

双线性非局部特征结合中继监督网络用于视网膜血管分割

杨东旭¹, 赵红东¹, 耿立新¹, 于快快²

1. 河北工业大学电子信息工程学院, 天津 300401; 2. 光电信息控制和安全技术重点实验室, 天津 300308

【摘要】对眼底图像中的视网膜血管精准分割是检测多种疾病的关键技术,在相关疾病自动筛查系统中发挥着重要的作用。针对现存方法追求分割精度时忽略对算法复杂性的考虑,导致在资源受限的医疗设备上部署困难的问题。本文通过进一步合理减少卷积层的特征通道数量来轻量化分割网络并提出了BNIS-Net。该网络采用多尺度图像作为输入融合到编码过程中,使得不同感受野之间建立良好的联系,并提出一种双线性非局部模块来增强相关上下文信息的捕捉能力。最后,在解码过程中采用中继监督的策略,为解码部分各级输出提供监督来约束网络的学习,这样可以有效改善收敛行为使浅层部分得到充分训练。BNIS-Net以0.41 M的参数量在DRIVE、STARE和CHASE_DB1 3个公开数据集上分别取得了81.02%、81.07%、78.15%的DSC值和0.983 3、0.986 1、0.985 9的AUC值。通过大量对比实验和消融研究证明,该方法能够更好地分割血管的边缘细节。

【关键词】糖尿病视网膜病变; 图像处理; 视网膜血管分割; 双线性非局部模块; 多尺度输入; 中继监督

【中图分类号】R318; TP391.41

【文献标志码】A

【文章编号】1005-202X(2022)12-1516-09

Bilinear non-local features combined with intermediate supervision network for retinal vessel segmentation

YANG Dongxu¹, ZHAO Hongdong¹, GENG Lixin¹, YU Kuaikua²

1. School of Electronics and Information Engineering, Hebei University of Technology, Tianjin 300401, China; 2. Key Laboratory of Photoelectric Information Control and Safety Technology, Tianjin 300308, China

Abstract: The accurate segmentation of retinal vessels in fundus images is of significance for the detection of various diseases and plays an important role in automated screening system for associated diseases. To address the problem that existing methods neglect to consider the complexity of the algorithm when pursuing segmentation accuracy, which leads to difficulties in deployment on resource-constrained medical devices, the number of feature channels in the convolutional layer is further reduced to lighten the segmentation network, and a bilinear non-local intermediate supervision network (BNIS-Net) is proposed. In BNIS-Net, the multi-scale images are taken as input and fused into the coding for establishing good connections between different receptive fields, and a bilinear non-local module is added to enhance the capture of relevant contextual information. During the decoding, an intermediate supervision strategy is adopted to constrain the learning of the network by providing supervision to the output of the decoding at all levels, which can effectively improve the BNIS-Net uses a parameter of 0.41 M on 3 public data sets of DRIVE, START and CHASE, and achieves DSC values of 81.02%, 81.07% and 78.15%, and AUC values of 0.983 3, 0.986 1 and 0.985 9, respectively. It was demonstrated by numerous comparative experiments and ablation studies that the method can better segment the edge details of vessels.

Keywords: diabetic retinopathy; image processing; retinal vessel segmentation; bilinear non-local module; multi-scale input; intermediate supervision

【收稿日期】2022-06-28

【基金项目】天津市科技计划(21YDTPJC00050); 光电信息控制和安全技术重点实验室基金(2021JCJQLB055008)

【作者简介】杨东旭, 博士, 研究方向: 医疗人工智能、医学图像处理, E-mail: tonshy@163.com

【通信作者】赵红东, E-mail: zhaohd@hebut.edu.cn

前言

世界卫生组织报道糖尿病视网膜病变(Diabetic Retinopathy, DR)是人类失明的主要原因^[1]。DR是一种具有特殊眼底变化的糖尿病并发症。其主要特征是出现了视网膜新生血管。DR的症状可以通过视网膜眼底图像检测出来,但很多患者忽略了对早期症状的关注。视网膜眼底图像不仅可以检测出

DR,也是诊断青光眼、高血压、关节硬化和老年性黄斑变性等与视网膜血管结构异常相关疾病的一种有效方法^[2]。眼科医生可以通过观察视网膜血管的形态特征,如形状、曲率、长度和分支来诊断疾病^[3]。因此,视网膜血管眼底图像的自动分割技术可以有效提高病理疾病检查的准确性,帮助眼科医生快速诊断和定制治疗方法。

近几十年来,随着计算机性能的不断提高,越来越多的算法被应用于视网膜血管的分割^[4-8]。根据学习模式的不同,可概括为基于规则的方法和深度学习方法^[9]。基于规则的方法通常采用匹配滤波、数学形态学、血管跟踪或基于模型的技术^[10]。匹配滤波技术利用血管本身的特性,通过最大化滤波响应来模拟血管的轮廓。数学形态学的方法主要考虑图像的内部形状,不考虑像素的强度,也就是说,它忽略了图像内容的细节。最后,基于模型的方法主要分为轮廓模型和变形模型^[11]。视网膜血管具有分支和交叉点过多的特点,这使得血管轮廓模型更加复杂。然而,变形模型通常只考虑图像的边缘信息,而忽略了其他特征,这就导致了最终分割的准确性不高。

