



GCFF-Net: 一种面向视网膜血管精细分割的多层级图卷积特征融合神经编解码网络

崔少国*, 张乐迁, 文浩

(重庆师范大学计算机与信息科学学院, 重庆 401331)

摘要 眼底视网膜血管分割可以帮助医生对于眼科疾病以及心脑血管疾病进行辅助诊断。然而, 由于血管拓扑结构复杂, 边界不清晰, 大大增加分割难度。针对这些问题, 在U型结构的基础上提出了一种基于图卷积特征融合网络。该网络使用图卷积模块对编码器特征中像素之间的全局上下文信息进行建模, 弥补普通卷积缺少全局建模的能力; 然后使用多尺度特征融合模块对编码器特征与解码器特征进行融合, 来减少特征层中噪声信息对分割结果的影响; 最后使用多层次特征融合模块将解码器每一层特征融合输出, 减少下采样过程中空间信息的缺失以及对深层特征的复用。在公开数据集 DRIVE、CHASEDB1 以及 STARE 上进行验证, F1 值分别为 82.90%、83.79% 和 80.63%, AUC 分别为 98.13%、98.78% 和 98.36%。实验结果表明, 所提方法优于近些年视网膜分割算法。

关键词 医学图像分割; 视网膜血管; U型结构; 图卷积; 特征融合

中图分类号 TP391.41

文献标志码 A

DOI 10.12178/1001-0548.2023131

GCFF-Net: A Multilayer Graph Convolutional Feature Fusion Neural Encoding and Decoding Network for Fine Segmentation of Retinal Vessels

CUI Shaoguo*, ZHANG Leqian, and WEN Hao

(College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China)

Abstract Fundus retinal vessels segmentation can assist doctors in the diagnosis of ophthalmic diseases and cardiovascular and cerebrovascular diseases. However, due to the complex topological structure of blood vessels and unclear boundaries, it greatly increases the difficulty of segmentation. A graph convolutional feature fusion network is proposed based on the U-shaped structure to address these issues. This network uses a graph convolution module to model the global contextual information between pixels in encoder features, making up for the lack of global modeling ability in ordinary convolutions; Then, a multi-scale feature fusion module is used to fuse the encoder features and decoder features to reduce the impact of noise information in the feature layer on the segmentation results; Finally, a multi-level feature fusion module is used to fuse and output the features of each layer of the decoder, reducing the loss of spatial information and the reuse of deep features during the downsampling process. Verified on the public datasets DRIVE, CHASEDB1, and STARE, the F1 values were 82.90%, 83.79%, and 80.63%, respectively, and the AUC values were 98.13%, 98.78%, and 98.36%, respectively. The experimental results show that the proposed method is superior to recent retinal segmentation algorithms.

Key words medical image segmentation; retinal vessels; U-shaped architecture; graph convolution; feature fusion

眼底是人类唯一可以直接观察动脉、静脉和毛细血管的部位。糖尿病、高血压以及各种眼部疾病都会造成眼底血管几何形态的变化^[1]。如糖尿病会导致视网膜血管病变与新生血管, 高血压会导致血

管异常增生等。因此在临床研究中, 视网膜血管分割对帮助医生诊断相关疾病具有重要意义。然而由于视网膜血管拓扑结构复杂以及本身成像的限制, 人为分割血管的难度较高, 因此, 提出一种高效、

收稿日期: 2023-05-10; 修回日期: 2024-03-21

基金项目: 国家自然科学基金(62003065); 重庆市科技局自然科学基金面上项目(CSTB2022NSCQ-MSX1206); 重庆市技术预见与制度创新项目(CSTB2022TFII-OFX0042); 教育部人文社科规划基金(22YJA870005); 重庆市教委重点项目(KJZD-K202200510); 重庆市社会科学规划项目(2022NDYB119); 重庆市教委人文社科项目(23SKGH072); 重庆师范大学人才基金(20XLB004)

作者简介: 崔少国, 博士, 教授, 主要从事大数据与人工智能等方面的研究。

*通信作者 E-mail: ailabcui@163.com

准确的视网膜血管分割方法非常重要。

近些年,视网膜血管自动分割算法从早期的无监督方法慢慢过渡到如今的有监督方法,这都归功于深度学习领域中卷积神经网络的飞速发展。在无监督算法中,主要是依据视网膜血管的一些物理特征(形状、颜色与纹理)或者人为定义的特征来对血管进行分割。文献[2]首先确定血管中心线,用作血管填充的依据,然后依据四个方向的差分算子来判断周围像素点是否作为中心线像素的候选点,最后通过迭代区域生长方法得到分割结果。文献[3]提出了一种基于水平集和区域生长的视网膜血管分割方法,首先通过自适应直方图均衡与2D Gabor小波变换进行数据增强,然后使用各向异性扩散滤波器提取血管边界特征并平滑图像,最后使用区域生长方法和基于区域的活动轮廓模型分割视网膜血管。

相较于无监督方法,有监督方法虽然需要人工对数据进行标注,但也取得了更高的分割准确度。文献[4]将典型的视网膜血管分割问题从全尺寸图像分割简化为区域血管元素识别,并且使用经过预训练的 AlexNet 作为 FCN^[5] 的特征提取骨干网络,最终很好地实现了血管像素的端到端分割。虽然 FCN 不限制输入网络的图像大小,但由于下采样操作造成空间信息的损失以及对卷积平移不变性的破坏,导致分割的血管出现断裂。文献[6]通过迭

代 mini-UNet 增加网络深度,提升了模型分割血管细节的能力。文献[7]等人对金字塔池化进行改进,并在跳跃连接处加入注意力机制,有效地减少特征信息的损失。文献[8]在生成对抗网络 GAN^[9] 的基础上,首先对眼底图像进行非均匀光照移除和主成分分析,然后将带偏移量的深度可分离卷积^[10] 和 SE 模块^[11] 同时加到生成器与判别器中,有效的减少了分割结果中血管断裂的现象。文献[12]将 Transformer^[13] 架构引入到 U-net 中,通过对编码阶段特征的通道信息和位置信息联合关注增强编码局部细节特征,提高了网络对血管特征的敏感度。上述基于 U 型架构的网络模型受限于卷积核大小的限制,不能对全局信息进行建模,并且在下采样过程中会造成空间信息的损失,从而造成分割结果中出现血管边界不清晰或出现噪声等现象。

