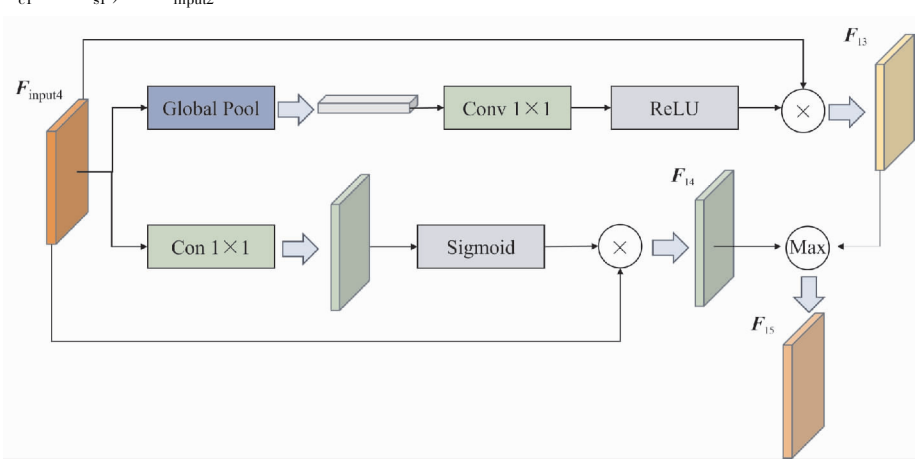


图 4　通道空间注意力模块

Fig. 4　Channel space attention module

$$\boldsymbol{F}_{14} = \boldsymbol{F}_{\text{input4}} \odot \text{Sigmoid}(\text{Con}_{1\times1}(\boldsymbol{F}_{\text{input4}})) \tag{17}$$

$$\boldsymbol{F}_{15} = \max(\boldsymbol{F}_{13}, \boldsymbol{F}_{14}) \tag{18}$$

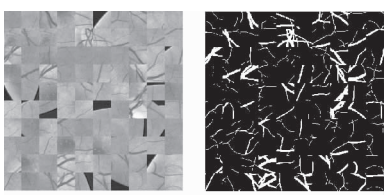
式中:⊙ 表示逐元素乘法;GAve 表示全局平均池化;$\text{Con}_{1\times1}$ 表示 1 × 1 卷积操作;Max 表示逐元素取最大值操作;$\boldsymbol{F}_{13}$ 代表通道注意力生成的特征图;$\boldsymbol{F}_{14}$ 表示空间注意力生成的特征图;$\boldsymbol{F}_{15}$ 为通道和空间特征图融合后的特征图。

2 实验分析

2.1 数据集

本文实验采用 DRIVE、CHASE_DB1 和 STARE 公共眼底图像数据集。DRIVE 数据集包含 40 张带有专业标注的眼底图像,分辨率为 565 ×584 像素,其中,20 张图像被分配用于训练,剩余 20 张用于测试。CHASE_DB1 数据集包含 28 张带有标注的眼底图像,分辨率为 999 × 960 像素,其中,20 张图像用于训练,8 张用于测试。STARE 数据集包含 20 张眼底图像,分辨率为 605 ×700 像素,其中,14 张图像用于训练,6 张用于测试。

鉴于视网膜图像数据集的规模较小,在训练阶段对数据集进行图像预处理和增强。具体来说:首先,采用滑动窗口对视网膜血管图像及其对应的标注图像进行裁剪;然后,对裁剪后的视网膜血管图像和标注图像进行随机反转等数据增强操作;最终,训练集扩充至 150 000 张图像。图像增强后的数据集如图 5 所示。