基于深度学习的方法得到了快速的发展并广泛应用于视网膜血管分割^[12-17],这些方法通过构建神经网络来实现端到端的学习,已成为解决该任务的主流研究方法。然而,他们普遍忽略了对算法复杂度的考虑,导致将算法在实际的医疗设备上应用遇到阻碍。本文通过合理减少卷积层的特征通道数量来轻量化分割网络并提出了BNIS-Net(Bilinear Non-local Intermediate Supervision Network)。针对图像单尺度输入和单标签监督的方式,从不同程度上限制了网络的特征学习能力。BNIS-Net采用多尺度图像作为输入融合到编码过程中,使得不同感受野之间建立良好的联系,并提出一种双线性非局部模块来增强相关上下文信息的捕捉能力。最后,在解码过程中采用中继监督的策略,为解码部分各级输出提供监督来约束网络的学习,这样可以有效改善收敛行为使浅层部分得到充分训练。在DRIVE、STARE和CHASE_DB1 3个公开数据集上通过大量实验表明,该方法在分割血管的边缘细节上具有更好的表现。

1 方法

1.1 BNIS-Net 网络结构

提出的BNIS-Net致力于实现一个轻量级且被生物医学领域广泛应用的分割模型。如图1所示,BNIS-Net是一种可以端到端学习的网络模型,它与U-Net的主要区别有3个。首先,采用了多尺度图像输入结合空间全局注意模块提取更丰富的语义信息来弥补编码带来的信息损失,而不是只有原图一种尺寸作为输入。其次,在解码器部分采用了双线性非局部模块,以增强网络的特征编码能力。最后采用中继监督策略,将网络中间层的局部输出与全局输出关联起来,进一步提升对目标边缘细节的分割精度。

1.2 多尺度特征输入

U-Net中的下采样操作导致了部分空间信息的丢失,尤其是对于像视网膜血管这类细小边界的分割任务,仅使用跳跃连接不足以更好地恢复原有特征。以往工作是将不同尺度的图像输入到不同的网络中,汇聚所有网络的输出得到最终的输出映射。而本文采用反混叠池化对图像进行下采样得到多种尺度的图像作为编码器的输入。这样操作的优点是融合多尺度特征丰富了初始语义信息,避免了参数的激增和网络复杂性提高。图1详细展示了所提出的多尺度特征输入方法。

1.3 双线性非局部模块

医学图像中往往包含大量决策不相关信息。特别是眼底图像中,很多背景信息是多余的,我们更关注视网膜血管部分,因为它能提供诊断相关疾病所

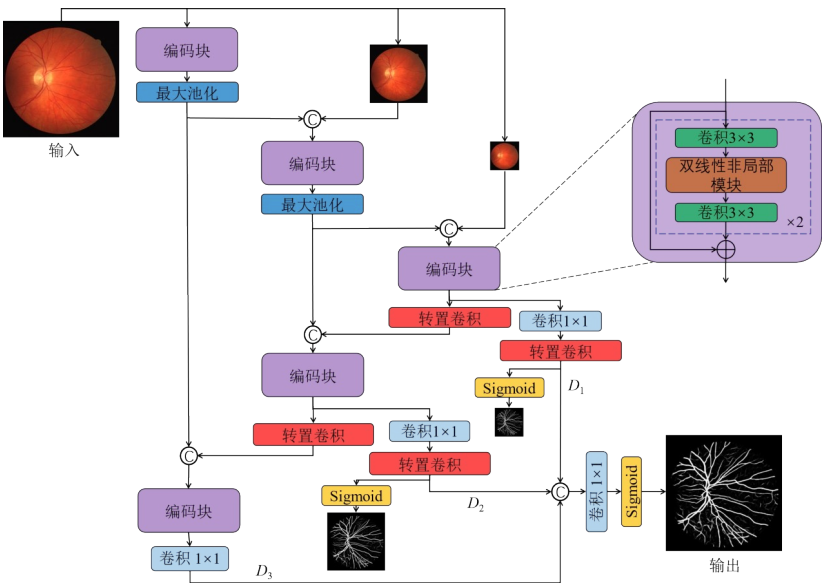


图1 BNIS-Net网络结构图
Figure 1 BNIS-Net structure

需要的信息。因此,尝试采用注意力机制,在有效增强对感兴趣部分关注的同时,减少背景中冗余信息的影响。近年来,注意力机制在许多领域得到了广泛的应用,并产生了许多基于注意力机制的方法。Vaswani等^[18]证明,使用自注意力模块关注所有位置并在嵌入空间中对其加权平均,来计算序列中某个位置(例如句子)的响应,并在机器翻译任务中获得最先进的结果。Wang等^[19]在自注意力方法基础上进行简化,形成更为通用的非局部过滤操作用于处理计算机视觉任务。双线性策略是一种由两个特征提取器组成的识别体系结构^[20],它们的输出在图像的每个位置使用外积操作进行组合,以获得一个图像描述符。该结构能够以平移不变的方式将局部特征之间的交互进行建模,这在细粒度分类任务中是特别有效的,双线性形式上是允许特征提取器 P_A 和 P_B 的输出进行相互调整的。

受上述工作启发,本文设计一种双线性非局部模块,如图2所示。

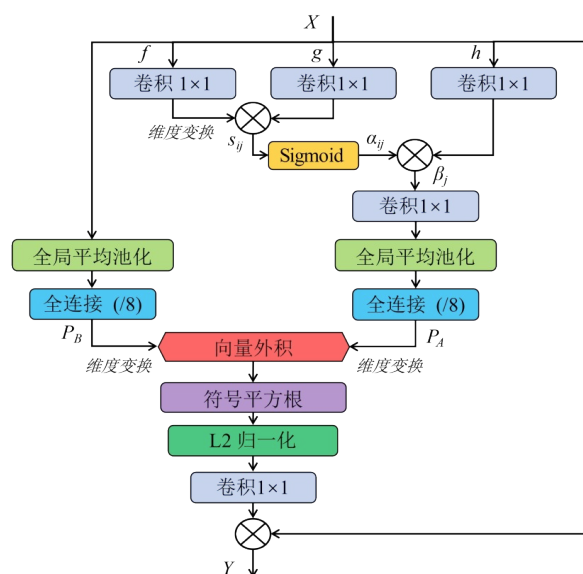


图2 双线性非局部模块

Figure 2 Bilinear non-local module

首先将来自特征提取器的特征 X 作为输入,然后将其转换为 f 、 g 和 h 3个特征空间来计算注意力。相似性函数 s_{ij} 如式(1)所示:

$$s_{ij} = f(x_i)^T g(x_j) = (W_f x_i)^T W_g x_j \quad (1)$$

这里 $f(x) = W_f x$, $g(x) = W_g x$, $h(x) = W_h x$ 。Softmax 归一化操作 α_{ij} 表示模型在合成第 j 个区域时对第 i 个位置的关注程度,可表示为:

$$\alpha_{ij} = \frac{\exp(s_{ij})}{\sum_{j=1}^N \exp(s_{ij})} \quad (2)$$

把 β_i 表示为上述产出 α_{ij} 和 $h(x)$ 的外积,它可以表示为:

$$\beta_i = \sum_{j=1}^N \alpha_{ij} h(x) = \sum_{j=1}^N \alpha_{ij} W_h x_j \quad (3)$$

在上述过程中,学习到的权重矩阵 $W_f \in R^{L \times C}$ 、 $W_g \in R^{L \times C}$ 和 $W_h \in R^{L \times C}$,可以通过线性投影或 1×1 卷积操作实现。 C 是通道的数量, N 是来自特征提取器的特征数量。将 L 的通道数减少到 $C/8$ 时,总的计算成本与全维度相似。该结构中的两个特征函数分支分别为 P_A 和 P_B ,特征函数是一个映射,它取一个图像 $i \in I$ 和一个位置 $l \in L$,输出一个大小为 $K \times D$ 的特征。在每个位置上的特征输出使用矩阵外积进行组合,即在一个位置上 P_A 和 P_B 的双线性组合由以下公式给出:

$$F_{\text{bilinear}}(l, i, P_A, P_B) = P_A(l, i)^T P_B(l, i) \quad (4)$$

其中, P_A 和 P_B 必须具有相同的维度 K ,其值取决于之前的特征提取器。随后将得到的双线性特征执行有符号的平方根操作和L2归一化处理。将结果与开始的输入特征 X 相乘就得到了最终的输出 Y 。

1.4 中继监督策略

研究人员对基于U型网络的改进,更多的聚焦于对网络组件模块的修改,这些工作被证明是有效的,但是忽略了对网络中间层的输出进行监督。模型训练过程中随着层数的增加,反向传播的梯度逐渐变小,模型学习速度降低,仅在最后一层使用单一的监督方式,中间层难以更好地实现对低级语义特征的学习,导致网络性能遭遇瓶颈。

为缓解梯度消失问题,增强中间层特征学习能力,本文设计了一种中继监督机制。具体做法如图1所示,在解码器中每个卷积层后接入一个卷积核为 1×1 的单通道卷积层,然后采用Sigmoid激活函数来输出预测概率图。由于每个概率图输出具有不同的分辨率,为此笔者将标注图像进行相应的缩放,来保持和各个分辨率一致便于计算他们之间的损失值。最后将各级输出融合到一起得到最终的混合损失,为了确保中继监督的输出和输入图像具有相同的分辨率,融合前对低分辨率输出接入一个上采样层。例如, D_1 的输出分辨率是输入原图的 $1/4$, D_2 的输出分辨率是原图的 $1/2$, D_3 的输出分辨率与原图是一致的。因此对 D_1 采用步长为4的上采样操作,对 D_2 采用步长为2的上采样操作将分辨率恢复到与输入图像分辨率相同,而 D_3 无需额外操作。

将训练数据记为 $T = \{(X_n, Y_n), n = 1, \dots, N\}$,这里 $X_n = \{x_i^n, i = 1, \dots, |X_n|\}$ 表示生物医学训练图像, $Y_n = \{y_i^n, i = 1, \dots, |X_n|\}$, $y_i^n \in \{0, 1\}$ 表示与之对应的标注值。将网络中参数的集合表示为 W ,并将中继监

督总数记为 M 。每个中继监督网后面接入一个 sigmoid 激活函数,其对应的权值记为 $\omega = (\omega^1, \dots, \omega^M)$ 。这里采用 sigmoid 函数是因为本文处理的生物学图像属于二值预测,当类别数大于2时则应采用 Softmax 激活函数。因此,网络整体的代价函数可以定义为:

$$L_{is} = (W, \omega) = \sum_{m=1}^M \alpha_m l_{is}^m(W, \omega^m) \quad (5)$$

其中, l_{is} 表示在网络中间层的中继监督策略所采用的损失函数, α_m 表示第 m 个输出损失占总损失的权重。

2 实验结果与分析

2.1 数据集与评估指标

DRIVE 数据集包含 40 幅分辨率为 584×560 的不同疾病进展程度的眼底图像^[21],其中 20 幅用于训练,20 幅用于测试。同时,为每张图像提供了人工分割结果作为标签值。STARE 数据集有 20 张图像,提供了人工分割的结果,图像分辨率为 605×700 ^[22]。CHASE_DB1 数据集有 28 张图像,包括人工分割的结果,分辨率为 960×999 ^[23]。对于 STARE 和 CHASE_DB1 两个数据集,实验中将一半数据进行训练,另一半进行测试。

