计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2022,31(10):368−374 [doi: 10.15888/j.cnki.csa.008738]
©中国科学院软件研究所版权所有.
E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于多尺度高阶注意力机制的视网膜血管分割①

姜璐璐[1], 李思聪[2], 曹加旺[1], 孙司琦[3], 冯　瑞[1,3], 邹海东[2,4]

[1](复旦大学 工程与应用技术研究院, 上海 200433)
[2](上海交通大学附属第一人民医院, 上海 200080)
[3](复旦大学 计算机科学技术学院, 上海 200433)
[4](苏州市产业技术研究院, 苏州 215011)
通信作者: 邹海东, E-mail: zouhaidong@sjtu.edu.cn

摘　要: 视网膜血管分割对于辅助医生诊断糖尿病性视网膜病变、黄斑萎缩、青光眼等眼科疾病具有重要意义. 注意力机制被广泛用于 U-Net 及其变体中以提高血管分割模型的性能. 为进一步提高视网膜血管的分割精度, 挖掘视网膜图像中的高阶及全局上下文信息, 本文提出基于多尺度高阶注意力机制的模型 (multi-scale high-order attention network, MHA-Net). 首先, 多尺度高阶注意力 (multi-scale high-order attention, MHA) 模块从深层特征图中提取多尺度和全局特征计算初始化注意力图, 从而改进模型处理医学图像分割时尺度不变的缺陷. 接下来, 该模块通过图的传递闭包构建注意力图, 进而提取高阶的深层特征. 通过将多尺度高阶注意力模块应用于编码器-解码器结构中, 在彩色眼底图像数据集 DRIVE 上进行血管分割, 实验结果表明, 基于多尺度高阶注意力机制的视网膜血管分割方法有效地提高了分割的精度.

关键词: 视网膜血管分割; 注意力机制; 多尺度高阶注意力; 空洞卷积; 深度学习; 图像分割

引用格式: 姜璐璐,李思聪,曹加旺,孙司琦,冯瑞,邹海东.基于多尺度高阶注意力机制的视网膜血管分割.计算机系统应用,2022,31(10):368–374. http://www.c-s-a.org.cn/1003-3254/8738.html

Retinal Vessel Segmentation Based on Multi-scale High-order Attention Mechanism

JIANG Lu-Lu[1], LI Si-Cong[2], CAO Jia-Wang[1], SUN Si-Qi[3], FENG Rui[1,3], ZOU Hai-Dong[2,4]

[1](Academy for Engineering and Technology, Fudan University, Shanghai 200433, China)
[2](Shanghai General Hospital, Shanghai 200080, China)
[3](School of Computer Science, Fudan University, Shanghai 200433, China)
[4](Suzhou Industrial Technology Research Institute, Suzhou 215011, China)

Abstract: Retinal vessel segmentation is vital for assisting doctors in diagnosing ophthalmic diseases, including diabetic retinopathy, macular atrophy, and glaucoma. The attention mechanism is widely used in U-Net and its variants to improve the vessel segmentation performance. For more accurate retinal vessel segmentation and exploration of high-order and global context information, we propose a multi-scale high-order attention network (MHA-Net). The multi-scale high-order attention (MHA) module first extracts multi-scale and global features from the high-level feature maps to compute the initial attention map, enabling the model to handle medical image segmentation with variable scales. Then the high-order attention constructs the attention map through graph transduction followed by the extraction of high-level features at high order. We further embed the proposed MHA module into an efficient encoder-decoder structure for retinal vessel segmentation. Comprehensive experiments are conducted on the color fundus image dataset DRIVE, which indicates that the proposed method improves the accuracy of retinal vessel segmentation effectively.

Key words: retinal vessel segmentation; attention mechanism; multi-scale high-order attention; dilated convolution; deep learning; image segmentation

① 基金项目: 上海市科委项目 (20DZ1100200)
收稿时间: 2022-01-07; 修改时间: 2022-02-24; 采用时间: 2022-03-01; csa 在线出版时间: 2022-07-07

© 中国科学院软件研究所 http://www.c-s-a.org.cn

1 引言

血管系统是视网膜的基本结构, 其形态学和拓扑结构的变化可以用来识别和分类系统性代谢和血液疾病的严重程度, 例如糖尿病和高血压[1]. 糖尿病性视网膜病变 (DR) 是糖尿病的一种常见并发症, 是由视网膜微血管渗漏和阻塞导致的一系列眼底病变. DR 可引起新血管的生长, 是否有异常新生血管也是判断增殖性 DR 与非增殖性 DR 的标准[2]. 高血压视网膜病变 (HR) 是另一种常见的由高血压引起的视网膜疾病[3]. 在高血压患者中, 可以观察到血管弯曲度增加或血管狭窄[4]. 通过视网膜血管获得的血管形状和分叉的信息, 可以增强对 DR 或者 HR 的监测. 因此, 分割视网膜血管对于一些严重疾病的早期诊断与治疗具有重要意义.