(a)眼底血管图像 (b)标签图

图 5 数据增强后的数据集

Fig. 5 The data enhanced dataset

2.2 实验设置

实验基于 Python 3. 12 环境,采用 PyTorch 框架 2. 3. 0 版本,Intel(R) Xeon(R) Silver 4214R CPU 和 NVIDIA RTX 3080 Ti 12 GiB GPU 服务器,操作系统为 Ubuntu 22. 04,配备 90 GiB 内存。模型训练参数设置如下:总训练轮数(epochs)为 50;批量大小(batch size)为 48;滑动窗口大小为 48 ×48;使用 Adam 优化器。训练过程中,使用二分类交叉熵损失函数,并结合余弦退火策略调整学习率[16],初始学习率设定为 0. 000 1。为防止模型在训练集上过拟合,实验引入了早停法(early stopping)机制,当验证集上的 AUC 值在连续 17 个 epoch 中未见提升时,训练过程将被终止。

2.3 评价指标

在视网膜血管分割研究领域,通常将眼底图像中的血管像素定义为正样本(positive class),而非血管像素则定义为负样本(negative class)。具体而言:模型正确识别的血管像素被称为真阳性(TP);实际上为血管但被模型错误分类为非血管的像素被称为假阴性(FN);模型正确识别的非血管像素被称为真阴性(TN);实际为非血管但被错误分类为血管的像素为假阳性(FP)。本文采用以下 5 个主要指标评估模型 EFE U-Net 的性能。

准确度(Accuracy, 式中简记 P_{Acc})衡量模型整体正确分类的比例,

$$P_{\text{Acc}} = \frac{P_{\text{TP}} + P_{\text{TN}}}{P_{\text{TP}} + P_{\text{TN}} + P_{\text{FP}} + P_{\text{FN}}} \tag{19}$$

灵敏度(Sensitivity, 式中简记 P_{Sen})也称为召回率,衡量模型识别正类的能力,

$$P_{\text{Sen}} = \frac{P_{\text{TN}}}{P_{\text{TN}} + P_{\text{FP}}} \tag{20}$$

特异性(Specificity, 式中简记 P_{Spe})衡量模型正确识别负类的能力,

$$P_{\text{Spe}} = \frac{P_{\text{TN}}}{P_{\text{TN}} + P_{\text{FP}}} \tag{21}$$

F1 分数(F1 score, 式中简记 F_1)是精确率和召回率的调和平均数,提供模型精确性和完整性的平衡视图,

$$F_1 = \frac{2P_{\text{TP}}}{2P_{\text{TP}} + P_{\text{FP}} + P_{\text{FN}}} \tag{22}$$

AUC(area under the curve)衡量模型在所有可能的分类阈值上的性能,AUC 值越高,表明模型的区分能力越强。

这些指标共同提供了模型性能的全面评估,包括其准确性、对血管和非血管区域的识别能力,以及整体的区分能力。

TP_NCC 表示正确分割血管连通区域数量,式中简记 N,

$$N = N_{\text{TP_total}} - 1 \tag{23}$$

式中,$N_{\text{TP_total}}$ 表示对 TP 区域二值图做 8 连通域分析得到的所有连通区域总数,包括背景连通区域,

默认背景连通区域为 1。TP_NCC 越小说明血管分割的连通性越好,即血管断裂少。

TP_VLC 表示正确分割血管连通区域的平均像素,式中简记 V,

$$V = \frac{V_{\text{TP_vessel}}}{N} \tag{24}$$

式中:$V_{\text{TP_vessel}}$ 表示正确分割血管区域的二值分割图中血管区域的总像素,也就是 TP 区域值为 255 的像素的总数。TP_VLC 越大说明每个血管片段的平均长度越长,分割结果的连续性越好。

2.4　实验结果与分析

为了准确评估 EFEU-Net 在视网膜血管分割任务中的表现,分别在 DRIVE、CHASE_DBI、STARE 数据集上,与 U-Net、U-Net + +[17]、Attention U-Net、Dense U-Net[18] 和 DSCNet[19] 等算法进行比较,实验结果见表 1 ~ 3。