针对以上问题,本文提出了一种基于图卷积的多层级特征融合网络的视网膜血管分割算法。

1 本文算法

1.1 模型总述

本文在 U 形编码器-解码器^[14] 架构的基础上,提出一种图卷积多尺度特征融合的眼底视网膜血管分割网络(Retinal Vessels Segmentation Algorithm Based on Graph Convolutional Feature Fusion Network, GCFF-Net)。网络结构如图 1 所示。

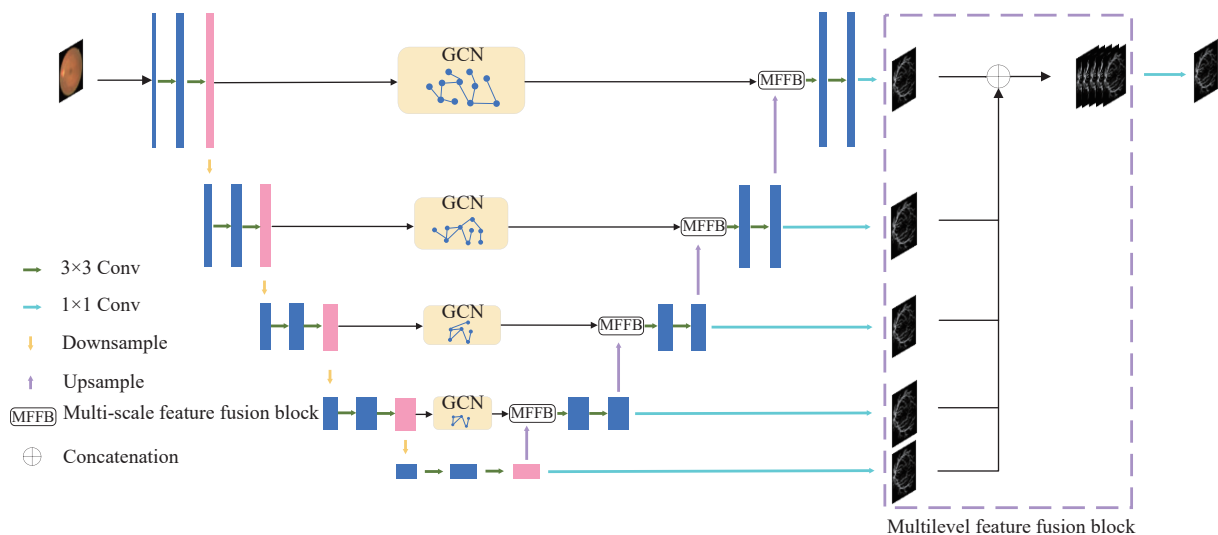


图 1 方法结构图

该网络主要由 4 部分组成,分别为 U 型编码器,图卷积模块,多尺度特征融合模块以及多层次特征融合模块。编码器的每一层主要由两个

3×3 的卷积层提取特征,在两层之间使用最大池化层进行下采样,在降低图像分辨率的同时增加特征图所蕴含的语义信息。在编码器与解码器之间的跳

跃连接处加入图卷积模块, 通过 K-近邻算法将编码器提取到的特征图转化为图结构, 并使用图卷积进行特征变换, 对图像的全局上下文信息进行建模, 并捕获像素之间的长距离依赖关系, 从而减少分割结果中出现血管断裂的现象。解码器的每一层主要由一个多尺度特征融合模块与两个 3×3 的卷积层组成, 多尺度特征融合模块通过计算图卷积特征层与普通卷积层特征的通道注意力权重对两种特征进行融合, 对特征层中的噪声信息进行抑制并注重血管的特征信息, 可以有效地减少分割结果中噪声的出现。解码器的每一层之间使用双线性插值算法代替转置卷积进行上采样, 来减少模型参数量。多层次特征融合模块将解码器每一层首先通过空洞卷积层提取不同尺度特征, 然后经过 sigmoid 激活与 1×1 卷积层变换通道数, 最后使用双线性插值将特征图上采样至输入图像大小并在通道维拼接, 以此来减少编码器下采样过程中空间信息的损失, 并对深层特征加以复用。模型的输出是将融合后的多层次特征经过 1×1 卷积与 softmax 激活得到最后的分割概率图。

1.2 图卷积模块

图卷积网络 (Graph Convolutional Networks, GCN) 是一种针对图结构数据的深度学习模型, 可以用于节点分类、图分类、链接预测等任务。GCN^[15] 的核心思想是将卷积运算扩展到图结构数据, 通过对邻居节点的信息进行聚合, 实现节点的特征更新。文献 [16] 提出一种新型视觉 GNN (ViG) 架构。ViG 首先将图像分割成一组形状规则的补丁, 并将每个补丁视为一个图节点, 通过 K 近邻算法找到每个节点的 K 个最近邻居作为图的边, 然后在每个节点上进行图卷积操作学习节点特征。基于此, 本文在编码器与解码器之间加入 GCN 模块, 将编码器特征图的每个像素都当作一个图节点, 使用 K 邻近算法构图, 通过图卷积操作捕获全局特征信息。GCN 模块架构如下:

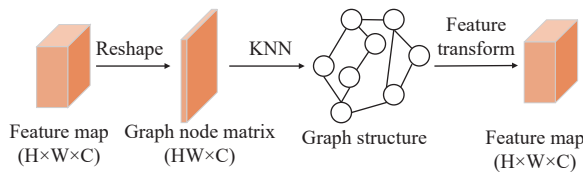


图2 图卷积模块

GCN 模块包含两个部分, 分别为图构造与图卷积计算。

1) 图的构造。首先使用 reshape 函数将编码器层的特征 $X \in \mathbb{R}^{H \times W \times C}$ 转换为图节点矩阵 $X' \in \mathbb{R}^{HW \times C}$ 。然后使用 K 近邻算法得出两两节点之间的距离, 选择每个节点距离最近的 K 个节点作为该节点的相邻节点, 由此可以得到邻接矩阵 A 。