现有的眼底视网膜成像技术有以下几类: 彩色眼底照相 (FP) 技术、眼底荧光素血管造影 (FFA)、光学相干断层扫描 (OCT) 以及眼底相干光层析血管成像 (OCTA). 彩色眼底照相是最常用的视网膜成像技术, 其优点是获取方式简单、图像易于观察.

传统的无监督方法一般包括: 滤波匹配法、区域生长、血管跟踪、阈值分割和图像形态学处理等. 这些传统的无监督方法不需要人工标注, 但这些方法依赖于手工提取特征进行血管表示与分割. 此外, 此类算法存在分割精度不够、泛化性较差等局限性.

与传统的无监督方法相比, 深度卷积神经网络方法具有更强大的特征表征和学习能力, 在医学图像分割任务中取得了最高水平[5]. 自 2015 年引入 U-Net[6] 以来, 它已成为医学影像分割中最具影响力的深度学习框架[7–10]. 其整体网络采用编码器-解码器的结构, 通过“跳跃连接”将不同分辨率的特征图进行通道融合产生较好的分割效果. 尽管 U-Net 具有良好的表示能力, 但它依赖于多级级联卷积神经网络. 这种方法在重复提取低层特征时会导致计算资源的过度和冗余使用[11].

注意力机制被提出用于解决以上问题, 其模仿了人类视觉所特有的大脑信号处理机制, 令网络从大量信息中重点关注对任务结果更重要的区域, 而抑制其他不重要的部分[12]. 在视网膜血管分割任务中, 背景像素占比较大, 而血管像素的占比小, 因此可以采用注意力机制关注血管区域. 卷积神经网络可以利用不同类型的注意力机制以关注重要的区域或者特征通道[13–18]. 例如, 空间注意力机制[11,18] 利用特征的空间关系生成空间注意力图从而使网络关注具有丰富信息的区域, 通道注意力机制[13] 通过显式建模通道间的依赖关系来提高模型的性能. 空间注意力和通道注意力的融合[15] 也已成功地应用于医学分割领域.

然而, 这些常用的方法是一阶注意力机制, 难以提取图像中一些更为抽象的高阶语义信息且不能充分利用到全图像的信息, 导致在处理形状和结构复杂的目标时发生退化[19]. 尤其在视网膜血管分割任务中, 由于血管形态结构多变, 以上方法仍欠缺对复杂和高阶特征信息的捕获能力.

本文提出了一种基于多尺度高阶注意力机制的视网膜图像分割方法 (MHA-Net), 可以明显提高视网膜血管的分割精度. 该方法采用改进的 U-Net 结构, 并引入多尺度高阶注意力模块, 对编码器提取到的深层特征进一步处理, 聚焦于图像的高阶语义信息, 从而改进模型处理医学图像分割时尺度不变的缺陷. 经过在 DRIVE[20] 数据集上的实验证明, 该方法有效地提高了分割的精度, 同时对细小血管的分割也更为精细.

2 相关工作

2.1 空洞卷积在图像分割中的应用

空洞卷积 (dilated convolution)[21,22] 通过在卷积核相邻两个元素之间插入零值, 在不增加参数量和计算成本的同时扩大了感受野. 受空洞空间金字塔池化 (ASPP)[23] 在语义图像分割中的应用启发, 空洞卷积在医学图像分割中同样得到了广泛的应用[17,24]. 但是, 基于空洞卷积的分割方法都存在一个共同问题, 稀疏采样会造成详细信息的丢失, 从而导致像素级分类不准确. D-LinkNet[25] 利用“短路连接 (shortcut)”结合了文献 [21] 的级联模型与文献 [1] 的并行模型.