表 1　EFE U-Net 与其他方法在 DRIVE 数据集上的比较

Tab. 1　Comparison of EFE U-Net and other methods on the DRIVE dataset

网络	F1 分数	Sensitivity	Specificity	Accuracy	AUC
U-Net	0. 816 09	0. 791 78	0. 978 31	0. 954 57	0. 974 64
U-Net + +	0. 812 84	0. 782 12	0. 979 24	0. 954 15	0. 972 86
Attention U-Net	0. 815 04	0. 792 52	0. 977 80	0. 954 21	0. 974 31
Dense U-Net	0. 812 63	0. 780 76	0. 979 46	0. 954 17	0. 974 23
DSCNet	0. 827 36	0. 804 04	0. 979 63	0. 957 28	0. 981 02
EFE U-Net	0. 830 91	0. 826 49	0. 976 24	0. 957 18	0. 981 55

表 2　EFE U-Net 与其他方法在 CHASE_DB1 数据集上的比较

Tab. 2　Comparison of EFE U-Net and other methods on the CHASE_DB1 dataset

网络	F1 分数	Sensitivity	Specificity	Accuracy	AUC
U-Net	0. 831 89	0. 805 41	0. 986 24	0. 969 04	0. 984 27
U-Ner + +	0. 831 36	0. 814 22	0. 984 81	0. 968 58	0. 984 92
Attention U-Net	0. 830 89	0. 807 13	0. 985 74	0. 968 76	0. 984 35
Dense U-Net	0. 834 80	0. 815 85	0. 985 42	0. 969 29	0. 985 42
DSCNet	0. 838 79	0. 817 60	0. 986 14	0. 970 11	0. 987 82
EFE U-Net	0. 844 98	0. 844 19	0. 983 92	0. 970 54	0. 988 72

表 3　EFE U-Net 与其他方法在 STARE 数据集上的比较

Tab. 3　Comparison of EFE U-Net and other methods on the STARE dataset

网络	F1 分数	Sensitivity	Specificity	Accuracy	AUC
U-Net	0. 799 49	0. 770 07	0. 983 72	0. 963 57	0. 977 38
U-Net + +	0. 808 68	0. 787 90	0. 983 26	0. 964 82	0. 978 11
Attention U-Net	0. 804 22	0. 778 65	0. 983 57	0. 964 25	0. 978 38
Dense U-Net	0. 791 74	0. 743 80	0. 985 93	0. 963 10	0. 975 37
DSCNet	0. 820 74	0. 793 99	0. 985 34	0. 967 29	0. 986 67
EFE U-Net	0. 828 49	0. 844 59	0. 979 77	0. 96702	0. 986 96

表 1 展示了各算法在 DRIVE 数据集上的性能指标,可以看出,EFE U-Net 在 F1 分数、灵敏度、准确度和 AUC 等关键指标上相较于 U-Net 分别提升了 0.014 8、0.034 7、0.002 6 和 0.006 9。EFE U-Net 在灵敏度和 F1 分数上取得了最佳性能,而在 AUC 上则次之。这表明 EFE U-Net 在识别视网膜血管方面具有较高的准确性和鲁棒性,能够有效地识别血管像素,并在召回率和精确率之间取得良好的平衡。尽管 EFE U-Net 的特异性(Spe)略低于其他算法,但其在其他指标上的优越性能表明了该算法在视网膜血管分割任务中的总体稳健性。

表 2 展示了各算法在 CHASE_DB1 数据集上的对比结果。EFE U-Net 在除特异性外的所有指标上均优于其他算法。与 U-Net 相比,EFE U-Net 在 F1 分数、灵敏度、准确度和 AUC 上分别提升了 0.013 1、0.038 8、0.001 5 和 0.004 5,进一步证实了 EFEU-Net 在不同数据集上的性能提升是显著的,能够有效地分割视网膜血管。

表 3 展示了各算法在 STARE 数据集上的对比结果。EFE U-Net 相比与 U-Net 在 F1 分数、灵敏度、准确度和 AUC 分别提升了 0.029 0、0.074 5、0.003 5、0.009 6。

综合表 1 ~ 3 的结果,EFE U-Net 在关键性能指标上的表现优于现有的主流算法,证明了其在视网膜血管分割任务中具有一定的有效性和优越性。能够在一定程度上减少对视网膜血管分割过程中产生的断裂和对于细小血管分割遗漏的现象。