2) 图卷积计算。本文使用拉普拉斯矩阵来定义图卷积。对于一个无向图 $\mathcal{G} = \{v, e\}$, A 代表邻接矩阵, D 代表度矩阵, $D_{ii} = \sum_j A_{ij}$ 。 $L = I - D^{-1/2} A D^{-1/2}$ 表示对 \mathcal{G} 进行归一化操作的拉普拉斯矩阵。 L 可以分解为 $L = U \Lambda U^T$, 其中 U 为特征向量矩阵, $\Lambda = \text{diag}[\lambda_1, \dots, \lambda_n]$ 为特征值矩阵。图卷积网络 (GCN) 引入 ChebNet 的一阶近似 ($K = 1$)。GCN 迭代的聚合来自邻居节点的信息。关于节点 v_i 的前向传播过程为:

$$h_i^{(l+1)} = \sigma \left(\sum_{v_j \in \mathcal{N}(v_i) \cup \{v_i\}} \hat{a}(v_i, v_j) W^{(l)} h_j^{(l)} \right) \quad (1)$$

式中, $\sigma(\cdot)$ 为非线性激活函数; $\hat{A} = (\hat{a}(v_i, v_j))$ 表示重新归一化的邻接矩阵 A ; $W^{(l)}$ 表示第 l 层的可学习变换矩阵。

1.3 多尺度特征融合模块

在 U-Net 中, Skip-connection 解决的问题主要是如何在编码器和解码器之间进行信息传递和整合, 以提高模型对于目标的分割精度。然而, 仅仅使用简单的跳跃连接无法去除编码器与解码器特征中的噪声信息, 并且还会丢失一部分血管信息。因此, 本文提出一种多尺度特征融合模块 (Multi-scale Feature Fusion Block), 增加特征图中血管特征的权重, 并减少噪声的权重, 使编码器与解码器之间特征更好的融合。其结构如下图 3 所示。

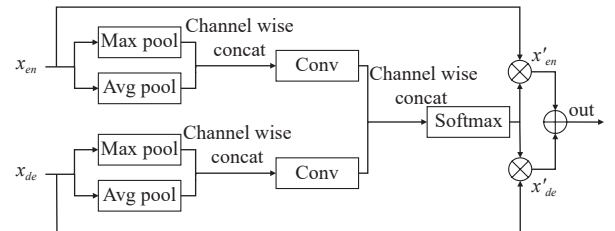


图3 多尺度特征融合模块

其中 x_{en} 与 x_{de} 分别为编码器与解码器的特征图, x'_{en} 与 x'_{de} 是经过权重计算后的特征图。首先, x_{en} 和 x_{de} 分别通过最大池化和的平均池化聚合特征图的空间信息^[17], 然后经过 1×1 卷积变换通道维度, 最后拼接在一起经过 softmax 层得到权重系数矩阵 α 。最后得到的编码器特征与解码器特征计算

过程如下:

$$x'_{en} = \alpha_0 \circ x_{en} \quad (2)$$

$$x'_{de} = \alpha_1 \circ x_{de} \quad (3)$$

式中, \circ 为哈达玛积; α_0 与 α_1 分别为权重系数矩阵 α 的第一、二个通道; α 的计算公式如下:

$$\alpha_0 = \text{Conv}[\text{cat}(\text{avgpool}(x_{en}), \text{maxpool}(x_{en}))] \quad (4)$$

$$\alpha_1 = \text{Conv}[\text{cat}(\text{avgpool}(x_{de}), \text{maxpool}(x_{de}))] \quad (5)$$

$$\alpha = \text{softmax}(\text{cat}(\alpha_0, \alpha_1)) \quad (6)$$

1.4 多层次特征融合模块

在 U-Net 中, 编码器通过多次下采样和卷积操作, 逐步提取出图像的高级特征, 而解码器则通过多次上采样或者转置卷积操作, 将这些特征转换成分割结果。然而, 在这个过程中, 随着特征图分辨率不断减小, 在增加语义信息的同时会造成空间信

息的损失, 因此, 解码器需要在重建图像时从编码器中获取更多的上下文信息, 以提高分割的准确性。基于此, 本文提出多层次特征融合模块 (Multilevel Feature Fusion Block), 其结构如下图 4 所示。

为了弥补下采样过程中空间信息的损失以及充分利用解码器每一层的特征信息, 该模块首先通过使用不同空洞率的卷积层^[18]来提取不同尺度的特征, 每一层的空洞率分别为 16, 8, 4, 2, 1, 再通过 sigmoid 激活函数得到特征权重 α , 特征层 x 与 α 相乘得到该层解码器输出, 然后通过 1×1 卷积与上采样将解码器每一层的输出投射到相同的特征空间, 每一层的上采样率分别为 1, 2, 4, 8, 16, 即:

$$\hat{x}_{de-out}^i = \text{upsampling}(x_{de-out}^i), i = 1, 2, 3, 4, 5 \quad (7)$$