之前的研究主要集中在通过增加在不同尺度特征图上的感受野, 从而直接提高分割网络的性能. 我们的工作与上述方法不同, 我们利用空洞卷积对不同尺度的特征图进行采样, 并通过聚合这些多尺度的特征图产生高阶注意力图, 从而进一步使网络聚焦于更加抽象和全面的语义信息.

2.2 高阶注意力

注意力机制的思想核心是通过计算权重矩阵而使网络有选择地关注具有重要信息的部分[12]. Okty 等人[11] 提出了用于医疗影像分割的注意力门控 (attention gate, AG) 模型, 该模型可以自动学习区分目标的外形和尺寸, 在小目标分割任务中效果尤其显著. 不同于在跳跃

© 中国科学院软件研究所 http://www.c-s-a.org.cn

连接中添加注意力门控 (AG) 的方法, SA-UNet[14] 引入了一个空间注意力模块, 通过在空间维度计算注意力权重矩阵并与输入的特征图相乘, 实现自适应地细化特征. 该方法是注意力模块在 U 形分割网络降采样后的深层特征图上的一种应用. Chen 等[19] 首先提出了高阶注意力模型, 并将其应用于行人重识别建模. 该模型利用注意机制中形成的复杂高阶统计量, 捕捉行人之间的细微差异, 从而产生区别性的关注建议. Ding 等[26] 利用图的传递闭包进一步优化高阶注意力模块, 在此基础上提出具有自适应感受野和动态权重的 high-order attention (HA) 模块. HA 模块通过图的传递闭包构建注意力图, 从而捕获高阶的上下文相关信息.

之前的一些工作 (如文献 [13]) 通过在 U 型网络的底部引入注意力机制来进一步挖掘深层次的特征. 然而, 这些网络更多地关注了局部信息, 而忽略深层特征中的全局信息. 这导致尽管在提取深层特征时添加了几种不同类型的注意力模块, 也不能有效地提高医学图像分割任务的性能. 相反, 模型的性能甚至会略有下降.

本文的工作是在上述注意力机制[14,19,26] 上的改进. 在 U 形网络的多个降采样块之后所得的深层特征的噪声相对较小, 因此注意力模块需要尽可能地挖掘深层特征中的全局信息. 另一方面, 与浅层特征相比, 在深层特征中引入噪声会对整个模型造成更大的损害. 因此, 本文设计了多尺度高阶注意力 (MHA) 模块, 其在不引入噪声的前提下引导网络提取深层特征中的更为全局的信息, 有效提高了视网膜血管中分割性能.

3 基于多尺度高阶注意力机制的分割网络

3.1 网络框架

图 1 给出了基于多尺度高阶注意力机制的视网膜图像分割方法 (MHA-Net) 的网络架构, 其遵循了编码器-解码器的 U 型结构. 编码器包含若干个下采样块和 MHA 模块, 其中每个下采样块由 1 个 3×3 的卷积层、1 个批处理规范化层和一个 ReLU 激活函数层组成, 3 个下采样块连接在一起后紧跟一个 2×2 的最大池化操作. 在下采样完成之后, 将提取到的图像深层次特征输入到 MHA 模块进行细化, MHA 模块的位置放置于网络底部, 即 U 型收缩路径和扩张路径之间. 在此处加入 attention 模块的原因是在靠前位置采集到的为低层次结构信息, 包含有许多噪声. 此外, 加权的 shortcut 被引入以保留原本的上下文信息. 最后, 经过融合得到的特征图通过编码器产生最终的分割结果. 解码器部分使用反卷积[27] 进行上采样操作.

