

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

边界注意力辅助的动态图卷积视网膜血管分割

吕佳, 王泽宇, 梁浩城

引用本文:

吕佳, 王泽宇, 梁浩城. 边界注意力辅助的动态图卷积视网膜血管分割[J]. *光电工程*, 2023, **50**(1): 220116.

Lv J, Wang Z Y, Liang H C. Boundary attention assisted dynamic graph convolution for retinal vascular segmentation[J]. *Opto-Electron Eng*, 2023, **50**(1): 220116.

<https://doi.org/10.12086/oe.2023.220116>

收稿日期: 2022-06-08; 修改日期: 2022-09-26; 录用日期: 2022-09-27

相关论文

鬼影卷积自适应视网膜血管分割算法

梁礼明, 周珑颂, 陈鑫, 余洁, 冯新刚

光电工程 2021, **48**(10): 210291 doi: [10.12086/oe.2021.210291](https://doi.org/10.12086/oe.2021.210291)

PCNN与形态匹配增强相结合的视网膜血管分割

徐光柱, 王亚文, 胡松, 陈鹏, 周军, 雷帮军

光电工程 2019, **46**(4): 180466 doi: [10.12086/oe.2019.180466](https://doi.org/10.12086/oe.2019.180466)

结合极坐标建模与神经网络的IVUS图像分割

刘靖雨, 蔡怀宇, 郝文月, 左廷涛, 贾忠伟, 汪毅, 陈晓冬

光电工程 2023, **50**(1): 220118 doi: [10.12086/oe.2023.220118](https://doi.org/10.12086/oe.2023.220118)

Customized anterior segment photoacoustic imaging for ophthalmic burn evaluation *in vivo*

Huangxuan Zhao, Ke Li, Fan Yang, Wenhui Zhou, Ningbo Chen, Liang Song, Chuansheng Zheng, Zhicheng Liu, Chengbo Liu

Opto-Electronic Advances 2021, **4**(6): 200017 doi: [10.29026/oea.2021.200017](https://doi.org/10.29026/oea.2021.200017)

更多相关论文见光电期刊集群网站 



<http://cn.ojournal.org/oe>



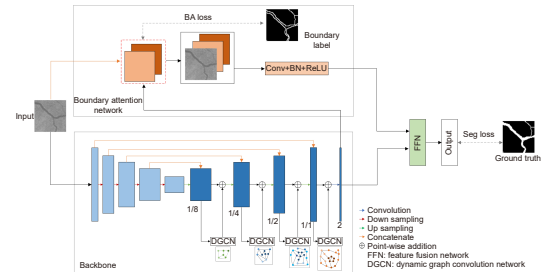
OE_Journal



Website

DOI: 10.12086/oe.2023.220116

边界注意力辅助的动态图卷积视网膜血管分割

吕佳^{1,2*}, 王泽宇¹, 梁浩城¹¹重庆师范大学计算机与信息科学学院, 重庆 401331;²重庆师范大学重庆国家应用数学中心, 重庆 401331

摘要: 针对视网膜血管分割任务中存在的毛细血管分割遗漏和断连的问题, 从最大限度地利用视网膜血管的特征信息的角度出发, 添补视网膜血管的全局结构信息和边界信息, 在 U 型网络的基础上, 提出边界注意力辅助的动态图卷积视网膜血管分割网络。本模型先将动态图卷积嵌入到 U 型网络中形成多尺度结构, 提升模型获取全局结构信息的能力, 以提高分割质量, 再利用边界注意力网络辅助模型, 增加模型对边界信息的关注度, 进一步提高分割性能。将模型在 DRIVE、CHASEDB1 和 STARE 三个视网膜图像数据集上进行实验, 均取得了较好的分割效果。实验结果证明, 该模型能较好地地区分噪声和毛细血管, 分割出结构较完整的视网膜血管, 具有泛化性和鲁棒性。

关键词: 医学图像处理; 视网膜血管分割; U 型网络; 动态图卷积网络; 边界注意力网络

中图分类号: TP183

文献标志码: A

吕佳, 王泽宇, 梁浩城. 边界注意力辅助的动态图卷积视网膜血管分割[J]. 光电工程, 2023, 50(1): 220116

Lv J, Wang Z Y, Liang H C. Boundary attention assisted dynamic graph convolution for retinal vascular segmentation[J]. *Opto-Electron Eng*, 2023, 50(1): 220116

Boundary attention assisted dynamic graph convolution for retinal vascular segmentation

Lv Jia^{1,2*}, Wang Zeyu¹, Liang Haocheng¹¹ College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China;² China National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China

Abstract: Aiming at the problem of missing and disconnected capillary segmentation in the retinal vascular segmentation task, from the perspective of maximizing the use of retinal vascular feature information, by adding the global structure information and retinal blood vessels boundary information, based on the U-shaped network, a dynamic graph convolution for retinal vascular segmentation model assisted by boundary attention is proposed. The dynamic graph convolution is first embedded into the U-shaped network to form a multi-scale structure, which improves the ability of the model to obtain the global structural information, and thus improving the segmentation quality. Then, the boundary attention network is utilized to assist the model to increase the attention to the boundary information, and further improve the segmentation performance. The proposed algorithm is tested on three retinal image datasets, DRIVE, CHASEDB1, and STARE, and good segmentation results are obtained. The

收稿日期: 2022-06-08; 修回日期: 2022-09-26; 录用日期: 2022-09-27

基金项目: 国家自然科学基金重大项目(11991024); 重庆市教委“成渝地区双城经济圈建设”科技创新项目(KJCX2020024); 重庆市教委重点项目(KJZD-K202200511); 重庆市科技局技术预见与制度创新项目(2022TFII-OFX0265)

*通信作者: 吕佳, lvjia@cqnu.edu.cn。

版权所有©2023 中国科学院光电技术研究所

experimental results show that the model can better distinguish the noise and capillary, and segment retinal blood vessels with more complete structure, which has generalization and robustness.

Keywords: medical image processing; retinal vessels segmentation; U-shaped network; dynamic graph convolution network; boundary attention network

1 引言

视网膜血管的健康状况是诊断眼部疾病乃至全身性疾病常考虑的因素之一, 尤其是视网膜血管的萎缩状况或病变层次等是明确疾病严重程度的关键。视网膜血管分割是获取这些关键信息的必要步骤, 良好的分割结果有助于准确地诊断眼部疾病^[1-2]。然而, 由于视网膜血管复杂的网状结构, 人工分割存在费时费力和易受主观因素影响等弊端, 因此自动分割视网膜血管受到了广泛关注^[3]。在视网膜血管自动分割任务中, 一方面, 由于视网膜图像具有血管和背景对比度低、背景噪声多等特性, 导致血管标注不正确或不完整等问题; 另一方面, 由于医学数据的收集受到个人隐私、收集成本等因素限制, 使得可用于训练的数据很少。这些均限制了模型的泛化性, 当面对临床噪声多的数据时, 模型容易遗漏血管像素或误分割噪声。而提升模型泛化性的关键在于能最大限度地学习血管的特征信息^[4]。