表 4 展示了不同算法在连通性和连续性的对比结果,可以看出本文提出的方法在 TP_NCC 和 TP_VLC 指标上均达到了最好,说明此方法相较于其他方法能够在一定程度上减少血管断裂和对细小血管遗漏的现象。

表 4 EFE U-Net 与其他方法在连通性和连续性的比较

Tab. 4 Comparison of EFE U-Net with other methods in connectivity and continuity

数据集	评价指标	U-Net	U-Net + +	Attention U-Net	Dense U-Net	DSCNet	EFE U-Net
DRIVE	TP_NCC	179.40	182.75	175.05	188.75	150.05	152.60
	TP_VLC	136.80	130.81	139.99	127.00	162.11	170.01
CHASE_DB1	TP_NCC	73.00	76.50	75.88	76.13	69.13	65.25
	TP_VLC	737.70	712.30	722.40	719.05	794.67	911.73
STARE	TP_NCC	61.56	64.00	58.00	67.44	61.22	42.33
	TP_VLC	358.97	347.04	375.52	320.34	373.10	576.92

2.5 可视化对比分析

如图 6 所示,通过对比分析不同算法在 DRIVE 数据集上分割结果的图像细节,可以观察到 EFE U-Net 算法在视网膜血管分割方面的优势,特别是在提取视网膜血管中的毛细血管方面展现了较好的性能。EFE U-Net 能够有效地捕捉到细小的血管结构,这些结构在其他算法的分割结果中往往难以辨识。可以看到,EFE U-Net 在保持血管连续性方面相较于其他算法有着更好的表现,一定程度上解决了细小血管断裂的问题。上述实验结果进一步验证了 EFE U-Net 在视网膜血管分割任务中的有效性,特别是在处理复杂的血管网络和细微结构时。EFE U-Net 通过其先进的网络结构和优化策略,显著提高了对视网膜血管的分割质量,这对于眼科疾病的早期诊断和治疗规划具有重要意义。