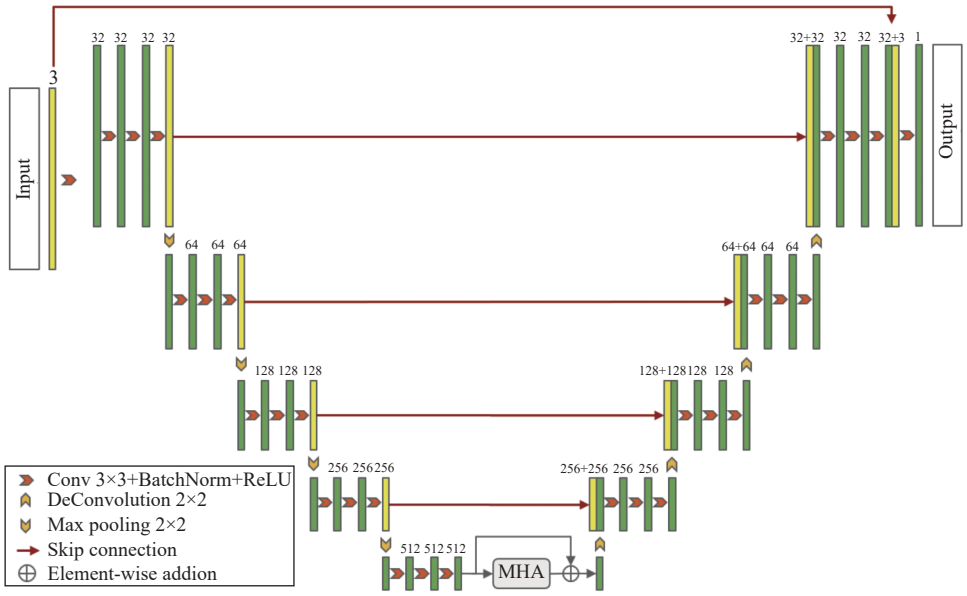


图 1 MHA-Net 架构图

3.2 多尺度高阶注意力模块

本文提出的多尺度高阶注意力模块如图 2 所示. 在编码器的底部, 原始的特征图$X^{\text{in}} \in \mathbb{R}^{H\times W\times C}$通过并行的共享权重的空洞卷积 (膨胀率 r 分别为 1, 2, 4, 8), 产生新的多尺度特征图分为为X_r (r=1, 2, 4, 8), 通过 1×1 卷积得到的特征图为X^*. 将这些多尺度特征图使用

© 中国科学院软件研究所 http://www.c-s-a.org.cn

式 (1) 计算得到融合的多尺度注意力矩阵A_r^*:

$$A_{r=1,2,4,8}^* = \frac{1}{C}\prod_{r=1,2,4,8} X_r \tag{1}$$

其中, $1/C$是用来控制数值爆炸的缩放因子. 之后, 利用图的传递闭包计算了多尺度高阶注意力矩阵A_m^*, $m \in \{1, 2, \cdots, n\}$. 具体计算$A_m^*$的细节将在第 3.3 节讨论. 最后, 将特征图$X^*$与归一化的高阶注意力矩阵相乘得到细化的特征图$X_m$, 如式 (2):

$$X_m = \Gamma_\theta\left(\frac{\exp(A_m^*[i,j])}{\sum_{j=1}^{N}\exp(A_m^*[i,j])}\cdot X^*\right) \tag{2}$$

Γ_θ代表 1×1 卷积. 在多尺度高阶注意力模块之后, 将细化后的特征图X_m乘上自适应因子α以抵消缩放因子$1/C$的偏移影响, 如式 (3):

$$X^{\text{out}} = \alpha\cdot\Gamma_\theta\left(\|_{h=1}^{n}(X_m)\right) + X^{\text{in}} \tag{3}$$

其中, $\|_{h=1}^{n}$代表通道连接.