最后将每一层输出在通道维度拼接起来得到多层次特征融合层。

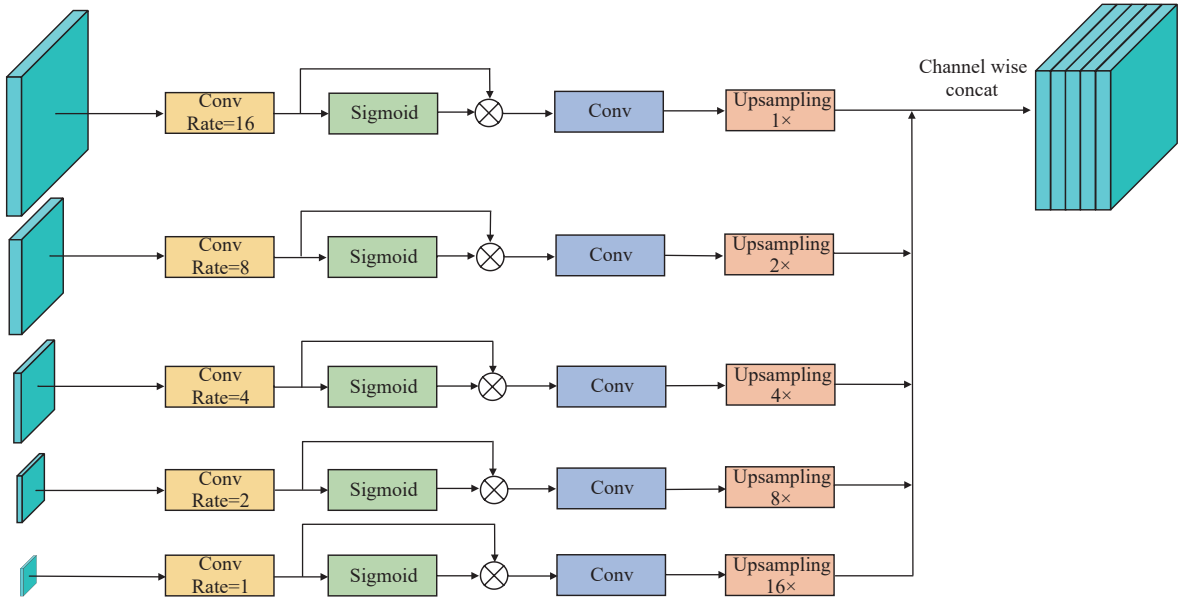


图 4 多层次特征融合模块

2 实验结果及分析

2.1 实验设置

本文实验环境为基于 12th Gen Intel(R) Core(TM) i7-12700 @2.10 GHz CPU 16G 内存, 显卡为 NVIDIA RTX A4000 GPU 16 G, 采用 Windows 10 系统。该模型代码基于 Pytorch 深度学习框架, 训练过程中初始学习率为 0.0005, 选用的优化器为 Adam, 损失函数设置为交叉熵损失加上 DICE 损失的联合损

失, 根据显卡性能选择送入网络的图片批量大小 batch_size 为 16, 使用余弦退火算法作为学习率衰减策略。此外若连续 10 个 epoch 未取得最优, 则停止训练。

2.2 数据集与预处理

本实验采用的数据集包括 DRIVE、STARE 以及 CHASEDB1。DRIVE 数据集的照片来自荷兰的糖尿病视网膜病变筛查项目, 其中包含 40 幅图像, 每幅图像大小为 768×584 。STARE 数据集来

自加州大学圣地亚哥分校, 包含 20 幅大小为 605×700 的图像。CHASEDB1 数据集包含 28 幅大小为 999×960 像素的彩色视网膜图像, 这些图像采集自 14 名学童的左眼和右眼。每幅图像都由两位独立的专家注释。

由于在视网膜血管图像中, 绿色通道图像的血管与背景对比度较高, 因此将原图分离为 RGB 三通道特征图并进行加权处理, 然后使用直方图均衡化将视网膜图像的直方图近似均匀分布, 从而增强图像的对比度, 最后进行 Gamma 变换, 更有效的保留图像亮度信息。为了防止神经网络出现欠拟合现象, 在本次研究中, 3 个数据集都只有几十张图片, 因此需要对数据集进行扩充: 对输入图像进行随机水平、垂直翻转, 分块。本实验将训练图像以及标签切分为 64×64 大小的像素块, 每个像素块间隔 16 个像素点。最终预处理结果如下图 5 所示。

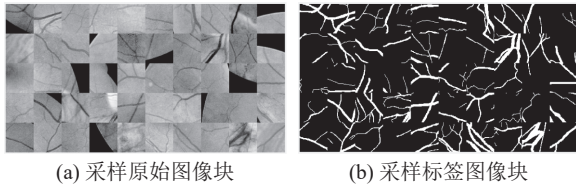


图 5 数据预处理最终结果

2.3 评价指标

为了定量评估本文模型的分割性能, 使用 F1-score、准确率 ACC、ROC 曲线下的 AUC 面积、特异性 SP (正确分类非血管像素占总体非血管像素的比例)、灵敏性 SE (正确分类血管像素占总体血管像素的比例) 作为评价指标:

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (8)$$

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

$$Se = \frac{TP}{TP + FN} \quad (10)$$

$$Sp = \frac{TN}{TN + FP} \quad (11)$$

式中, TP, TN, FP, FN 分别为真阳性, 真阴性, 假阳性, 假阴性。

2.4 不同模型分割性能对比

为了验证本文算法在视网膜血管图像上的分割性能, 将 GCFF-Net 与 U-Net、Attention U-Net^[19] 两个模型在 DRIVE、CHASEDB1、STARE 三个数据集上进行比较。为公平比较 3 个模型分割性能, 对比实验采用相同的实验环境。性能指标对比与分割效果对比分别如表 1 与表 2 所示。

表 1 不同模型在 DRIVE、CHASEDB1 和 STARE 上的性能指标比较

Dataset	Model	F1	SE	SP	ACC	AUC(ROC)
DRIVE	U-Net	0.824 8	0.821 3	0.975 1	0.955 5	0.977 3
	Attention U-Net	0.825 4	0.838 9	0.971 7	0.954 8	0.980 5
	GCFF-Net(ours)	0.829 0	0.834 5	0.973 9	0.957 3	0.981 3
CHASEDB1	U-Net	0.822 4	0.821 3	0.978 2	0.960 9	0.983 7
	Attention U-Net	0.832 8	0.832 0	0.977 8	0.963 1	0.987 3
	GCFF-Net(ours)	0.837 9	0.842 7	0.979 1	0.964 0	0.987 8
STARE	U-Net	0.800 2	0.728 5	0.989 6	0.963 2	0.980 7
	Attention U-Net	0.803 2	0.754 6	0.986 0	0.962 6	0.982 1
	GCFF-Net(ours)	0.806 3	0.778 6	0.981 4	0.959 5	0.983 6

表 2 不同模型复杂度对比

model	Params(M)	Times(ms)
U-Net	34.525 9	3.400 7
Attention-U-net	34.877 4	6.447 8
GCFF-Net	34.879 3	10.155 9

2.4.1 分割性能指标对比

由表 1 可知, 由于 Attention U-Net 在 U-Net 的基础上加入基于网格的注意力门机制, 增加了

模型对前景像素的敏感度, 在表 1 中 DRIVE 和 CHASEDB1 数据集体现为 SE 指标的上升, 在一定程度上提高了提取血管像素的能力, 然而 Attention U-Net 过于注重前景像素会造成过度分割的问题, 即将背景像素错误分类为前景像素, 从而导致 SP 指标的下降。本文算法在 SE 与 SP 指标上都取得了较高的得分, 且 F1 值与 AUC 值在 3 个数据集上均由于其他方法, 说明本文算法的分割结果与

金标准的相似程度最高。

图 6、图 7 与图 8 分别为上述 3 个网络模型在 DRIVE、CHASEDB1 和 STARE 数据集上的 ROC 曲线与 PR 曲线，可以看出本文算法较其他模型的 AUC(ROC) 值与 AUC(PR) 值分别平均高出 0.24%、0.63%、0.22% 和 0.23%、1.13%、0.93%，说明本文算法分割血管的能力优于其他网络。