图 7 展示了本文模型与其他对比模型的 ROC 曲线和 PRC 曲线对比图,可以发现,本文所提方法在 ROC 和 PRC 曲线上均高于其他方法。

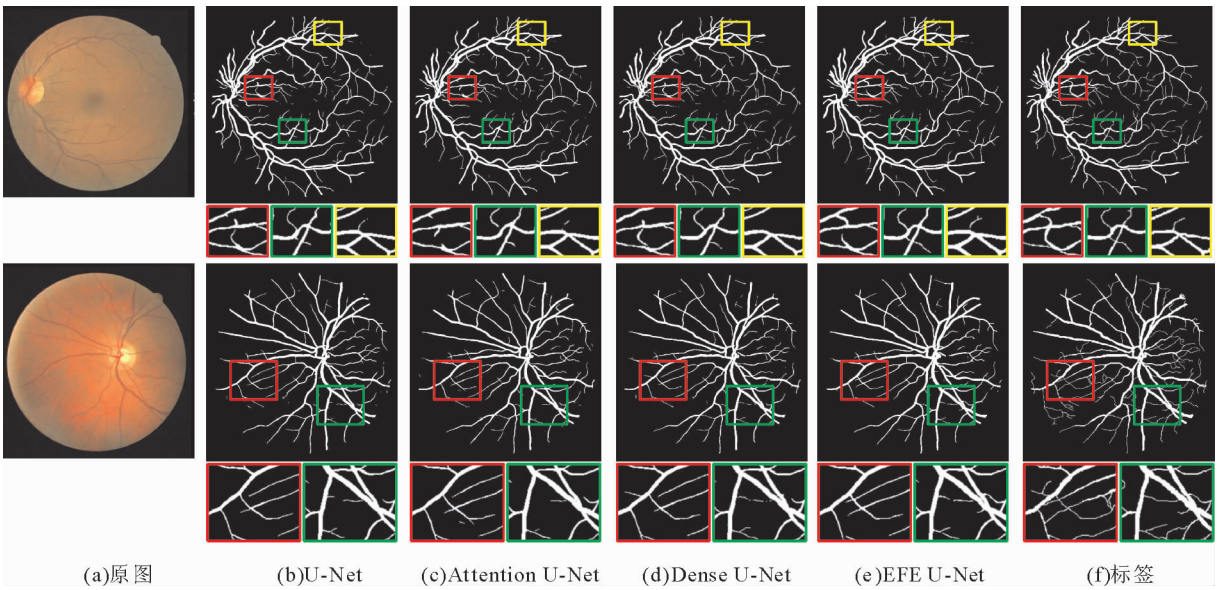


图 6　EFE U-Net 在 DRIVE 数据集上与其他方法可视化对比

Fig. 6　Visual comparison of EFE U-Net with other methods on the DRIVE dataset

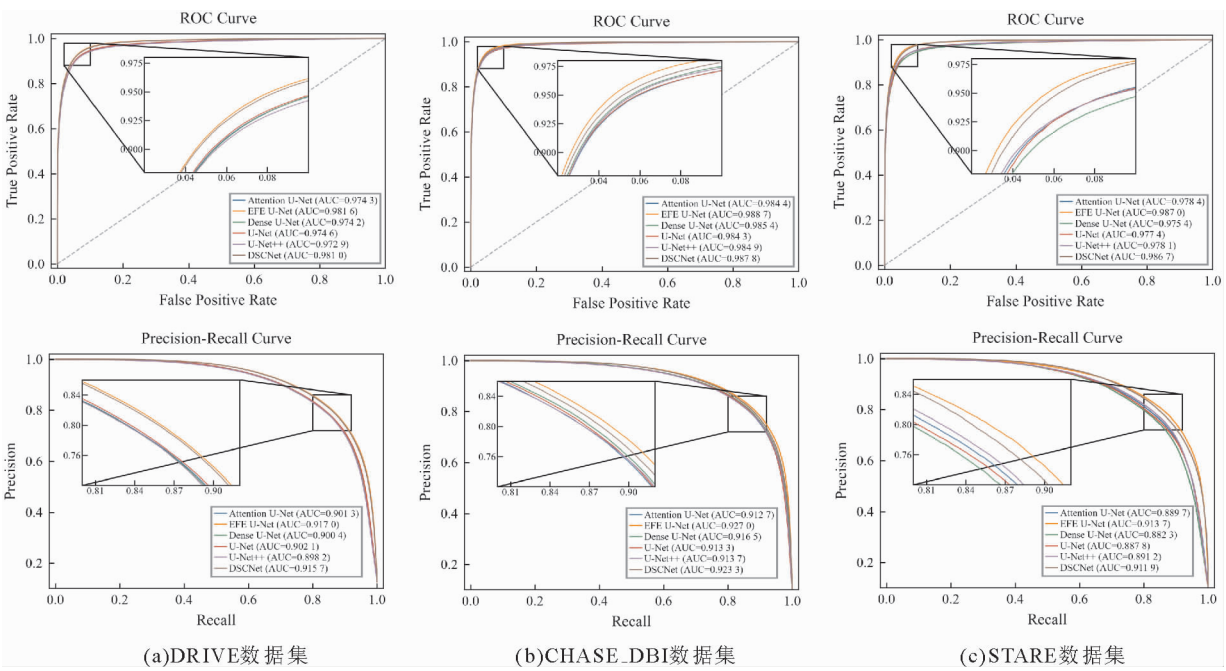


图 7　不同模型 ROC 和 PRC 曲线对比

Fig. 7　Comparison of ROC and PRC curves of different models

通过在不同数据集上的测试可以发现,EFE U-Net 分割结果中的特异性低于有些方法,这说明本文方法在分割过程中对于非血管像素的识别还存在一定的不足,如图 8 所示。