近年来, U型网络 (U-shape network, U-Net) 采用跳跃连接融合多尺度特征的方式, 在数据量少的医学图像处理任务中取得不错的效果^[5], 自然地应用于视网膜血管分割任务。然而, U-Net 在训练过程中会忽略视网膜血管的特性, 无法充分提取血管的特征信息, 导致分割结果易产生较多断连的毛细血管。为此, 研究者针对视网膜血管分割任务对 U-Net 进行了诸多改进。Jin 等人^[6]将可变形卷积引入到 U-Net 中, 利用可变形卷积捕获多尺度局部上下文依赖, 进一步丰富了特征信息, 改善了视网膜血管分割的效果; Wu 等人^[1]利用视网膜血管尺度变化大的特点, 设计了尺度感知特征聚合模块和自适应特征融合模块, 有效提取了视网膜血管的多尺度特征, 并引导了相邻层次特征之间的融合; Zhang 等人^[7]结合多核池化和空洞卷积来改进 U-Net, 通过选择不同的扩张速率, 得到了更大的感受野, 获取到更多的局部上下文信息。尽管这些方法从不同的角度利用了视网膜血管的特征信息, 却忽略了视网膜血管的全局结构信息, 而获取结构信息已被证明是提高图像分割性能的有效方法^[8]。

图卷积网络 (Graph convolution network, GCN) 是一种基于图 (graph) 的特征提取网络, 具有较好的结构信息获取能力^[9]。Shin 等人^[10]将 GCN 与卷积神经网络 (convolutional neural network, CNN) 级联, 从 CNN 输出的分割图中提取血管骨架来构造图, 进一步获取血管的结构信息, 提高了视网膜血管的分割性能; Meng 等人^[11]利用注意力优化模块聚合 GCN 和 CNN, 选择视网膜图像中视盘和视杯的轮廓点构造图, 以此融合结构信息来提升视盘和视杯的分割性能。然而, 上述方法构造的图是独立于分类任务的静态图, 难以降低真实数据受噪声的干扰度, 导致图的质量较低。由于 GCN 的性能高度依赖于图的质量, 因此静态图限制了 GCN 的性能^[12]。本文采取动态图的图构造方法^[13]来克服静态图不能调整图的问题, 以此捕获视网膜图像更充分的全局结构信息。动态图的原理在于图的邻接矩阵权重能通过反向传播学习调整, 以此输出更灵活有效的图。

虽然采用动态图卷积丰富了视网膜血管的全局结构信息, 但在上下采样过程中, 血管的浅层边界信息仍会被稀释。为了强化边界信息, Zhang 等人^[14]将原始分割任务转换为多分类任务, 使得模型能更多地关注血管边界信息, 但该模型需要针对粗细血管的边界区域进行额外标注, 流程过于繁琐。Zhang 等人^[15]提出一种只关注血管边界的门控卷积来强调血管边界, 虽然模型省去了边界分类的流程, 但增加了计算复杂度。而本文采用了一种高效的边界注意力网络来强化视网膜血管的浅层边界信息, 其不仅轻量化还能提升视网膜血管的分割性能。

综上所述, 针对视网膜血管分割任务中毛细血管断连的问题, 本文利用动态图卷积为 U-Net 添补了全局结构信息, 并采取轻量化的边界注意力网络保留了边界信息, 从而达到充分提取视网膜血管特征信息的目的, 提升了视网膜血管的分割质量。本文主要贡献为: 1) 提出一种边界注意力辅助的动态图卷积视网膜血管分割网络, 通过捕获视网膜血管的全局结构信息和边界信息, 改善毛细血管断连的情况; 2) 提出基于 U-Net 的多尺度动态图卷积, 增强动态图卷积捕获

全局结构信息的能力; 3) 引入轻量化的边界注意力网络, 使模型更加关注视网膜血管的边界区域, 以减少边界信息的丢失。

2 本文方法

2.1 网络结构

视网膜图像因受光照不均匀以及各种病变因素的影响, 背景噪声多且部分血管像素模糊, 导致利用简单跳跃连接以及上下采样操作的 U-Net 丢失了视网膜血管的大部分全局结构信息和边界信息, 无法更有效地区分视网膜图像中的噪声和血管像素, 因而造成毛细血管分割遗漏或断连的问题。针对上述问题, 本文提出边界注意力辅助的动态图卷积视网膜血管分割网络——边界注意力辅助的动态图卷积 U 型网络 (boundary attention assisted dynamic graph convolution U-Net, BDGU-Net)。该网络利用动态图卷积, 将血管的特征信息存储在动态图中进行推理, 这样不仅有效地区分了高度相似的噪声和毛细血管, 而且为 U-Net 提供了更细致的全局结构信息。此外, 该模型借助边界注意力网络来加强上下采样过程中被稀释的浅层边界信息, 以提高模型的分割能力。BDGU-Net 的网络结构如图 1 所示, 主要包括主干网络、边界注意力网络和特征融合网络。

主干网络包括编码器以及解码器。在编码器中,

特征图每经过一次下采样, 其尺度减小一半、通道数增加一倍。在解码器中, 特征图每经过一次上采样, 其尺度增加一半、通道数减少一倍, 并与编码器相同尺度的特征图进行跳跃连接, 再经过该尺度动态图卷积获取其结构信息。文献 [16] 证明, 小尺度的特征图包含了较丰富的语义信息, 大尺度的特征图保留了足够的结构信息, 故不同尺度的特征图之间可以相互补充。因而 BDGU-Net 将动态图卷积自底向上地应用于解码器中不同尺度的血管特征图。具体地, 本模型将 1/8 大小的特征图输入动态图卷积, 加强该尺度的结构信息, 再将其与原始特征信息相加融合, 执行上采样操作, 以使融合结构信息的特征图传递至下一解码层中, 从而丰富大尺度特征图的结构信息。1/4 大小、1/2 大小以及原图大小的特征图同样采取上述操作, 最终捕获到视网膜血管更充分的全局结构信息。特别地, 原图大小的特征图不再进行上采样操作, 而是直接利用 1×1 卷积操作输出对应类别数。边界注意力网络主要借助边界注意力损失引导主干网络定位边界区域, 可以被视为一种中间监督。特征融合网络可以更好地融合主干网络和边界注意力网络分别提供的语义信息和边界信息, 模型得以保留更多的血管像素。