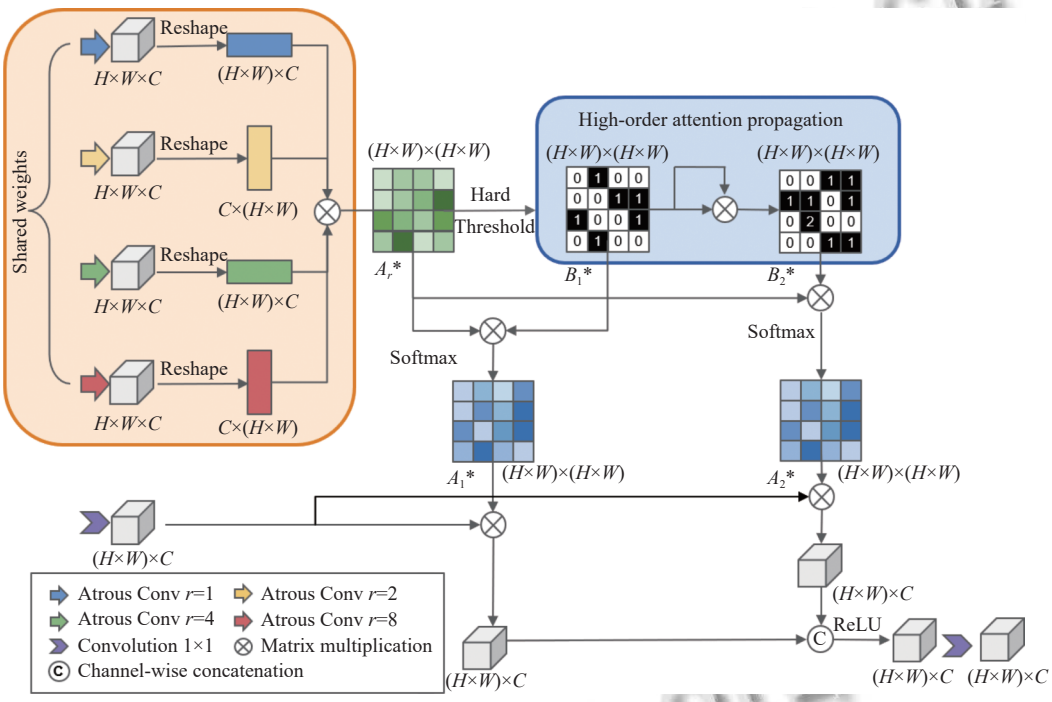


图 2 多尺度高阶注意力 (MHA) 模块

深层特征图在通过多尺度高阶注意力模块之后, 提取了更加高阶抽象的语义特征, 也更具有区分力, 从而更聚焦于血管的分割. 之后, 再通过解码器模块, 逐渐从低分辨率重构至高分辨率.

3.3 高阶注意力传播机制

根据文献 [26], 最初的多尺度注意力融合矩阵A_r^*可以看做图的邻接矩阵, 图中的边表示连接的两个节点属于同一类. 如图 3 所示, 给定注意力图A_r^*, 通过阈值化删去低置信度的边后形成下采样的图B_1^*, 如式 (4):

$$B_1^*[i,j] = \begin{cases} 1, A_r^*[i,j] \geqslant \delta \\ 0, A_r^*[i,j] < \delta \end{cases} \tag{4}$$

其中, δ代表阈值, 设置为 0.5. 如图 4 所示, 根据图的传递闭包, B_m^*可以通过邻接矩阵B_1^*自乘 m–1 次得到:

$$B_m^* = \underbrace{B_1^* B_1^* \cdots B_1^*}_{m} \tag{5}$$

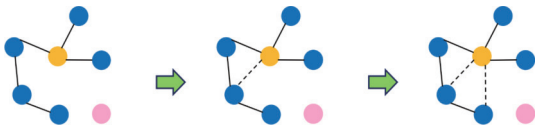


图 3 三阶高阶注意力传播原理图: 以黄色点为中心点通过图的传递闭包进行传播

高阶图B_i^*可以通过式 (5) 得到. 最后, 通过哈达玛积得到高阶注意力矩阵, 如式 (6) 所示:

$$A_m^*[i,j] = A_r^*[i,j] \times B_m^*[i,j] \tag{6}$$

其中, m 表示邻接矩阵幂次的整数, 代表注意力传播的阶数. 因此, 不同层次的注意力信息通过B_m^*解耦成不同

© 中国科学院软件研究所 http://www.c-s-a.org.cn

的注意图并得到高度相关的邻居. 生成的高阶注意图 A_m^* 用于聚合多层次的上下文信息.

4 实验验证与结果分析

4.1 数据集与实验参数

本文使用的数据集是 DRIVE (digital retinal image for vessel extraction)[20]. 该数据集包含 40 张图像像素尺寸为 584×565 的彩色眼底图像, 其中训练集与测试集各 20 张. 为扩充数据, 避免训练样本过少可能造成的过拟合问题, 我们对训练样本随机采样 256×256 的 patch. 此外, 使用随机翻转、随机旋转、弹性形变等方法进行数据增强. 本文使用 PyTorch 框架[28], 批量设置为 16, 采用 Adam 算法[29] 优化模型, 学习率设置为 0.0001. 动量和权重衰减因子分别设置为 0.9 和 0.999.