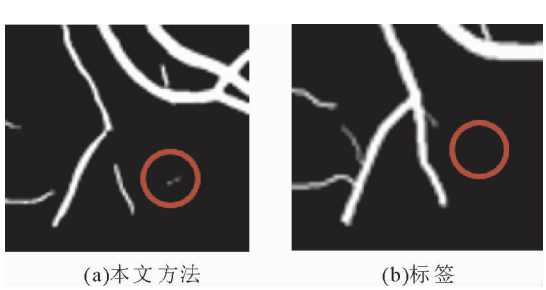


图 8　分割错误示例

Fig. 8　Example of segmentation error

2.6 消融实验

为了验证 EFE U-Net 模型中各组件对模型整体的影响,分别在 DRIVE、CHASE_DB1 和 STARE 数据集上进行消融研究,实验结果如表 5 ~7 所示。其中:M1 表示基准模型也就是 U-Net;M2 表示没有使用边缘特征融合模块(EFFM)和通道空间注意力模块(CSAM)的边缘特征增强模块的网络;M3 表示没有使用边缘特征融合模块(EFFM)的边缘特征增强模块的网络;M4 表示没有使用通道空间注意力模块(CSAM)的边缘特征增强模块的网络;M5 是使用了边缘特征融合模块(EFFM)和通道空间注意力模块(CSAM)的边缘特征增强模块的网络,也就是本文提出的 EFE U-Net。从表 5 ~7 可以看出,在视网膜血管分割任务中,本文算法具有良好的分割效果,可以有效改进 U-Net 在视网膜血管分割任务中的不足。

表 5 DRIVE 数据集上的消融实验结果

Tab. 5 Results of ablation experiments on the DRIVE dataset

网络	F1 分数	Sensitivity	Specificity	Accuracy	AUC
M1	0.816 09	0.791 78	0.978 31	0.954 57	0.974 64
M2	0.825 15	0.800 20	0.979 68	0.956 83	0.979 71
M3	0.828 54	0.817 74	0.977 22	0.956 92	0.980 68
M4	0.825 44	0.801 58	0.979 49	0.956 84	0.979 75
M5	0.830 91	0.826 49	0.976 24	0.957 18	0.981 55

表 6 CHASE_DB1 数据集上的消融实验结果

Tab. 6 Results of ablation experiments on the CHASE_DB1 dataset

网络	F1 分数	Sensitivity	Specificity	Accuracy	AUC
M1	0.831 89	0.805 41	0.986 24	0.969 04	0.984 27
M2	0.835 24	0.805 82	0.986 99	0.969 76	0.986 60
M3	0.823 68	0.827 52	0.981 07	0.966 59	0.986 73
M4	0.810 81	0.777 45	0.985 39	0.965 78	0.983 35
M5	0.844 98	0.844 19	0.983 92	0.970 54	0.988 72

表 7 STARE 数据集上的消融实验结果

Tab. 7 Results of ablation experiments on the STARE dataset

网络	F1 分数	Sensitivity	Specificity	Accuracy	AUC
M1	0.799 49	0.770 07	0.983 72	0.963 57	0.977 38
M2	0.809 80	0.785 22	0.983 95	0.965 21	0.984 06
M3	0.823 68	0.827 52	0.981 07	0.966 59	0.986 73
M4	0.810 81	0.774 51	0.985 39	0.965 78	0.983 35
M5	0.828 49	0.844 59	0.979 77	0.967 02	0.986 96

3 结语

针对视网膜血管结构复杂、在分割过程中易对微小血管分割忽视和断裂的问题,本文提出基于边缘特征增强的 U-Net 视网膜分割方法(EFE U-Net)。EFEM 模块增强了图像中的边缘特征信息,从而改善传统卷积方式不能有效提取和利用边缘特征的问题,进而减少在分割过程中会产生的视网膜细小血管丢失和部分血管分割结果断裂的问题,使模型能更好地关注视网膜血管的结构特征,利用边缘特征融合模块对图像中的边缘信息进一步增强,使用通道空间注意力抑制不相关信息。通过对比实验验证,本文方法相较于其他方法有明显的优势,实现了对于视网膜血管的有效分割。未来的研究重点将放在模型轻量化上,在保持高准确度的同时,减少模型的计算负担,使其更适合在资源受限的环境中部署,并可以对模