2.2 动态图卷积

U-Net 通过简单地融合编码器和解码器相同尺度

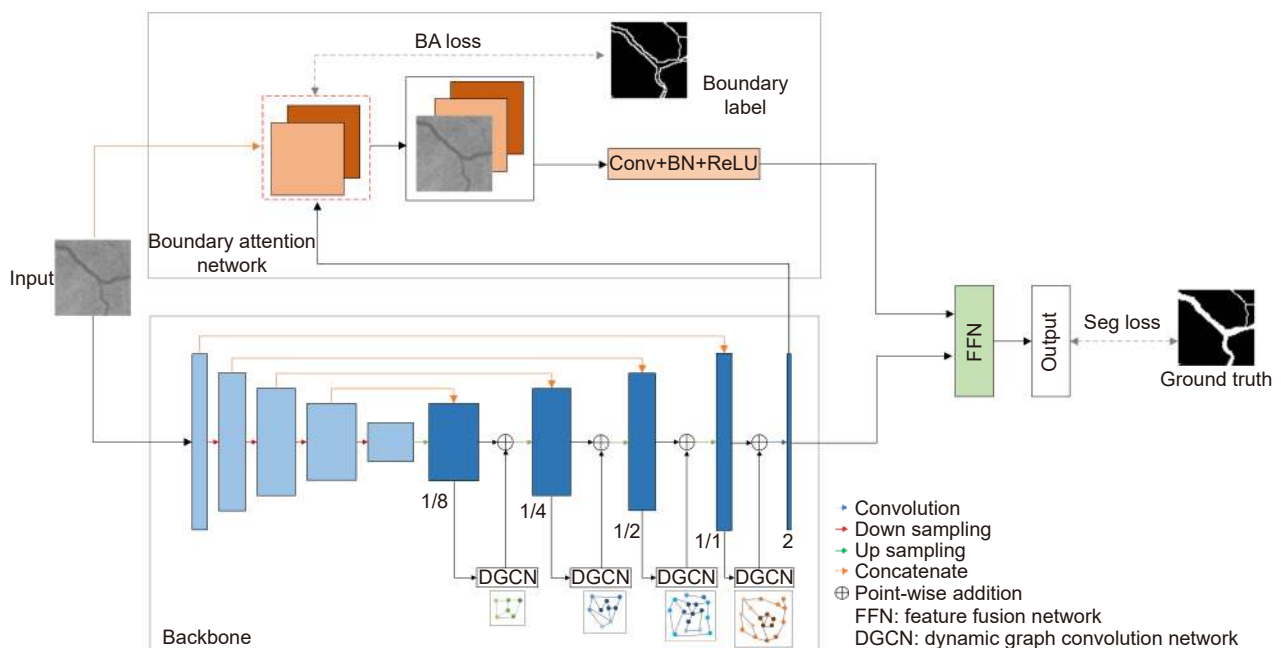


图 1 边界注意力辅助的动图卷积 U 型网络

Fig. 1 Boundary attention assisted dynamic graph convolution U-shaped network

特征图的方式,整合了视网膜血管粗粒度的全局结构信息,即丢失了毛细血管或复杂形态的血管的结构信息,导致 U-Net 区分噪声和血管像素的能力较差。本方法在 U-Net 解码器中嵌入动态图卷积来增强模型辨别背景和血管像素的能力,使整体的分割性能提高。动态图的原理^[17]如图 2 所示。首先将解码器提取的血管特征信息构造为加权邻接矩阵表示的图结构,在前向传播的过程中,聚合不同关联度的邻居信息,以此更新中心节点的特征。后在损失函数反向传播的过程中,利用梯度更新自动学习节点之间的依赖关系,调整邻接矩阵,在增加同类别节点特征关联度的同时,降低噪声节点和血管节点间的特征关联度,这样能更容易区分血管像素和背景像素,从而为模型提供细粒度的全局结构信息。动态图卷积分为动态图构造和图卷积计算两个步骤,如图 3(a) 所示。

第一步,动态图构造。给定图 $G=(V,A)$,其中 $V \in \mathbb{R}^{N \times C}$ 为图节点矩阵, N 表示图节点数目, C 表示图节点特征的维数; $A \in \mathbb{R}^{N \times N}$ 为邻接矩阵,表示各节点间的连接关系。给定 U-Net 解码器输出的原始特征 $X \in \mathbb{R}^{W \times H \times C}$, W 、 H 和 C 分别为特征图的宽、高和通道数,首先利用 Reshape 函数将 X 转换为图节点矩阵 $X' \in \mathbb{R}^{WH \times C}$,再利用 1×1 卷积和点积运算 (Dot-product) 构造邻接矩阵。邻接矩阵的计算公式^[13]为

$$A = \theta(X; W_\theta) \otimes \text{diag}(\rho(\bar{X}; W_\rho)) \otimes \theta(X; W_\theta)^T, \quad (1)$$

式中: \otimes 表示点积; $\theta(\cdot)$ 包含 1×1 卷积和 Reshape 函数, $\theta(X) \in \mathbb{R}^{WH \times C}$; W_θ 和 W_ρ 表示邻接矩阵的权重; $\text{diag}(\rho(\bar{X}))$ 是对角矩阵,表示二维特征图的通道注意力系数, $\bar{X} \in \mathbb{R}^{1 \times 1 \times C}$ 是 X 经过全局平均池化后的特征,

ρ 包括 1×1 卷积和激活函数 Sigmoid。最后利用 Softmax 函数归一化邻接矩阵。

第二步,图卷积计算。构造动态图 (A, X') 后,利用图卷积公式推理对应的结构信息。图卷积公式^[9]为

$$Z = \sigma(A X' W), \quad (2)$$

式中: σ 为激活函数, $W \in \mathbb{R}^{D \times D}$ 为原始空间的参数矩阵。

2.3 边界注意力网络

虽然本模型利用动态图卷积扩展了 U-Net 提取的特征信息,但在主干网络上下采样的过程中,血管的边界信息仍被逐渐弱化,限制了模型的分割性能。文献 [18] 提出了一种边界注意力网络,其不仅轻量化且能输出高质量的人像分割图像,故本文遵循其思路设计了一种边界注意力网络来加强视网膜血管的边界信息。该网络通过卷积操作将主干网络的输出投影设为 2 通道,将其作为边界注意力图。边界注意力损失 (Boundary attention loss, BA Loss) 引导边界注意力图定位边界区域,边界区域的标注利用传统的无监督形态学方法提取。然后将输入的图像与边界注意力图拼接成 3 通道图 (原始视网膜图像经过图像预处理后变为单通道),再经过一个卷积序列提取浅层的细节信息,卷积序列包括一个 1×1 卷积、批归一化以及激活函数。