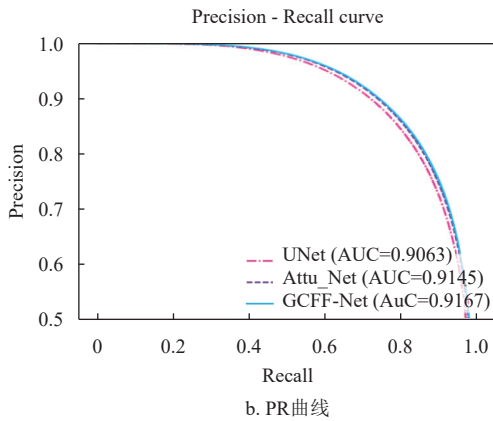
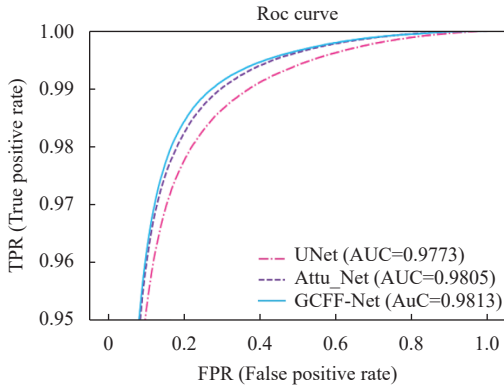


图 6 不同模型在 DRIVE 数据集上 ROC 曲线与 PR 曲线对比图

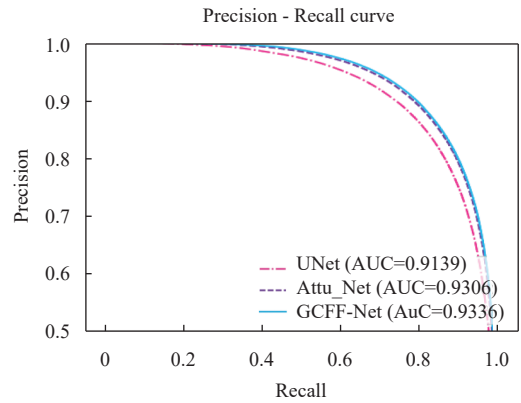
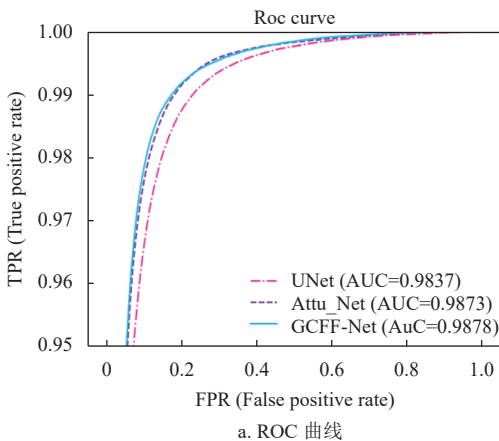


图 7 不同模型在 CHASEDB1 数据集上 ROC 曲线与 PR 曲线对比图

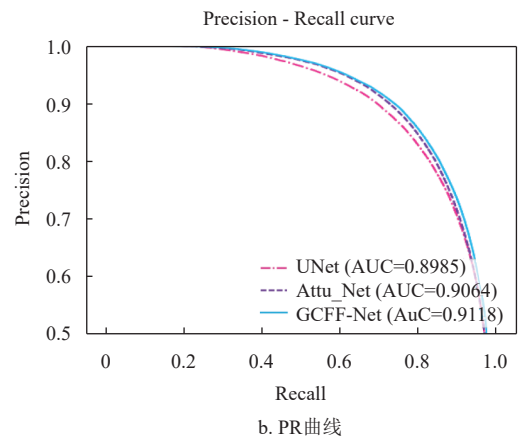
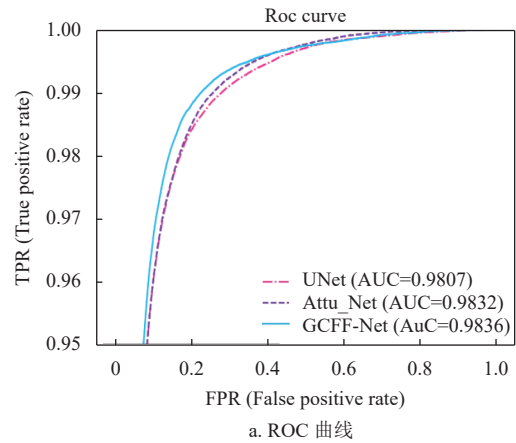


图 8 不同模型在 STARE 数据集上 ROC 曲线与 PR 曲线对比图

2.4.2 分割效果图对比

图 9 从 DRIVE、CHASEDB1 和 STARE 三个数据集中分别挑选出一张图片进行分割效果对比，每幅图片都有两个对比的矩形局部区域。表中第一行 DRIVE 数据集的分割结果中，由于血管末端的小血管太薄，U-Net 与 Attention U-Net 都将部分背

景像素分类成血管像素, 导致局部分割图中出现些许噪声, 这是因为这两种算法直接将编码器特征与解码器特征拼接, 无法去除卷积层提取到的特征中的噪声信息, 而本文算法在跳跃连接处加入的多尺度特征融合模块, 可以有效减少编码器与解码器特征中的噪声信息, 同时增加血管特征的权重, 从而避免分割结果中噪声的出现。表中第二行 CHASEDB1 数据集的分割结果中, 因为光照不足或者过度曝光, 使得图像对比度降低, 进而造成视网膜血管边界不清晰和血管信息丢失, 从矩形局部分割图中可以看出, U-Net 与 Attention U-Net 使用普通卷积只能提取局部特征, 虽然 Attention U-Net 在卷积层中加入注意力机制提升了提取血管像