4.2 评价指标

为了对实验结果进行客观的定量分析, 选取以下指标进行计算: Dice 系数 (*DSC*)、准确率 (*ACC*)、敏感度 (*SE*)、特异性 (*SP*) 和 ROC 曲线下面积 AUC. AUC 的范围在 0–1 之间, AUC 越逼近 1, 其模型预测能力越高. 评价指标的计算方式如下:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \tag{7}$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

$$SE = \frac{TP}{TP + FN} \tag{9}$$

$$SP = \frac{TN}{TN + FP} \tag{10}$$

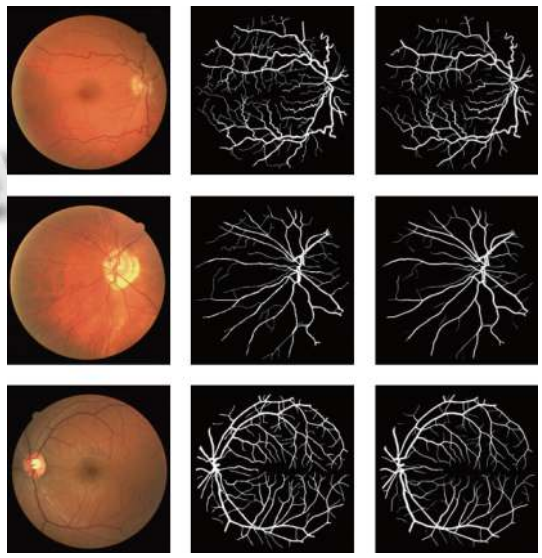
其中, X 代表金标准, Y 代表预测结果. 真阳性 TP 为正确分类的血管像素个数, 真阴性 TN 正确分类的背景点像素个数, 假阳性 FP 为背景像素误分成血管像素的个数, 假阴性 FN 为血管像素误分成背景像素的个数.

4.3 实验结果分析

本文算法性能在 DRIVE 数据集上评估, 图 4 展示了部分分割结果. 图 4(a) 为原始图像, 图 4(b) 为金标准图像, 图 4(c) 为本文算法的分割结果, 从结果可以看出, 本文算法整体分割效果良好, 平滑度也优于金标准. 同时, 本文算法细节上表现优秀, 保持了微血管的连通性, 说明本文中采取的注意力机制能够关注到重要的血管区域.

为了验证本文所提出的模型性能的优越性, 表 1 将本文算法与近两年最先进的血管分割算法的各项指标进行对比, 其中加粗字体部分为每项最优指标.

结果表明, 本文提出的多尺度高阶注意力方法 MHA-Net 取得了优异的表现, 其 Dice 系数、灵敏度和 AUC 分别达到了 0.8266、0.8312 和 0.9883, 在所有方法中表现最优. 本文算法在保证高准确率的同时, 有着良好的敏感度, 这意味着分割结果尽可能地保留血管信息, 分割得到的血管连续完整. 综上所述, 本文算法整体性能优于现有算法.



(a) 原始图 (b) 金标准 (c) 本文分割结果

图 4 DRIVE 数据集分割结果

表 1 DRIVE 数据集上不同算法分割性能比较

算法	年份	*DSC*	*ACC*	*SE*	*SP*	AUC
Vessel-Net[8]	2019	—	0.9578	0.8083	0.9802	0.9821
AG-Net[18]	2019	0.8211	0.9692	0.8100	**0.9848**	0.9856
WU-Net[9]	2020	—	0.9689	0.8213	0.9837	0.9869
IterNet[10]	2020	0.8218	0.9574	0.7791	0.9831	0.9813
SA-UNet[14]	2020	0.8263	0.9698	0.8218	0.9840	0.9864
HANet[26]	2020	—	**0.9712**	0.8297	0.9843	—
MHA-Net	2021	**0.8266**	0.9697	**0.8312**	0.9833	**0.9883**

为了证明提出的多尺度高阶注意力 (MHA) 模块的有效性, 在 DRIVE 数据集上还进行了消融实验. 表 2 展示了 U-Net、U-Net+MHA、Backbone、Backbone+HA 以及 MHA-Net 的分割性能. 其中 U-Net+MHA 表示在 U-Net 基础上引入 MHA 模块的网络, Backbone 表示在本文使用的骨干网络, Backbone+HA 表示与本文相同的骨干网络上引入原始的高阶注意力 HA 模块, MHA-Net 为本文算法, 相当于 Backbone+MHA, 即在本文使用的骨干网络上引入多尺度高阶注意力 MHA 模块.