2.4 特征融合网络

语义特征是高层次的特征,边界特征是低层次的特征,它们之间在特征表示层次上存在差异,不能简

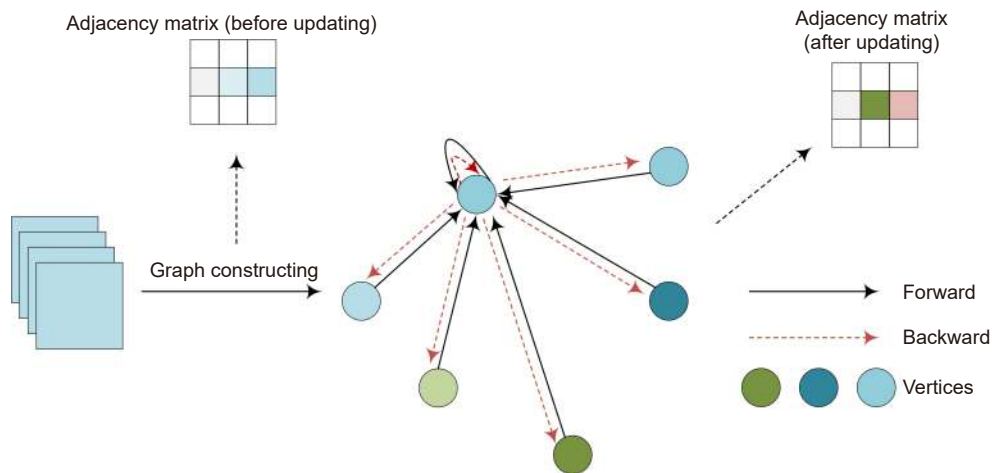


图 2 动态图原理

Fig. 2 Principle of dynamic graph

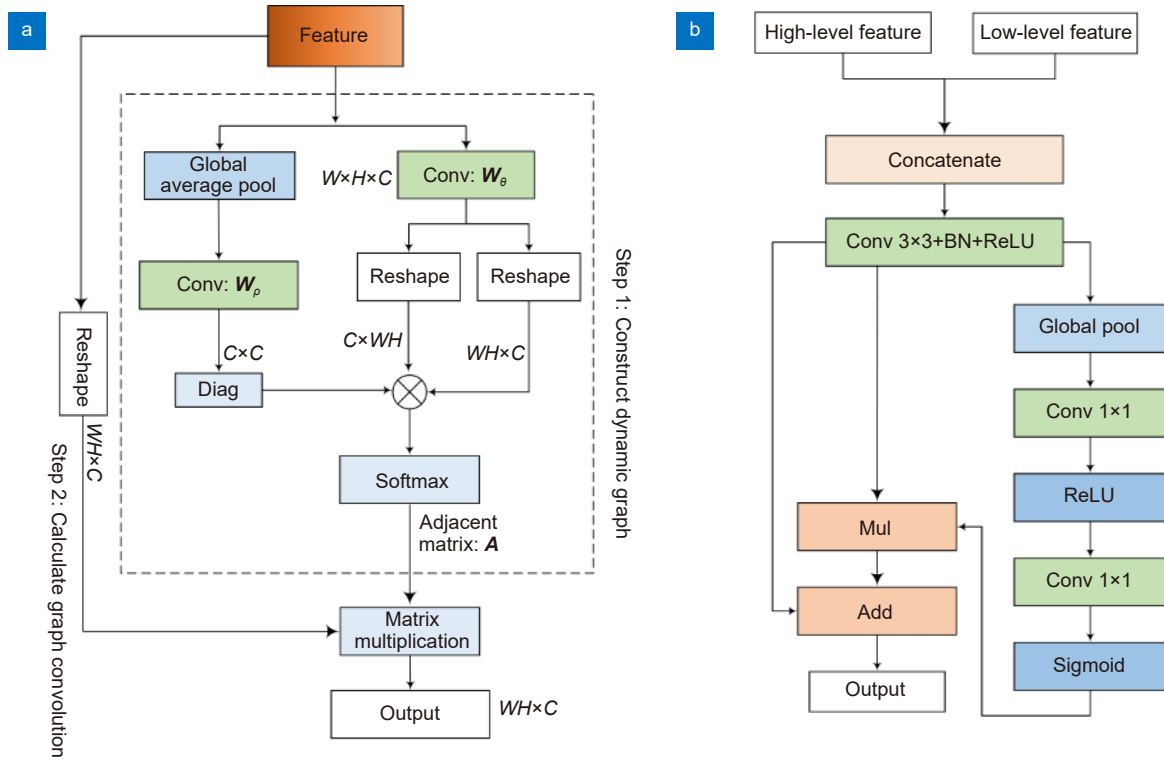


图 3 动态图卷积计算过程和特征融合网络。(a) 动态图卷积计算过程; (b) 特征融合网络

Fig. 3 Dynamic graph convolution calculation process and Feature fusion network.

(a) Dynamic graph convolution calculation process; (b) Feature fusion network

单地通过元素值的累加或通道的拼接来组合它们。故引入文献 [19] 的特征融合方法, 对于不同层次的两特征, 先将它们的通道拼接, 后传递到一个卷积序列中, 再利用全局池化、两个 1×1 卷积以及激活函数来计算卷积序列输出特征的权重向量, 然后将权重向量和卷积序列的输出相乘后相加, 起到血管像素特征选择和结合的作用, 让模型的分割目标聚焦在血管像素上。该模块结构如图 3(b) 所示。

2.5 损失函数

考虑到视网膜血管分割是二分类问题, 故采用二分类交叉熵作为 BDGU-Net 的损失函数:

$$L_{\text{seg}} = -[y_i^s \cdot \log(y_p^s) + (1 - y_i^s) \cdot \log(1 - y_p^s)], \quad (3)$$

式中: y_p^s 表示预测结果, y_i^s 表示真实标注。边界注意力图的损失函数也是二分类交叉熵:

$$L_{\text{bound}} = -[y_i^b \cdot \log(y_p^b) + (1 - y_i^b) \cdot \log(1 - y_p^b)], \quad (4)$$

式中: y_p^b 表示边界区域的预测值, y_i^b 表示边界区域的标注。

本文的损失函数是分割损失 (segmentation loss, Seg Loss)、边界注意力损失的加权和, 如式 (5) 所示, 实验表明 $\alpha = 0.6$, $\beta = 0.4$ 时效果最好。

$$L = \alpha L_{\text{seg}} + \beta L_{\text{bound}}. \quad (5)$$

3 实验结果与分析

3.1 实验环境与参数设置

本实验采用 Python 语言编程, 基于 PyTorch 框架实现, 模型在配备 64 G 内存的 Intel Xeon Silver 4114 CPU 和 NVIDIA TITAN V 显卡的计算机上运行, 部署环境为 Windows10 操作系统。网络训练轮次总数设置为 50 轮, batch size 设置为 32, 采用优化算法 Adam 训练整体网络结构, 初始学习率设置为 0.0005, 并利用余弦退火策略辅助训练。

3.2 数据集与数据预处理

为了验证 BDGU-Net 的有效性, 选用 DRIVE、CHASEDB1 和 STARE 三个公开数据集进行实验, 数据集的示例如图 4 所示。DRIVE 数据集来自荷兰的一个糖尿病视网膜病变筛查项目, 包含 40 张分辨率为 $565 \text{ pixels} \times 584 \text{ pixels}$ 的 RGB 图像, 其中 33 张图像未显示任何糖尿病视网膜病变迹象, 7 张图像显示轻度早期糖尿病视网膜病变迹象; CHASEDB1 数据集是由 Kinston 大学公开的小规模数据集, 从 14 名儿

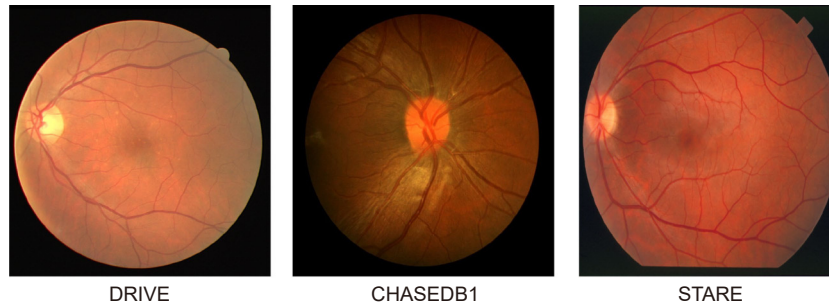


图 4 视网膜图像
Fig. 4 Retina image

童的左右眼拍照采集获得, 共包含 28 张分辨率为 999 pixels×960 pixels 的 RGB 图像; STARE 数据集是由加州大学圣地亚哥分校提供的, 包含 20 张分辨率为 605 pixels×700 pixels 的 RGB 图像, 其中 10 张图像为正常眼底图像, 其余 10 张图像存在不同程度的病变。

上述三种视网膜图像受光照或眼部疾病的影响, 图像中的血管像素不清晰, 本文采用图像预处理策略^[6]来突显血管特征, 这样更有利于模型提取血管像素的特征。先把原始的 RGB 图像转换为灰度图像, 对灰度图像归一化后再对其采用对比度受限自适应直方图均衡化 (CLAHE) 来增强血管与背景之间对比度, 最后采用伽马校正增强图像中亮度较低的血管区域。此外, 实验中还对训练数据集进行增强和扩充, 将训练集中的输入图像、图像真实标注以及边界标注对应随机裁剪为尺寸为 48 pixels×48 pixels 的切片, 再对图像随机旋转, 最终预处理结果示例见图 5。两种数据集分别切片 20000 张, 前 90% 用于训练, 后 10% 用于验证。

3.3 评价指标

为了定量评估本文的模型, 使用 F1 (F1-Score, F1)、灵敏度 (SEnsitivity, SE)、特异性 (SPecificity, SP)、准确性 (ACCuracy, ACC)、AUC (Area Under Curve, AUC) 这 5 个常用指标来评价模型的性能, 前 4 个指标计算公式如表 1 所示。表 1 中的 TP、FP、FN 和 TN 分别是真阳性 (正确分割的血管像素), 假阳性 (错误分割的血管像素)、假阴性 (错误分割的背景像素) 和真阴性 (正确分割的背景像素)。F1 计算公式中的 PR 表示 TP 占 TP 和 FP 总和的比例。AUC 是 ROC 曲线下的面积, 以 SP 为横坐标, SE 为纵坐标, 其值越接近 1, 表明模型真实性越高。