素的能力, 缓解了血管断裂的问题, 但也引入了少许噪声, 而本文算法在编码器与解码器之间加入图卷积模块, 来捕获血管像素之间的长距离依赖关系, 从而有效减少了分割结果中血管断裂的现象。表中第三行 STARE 数据集中, 对比区域为视盘区域, 是整幅图像中血管最集中的地方, 该处的血管分割难度更大。虽然 U-Net 与 Attention U-Net 可以很好的分割出血管轮廓, 但也出现了血管边界不清晰的问题, 这是由于编码器在下采样过程中会丢失空间信息, 而本文算法提出的多层次特征融合模块可以有效弥补空间信息的损失以及充分利用编解码器每一层的特征, 在血管拓扑结构复杂的情况下, 依然能够准确地分割出血管结构。

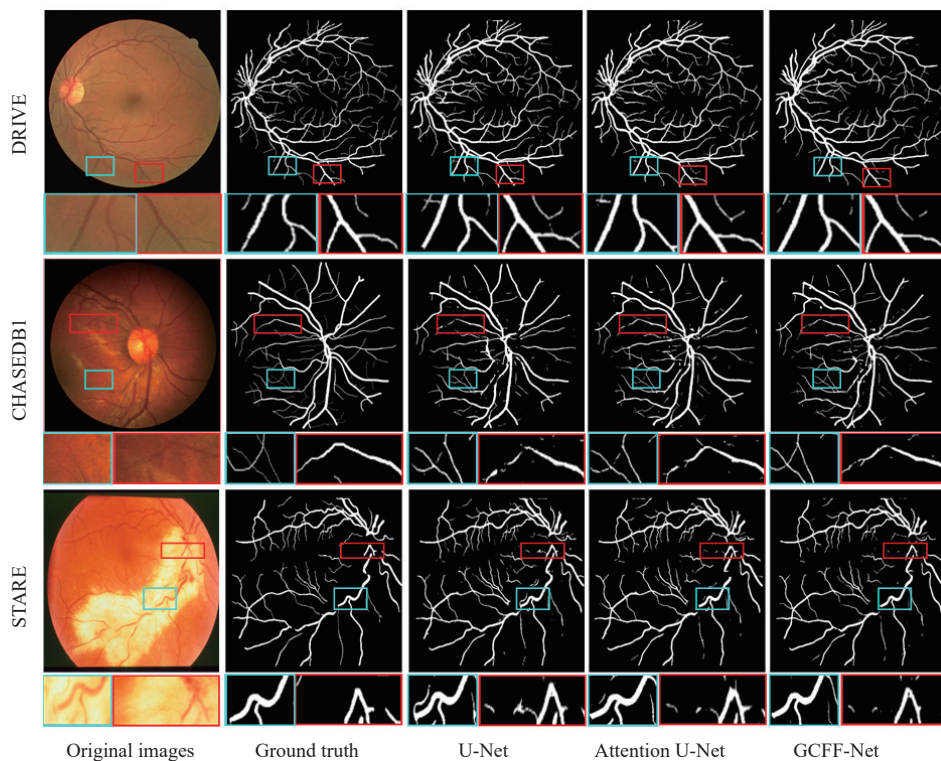


图9 不同模型分割效果对比图

综上所述, 本文算法在 DRIVE、CHASEDB1 与 STARE 数据集上都取得了优良的分割结果。相较于其他分割算法, 本文算法分割出的血管结构更加清晰, 不易出现血管断裂的现象, 位于血管末端的小血管也能实现较为精确的分割。同时在由于光照不足造成图像对比度低的情况下, 本文算法也能避免伪影的干扰, 捕获完整的血管。实验结果证明, 本文算法能出色的完成视网膜血管分割的任务。

2.4.3 模型复杂度对比

本节主要从时间复杂度与空间复杂度两个方面

对模型的复杂度进行分析。其中空间复杂度主要通过模型的参数量来衡量; 时间复杂度通过将 (1, 1, 64, 64) 大小的张量输入模型, 并迭代 100 次, 最后计算平均每次迭代所使用的时间。

由表 2 可知, 本文提出的模型相较于 U-Net 参数量仅增加了 1.02%, 且与 Attention-UNet 的参数量相差无几。在模型的推理速度方面, 由于本模型的图卷积模块中使用 K 近邻算法计算每个像素之间的距离来达到构图的目的, 因此推理时间较 UNet 与 Attention-UNet 有所增加。综上所述, 由

于本文模型的优异性能,因此模型在时间复杂度与空间复杂度上稍许增加是可以接受的。

2.4.4 与其他先进算法对比

为了进一步验证本文算法在视网膜血管上的分割能力,先择了近几年的先进算法在 F1、SE、SP、ACC、AUC 这 5 个指标上进行对比,对比结果如表 3-5 所示。MLA-Net 提出自适应通道注意力模块对特征通道赋以权重,以此来减少特征中冗余的信息,但其 SE 值在 3 个数据集上均低于本文,且在 CHASEDB1 与 STARE 数据集上的 F1 值与 AUC 值也低于本文,其中 F1 值分别比本文低 0.11% 和 1.31%,AUC 值分别比本文低 0.66% 和 0.10%,说明多尺度特征融合模块对特征图血管噪声的抑制要强于自适应通道注意力模块。DU-Net 使用可变形卷积代替 U-Net 中的传统卷积,从而学习局部、密集和自适应感受野来模拟各种形状的和尺度的视网膜血管,虽然可变形卷积可以依据输入图像改变卷积核的感受野,但依然很难关注到

血管这样细小的结构,而本文所使用的图卷积模块通过捕获血管像素之间的长距离依赖关系,可以更好地关注血管的结构信息,因此本文算法在 DRIVE 与 CHASEDB1 数据集上的各指标均高于 DU-Net。ConvMixer 借用 ViT 的思想,将输入图像划分为补丁,并使用深度卷积与点卷积分别在空间和通道维度上混和提取的特征,然而其 F1 值仍低于本文算法,且该模型在整个网络中保持相同的大小和分辨率,虽然有利于减少上采样与下采样所造成的信息损失,但也增加了模型的参数量与计算开销。SA-Unet 引入空间注意力模块使特征图自适应特征细化,是模型更加关注血管特征,但其 SE 值比本文低 1.33%,说明其提取血管特征的能力弱于本文所提出的多尺度特征融合模块。综合表 3~5 可知,本文算法的分割性能优于其他网络,尤其是 F1 值在三个数据集上都达到最优,体现了本文算法更加稳健。