结果表明: (1) U-Net+MHA 比 U-Net 有更好的性

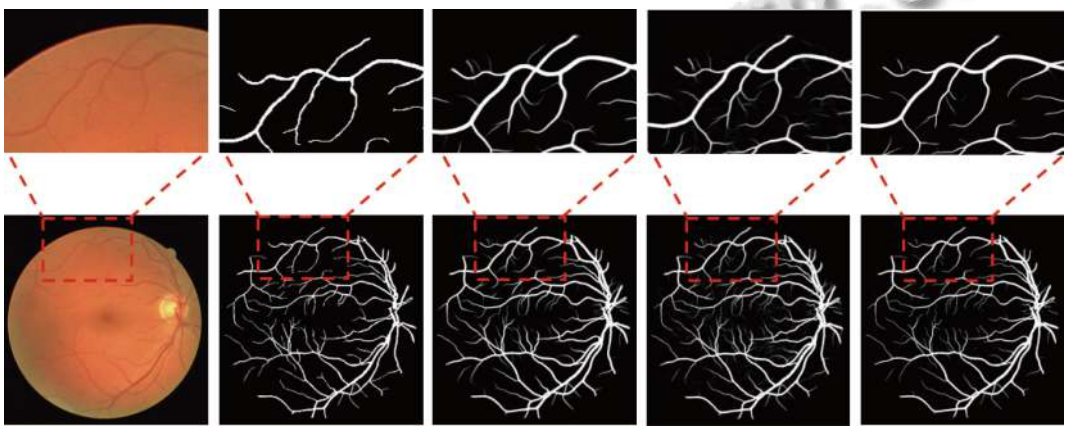
© 中国科学院软件研究所 http://www.c-s-a.org.cn

能, 准确率提高 0.07%, 敏感度提高 1.56%, AUC 提高 0.20%, 这证明了本文提出的多尺度高阶注意 (MHA) 模块的有效性. (2) MHA-Net 在准确率、灵敏度和 AUC 指标上都优于 Backbone+HA, 这表明多尺度高阶注意力模块对多尺度上下文特征信息捕捉能力更强, 对复杂结构的血管图像有更强的特征提取能力. (3) 本文提出的 MHA-Net 在大多数指标上都表现最好, 在视网膜血管分割领域全面优于 U-Net, 说明该网络模型的合理性和优越性.

此外, 对实验结果进行了可视化分析, 如图 5 所示, 从左至右依次是原始图、金标准、U-Net 分割结果、Backbone+HA 分割结果以及 MHA-Net (本文) 分割结果. 我们放大了微血管, 本文提出的 MHA-Net 分割结果更加精细, 在血管末也未出现粘连或者断裂的情况.

表 2 DRIVE 数据集上的消融实验

方法	*ACC*	*SE*	*SP*	AUC
U-Net[6]	0.9666	0.7677	0.9857	0.9789
U-Net+MHA	0.9673	0.7833	0.9848	0.9809
Backbone	0.9671	0.8046	0.9832	0.9862
Backbone+HA	0.9689	0.8297	**0.9860**	0.9876
MHA-Net	**0.9697**	**0.8312**	0.9833	**0.9883**



(a) 原始图 (b) 金标准 (c) U-Net (d) Backbone+HA (e) MHA-Net (本文)

图 5 DRIVE 数据集分割结果对比

5 结论与展望

本文针对视网膜血管分割任务中血管粗细不匀、形状多变、微小血管易断裂等问题, 本文提出多尺度高阶注意力 (MHA) 机制以自适应地挖掘深层次特征. MHA-Net 以端到端方式进行视网膜血管分割训练, 并通过 MHA 模块学习到具有鉴别性的特征. 在 DRIVE 上的实验表明, 本文提出的算法取得了优越的分割性能. 同时, MHA 模块可以即插即用, 在各种医学影像分割任务中适用. 后续的工作将尝试把多尺度高阶注意力机制运用到三维的影像分割中.

参考文献

1 Smart TJ, Richards CJ, Bhatnagar R, *et al*. A study of red blood cell deformability in diabetic retinopathy using optical tweezers. Proceedings of SPIE 9548, Optical trapping and optical micromanipulation XII. San Diego: SPIE, 2015. 954825.

2 Laibacher T, Weyde T, Jalali S. M2U-Net: Effective and efficient retinal vessel segmentation for resource-constrained environments. arXiv: 1811.07738, 2018.

3 Irshad S, Akram MU. Classification of retinal vessels into arteries and veins for detection of hypertensive retinopathy. 2014 Cairo International Biomedical Engineering Conference (CIBEC). Giza: IEEE, 2014. 133–136.