一些评价标准被广泛用于图像分割任务中,以评价不同算法的性能。准确度(ACC)指的是像素的准确度,它表示正确的像素数占总像素数的比例。精度(Pr)表示实际正确的样本与总的预测正确样本的比例。灵敏度(Se)表示在样本中,有多少阳性例子被预测正确。特异性(Sp)表示识别的阴性样本占所有阴性样本的比例。交并比(Intersection of Union, IoU)反映了人工分割区域面积和算法自动分割区域面积的交集占二者并集的比例。戴斯相似系数(Dice Similarity Coefficient, DSC),该指标反映两倍的重合区域面积占人工分割区域面积和算法自动分割区域面积之和的比例。另外还有接受者操作特征曲线(ROC)的曲线下面积(AUC)。这些指标的具体表达方式如下:

$$Se = \frac{TP}{TP + FN} \quad (6)$$

$$Pr = \frac{TP}{TP + FP} \quad (7)$$

$$Sp = \frac{TN}{TN + FP} \quad (8)$$

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

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

$$DSC = 2 \times \frac{Se \times Pr}{Se + Pr} \quad (11)$$

其中,TP和TN分别表示真阳性样本和真阴性样本,而FP和FN分别代表假阳性样本和假阴性样本。

2.2 实验环境与设置

硬件环境为 Intel(R)Core(TM)i7-8700 CPU, Nvidia RTX2070 GPU。开发环境基于 Windows10 操作系统, CUDA10.0+cuDNN10.0+TensorFlow1.14.0+keras2.2.4,编程语言为 Python3.7.3。将图像大小统一为 576×576 ,并选择 Adam 优化器来最小化混合损失函数,学习率为 0.001。在训练过程中,批次大小被设置为 2,训练最大迭代次数为 300。为了防止过度拟合,实验中采用了早停策略。

2.3 消融研究

为了验证多尺度输入、双线性非局部模块和中继监督策略的有效性,进行了一系列的消融实验。首先,将类似 U-Net 的轻量化浅层网络作为 Baseline,该网络未添加任何额外的模块。在 Baseline 结构中单独引入多尺度输入,表示为 Baseline+MSI。在 Baseline 上仅添加双线性非局部模块,用 Baseline+BNL 表示。让 Baseline 与中继监督策略结合,将其表示为 Baseline+IS。最后将多尺度输入、双线性非局部模块和中继监督策略同时应用在 Baseline 上构成了本文模型 BNIS-Net。

如表 1 所示,本文提出的多尺度输入方法、双线性非局部模块和中继监督策略在分割性能的定量度量中有显著的提升。其中,与 Baseline 相比,采用多尺度输入在 DRIVE 数据集上的表现,在 DSC 和 ACC 上分别提升了 0.59% 和 1.43%。该改进表明,提取多尺度信息有利于捕获更多有用特征。在 Baseline 中添加双线性非局部模块 BNL 后,充分学习特征的长程依赖,评估指标 DSC 和 ACC 相比 Baseline 分别提升了 0.90% 和 1.51%。随后,将中继监督机制引入 Baseline,将中间层的局部输出与全局输出关联起来,评估指标 DSC 和 ACC 相比 Baseline 分别提升了 1.08% 和 1.49%。最后,将上述策略同时用于 Baseline,性能有了较大的提升,即 DSC 和 ACC 分别提高了 2.30% 和 1.61%。因此,上述实验结果极大地证明了本文提出的 BNIS-Net 可以减少语义差距,增强特征之间的语义关系,有效地细化多层次的特征表示。为了更直观的对比各改进方法的表现,可视化结果如图 3 所示。在整体插图中选择并放大了一些细节,以便于进一步观察微小血管的分割结果。

2.4 对比不同分割方法

这部分实验旨在与典型的深度学习算法在图像分割任务中进行比较,如 SegNet^[24]、DeepLabV3+^[25]、ENet^[26]、U-Net^[13]和 Att-UNet^[27],从而验证提出的算法在视网膜血管分割中的有效性。如表 2 所示,从定

表1 对BNIS-Net进行消融研究
Table 1 Ablation studies on BNIS-Net

模型	Se/%	Sp/%	DSC/%	ACC/%	AUC
Baseline	72.92	98.49	78.72	95.11	0.979 4
Baseline+MSI	74.91	98.66	79.31	96.54	0.980 6
Baseline+BNL	75.56	98.64	79.62	96.62	0.981 0
Baseline+IS	76.83	98.49	79.80	96.60	0.981 1
BNIS-Net	82.00	98.38	81.02	96.72	0.983 3

量分析来看,本文模型 BNIS-Net将边缘血管的结构信息更好地集成到语义分割特征中。在 DRIVE、

STARE 和 CHASE_DB1 3 个数据集上,BNIS-Net 在 IoU、DSC 和 ACC 3 个评估指标上都取得了最佳的表现。这表明该方法相比其他卷积神经网络 (Convolutional Neural Network, CNN)具有更好的量化性能。

在监督学习中,敏感度和特异性两个评估指标通常需要平衡。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。因此,本文适当地牺牲了特异性来获得更高的 IoU、DSC 和 ACC。