表 3 不同算法在 DRIVE 数据集上的性能指标

Algorithm	F1	SE	SP	ACC	AUC
R2U-Net ^[20]	0.815 0	0.773 5	0.981 8	0.955 3	0.978 4
DUNet ^[21]	0.823 7	0.796 3	0.972 1	0.956 6	0.982 0
MResU-Net ^[22]	0.823 7	0.796 9	0.979 9	--	0.979 9
RSAN ^[23]	0.822 2	0.814 9	0.983 9	0.969 1	0.985 5
MLA-Net ^[24]	0.827 9	0.804 6	0.980 5	0.958 1	0.982 7
SA-Unet ^[25]	0.826 3	0.821 2	0.984 0	0.969 8	0.986 4
ConvMixer ^[26]	0.824 5	-	-	-	-
GCFE-Net	0.829 0	0.834 5	0.973 9	0.957 3	0.981 3

表 4 不同算法在 CHASEDB1 数据集上的性能指标

Algorithm	F1	SE	SP	ACC	AUC
R2U-Net ^[20]	0.792 8	0.775 6	0.982 0	0.963 4	0.981 5
DU-Net ^[21]	0.800 1	0.811 5	0.975 2	0.961 0	0.980 4
RSAN ^[23]	0.811 1	0.848 6	0.983 6	0.975 1	0.980 4
MLA-Net ^[24]	0.824 8	0.804 2	0.980 1	0.967 3	0.982 4
SA-Unet ^[25]	0.815 3	0.857 3	0.983 5	0.975 5	0.990 5
DEU-Net ^[27]	0.803 7	0.807 4	0.982 1	0.966 1	0.981 2
GCFE-Net	0.837 9	0.842 7	0.979 1	0.964 0	0.987 8

表 5 不同算法在 STARE 数据集上的性能指标

Algorithm	F1	SE	SP	ACC	AUC
MLA-Net ^[24]	0.796 5	0.792 5	0.981 2	0.964 9	0.982 6
MAU-Net ^[28]	0.782 6	0.753 6	0.980 8	0.958 1	0.972 1
Res2Unet ^[29]	0.785 8	0.739 2	0.986 5	0.965 0	0.970 9
Zhang ^[30]	0.787 8	0.738 1	0.987 1	0.965 5	0.982 3
Wang ^[31]	0.794 7	0.823 0	0.994 5	0.964 1	0.962 0
GCFE-Net	0.806 3	0.778 6	0.981 4	0.959 5	0.983 6

2.5 消融实验

2.5.1 GCN 模块 K 取值对比实验

在使用 K 近邻算法构造图的过程中, 邻居节点的数量 K 是一个控制聚集范围的超参数。邻居太少会降低信息交换的质量, 而邻居节点太多会降低每个节点特征的独特性。为了得到最优的 K 值, 在 DRIVE 数据集上将 K 设置从 1 调优到 7, 由图 10 可知, 当邻居节点 K 取 6 时, F1 值最佳。

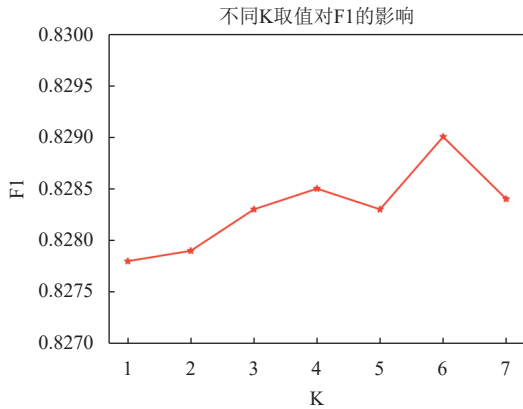


图 10 K 取值分析

2.5.2 算法各模块消融实验

为了验证本文算法所提出模块的有效性, 分别

在 DRIVE、CHASEDB1 和 STARE 三个数据集上进行消融实验。实验结果如表 2 所示。表中模型 A_1 表示基础模型 U-Net, 模型 A_2 表示在模型 A_1 的基础上加入 GCN 模块, 模型 A_3 表示在模型 A_2 的基础上加入多层次特征融合模块, 模型 A_4 为本文算法。

从表 6 可知, 由于 U-Net 使用普通卷积只能提取局部特征信息, 导致血管像素容易被分类成背景像素, A_2 模型在 U-Net 的基础上加入图卷积模块, 可以充分提取图像的全局特征信息, 并且捕获血管像素之间的长期依赖关系, 从而提升了模型提取血管像素的能力, 在 DRIVE 数据集和 CHASEDB1 数据集上可以看到 SE 指标有所上升, 由于 STARE 数据集中大部分图像血管像素占比较低, SE 指标有所下降, 但 SP 有所提升。模型 A_3 加入多层次特征融合模块, 各项指标均有所上升, 说明多层次特征融合模块可以有效减少 U-Net 下采样过程中空间信息的损失, 提升模型分割的准确度。模型 A_3 加入多尺度特征融合块, 减少了特征图中的噪声信息的同时且保留语义信息, 达到了最好的分割效果。综上所述, 本文提出的模块具有一定的有效性和合理性。

表 6 消融实验结果

Dataset	model	F1	SE	SP	ACC	AUC
DRIVE	A_1	0.824 8	0.821 3	0.975 1	0.955 5	0.977 3
	A_2	0.826 2	0.831 3	0.973 0	0.956 0	0.977 9
	A_3	0.827 8	0.834 2	0.974 1	0.957 1	0.981 0
	A_4	0.829 0	0.834 5	0.973 9	0.957 3	0.981 3
CHASEDB1	A_1	0.822 4	0.821 3	0.978 2	0.960 9	0.983 7
	A_2	0.835 5	0.845 5	0.977 9	0.962 9	0.986 5
	A_3	0.836 4	0.846 4	0.978 0	0.963 5	0.987 2
	A_4	0.837 9	0.842 7	0.979 1	0.964 0	0.987 8
STARE	A_1	0.800 2	0.728 5	0.989 6	0.963 2	0.980 7
	A_2	0.801 5	0.769 2	0.981 7	0.958 7	0.982 2
	A_3	0.803 8	0.775 4	0.981 3	0.959 0	0.982 9
	A_4	0.806 3	0.778 6	0.981 4	0.959 5	0.983 6

3 结束语

本文针对视网膜血管分割中出现的血管断裂、边界不清晰等问题, 在基于编码器-解码器的架构上进行改进, 提出了一种图卷积多层次特征融合的视网膜血管分割算法。为了提升模型提取血管特征的能力, 本文算法在编码器与解码器之间加入图卷积模块捕获血管像素之间的长距离依赖关系; 同时