3.4 实验结果

3.4.1 消融实验

为了验证本文采取的多尺度动态图卷积的有效性以及边界注意力网络的辅助作用, 分别在 DRIVE、CHASEDB1 和 STARE 三个数据集上设计了 3 组消融实验, 并用 U-Net 对这些方法的分割性能进行定量比较, 实验结果见表 2。图 6 对 3 组消融实验的实验结

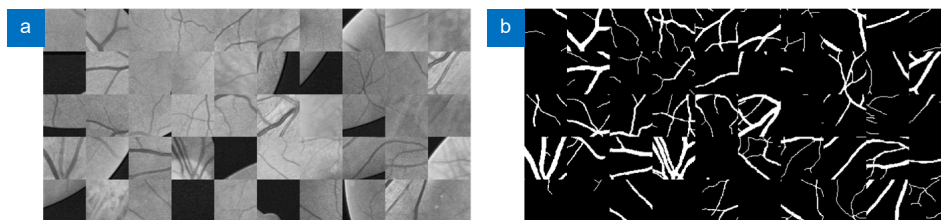


图 5 数据预处理结果图。(a) 预处理图像切片; (b) 真实标注切片
Fig. 5 Data preprocessing results. (a) Pre-processed image slices; (b) Ground truth slices

表 1 评价指标

Table 1 Evaluation indexes

F1	SE	SP	ACC
$\frac{2 \times PR \times SE}{PR + SE}$	$\frac{TP}{TP + FN}$	$\frac{TN}{TN + FP}$	$\frac{TP + TN}{TP + TN + FP + FN}$

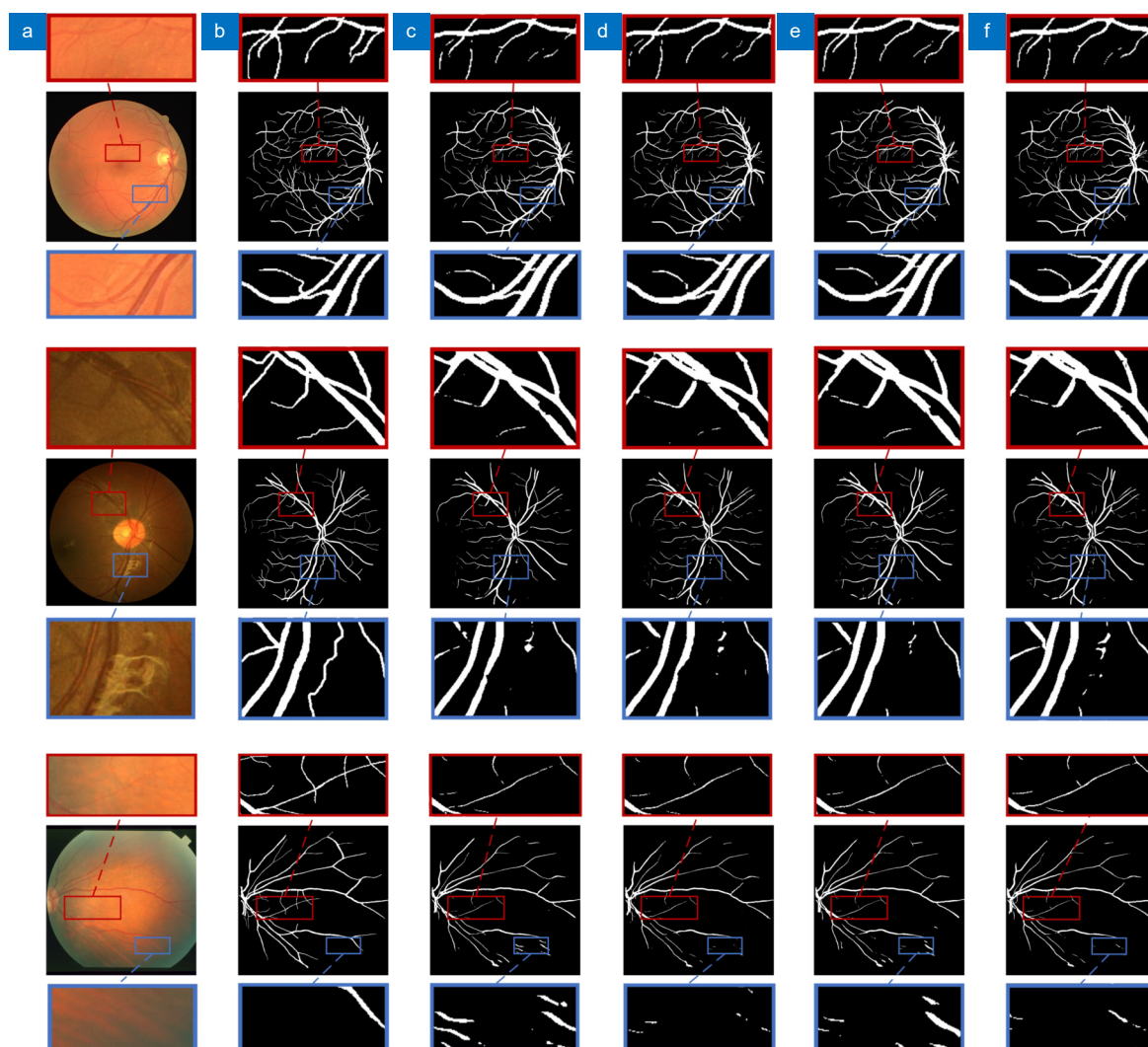


图 6 消融实验效果对比图。(a) 原图及原图细节; (b) 真实标注; (c) U-Net; (d) DGU-Net; (e) BU-Net; (f) BDGU-Net
 Fig. 6 Comparison of ablation results. (a) Original image and details; (b) Ground truth; (c) U-Net; (d) DGU-Net; (e) BU-Net; (f) BDGU-Net

表 2 消融实验结果

Table 2 Ablation experiments results

Dataset	Network	F1	SE	SP	ACC	AUC	Time/s
DRIVE	U-Net	0.8353	0.8330	0.9777	0.9601	0.9848	1.00
	DGU-Net	0.8357	0.8275	0.9789	0.9605	0.9850	1.22
	BU-Net	0.8355	0.8352	0.9773	0.9600	0.9849	1.00
	BDGU-Net	0.8359	0.8300	0.9785	0.9604	0.9851	1.24
CHASEDB1	U-Net	0.8108	0.8052	0.9820	0.9660	0.9854	1.31
	DGU-Net	0.8053	0.7812	0.9842	0.9658	0.9846	1.73
	BU-Net	0.8133	0.8210	0.9803	0.9659	0.9858	1.31
	BDGU-Net	0.8136	0.8146	0.9813	0.9662	0.9856	1.76
STARE	U-Net	0.7845	0.7380	0.9863	0.9648	0.9810	2.21
	DGU-Net	0.7824	0.7045	0.9908	0.9660	0.9825	2.85
	BU-Net	0.7845	0.7352	0.9868	0.9650	0.9806	2.21
	BDGU-Net	0.7934	0.7273	0.9900	0.9672	0.9834	2.85

果进行可视化, 同时在整张图中选择并放大了一些细节, 以便更好地观察微小血管和噪声的分割结果。

从表 2 和图 6 可得到如下观察结果:

1) 本方法在 STARE 和 CHASEDB1 两个数据集上性能提升较明显。与 STARE 数据集上的 U-Net 相比, 仅引入多尺度动态图卷积的 U-Net (multiscale dynamic graph convolution U-Net, DGU-Net) 的 SP、ACC 和 AUC 指标分别增加了 0.45%、0.12%、0.15%, 仅引入边界注意力的 U-Net (boundary attention U-Net, BU-Net) 相比 U-Net 虽无显著的提升, 但其 F1、SE 指标相比 DGU-Net 分别增加了 0.21%、3.07%。此外, 将两者结合的 BDGU-Net 平衡了 DGU-Net 和 BU-Net 的 SE、SP 指标之间的差异, 相比 U-Net 的 F1、SP、ACC、AUC 指标分别增加了 0.89%、0.37%、0.24%、0.24%。在 CHASEDB1 数据集上, BDGU-Net 比 U-Net 的 F1、SE、ACC、AUC 指标分别提升了 0.28%、0.94%、0.02%、0.02%。

2) 本方法在 DRIVE 数据集上性能相较于 U-Net 提升有限。DGU-Net 的 SP 指标增加了 0.12%, BU-Net 的 SE 指标增加了 0.22%, BDGU-Net 除 SE 指标比 U-Net 低 0.3% 外, 其它指标均略高于 U-Net。STARE 和 CHASEDB1 数据集与 DRIVE 数据集相比, 噪声较多, 有用的特征更难提取。因此 BDGU-Net 在复杂的、噪声较多的数据集上拥有更好的表现效果。

3) 单独引入动态图卷积或边界注意力的网络存在局限性。DGU-Net 在三个数据集上的 SE 指标都相对较低, BU-Net 则是在 SP 指标上相对较低。因为在监督训练中, 通常难在 SE 指标和 SP 指标之间取得平衡, 如果提高分割的灵敏度, 可能会有更多背景中的非血管像素被识别为血管像素, 这是以降低特异性和影响整体准确性为代价的。反之, 如果提高分割的特异性, 可能会有较多的血管像素被遗漏, 也会影响整体的分割准确性。将动态图卷积和边界注意力结合起来的 BDGU-Net 则能平衡分割的灵敏度和特异性, 使分割的整体准确性得到提高。如图 6(d)、图 6(e) 和

图 6(f) 所示, DGU-Net 能较好地识别噪声而不能更好地保证血管的连通性; BU-Net 则相反, 其能分割出更多的血管, 而噪声的判别能力弱; BDGU-Net 通过结合动态图卷积和边界注意力, 使得分割出的结果既能较好地保证血管的连通性又能区分出噪声。

4) 本文还比较了消融实验中不同网络对单张眼底图像分割的平均处理时间, BDGU-Net 在可接受的时间消耗范围内进一步提高了分割的精度。

3.4.2 损失函数权重对比实验

由于本文方法的损失函数包括 BA loss 和 Seg loss, 为了得到一组较好的损失权重 α 和 β , 在已有的经验基础上, 设定了 4 组权重系数在 DRIVE 数据集上进行对比实验。由表 3 可以看出, $\alpha=0.4$, $\beta=0.6$ 时, 网络的分割灵敏度和特异性较平衡, 使该模型在 F1 和 AUC 指标上表现略好, 因此本文将第 3 组权重系数作为实验超参数。

3.4.3 与其它网络的性能对比

为了综合地评估 BDGU-Net 的分割性能, 将 BDGU-Net 与其它网络进行比较, 其中包括非专家观测者在三个数据集上的结果, 以便显示自动分割算法的优势。结果分为两个方面进行展示, 一个是指标对比, 见表 4, 另一个是可视化对比, 见图 7。

从表 4 可知, 在 DRIVE 数据集上, Iternet 通过 mini-UNet 的多次迭代来加深网络, 有效提高了血管分割结果的连通性, SE 提高至 83.70%, 而 SP 比本方法低 0.15%, 且 ACC 也略低于本方法; MLADUNet 采用多层次注意力模块来利用底层详细信息, 有效提高了网络区别噪声的能力, SP 提高至 97.88%, 而 SE 比本方法低 0.31%。在 CHASEDB1 数据集上, Res2Unet 结合多尺度策略和通道注意力机制, 并使用后处理方式, 有效提高 SE 至 84.44%, 而本方法的 SP 相较其高了 0.6%, ACC 和 AUC 也略高于 Res2Unet。在 STARE 数据集上, Yang 等^[22]提出有效的损失函数针对粗细血管进行精确分割, 使 SE 提高至 79.46%, 而 SP 比本方法低 0.79%, ACC 也低

表 3 不同 α 和 β 系数取值分析

Table 3 Value analysis of different α and β coefficients

α	β	F1	SE	SP	ACC	AUC
0.8	0.2	0.8348	0.8252	0.9790	0.9604	0.9848
0.7	0.3	0.8351	0.8230	0.9795	0.9605	0.9848
0.6	0.4	0.8359	0.8300	0.9785	0.9604	0.9851
0.5	0.5	0.8351	0.8284	0.9785	0.9603	0.9849

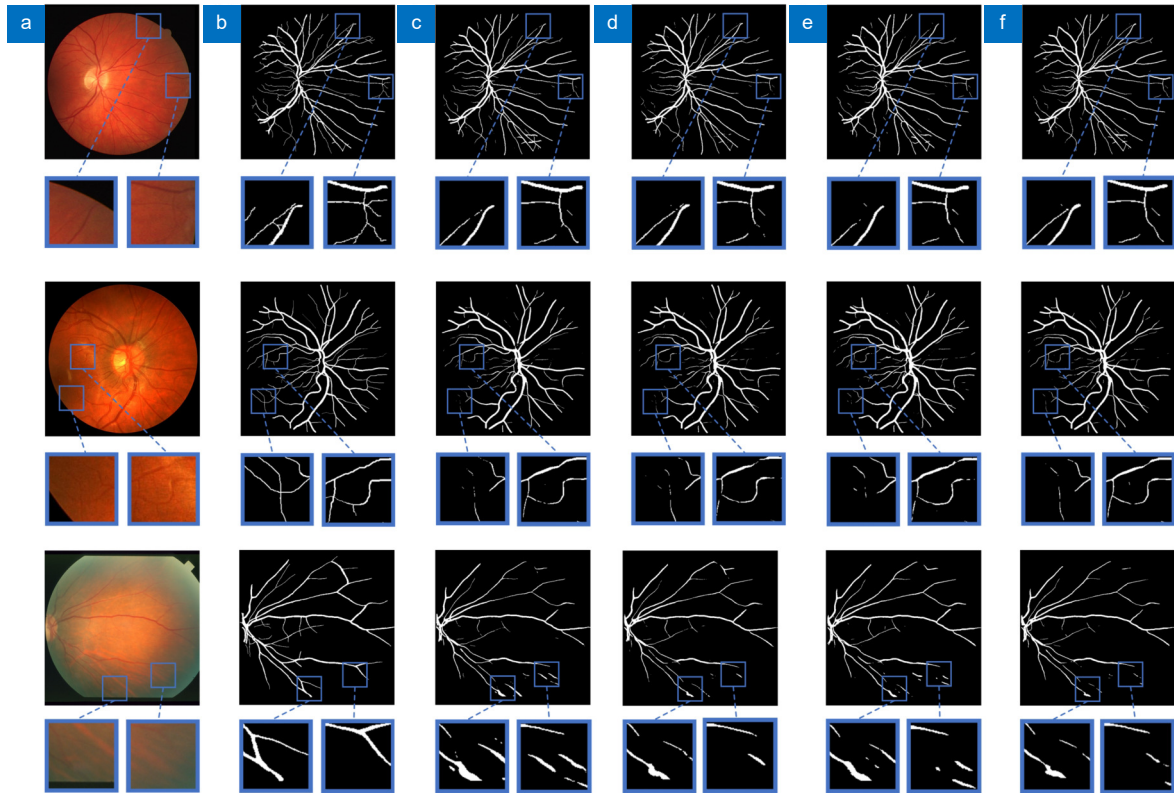


图 7 不同网络效果对比图。(a) 原图及原图细节; (b) 真实标注; (c) Iternet; (d) MLA-DU-Net; (e) Res2Unet; (f) BDGU-Net