4 Cheung CYL, Zheng YF, Hsu W, *et al*. Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. Ophthalmology, 2011, 118(5): 812–818. [doi: 10.1016/j.ophtha.2010.08.045]

5 Litjens G, Kooi T, Bejnordi BE, *et al*. A survey on deep learning in medical image analysis. Medical Image Analysis, 2017, 42: 60–88. [doi: 10.1016/j.media.2017.07.005]

6 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015. 234–241.

7 Zhou ZW, Siddiquee MMR, Tajbakhsh N, *et al*. Unet++: A nested U-Net architecture for medical image segmentation. Deep Learning in Medical Image Analysis And Multimodal Learning for Clinical Decision Support. Granada: Springer, 2018. 3–11.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

8 Wu YC, Xia Y, Song Y, *et al*. Vessel-Net: Retinal vessel segmentation under multi-path supervision. 22nd International Conference on Medical Image Computing and Computer Assisted Intervention. Shenzhen: Springer, 2019. 264–272.
9 Li Y, Wang Y, Leng T, *et al*. Wavelet U-Net for medical image segmentation. 29th International Conference on Artificial Neural Networks and Machine Learning. Bratislava: Springer, 2020. 800–810.
10 Li LZ, Verma M, Nakashima Y, *et al*. IterNet: Retinal image segmentation utilizing structural redundancy in vessel networks. Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV). Snowmass: IEEE, 2020. 3645–3654.
11 Oktay O, Schlemper J, Le Folgoc L, *et al*. Attention U-Net: Learning where to look for the pancreas. arXiv: 1804.03999, 2018.
12 Vaswani A, Shazeer N, Parmar N, *et al*. Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: ACM, 2017. 6000–6010.
13 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
14 Guo CL, Szemenyei M, Yi YG, *et al*. SA-UNet: Spatial attention U-Net for retinal vessel segmentation. 2020 25th International Conference on Pattern Recognition (ICPR). Milan: IEEE, 2021. 1236–1242.
15 Roy AG, Navab N, Wachinger C. Concurrent spatial and channel 'squeeze & excitation' in fully convolutional networks. 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Granada: Springer, 2018. 421–429.
16 Wang Y, Deng ZJ, Hu XW, *et al*. Deep attentional features for prostate segmentation in ultrasound. 21st International Conference on Medical Image Computing and Computer Assisted Intervention. Granada: Springer, 2018. 523–530.
17 Lv Y, Ma H, Li JN, *et al*. Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation. IEEE Access, 2020, 8: 32826–32839. [doi: 10.1109/ACCESS.2020.2974027]
18 Zhang SH, Fu HZ, Yan YG, *et al*. Attention guided network for retinal image segmentation. 22nd International Conference on Medical Image Computing and Computer Assisted Intervention. Shenzhen: Springer, 2019. 797–805.
19 Chen BH, Deng WH, Hu JN. Mixed high-order attention network for person re-identification. Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 371–381.
20 Staal J, Abràmoff MD, Niemeijer M, *et al*. Ridge-based vessel segmentation in color images of the retina. IEEE Transactions on Medical Imaging, 2004, 23(4): 501–509. [doi: 10.1109/TMI.2004.825627]
21 Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. arXiv: 1511.07122, 2015.
22 Chen LC, Zhu YK, Papandreou G, *et al*. Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich: Springer, 2018. 833–851.
23 Chen LC, Papandreou G, Kokkinos I, *et al*. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834–848. [doi: 10.1109/TPAMI.2017.2699184]
24 Yi LP, Zhang JS, Zhang R, *et al*. SU-Net: An efficient encoder-decoder model of federated learning for brain tumor segmentation. 29th International Conference on Artificial Neural Networks and Machine Learning. Bratislava: Springer, 2020. 761–773.
25 Zhou LC, Zhang C, Wu M. D-LinkNet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018. 191–1924.
26 Ding F, Yang G, Wu J, *et al*. High-order attention networks for medical image segmentation. 23rd International Conference on Medical Image Computing and Computer Assisted Intervention. Lima: Springer, 2020. 253–262.
27 Zeiler MD, Taylor GW, Fergus R. Adaptive deconvolutional networks for mid and high level feature learning. 2011 International Conference on Computer Vision. Barcelona: IEEE, 2011. 2018–2025.
28 Paszke A, Gross S, Massa F, *et al*. PyTorch: An imperative style, high-performance deep learning library. Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, 2019. 721.
29 Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv: 1412.6980, 2014.

(校对责编: 牛欣悦)

© 中国科学院软件研究所 http://www.c-s-a.org.cn