在跳跃连接处加入多尺度特征融合模块以减少特征层中的噪声信息, 并且使模型更加关注血管特征。此外, 本文提出的多层次特征融合模块可以有效的减少因下采样造成的空间信息的损失。试验结果表明, 本文算法的 F1 值在 3 个数据集上均高于其他算法, 分割结果可视化图也显示出本文算法的优越性。然而, 本文算法的图卷积模块中构图所使用的 K -近邻算法比较耗时, 增加了模型的训练时长

和推理时长, 在后续的研究中, 主要探索使用注意力矩阵计算像素之间的相似性代替 K 近邻算法来构图, 并且使用部分卷积代替传统卷积来减少模型的参数量, 从而减少模型的时间复杂度与空间复杂度。

参考文献

- [1] 李兰兰, 张孝辉, 牛得草, 等. 深度学习在视网膜血管分割上的研究进展[J]. *计算机科学与探索*, 2021, 15(11): 2063-2076.
LI L L, ZHANG X H, NIU D C, et al. Research progress of deep learning in retinal vessel segmentation[J]. *Journal of Frontiers of Computer Science and Technology*, 2021, 15(11): 2063-2076.
- [2] MENDONÇA A M, CAMPILHO A. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction[J]. *IEEE Trans Med Imaging*, 2006, 25(9): 1200-1213.
- [3] YU Q Z, WANG X H, WANG X F, et al. Retinal vessels segmentation based on level set and region growing[J]. *Pattern Recognition*, 2014, 47(7): 2437-2446.
- [4] JIANG Z, ZHANG H, WANG Y, et al. Retinal blood vessel segmentation using fully convolutional network with transfer learning[J]. *Comput Med Imaging Graph*, 2018, 68: 1-15.
- [5] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2015: 3431-3440.
- [6] LI L Z, VERMA M, NAKASHIMA Y, et al. IterNet: Retinal image segmentation utilizing structural redundancy in vessel networks[C]// Proceedings of the IEEE Winter Conference on Applications of Computer Vision. New York: IEEE, 2020: 3656-3665.
- [7] DU X F, WANG J S, SUN W Z. UNet retinal blood vessel segmentation algorithm based on improved pyramid pooling method and attention mechanism[J]. *Physics in Medicine and Biology*, 2021, 66(17): 175013.
- [8] 梁礼明, 蓝智敏, 盛校棋, 等. 融合轮廓信息与条件生成对抗的视网膜血管分割[J]. *生物医学工程学杂志*, 2021, 38(2): 276-285.
LIANG L M, LAN Z M, SHENG X Q, et al. Segmentation of retinal vessels by fusing contour information and conditional generative adversarial[J]. *Journal of Biomedical Engineering*, 2021, 38(2): 276-285.
- [9] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [10] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 1251-1258.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 7132-7141.
- [12] 蒋芸, 刘文欢, 梁菁. 联合注意力和 Transformer 的视网膜血管分割网络[J]. *计算机工程与科学*, 2022, 44(11): 2037-2047.
JIANG Y, LIU W H, LIANG J. Retinal vessel segmentation network with joint attention and Transformer[J]. *Computer Engineering & Science*, 2022, 44(11): 2037-2047.
- [13] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017, 30.
- [14] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[M]// Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241.
- [15] KIPF T, WELING M. Semi-supervised classification with graph convolutional networks[C]// International Conference on Learning Representations. 2016.
- [16] HAN K, WANG Y, GUO J, et al. Vision GNN: An image is worth graph of nodes[C]// Advances in Neural Information Processing Systems.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]// European Conference on Computer Vision. Cham: Springer, 2018: 3-19.
- [18] 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 Trans Pattern Anal Mach Intell*, 2018, 40(4): 834-848.
- [19] OKTAY O, SCHLEMPER J, LE FOLGOC L, et al. Attention U-Net: Learning where to look for the pancreas[C]// Medical Imaging with Deep Learning.
- [20] ALOM M Z, YAKOPCIC C, HASAN M, et al. Recurrent residual U-Net for medical image segmentation[J]. *Journal of Medical Imaging*, 2019, 6(1): 014006.
- [21] 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(C): 149-162.
- [22] LI D, DHARMAWAN D A, NG B P, et al. Residual U-net for retinal vessel segmentation[C]// Proceedings of the IEEE International Conference on Image Processing. New York: IEEE, 2019: 1425-1429.
- [23] GUO C L, SZEMENYEI M, YI Y G, et al. Residual spatial attention network for retinal vessel segmentation[C]// Proceedings of the Neural Information Processing: 27th International Conference. New York: ACM, 2020: 509-519.
- [24] YUAN Y, ZHANG L, WANG L, et al. Multi-level attention network for retinal vessel segmentation[J]. *IEEE J Biomed Health Inform*, 2022, 26(1): 312-323.
- [25] GUO C L, SZEMENYEI M, YI Y G, et al. SA-UNet: Spatial attention U-net for retinal vessel segmentation [C]// Proceedings of the 25th International Conference on Pattern Recognition. New York: IEEE, 2021: 1236-1242.
- [26] SOLANO A, DIETRICH K N, MARTÍNEZ-SOBER M, et al. Deep learning architectures for diagnosis of diabetic

- retinopathy[J]. [Applied Sciences](#), 2023, 13(7): 4445.
- [27] WANG B, QIU S, HE H G. Dual encoding U-net for retinal vessel segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2019: 84-92.
- [28] LI H, WANG Y K, WAN C, et al. MAU-net: A retinal vessels segmentation method[C]//Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society. New York: IEEE, 2020: 1958-1961.
- [29] LI X J, DING J Q, TANG J J, et al. Res2Unet: A multi-scale channel attention network for retinal vessel segmentation[J]. [Neural Computing and Applications](#), 2022, 34(14): 12001-12015.
- [30] ZHANG Y, FANG J, CHEN Y, et al. Edge-aware U-net with gated convolution for retinal vessel segmentation[J]. [Biomedical Signal Processing and Control](#), 2022, 73: 103472.
- [31] WANG C, ZHAO Z Y, YU Y. Fine retinal vessel segmentation by combining Nest U-net and patch-learning[J]. [Soft Computing](#), 2021, 25(7): 5519-5532.

编辑 叶芳