

基于改进 DeepLabv3+ 的 OCTA 图像中央凹无血管区分割算法

李 薇¹ 曹 丽¹
LI Wei CAO Li

摘 要

针对传统 DeepLabv3+ 模型在 OCTA 图像分割中边界定位不准确以及计算资源消耗大的问题,提出了一种改进的 DeepLabv3+ 网络模型用于分割 OCTA 图像的中央凹无血管区,以提升中央凹无血管区的分割效果。首先,用 MobileNetv2 替代 Xception 网络,以降低模型的计算量;其次,用深度可分离卷积对空洞空间金字塔池化模块进行优化,以减少参数量、提高网络多尺度特征提取的能力。同时增加坐标注意力模块,帮助模型更好地捕捉 OCTA 图像中复杂的空间结构信息。实验结果表明,在 OCTA 图像的 FAZ 分割任务中,改进型 DeepLabv3+ 网络模型能够在保持原有分割精度的基础上显著减少模型的参数量、降低模型的计算复杂度。

关键词

OCTA 图像; 中央凹无血管区; DeepLabv3+; 深度可分离卷积; 坐标注意力

doi: 10.3969/j.issn.1672-9528.2025.05.007

0 引言

随着医学影像技术的飞速发展和日益广泛的普及,影像学技术已经成为临床实践中不可或缺的辅助工具,在疾病诊断、手术规划以及预后评估等方面发挥着至关重要的作用。光学相干断层血管成像(OCTA)^[1]是一种新兴的非侵入性眼底成像技术,通过光的干涉原理来捕捉血管的结构和血流信息,从而提供具有高分辨率的血管成像。此前,OCTA 技术已经广泛被应用于各种眼科疾病的诊断之中,包括黄斑区病变、糖尿病视网膜病变、视网膜血管阻塞等。这些眼病与中央凹无血管区(FAZ)^[2]的大小和形态变化有着密切的关联。FAZ 是视网膜中视觉最敏感的部位,对于保持视觉的清晰度和精确性起着至关重要的作用。在正常眼、糖尿病眼和近视眼 3 种不同的情况下,糖尿病眼的 FAZ 面积相较于正常眼有显著增加^[3]。同样近视眼,特别是高度近视眼,其 FAZ 面积会增大,而血管直径则会减小。因此,FAZ 的面积和形态变化可以为糖尿病和近视的临床诊断提供重要的参考依据。在 OCTA 图像中,FAZ 的精确分割对于眼底疾病的诊断具有极其重要的意义。

早期有许多经典的 FAZ 分割方法。例如,阈值分割^[4]是基于像素强度阈值分割 FAZ 的一种简单且常用的方法。该方法将图像的像素值与预定义的阈值进行比较,一旦某个像素值大于阈值,则将其标记为属于 FAZ。阈值分割的结果可以

通过后续形态学操作做进一步的优化。基于区域生长的 FAZ 分割方法^[5]通过选择一个种子点,然后比较相邻像素与种子点的相似性逐渐扩大 FAZ,该方法需要适当的相似性测量和种子点选择。基于形态学的 FAZ 分割方法^[6]则是基于腐蚀、膨胀、开合等形态学操作对包含 FAZ 的图像进行处理,提取感兴趣的结构,可对形态学特征明显的 FAZ 实现良好的分割效果。然而,这些经典的分割方法虽然取得了一定的成效,但仍然存在局限性,如由于图像质量差导致的边界不准确、FAZ 与非灌注区混淆、层投影错误导致的分割误差以及不能很好地适应复杂的图像场景和形状等。

近年来,随着深度学习技术的发展,基于深度学习的眼底图像分割方法取得了很大的成功。U-Net^[7]网络结构以其对称的编码-解码器设计,在医学图像分割中表现优异,能够保留更多的细节信息。此外,DeepLabv3+^[8]网络通过结合空洞卷积技术和空间金字塔池化模块,能捕获丰富的上下文信息。这种创新的方法可显著提高图像分割的精度,使得网络模型在处理复杂场景时能够更加精确地识别和区分不同的对象。这些深度学习模型通过利用大量的眼底图像数据进行训练,能够学习到复杂的图像特征,实现对 FAZ 等关键结构的精确分割。相较于传统的医学图像分割方法,基于深度学习的方法在处理复杂图像场景、提高分割精度和减少误差方面展现出了显著的优势。Guo 等人^[9]提出的 MEDNet 模型,是首个专门检测 OCTA 图像中的 FAZ 的深度学习网络,该网络架构由一系列多尺度的编码器和解码器块构成,其设计目的是能够有效地捕捉和处理在不同尺度上的特征信息。随后,