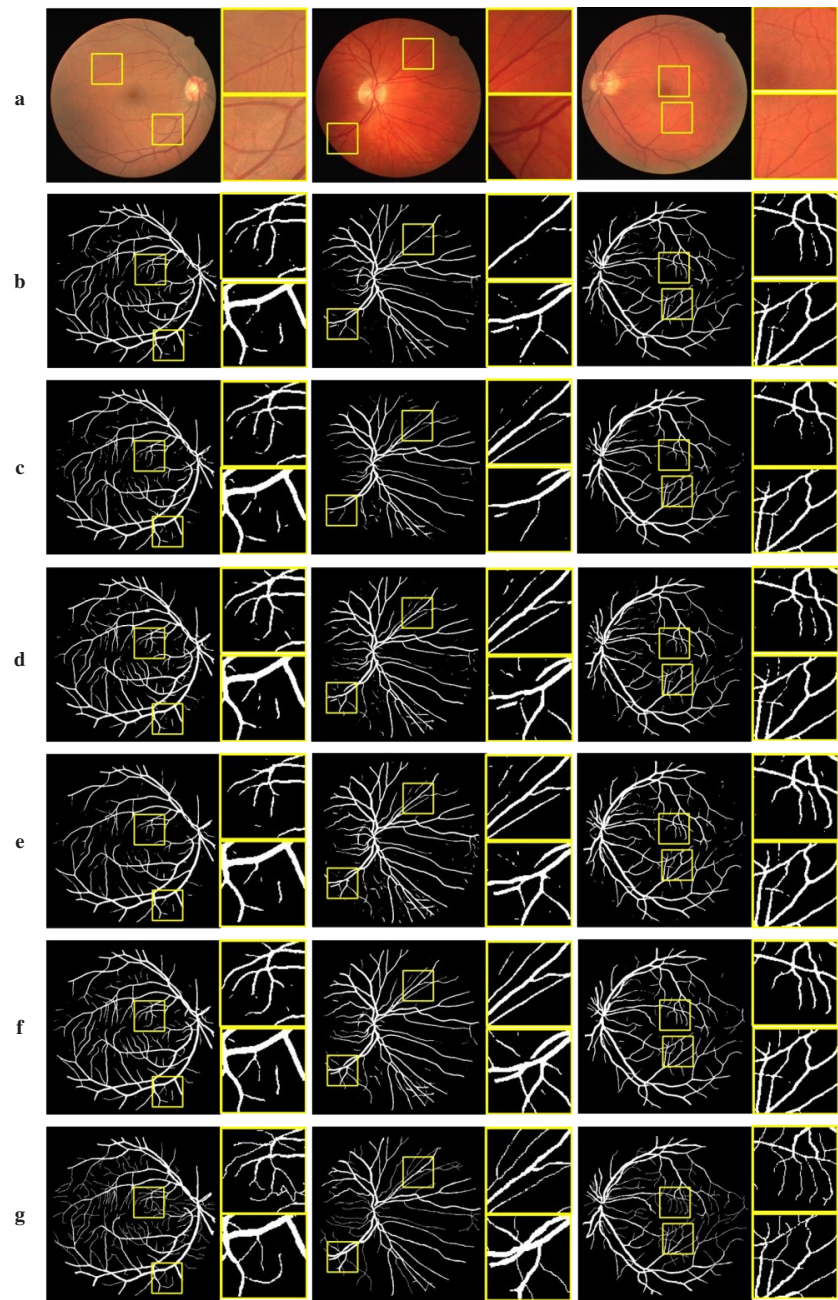


图3 从DRIVE、STARE和CHASE_DB1 3个数据集中任取1张原始图像及各方法分割结果

Figure 3 An original image taken from the data sets of DRIVE, STARE and CHASE_DB1 and the segmentation results of different methods
a:原始彩色图像,b:Baseline,c:Baseline+MSI,d:Baseline+BNL,e:Baseline+IS,f:BNIS-Net,g:标注图像;第1~3列分别为DRIVE、STARE和CHASE_DB1数据集图像

表 2 对比不同方法在 3 个数据集上的表现

Table 2 Comparison of the performance of different methods on 3 data sets

数据集		Se/%	Sp/%	IoU/%	DSC/%	ACC/%
DRIVE	SegNet ^[24]	86.90	82.21	32.81	49.41	82.62
	DeepLabV3+ ^[25]	66.88	98.08	55.62	71.51	95.35
	ENet ^[26]	71.16	97.89	58.32	73.67	95.56
	U-Net ^[13]	75.71	98.30	64.18	78.18	96.32
	Att-UNet ^[27]	74.15	98.75	65.59	79.22	96.60
	BNIS-Net	82.00	98.38	68.09	81.02	96.72
STARE	SegNet ^[24]	77.38	88.56	32.89	49.50	87.70
	DeepLabV3+ ^[25]	57.69	98.04	46.51	63.49	94.94
	ENet ^[26]	40.98	99.04	36.92	53.93	94.38
	U-Net ^[13]	66.53	98.90	58.69	73.97	96.42
	Att-UNet ^[27]	72.29	98.80	63.16	77.42	96.77
	BNIS-Net	81.81	98.34	68.17	81.07	97.07
CHASE_DB1	SegNet ^[24]	85.50	74.16	18.47	31.18	74.86
	DeepLabV3+ ^[25]	67.19	96.78	44.96	62.03	94.97
	ENet ^[26]	56.89	98.30	45.15	62.21	95.77
	U-Net ^[13]	78.11	97.99	59.73	74.79	96.77
	Att-UNet ^[27]	69.13	98.65	57.26	72.82	96.84
	BNIS-Net	83.99	97.98	64.13	78.15	97.12

此外,笔者还使用ROC曲线对所有方法进行了评估,如图4所示。ROC曲线的横坐标是假阳性率,纵坐标是真阳性率。当曲线更接近左上边界时,该方法的效果更好。可以看出,对于DRIVE、STARE和

CHASE_DB1数据集,BNIS-Net获得的AUC最大,分别为0.983 3、0.986 1、0.985 9。图5提供了不同方法在3个数据集上的可视化结果细节。