型进行优化,使用模型剪枝技术,去除冗余或对模型性能影响较小的部分,从而减少参数量,或使用知识蒸馏方法,将复杂的教师模型的知识迁移到一个较小的学生模型中。

参考文献

[1] Abdulsahib A A, Mahmoud M A, Abed M M, et al. Comprehensive review of retinal blood vessel segmentation and classification techniques: Intelligent solutions for green computing in medical images, current challenges, open issues, and knowledge gaps in fundus medical images[J]. Network Modeling Analysis in Health Informatics and Bioinformatics, 2021, 10(1): 20.

[2] Li L Z, Verma M, Nakashima Y, et al. IterNet: Retinal image segmentation utilizing structural redundancy in vessel networks[C]// 2020 IEEE Winter Conference on Applications of Computer Vision (WACV). Snowmass: IEEE, 2020: 3645-3654.

[3] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4):640-651.

[4] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention. Cham: Springer International Publishing, 2015: 234-241.

[5] Zhuang J T. LadderNet: Multi-path networks based on U-Net for medical image segmentation[EB/OL]. (2018-10-17)[2026-01-05]. https://arxiv.org/abs/1810.07810.

[6] Oktay O, Schlemper J, Le Folgoc L, et al. Attention U-Net: Learning where to look for the pancreas[EB/OL]. (2018-04-11)[2026-01-05]. https://arxiv.org/abs/1804.03999.

[7] Jin Q G, Meng Z P, Pham T D, et al. DUNet: A deformable network for retinal vessel segmentation[J]. Knowledge-Based Systems, 2019, 178:149-162.

[8] 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.

[9] Zhang Z J, Fu H Z, Dai H, et al. ET-Net: A generic edge-attention guidance network for medical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention. Cham: Springer International Publishing, 2019: 442-450.

[10] Pan J H, Zhai Y K, Lyu H L, et al. MSFA-Net: Multiple spatial-channel feature aggregation network for change detection and a UAV-CD dataset[C]// IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium. Athens: IEEE, 2024: 10328-10332.

[11] 张瑞, 唐乔湛,李斯卉, 等. 基于多尺度特征和反向注意力的肝脏肿瘤自动分割方法[J]. 西北大学学报(自然科学版), 2023,53(6):964-973.
Zhang R, Tang Q Z, Li S H, et al. Automatic liver tumor segmentation using multi-scale features and reverse attention[J]. Journal of Northwest University (Natural Science Edition), 2023, 53(6): 964-973.

[12] Ghiasi G, Lin T Y, Le Q V. DropBlock: A regularization method for convolutional networks[C]// Advances in Neural Information Processing Systems. Montreal: NeurIPS Foundation, 2018: 10727-10737.

[13] Zhao S J, Zhang X L, Xiao P F, et al. Exchanging dual-encoder-decoder: A new strategy for change detection with semantic guidance and spatial localization[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-16.

[14] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.

[15] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]// Computer Vision-ECCV 2018. Cham: Springer, 2018: 3-19.

[16] Huang G, Li Y X, Pleiss G, et al. Snapshot ensembles: Train 1, get M for free[EB/OL]. (2017-04-01)[2026-01-05]. https://arxiv.org/abs/1704.00109.

[17] Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]// Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3-11.

[18] Li X M, Chen H, Qi X J, et al. H-DenseUNet: Hybrid densely connected UNet for liver and liver tumor segmentation from CT volumes[J]. IEEE Transactions on Medical Imaging, 2018, 37(12): 2663-2674.

[19] Qi Y L, He Y T, Qi X M, et al. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation[C]// 2023 IEEE/CVF Interrational Conference on Computer Vision. Pairs: IEEE, 2023: 6047-6056.

(编　辑　李　静)