1. 武汉轻工大学电气与电子工程学院 湖北武汉 430023

Guo 等人^[10]提出的另一种基于编码器-解码器架构的网络则特别采用了边界对齐策略以及边界监督机制,以此来实现对视网膜图像中 FAZ 的自动分割和量化。Zhao 等人^[11]提出了一种基于多维度多区域特性的金字塔模型,实现对多尺度多区域特征的有效融合,从而有效地从整体上抽取重要的特征,能够随着特征图尺寸的减小不断增加特征图的感受野。马晓普等人^[12]提出了一种新的注意力机制—坐标注意力,通过一维全局池化在水平和垂直方向捕获长距离依赖和精确位置信息,以提高轻量级模型的表达能力。

本文对传统的 DeepLabv3+ 模型进行了改进,提出了一种用于分割 OCTA 图像中 FAZ 的深度学习算法,其网络模型由编码器和解码器两大部分组成,用 MobileNetv2 替换传统 DeepLabv3+ 模型中作为特征提取骨干的 Xception 网络,并对空洞空间金字塔池化 (ASPP) 模块进行细致的优化。在优化过程中,将传统的卷积单元用深度可分卷积单元替代,以大幅减少网络模型的参数量,提高模型的训练效率。此外,新模型中还集成了坐标注意力模块,这一模块可显著提升模型对输入图像空间结构的理解能力,进而增强模型在处理空间结构相关任务时的性能。

1 改进的 DeepLabv3+ 模型

1.1 模型总体结构

在编码器模块中,输入图像首先经过 MobileNetv2 主干网络,该网络利用深度可分离卷积层能高效地提取图像的特征信息。接着,得到的特征图会被送入优化后的 ASPP^[13] 模块。在该模块中,特征图分别通过 1×1 卷积层、多个具有不同空洞率 (6、12、18) 的空洞卷积层以及全局平均池化层,经过处理后将结果进行拼接与融合。随后,坐标注意力模块 (CA) 被用于进一步增强对图像重要区域的关注能力。最后,经过 1×1 卷积层的降维处理,以便进行后续的处理和预测。如图 1 所示。

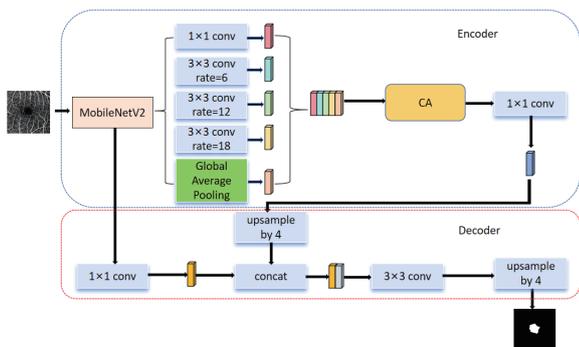


图 1 改进的 DeepLabv3+ 模型总体结构示意图

在解码器模块中,从 MobileNetv2 主干网络提取的低级特征图作为其中一个输入,这些低级特征图经过 1×1 卷积进行通道压缩,既能减少计算量,又能保留重要信息。接下

来,压缩后的低级特征图与编码器部分通过 4 倍上采样得到的高级特征进行融合。融合后的特征图结合了丰富的语义信息和细节。最终,通过 3×3 卷积和双线性插值上采样操作,以恢复空间信息并细化目标边界,从而生成最终的分割预测结果。

1.2 主干网络

MobileNetv2 是一个轻量级的卷积神经网络,旨在通过利用深度可分离卷积来减少参数的数量。这种方法允许在不影响计算效率的情况下维护不同类型数据的准确性。深度可分离卷积可以将传统卷积操作分解成深度卷积和逐点卷积,减少计算量和参数量。深度卷积对每个单独的通道分别进行卷积操作,而逐点卷积负责通过计算输入通道的线性组合来构建新特征。如图 2 所示。