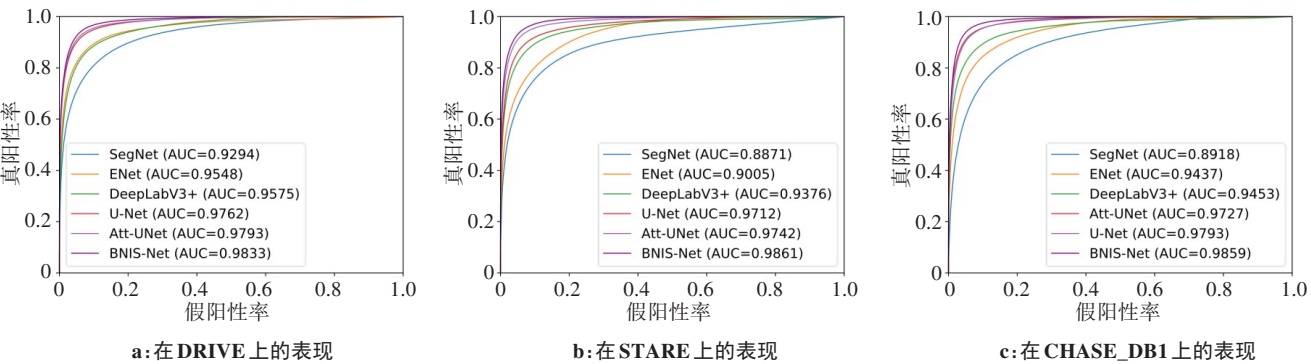


图4 对比各方法在3个数据集上所得到的ROC曲线和AUC值

Figure 4 Comparison of ROC curves and AUC values obtained by different methods on 3 data sets

2.5 模型复杂度分析

本小节考虑到模型的复杂度情况,在DRIVE数据集上对比了5种基于CNN的分割方法。模型参数数量、内存占用和计算量被列出以供比较,如表3所示。从表3可以得出结论,本文方法的参数数量和计算量微高于轻量化分割模型ENet^[26],而内存占用是最

低的仅为5.40 MB,AUC值也明显优于ENet。此外,笔者还对比了不同方法单张图像的推理耗时,BNIS-Net耗时为0.22 s。在可接受的时间消耗范围内本文方法进一步提高了分割精度,并支持其在临床视网膜筛查中的应用。

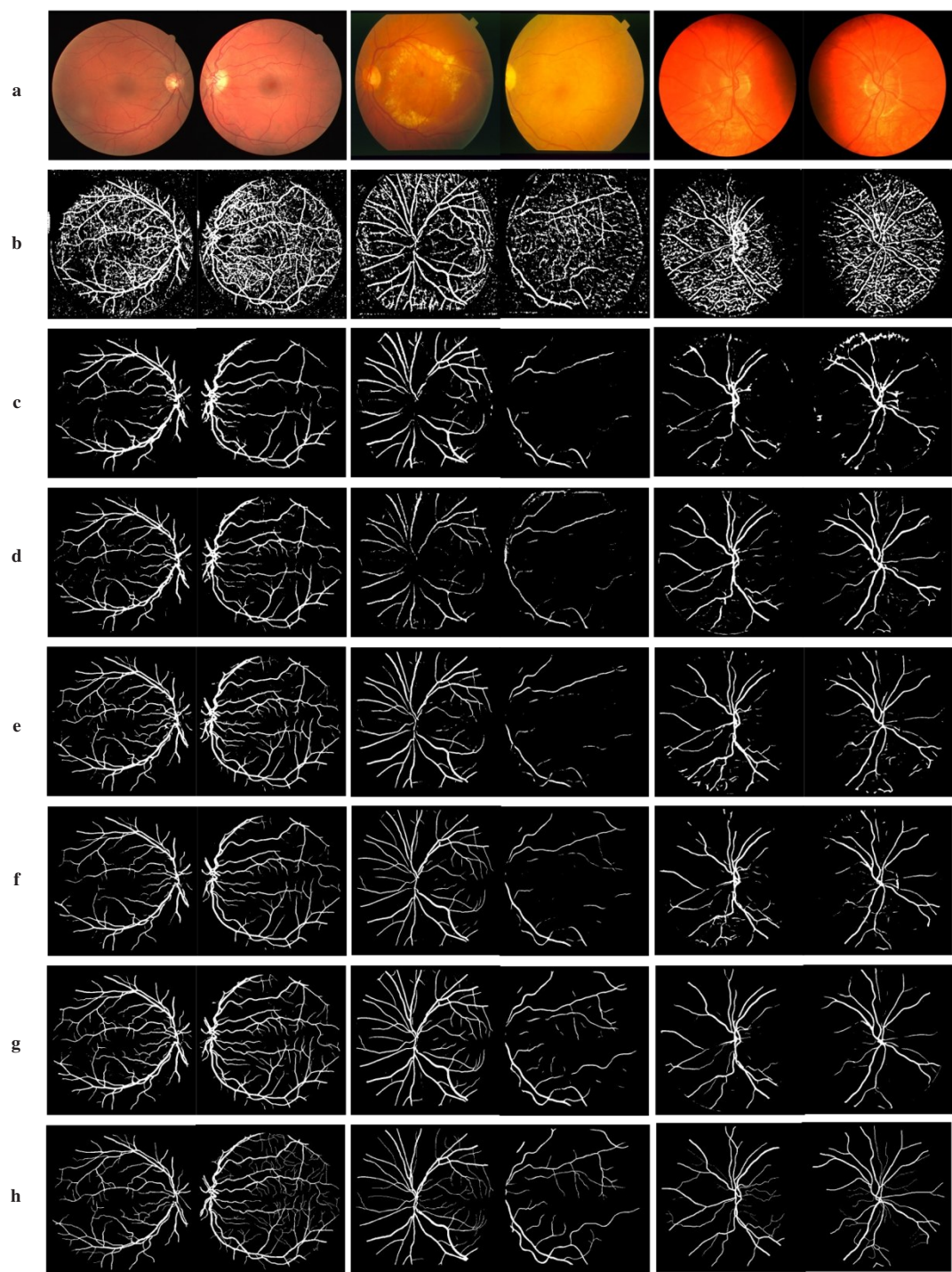


图 5 从 DRIVE、STARE 和 CHASE_DB1 3 个数据集中任取两张原始图像及各方法分割结果

Figure 5 Two original images taken from the data sets of DRIVE, STARE and CHASE_DB1 and the segmentation results of different methods
a: 原始彩色图像; b: SegNet; c: DeepLabV3+; d: ENet; e: U-Net; f: Att-UNet; g: NAUNet; h: 标注图像; 第 1~3 列分别为 DRIVE、STARE 和 CHASE_DB1 数据集图像