Fig. 7 Comparison of ablation results. (a) Original image and details; (b) Ground truth; (c) Iternet; (d) MLA-DU-Net; (e) Res2Unet; (f) BDGU-Net

表 4 不同网络在 DRIVE、CHASEDB1 和 STARE 数据集的指标对比

Table 4 Index comparison of different networks in DRIVE, CHASEDB1 and STARE datasets

Dataset	Network	Year	F1	SE	SP	ACC	AUC	Time/s
DRIVE	2 nd Human Observer ^[20]	—	—	0.7760	0.9724	0.9472	—	—
	Iternet ^[21]	2020	0.8353	0.8370	0.9770	0.9600	0.9847	1.37
	Yang ^[22]	2021	0.8297	0.8353	0.9751	0.9579	—	—
	MLA-DU-Net ^[23]	2021	0.8352	0.8269	0.9788	0.9603	0.9849	1.64
	Res2Unet ^[24]	2022	0.8292	0.8332	0.9756	0.9583	0.9782	1.21
	Zhang ^[15]	2022	0.8349	0.8345	0.9773	0.9600	0.9850	1.20
	BDGU-Net	2022	0.8359	0.8300	0.9785	0.9604	0.9851	1.24
CHASEDB1	2 nd Human Observer ^[20]	—	—	0.8105	0.9711	0.9545	—	—
	Iternet ^[21]	2020	0.8120	0.8144	0.9809	0.9658	0.9856	1.48
	Yang ^[22]	2021	0.7997	0.8176	0.9776	0.9632	—	—
	MLA-DU-Net ^[23]	2021	0.8063	0.8137	0.9796	0.9646	0.9841	2.13
	Res2Unet ^[24]	2022	0.8071	0.8444	0.9753	0.9634	0.9794	1.47
	Zhang ^[15]	2022	0.8050	0.8141	0.9792	0.9643	0.9841	1.72
	BDGU-Net	2022	0.8136	0.8146	0.9813	0.9662	0.9856	1.76
STARE	2 nd Human Observer ^[20]	—	—	0.8952	0.9384	0.9349	—	—
	Iternet ^[21]	2020	0.7962	0.7490	0.9874	0.9667	0.9838	2.76
	Yang ^[22]	2021	0.8155	0.7946	0.9821	0.9626	—	—
	MLA-DU-Net ^[23]	2021	0.7965	0.7925	0.9812	0.9649	0.9826	3.95
	Res2Unet ^[24]	2022	0.7858	0.7392	0.9865	0.9650	0.9709	2.77
	Zhang ^[15]	2022	0.7878	0.7381	0.9871	0.9655	0.9823	3.07
	BDGU-Net	2022	0.7934	0.7273	0.9900	0.9672	0.9834	2.85

了 0.46%。从整体来看, 本方法的 SP 指标表现较好, 而 SE 指标略逊于其它网络, ACC 和 AUC 指标与 Iternet 相差不大。结合图 7(c)、图 7(d)、图 7(e)、图 7(f) 的对比可知, Iternet 和 Res2Unet 的血管分割结果的连通性较好, 而 MLA-DUNet 和 BDGU-Net 区分背景中噪声的能力更强。值得注意的是, 在血管结构较复杂的区域, BDGU-Net 也能较好地保证分割血管的连通性。综上, BDGU-Net 既能保持分割血管的连通性又能更好地区分背景中的噪声, 可以有效地完成视网膜血管分割任务, 表现出较好的泛化性和抗干扰能力。

4 结 论

本文从 U-Net 丢失了大部分全局结构信息和边界信息, 难以进一步提取血管特征信息, 造成毛细血管分割遗漏和断连的问题出发, 提出一种基于改进 U-Net 的边界注意力辅助的动态图卷积视网膜血管分割模型。该模型在原始空间采用多尺度动态图卷积, 捕获特征图更细致的全局结构信息, 同时利用简单高效的边界注意力网络辅助主干网络来增强边界细节信息, 从而达到改善视网膜血管分割断连或噪声误分类的目的。然而, 在实验过程中, 为了使边界注意力网络轻量化, 只采用单层的神经网络提取浅层的细节信息, 使得边界注意力网络表达有限, 缺乏对血管边界信息更准确的把握。因此, 提高本模型分割视网膜血管边界的精准性将是我们下一步的工作方向。

参考文献

- [1] Wu H S, Wang W, Zhong J F, et al. SCS-net: a scale and context sensitive network for retinal vessel segmentation[J]. *Med Image Anal*, 2021, **70**: 102025.
- [2] Liang L M, Zhou L S, Chen X, et al. Ghost convolution adaptive retinal vessel segmentation algorithm[J]. *Opto-Electron Eng*, 2021, **48**(10): 210291.
梁礼明, 周珑颂, 陈鑫, 等. 鬼影卷积自适应视网膜血管分割算法[J]. *光电工程*, 2021, **48**(10): 210291.
- [3] Li L L, Zhang X H, Niu D C, et al. Research progress of deep learning in retinal vessel segmentation[J]. *J Front Comput Sci Technol*, 2021, **15**(11): 2063–2076.
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.
- [4] Zhou Y Q, Yu H C, Shi H. Study group learning: improving retinal vessel segmentation trained with noisy labels[C]// *Proceedings of the 24th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Strasbourg, 2021: 57–67. https://doi.org/10.1007/978-3-030-87193-2_6.
- [5] Ummadi V. U-net and its variants for medical image segmentation: a short review[Z]. arXiv: 2204.08470, 2022. <https://arxiv.org/abs/2204.08470v1>.
- [6] Jin Q G, Meng Z P, Pham T D, et al. DUNet: a deformable network for retinal vessel segmentation[J]. *Knowl-Based Syst*, 2019, **178**: 149–162.
- [7] Zhang T, Li J, Zhao Y, et al. MC-UNet multi-module concatenation based on u-shape network for retinal blood vessels segmentation[Z]. arXiv: 2204.03213, 2022. <https://arxiv.org/abs/2204.03213v1>.
- [8] Zhu X F, Gan J Z, Lu G Q, et al. Spectral clustering via half-quadratic optimization[J]. *World Wide Web*, 2020, **23**(3): 1969–1988.
- [9] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[Z]. arXiv: 1609.02907, 2016. <https://arxiv.org/abs/1609.02907>.
- [10] Shin S Y, Lee S, Yun I D, et al. Deep vessel segmentation by learning graphical connectivity[J]. *Med Image Anal*, 2019, **58**: 101556.
- [11] Meng Y D, Wei M, Gao D X, et al. CNN-GCN aggregation enabled boundary regression for biomedical image segmentation[C]// *Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention*, Lima, 2020: 352–362. https://doi.org/10.1007/978-3-030-59719-1_35.
- [12] Zhu Y H, Ma J B, Yuan C A, et al. Interpretable learning based Dynamic Graph Convolutional Networks for Alzheimer's Disease analysis[J]. *Inf Fusion*, 2022, **77**: 53–61.
- [13] Li X, Yang Y B, Zhao Q J, et al. Spatial pyramid based graph reasoning for semantic segmentation[C]// *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 2020: 8950–8959. <https://doi.org/10.1109/CVPR42600.2020.00897>.
- [14] Zhang Y S, Chung A C S. Deep supervision with additional labels for retinal vessel segmentation task[C]// *Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, 2018: 83–91. https://doi.org/10.1007/978-3-030-00934-2_10.
- [15] Zhang Y, Fang J, Chen Y, et al. Edge-aware U-net with gated convolution for retinal vessel segmentation[J]. *Biomed Signal Process Control*, 2022, **73**: 103472.
- [16] Yu F, Wang D Q, Shelhamer E, et al. Deep layer aggregation[C]// *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 2018: 2403–2412. <https://doi.org/10.1109/CVPR.2018.00255>.
- [17] Jin X, Lai Z H, Jin Z. Learning dynamic relationships for facial expression recognition based on graph convolutional network[J]. *IEEE Trans Image Process*, 2021, **30**: 7143–7155.
- [18] Chen X, Qi D L, Shen J X. Boundary-aware network for fast and high-accuracy portrait segmentation[Z]. arXiv: 1901.03814, 2019. <https://arxiv.org/abs/1901.03814>.
- [19] Yu C Q, Wang J B, Peng C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[C]// *Proceedings of the 15th European Conference on Computer Vision*, Munich, 2018: 325–341. https://doi.org/10.1007/978-3-030-01261-8_20.
- [20] Zhang Y, He M, Chen Z N, et al. Bridge-Net: context-involved U-net with patch-based loss weight mapping for retinal blood vessel segmentation[J]. *Expert Syst Appl*, 2022, **195**: 116526.
- [21] Li L Z, Verma M, Nakashima Y, et al. IterNet: retinal image segmentation utilizing structural redundancy in vessel networks[C]// *Proceedings of 2020 IEEE Winter Conference on*