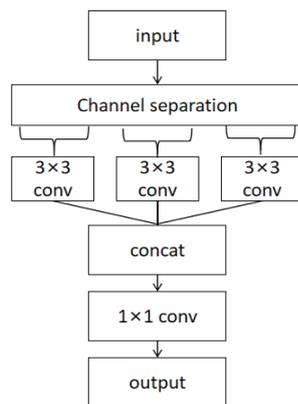


图 2 深度可分离卷积模块结构示意图

1.3 坐标注意力模块

坐标注意力模块 (CA) 专注于处理二维空间信息。其主要操作是通过对输入特征图的行和列进行自适应平均池化处理,计算行和列方向上每个像素点的重要性。其操作可以让模型更加关注图像中不同位置的相关性。在模型中加入 CA 可以增强网络对目标的感知和理解,从而提高模型在复杂场景下的性能。

如图 3 所示,CA 模块的输入特征层大小为 $C \times H \times W$,在水平方向平均池化后,得到的特征层大小为 $C \times H \times 1$,将特征映射到高度维度;在垂直方向平均池化后,得到的特征层大小为 $C \times 1 \times W$,将特征映射到宽度维度;然后将两个平行相位进行合并,将宽度和高度转置到同一维数,再进行叠加,将宽度和高度特征合并在一起,得到的特征层大小为 $C \times 1 \times (W + H)$,通过批标准化和非线性变换来编码垂直方向和水平方向的空间信息,这两种信息可以将其互补地应用于输入特征图,以增强关注对象的表示,再进行拆分,将宽度和高度分离为: $C \times H \times 1$ 和 $C \times 1 \times W$,然后进行转置,再各自通过 1×1 卷积得到输入特征图一样的通道数,其中使用 Sigmoid 激活函数,并归一化加权。

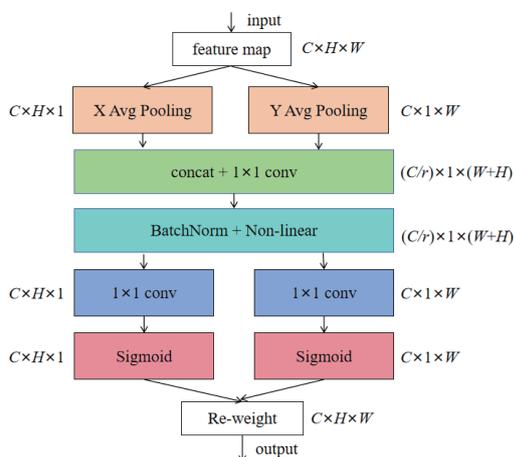


图3 坐标注意力模块结构示意图

2 实验与分析

2.1 数据集

为验证改进的 DeepLabv3+ 模型的有效性, 本文实验采用了 OCTAGON2.0^[14] 和 ROSE^[15] 数据集。OCTAGON2.0 数据集包含 213 张 OCTA 图像, 分辨率为 320 px × 320 px, 其中 144 张为健康的正常眼底图像, 视场大小为 6 mm × 6 mm; 69 张为糖尿病人眼底图像, 视场大小为 3 mm × 3 mm。ROSE 数据集包含 117 张 OCTA 图像, 来自 39 名受试者, 其中 26 名为阿尔茨海默病患者, 13 名为健康受试者, 分辨率为 304 px × 304 px, 视场大小为 3 mm × 3 mm。

2.2 实验设置

本文实验基于 Windows11 操作系统, 用 PyTorch 框架实现, Adam 优化器的学习率为 0.01, 模型训练 200 次, 批处理大小为 8。每个原始图像裁剪到 480 px × 480 px 的大小用于模型训练, 训练集与测试集的比值为 7:3。

2.3 评价指标

本文采用图像语义分割中的平均交并比 MIoU、每秒浮点操作数 FLOPs 和模型训练的参数 Parameter (Params) 来评估模型性能。用公式分别表示为:

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} \quad (1)$$

$$Params = n \times (h \times w \times c + 1) \quad (2)$$

$$FLOPs = H' \times W' \times n \times (h \times w \times c + 1) \quad (3)$$

式中: TP、FP、FN 分别表示真阳性、假阳性、假阴性像素的个数; k 表示分割类别的个数, 设为 1; c 代表输入的通道数; n 表示输出的通道数; $h \times w$ 表示卷积核的大小; $H' \times W'$ 表示输出特征图的尺寸。

2.4 实验结果分析

在 OCTAGON2.0 和 ROSE 数据集上, 与 3 种经典的分割方法进行定量比较的结果如表 1 和表 2 所示。本文算法

在 OCTAGON2.0 数据集上的 MIoU 略低于 U-Net, 但是在 ROSE 数据集上取得了最高的 MIoU 值。这表明, 改进后的模型能够更精确地分割 OCTA 图像中的 FAZ, 减少了误分割和漏分割的情况, 提升了分割的整体效果。

在 FLOPs 指标上, 本文算法表现尤为突出, OCTAGON2.0 数据集上仅为 16.19, ROSE 数据集上为 28.56, 显著低于其他方法。尤其是与传统 DeepLabv3 (65.97、55.40) 和 U-Net (53.44、50.09) 的计算量相比, 本文方法大幅降低了计算复杂度, 使得模型在保持较高分割精度的同时, 具备了更强的计算效率。

在参数量方面, 本文算法也表现优异, OCTAGON2.0 和 ROSE 数据集的参数量分别为 27.43 和 16.71, 均低于 DeepLabv3+ (31.70 和 24.33) 和 U-Net (48.77 和 50.83)。这表明尽管本文方法能够实现更高的分割精度, 但其模型的规模较小, 具备更好的存储和部署优势。

表 1 OCTAGON2.0 数据集上 FZA 分割算法的定量分割结果

方法	OCTAGON2.0		
	MIoU/%	FLOPs/10 ⁹	Params/10 ⁶
U-Net	80.29	53.44	48.77
DeepLabv3	71.33	65.97	40.16
DeepLabv3+	73.86	38.34	31.70
改进的 DeepLabv3+	76.58	16.19	27.43

表 2 ROSE 数据集上 FZA 分割算法的定量分割结果

方法	ROSE		
	MIoU/%	FLOPs/10 ⁹	Params/10 ⁶
U-Net	71.57	50.09	50.83
DeepLabv3	70.63	55.40	44.19
DeepLabv3+	73.05	34.78	24.33
改进的 DeepLabv3+	75.35	28.56	16.71

图 4 展示了使用经典分割方法和本文方法的分割结果。U-Net 仅大致勾勒出 FAZ 轮廓, 对于不规则 FAZ 轮廓的准确分割有难度。DeepLabv3 和 DeepLabv3+ 在 OCTAGON2.0 数据集上表现良好, 但在 ROSE 数据集上对 FAZ 轮廓的分割效果不佳。相比之下, 使用本文的方法分割的 FAZ 轮廓更清晰, 更接近真实情况。

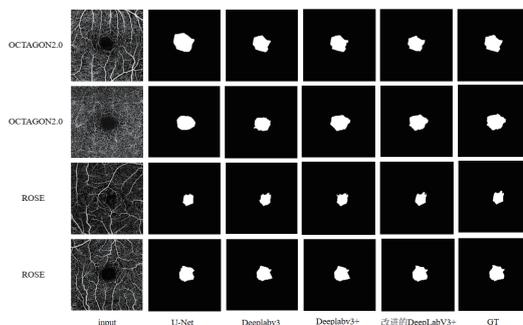


图4 FAZ 分割模型的定性分割结果

2.5 消融实验

在 ROSE 数据集上的消融实验结果如表 3 所示。引入 MobileNetv2 后, MIoU 略微提升至 73.68%, FLOPs 降低至 21.36×10^9 , 参数量也降低至 20.57×10^6 ; 引入 CA 模块后, 虽然 FLOPs 值增加至 36.17×10^9 , 参数量略微增加至 25.85×10^6 , 但是 MIoU 提升至 74.91%。同时引入 MobileNetv2 和 CA 模块后, MIoU 显著提升, FLOPs 和参数量都显著降低, 证明了该算法不仅能够提高分割精度而且可以实现轻量化。