2.6 交叉训练

泛化能力是评价深度学习模型性能的一个重要指标,特别是在医学图像分析任务中。因此,笔者采用交叉训练的方式评估所提模型的泛化能力。交叉训练表示在一个数据集上直接测试另一个数据集上训练的模型权重。表 4 显示了该方法与当前先进方法在泛化能力上的比较结果。与现有的其他方法相

比,本文模型在 STARE 训练和 DRIVE 测试中获得了良好的性能,ACC 和 AUC 取得最高值,分别为 96.35% 和 0.973 5。同样,对于在 DRIVE 的训练后用 STARE 来测试,其 ACC 值为 96.40%, AUC 值为 0.968 6。

2.7 局限性讨论

视网膜血管特别是微小血管的分析对相关眼科

表3 不同方法的复杂度分析

Table 3 Comparison of the complexity of different methods

模型	参数量/M	内存占用/MB	计算量/M	推理耗时/s	AUC
SegNet ^[24]	29.46	337.37	58.91	0.28	0.929 4
ENet ^[26]	0.37	5.51	0.74	0.16	0.954 8
DeepLabV3+ ^[25]	41.06	470.08	82.23	0.21	0.957 5
U-Net ^[13]	2.07	23.77	4.13	0.18	0.976 2
Att-UNet ^[27]	8.91	102.41	17.82	0.19	0.979 3
BNIS-Net	0.41	5.40	0.81	0.22	0.983 3

表4 交叉训练评估结果

Table 4 Results of the cross-training evaluation

测试数据集		Se/%	Sp/%	ACC/%	AUC
DRIVE(在 STARE 训练)	Yan 等 ^[17]	70.14	98.02	94.44	0.956 8
	Li 等 ^[28]	72.73	98.10	94.86	0.967 7
	DUNet ^[29]	65.05	99.14	94.81	0.971 8
	BNIS-Net	71.45	98.74	96.35	0.973 5
STARE(在 DRIVE 训练)	Yan 等 ^[17]	73.19	98.40	95.80	0.967 8
	Li 等 ^[28]	70.27	98.28	95.45	0.967 1
	DUNet ^[29]	70.00	97.59	94.74	0.957 1
	BNIS-Net	76.24	98.08	96.40	0.968 6

疾病的诊断、监测和治疗计划至关重要。这些疾病引起的血管阻塞症状可引起视网膜缺血,导致视网膜新生血管。在这些疾病的进展阶段,新生小血管的异常生长会导致视力受损甚至失明。血管分割时微小血管的丢失可能导致这些疾病的漏诊和延误治疗。经过统计,DRIVE 和 STARE 两个数据集的图像血管像素的数量通常不超过整个图像的 10%,一些微小血管的长度甚至小于 10 像素。

为了更好地分析结果存在的偏差,从 STARE 数据集中任取一种视网膜血管图像,如图 6 所示。其中图 6a 为原始的视网膜图像,图 6b 为数据集提供的标注图像,图 6c 显示了本文所提方法的分割结果。在数据集的介绍中了解到,原始图像包含对比度较差的图像和伴有眼部疾病的图像,对这些图像进行分割将会出现较差的结果。事实上,出现这种情况的原因不是本文方法的血管检测能力不足,而是算法高效且敏感,对标注图像中未标注出来的,但原图中又真实存在的细微血管也检测了出来。以 STARE 数据集中的一张图像为例,图 6d~f 对图 6a~c 图像中方框部分进行放大,便于更好地呈现图像细节信息。图 6e 是标注图像给出的结果,图 6f 本文所提方法的

分割结果。可以明显看出标注图像中遗漏了不显著的细小血管,而本文方法几乎将肉眼可以观察到的血管全部检测了出来。对于伴有病变和质量较差的图像,并没有影响本文所提方法很好的分割性能,但人工标记的标注图像导致了存在的偏差。笔者认为如果有更多人工标注更为准确的图像用于训练,那么最终的分割性能还会有所提升。

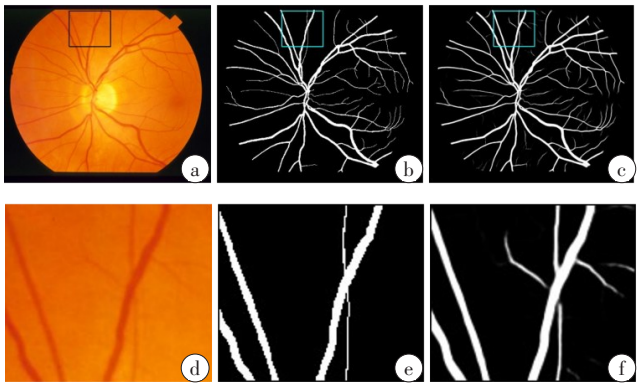


图6 通过局部放大观察细小血管

Figure 6 Observing the small blood vessels through local magnification

a: 原始图像; b: 标注图像; c: BNIS-Net 分割结果; d~f: 代表图 a~c 所对应的方框部分的放大图像

3 结 论

在这项研究中,提出一种新的医学图像分割框架,即BNIS-Net。其中采用多尺度图像作为输入融合到编码器中,使得不同感受野之间建立良好的联系,并提出一种双线性非局部模块来获取特征长程依赖增强相关上下文信息的捕捉能力。最后,在解码阶段采用中继监督方法,为解码器各级输出提供监督来约束网络学习,这样可以有效改善收敛行为使浅层部分得到充分训练。在3个公开视网膜血管数据集上,通过大量实验和定量定性的分析,表明本文所提方法能够更好的分割细节,具有较强的泛化能力和良好的应用前景。