Applications of Computer Vision, Snowmass, 2020: 3656–3665.
<https://doi.org/10.1109/WACV45572.2020.9093621>.

- [22] Yang L, Wang H X, Zeng Q S, et al. A hybrid deep segmentation network for fundus vessels via deep-learning framework[J]. *Neurocomputing*, 2021, 448: 168–178.
- [23] Yuan Y C, Zhang L, Wang L T, et al. Multi-level attention

network for retinal vessel segmentation[J]. *IEEE J Biomed Health Inform*, 2022, 26(1): 312–323.

- [24] Li X J, Ding J Q, Tang J J, et al. Res2Unet: a multi-scale channel attention network for retinal vessel segmentation[J]. *Neural Comput Appl*, 2022, 34(14): 12001–12015.

作者简介



【通信作者】吕佳(1978-), 女, 博士, 教授, 博士生导师, 主要从事机器学习、数据挖掘及其在医学图像处理等方面的研究。

E-mail: lvjia@cqnu.edu.cn



梁浩城(1998-), 男, 硕士, 主要从事深度学习及其在医学图像处理等方面的研究。

E-mail: 2021210516051@stu.cqnu.edu.cn。

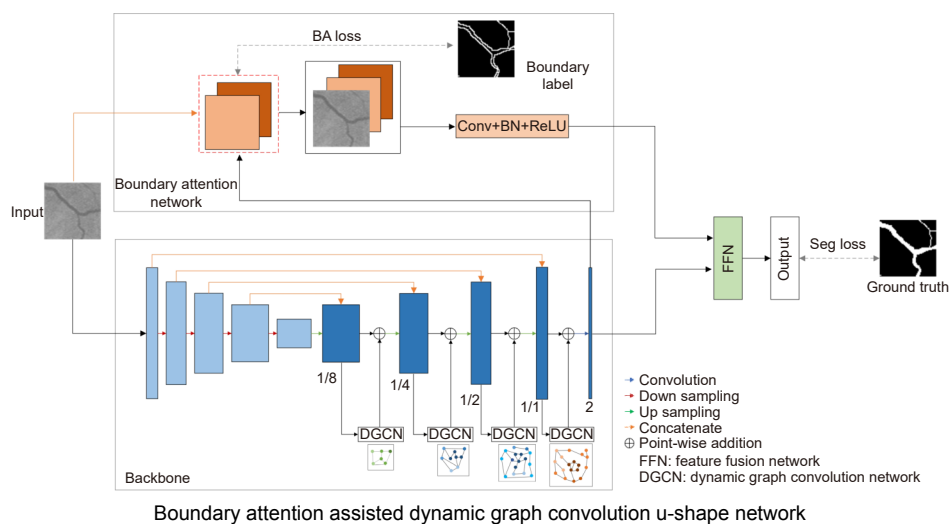


王泽宇(1997-), 女, 硕士, 主要从事深度学习及其在医学图像处理等方面的研究。

E-mail: 2021210516080@stu.cqnu.edu.cn

Boundary attention assisted dynamic graph convolution for retinal vascular segmentation

Lv Jia^{1,2*}, Wang Zeyu¹, Liang Haocheng¹



Overview: The state of retinal blood vessels is an important indicator for clinicians in the auxiliary diagnosis of eye diseases and systemic diseases. In particular, the degree of atrophy and pathological conditions of retinal blood vessels are the key indicators for judging the severity of the diseases. Automatic segmentation of retinal blood vessels is an indispensable step to obtain the key information. Good segmentation results are conducive to accurate diagnosis of the eye diseases. Due to the good characteristic of U-Net that can use skip connection to connect multi-scale feature maps, it performs well in segmentation tasks with small data volume, therefore, it could be applied to retinal vascular segmentation. However, U-Net ignores the features of retinal blood vessels in the training process, resulting in the inability to fully extract the feature information of blood vessels, while its segmentation results show that the vessel pixels are missing or the background noise is incorrectly segmented into blood vessels. Researchers have made various improvements on U-Net for the retinal vessel segmentation task, but the methods still ignore the global structure information and boundary information of retinal vessels. To solve the above problems, a boundary attention assisted dynamic graph convolution retinal vessel segmentation model based on U-Net is proposed in this paper, which supplements the model with more sufficient global structure information and blood vessel boundary information, and extracts more blood vessel feature information as much as possible. First, RGB image graying, contrast-limited adaptive histogram equalization, and gamma correction were used to preprocess the retinal images, which can improve the contrast between the vascular pixels and background, and even improve the brightness of some vascular areas. Then, rotation and slice were adopted to enhance the data. The processed images were input into the model to obtain the segmentation result. In the model, dynamic graph convolution was embedded into the decoder of U-Net to form multi-scale structures to fuse the structural information of feature maps with different scales. The method not only can enhance the ability of dynamic graph convolution to obtain global structural information but also can reduce the interference degree of the noise and the segmenting incorrectly background on the vascular pixels. At the same time, in order to strengthen the diluted vascular boundary information in the process of up-down sampling, the boundary attention network was utilized to enhance the model's attention to the boundary information for the sake of improving the segmentation performance. The presented model was tested on the retinal image datasets, DRIVE, CHASEDB1, and STARE. The experimental results show that the AUC of the algorithm on DRIVE, CHASEDB1 and STARE are 0.9851, 0.9856 and 0.9834, respectively. It is proved that the model is effective.

Lv J, Wang Z Y, Liang H C. Boundary attention assisted dynamic graph convolution for retinal vascular segmentation[J]. *Opto-Electron Eng*, 2023, 50(1): 220116; DOI: 10.12086/oe.2023.220116

Foundation item: National Natural Science Foundation Projects (11991024), Science and Technology Innovation Project of "Construction of Chengdu-Chongqing Twin Cities Economic Circle" (KJCX2020024), Chongqing Education Commission Key Project (KJZD-K202200511), and Technology Foresight and System Innovation Project of Chongqing Science and Technology Bureau (2022TFII-OFX0265)

¹College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China; ²China National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China

* E-mail: lvjia@cqnu.edu.cn