表 3 消融实验结果分析

网络模型	Xception	MobileNetv2	CA	MIoU /%	FLOPs / 10^9	Params / 10^6
DeepLabv3+	√			73.05	34.78	24.33
		√		73.68	21.36	20.57
			√	74.91	36.17	25.85
		√	√	75.35	28.56	16.71

3 总结

本文提出了一种改进的 DeepLabv3+ 模型用于 OCTA 图像的 FAZ 分割。该模型通过使用 MobileNetv2 骨干网络, 并用深度可分离卷积代替初始 ASPP 模块中的标准卷积, 显著减少了模型参数, 同时也提高了模型的训练速度。此外, 引入了坐标注意力机制, 有效地提升了模型的分割精度。一系列的定性和定量实验结果均证明改进的 DeepLabv3+ 模型相比于经典分割算法具有显著的优势。

参考文献:

[1] 王晗. 基于深度学习的 OCTA 图像视网膜血管和中央凹无血管区分割 [D]. 长沙: 中南大学, 2023.

[2] 许畅, 毛晓春. 正常人黄斑中央凹无血管区域面积分析 [J]. 国际眼科杂志, 2017, 17(3): 499-503.

[3] 蓝海磊. 面向糖尿病黄斑缺血分级的无灌注区和视网膜血管分割研究 [D]. 长沙: 中南大学, 2023.

[4] MEHTA N, BRAUN P X, GENDELMAN I, et al. Repeatability of binarization thresholding methods for optical coherence tomography angiography image quantification[J]. Scientific reports, 2020, 10(1): 15368.

[5] LIU J, YAN S X, LU N, et al. Automatic segmentation of foveal avascular zone based on adaptive watershed algorithm in retinal optical coherence tomography angiography images[J]. Journal of innovative optical health sciences, 2022, 15(1): 2242001.

[6] GONG D, YU W H, ZHANG X, et al. A morphological study

of the foveal avascular zone in patients with diabetes mellitus using optical coherence tomography angiography[J]. Graefes's archive for clinical and experimental ophthalmology, 2016, 254(5): 873-879.

[7] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [J]. Medical image computing and computer assisted intervention, 2015, 9351: 234-241.

[8] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[J]. Computer vision ECCV, 2018, 11211: 833-851.

[9] GUO Y K, WANG J, JIA Y L, et al. MEDNet, a neural network for automated detection of avascular area in OCT angiography [J]. Biomedical optics express, 2018, 9(11): 5147-5158.

[10] GUO M L, ZHAO M, CHEONG A M Y, et al. Automatic quantification of superficial foveal avascular zone in optical coherence tomography angiography implemented with deep learning [J]. Visual computing for industry, biomedicine, and art, 2019, 2(1): 21.

[11] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 6230-6239.

[12] 马晓普, 刘文涛, 李贺. 改进型 DeepLabV3+ 的糖尿病眼底病变分割 [J]. 华中科技大学学报 (自然科学版), 2024, 52(5): 90-97.

[13] 邝先验, 陈奕希, 刘平, 等. 基于双重注意力机制的皮肤病变图像分割算法 [J]. 计算机应用与软件, 2024, 41(11): 261-267.

[14] DIAZ M, NOVO J, CUTRIN P, et al. Automatic segmentation of the Foveal avascular zone in ophthalmological OCT-A images[J]. Public library of science one, 2019, 14(2):212364.

[15] MA Y H, HAO H Y, FU H Z, et al. ROSE: a retinal OCT-angiography vessel segmentation dataset and new model [J]. IEEE transactions on medical imaging, 2021, 40(3):928-939.

【作者简介】

李薇 (1999—), 女, 山西晋城人, 硕士研究生, 研究方向: 医学图像处理。

曹丽 (1987—), 女, 湖南益阳人, 博士, 讲师, 研究方向: 信号处理、人工智能和医学图像处理的快速算法。

(收稿日期: 2024-12-30 修回日期: 2025-04-28)