【参考文献】

- [1] Soomro TA, Afifi AJ, Gao J, et al. Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation[J]. Expert Syst Appl, 2019, 134: 36-52.
- [2] Rezaee K, Haddadnia J, Tashk A. Optimized clinical segmentation of retinal blood vessels by using combination of adaptive filtering, fuzzy entropy and skeletonization[J]. Appl Soft Comput, 2017, 52: 937-951.
- [3] 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.
- [4] 罗文劫, 韩国庆, 田学东. 多尺度注意力解析网络的视网膜血管分割方法[J]. 激光与光电子学进展, 2021, 58(20): 431-444.
Luo WJ, Han GQ, Tian XD. Retinal vessel segmentation method for multi-scale attention resolution networks[J]. Laser & Optoelectronics Progress, 2021, 58(20): 431-444.
- [5] 王亮, 陈春晓, 傅雪, 等. 基于FDMU-net的早产儿视网膜图像血管分割[J]. 激光与光电子学进展, 2021, 58(14): 475-481.
Wang L, Chen CX, Fu X, et al. FDMU-net based retinal image vascular segmentation in preterm infants[J]. Laser & Optoelectronics Progress, 2021, 58(14): 475-481.
- [6] 孟琳, 刘静, 曹慧, 等. 基于Frangi滤波器和Otsu视网膜血管分割[J]. 激光与光电子学进展, 2019, 56(18): 127-133.
Meng L, Liu J, Cao H, et al. Frangi filter-based and Otsu retinal vessel segmentation[J]. Laser & Optoelectronics Progress, 2019, 56(18): 127-133.
- [7] 王文斌, 李灿标, 郑楚君. 基于Hessian的方向自适应Gabor小波的视网膜血管分割[J]. 激光与光电子学进展, 2020, 57(8): 208-215.
Wang WB, Li CB, Zheng CJ. Hessian-based directional adaptive Gabor wavelets for retinal vessel segmentation [J]. Laser & Optoelectronics Progress, 2020, 57(8): 208-215.
- [8] 张欢, 仇大伟, 冯毅博, 等. U-Net模型改进及其在医学图像分割上的研究综述[J]. 激光与光电子学进展, 2022, 59(2): 1-17.
Zhang H, Chou DW, Feng YB, et al. A review of U-Net model improvement and its research on medical image segmentation[J]. Laser & Optoelectronics Progress, 2022, 59(2): 1-17.
- [9] Jin Q, Meng Z, Pham TD, et al. DUNet: a deformable network for retinal vessel segmentation[J]. Knowl-Based Syst, 2019, 178: 149-162.
- [10] Xiang Y, Gao X, Zou B, et al. Segmentation of retinal blood vessels based on divergence and bot-hat transform [C]//2014 IEEE International Conference on Progress in Informatics and Computing. IEEE, 2014: 316-320.
- [11] Lam BS, Gao Y, Liew AW. General retinal vessel segmentation using regularization-based multiconcavity modeling[J]. IEEE Trans Med Imaging, 2010, 29(7): 1369-1381.
- [12] Tang XL, Zhong B, Peng JP, et al. Multi-scale channel importance sorting and spatial attention mechanism for retinal vessels segmentation[J]. Appl Soft Comput, 2020, 93: 106353.
- [13] 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.
- [14] Palanivel DA, Natarajan S, Gopalakrishnan S. Retinal vessel segmentation using multifractal characterization [J]. Appl Soft Comput, 2020, 94: 106439.
- [15] Feng S, Zhuo Z, Pan D, et al. CcNet: a cross-connected convolutional network for segmenting retinal vessels using multi-scale features[J]. Neurocomputing, 2020, 392(2): 268-276.
- [16] Yan Z, Yang X, Cheng KT. Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation[J]. IEEE Trans Biomed Eng, 2018, 65(9): 1912-1923.
- [17] Yan ZQ, Yang X, Cheng KT. A three-stage deep learning model for accurate retinal vessel segmentation[J]. IEEE J Biomed Health Inform, 2019, 23(4): 1427-1436.
- [18] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. arXiv preprint arXiv: 1706.03762, 2017.
- [19] Wang X, Girshick R, Gupta A, et al. Non-local neural networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 7794-7803.
- [20] Lin TY, Roy Chowdhury A, Maji S. Bilinear CNN models for fine-grained visual recognition[C]//2015 IEEE International Conference on Computer Vision (ICCV). IEEE, 2015: 1449-1457.
- [21] Staal J, Abramoff MD, Niemeijer M, et al. Ridge-based vessel segmentation in color images of the retina[J]. IEEE Trans Med Imaging, 2004, 23(4): 501-509.
- [22] Soares JV, Leandro JJ, Cesar RM, et al. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification[J]. IEEE Trans Med Imaging, 2006, 25(9): 1214-1222.
- [23] Owen CG, Rudnicka AR, Mullen R, et al. Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (CAIAR) program[J]. Invest Ophthalmol Vis Sci, 2009, 50(5): 2004-2010.
- [24] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(12): 2481-2495.
- [25] Chen LC, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 833-851.
- [26] Paszke A, Chaurasia A, Kim S, et al. ENet: a deep neural network architecture for real-time semantic segmentation[J]. arXiv preprint arXiv: 1606.02147, 2016.
- [27] Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: learning where to look for the pancreas[J]. arXiv preprint arXiv: 1804.03999, 2018.
- [28] Li QL, Feng BW, Xie LP, et al. A cross-modality learning approach for vessel segmentation in retinal images[J]. IEEE Trans Med Imaging, 2016, 35(1): 109-118.
- [29] Jin Q, Meng Z, Pham TD, et al. DUNet: a deformable network for retinal vessel segmentation[J]. Knowl-Based Syst, 2019, 178: 149-162.

(编辑:薛泽玲)