

基于多尺度偶数卷积注意力 U-Net 的医学图像分割

杨真真^{1,2} 孙 雪¹ 邵 静¹ 杨永鹏^{1,3}

(1. 南京邮电大学宽带无线通信与传感网技术教育部重点实验室, 江苏南京 210003;

2. 南京邮电大学理学院, 江苏南京 210023; 3. 南京信息职业技术学院网络与通信学院, 江苏南京 210023)

摘 要: 为了提高 U-Net 网络性能的同时尽可能减少额外计算量, 本文提出了一种新的多尺度偶数卷积注意力 U-Net (Multiscale Even Convolution Attention U-Net, MECAU-Net) 网络。该网络在编码端采用 2×2 偶数卷积代替 3×3 卷积进行特征提取, 并借鉴多尺度思想, 采用 4×4 偶数卷积将得到的信息直接传递给主干部分, 以获取更全面的图像信息并减少额外计算开销, 同时还采用对称填充解决偶数卷积提取信息过程中产生的偏移问题。此外, 在 2×2 偶数卷积模块后加入卷积注意力模块, 结合空间和通道注意力, 在提取更丰富的信息的同时几乎不增加额外开销。最后, 在两个医学图像数据集上进行仿真实验, 实验结果表明提出的 MECAU-Net 网络相对于 U-Net 在稍微增加计算成本的情况下, 分割性能得到了较大的提升, 并比其他对比网络取得更好的分割性能的同时还降低了参数量。

关键词: 卷积神经网络; 医学图像分割; 偶数卷积; 多尺度; 注意力模块

中图分类号: TN911.73 **文献标识码:** A **DOI:** 10.16798/j.issn.1003-0530.2022.09.014

引用格式: 杨真真, 孙雪, 邵静, 等. 基于多尺度偶数卷积注意力 U-Net 的医学图像分割[J]. 信号处理, 2022, 38(9): 1912-1921. DOI: 10.16798/j.issn.1003-0530.2022.09.014.

Reference format: YANG Zhenzhen, SUN Xue, SHAO Jing, et al. Medical image segmentation based on multiscale even convolution attention U-Net[J]. Journal of Signal Processing, 2022, 38(9): 1912-1921. DOI: 10.16798/j.issn.1003-0530.2022.09.014.

Medical Image Segmentation Based on Multiscale Even Convolution Attention U-Net

YANG Zhenzhen^{1,2} SUN Xue¹ SHAO Jing¹ YANG Yongpeng^{1,3}

(1. Key Laboratory of Ministry of Education in Broadband Wireless Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China; 2. School of Science, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China; 3. School of Network and Communication, Nanjing Vocational College of Information Technology, Nanjing, Jiangsu 210023, China)

Abstract: In order to improve the performance of U-Net and reduce the additional computational complexity as much as possible, a new multiscale even convolution attention U-Net (MECAU-Net) network is proposed in this paper. This network uses 2×2 even convolution instead of 3×3 convolution for feature extraction at the encoder. And inspired by the idea of the multiscale idea, 4×4 even convolution is used to directly transfer the obtained information to the backbone, so as to obtain more comprehensive image information and reduce additional computational cost. At the same time, the symmetric padding is used to solve the shift problem in the process of extracting information from even convolution kernels. In addi-

收稿日期: 2022-02-14; 修回日期: 2022-03-19

基金项目: 国家自然科学基金(61501251, 62171232); 南京邮电大学宽带无线通信与传感网技术教育部重点实验室开放研究基金(JZNY202113); 南京邮电大学科研项目(NY220207)

tion, the convolution block attention module is added to combine the spatial and channel attention modules after the 2×2 even convolution module, which can extract richer information without adding additional computational complexity. Finally, simulation experiments are carried out on two medical image datasets. The experimental results show that our proposed MECAU-Net network greatly improves the segmentation performance with slight additional computational cost. In addition, it achieves better segmentation performance than other comparison networks and reduces the amount of parameters.

Key words: convolutional neural networks; medical image segmentation; even convolution; multiscale; attention module

1 引言

深度学习在图像分类、语义分割和目标检测等许多场景都实现了非常成功的应用^[1]。医学图像分割以其独特的应用场景受到研究者的广泛关注,其数据需要专业人士进行标注,数据集的获取相对困难,提升了医学图像分割的难度。随着深度学习的迅猛发展,研究者设计了各种网络来提高医学图像分割的性能。Ronneberger等^[2]提出U-Net网络,首次使用U型网络结构对医学图像进行分割,取得了很好的性能,其特有的U型结构很好的传递了提取的特征信息,非常适合小数据集的医学图像分割。

为了进一步提高分割性能,研究者相继提出了一些U-Net网络的改进网络,这些网络主要通过增加参数和计算量获得更深更宽的网络来增强特征提取能力。针对U-Net只能对二维图像进行分割,但临床诊断的图片大都是三维的问题,V-Net网络^[3]提出了一种基于体积、全卷积神经网络的三维图像分割方法,并引入一个新的目标函数,对训练过程进行优化。V-Net结构与U-Net几乎没有差别,分割精度上没有太大提升;为了避免U-Net网络编码部分提取出的浅层特征信息被每一层的池化操作损耗,Zhou等^[4]提出了一种新的嵌套结构U-Net++,该网络关注深浅层特征信息的融合,对U-Net跳跃连接部分进行改进,改进后的网络将编码部分每一层提取出的信息传递给对应的解码部分,并使用剪枝操作平衡过量的参数和网络深度之间的关系。但是U-Net++忽略了编码部分提取信息的重要性,只是利用嵌套和密集连接的跳跃结构加强对已有信息的传递过程。为了进一步提高网络分割精度,Huang等^[5]又提出了UNet 3+网络,该网络利用全尺度跳跃连接和深度监督融合不同尺度不同层次的特征信息,在图像分割领域得到了更好的性能,但同时产生了远超U-Net和深度监督的网络结构的参

数量。

残差网络(Residual Network, ResNet)^[6]的出现为U-Net的改进带来了另一种思路,该网络首次将残差连接嵌入到卷积神经网络,使堆积层在输入特征的基础上学习到新的特征,在一定程度上改善了网络层数增加带来的梯度消失问题;上下文编码网络(Context Encoder Network, CE-Net)^[7]提出了一种上下文编码模块,将残差连接与扩展卷积相结合,捕获更多高级信息并保留空间信息,使编码部分获取到更全面的特征信息。此外,在残差连接的基础上,又出现了多尺度残差U-Net (Multi-scale Residual U-Net, MultiResUNet)^[8]和多尺度深度连接U-Net (Multi-scale Densely connected U-Net, MDU-Net)^[9]。MultiResUNet提出了一个MultiRes模块与U-Net结合,MultiRes模块是残差连接的扩展,该模块将3个 3×3 卷积后的结果拼接起来作为一个组合特征图,再与输入特征图经过 1×1 卷积后的结果相加得到最后的输出特征图;MDU-Net与MultiResUNet非常类似,除了利用多尺度特征融合采集更多语义信息以外,都通过建立残差连接提高网络分割精度。但CE-Net、MultiResUNet和MDU-Net中残差连接的使用直接导致了网络计算量的急剧增加,限制了改进后的网络的使用环境。

此外,随着深度学习在图像处理领域的进一步深入研究,注意力模块在Seq2Seq网络^[10]翻译模型中首次被提出以后,就得到了快速发展和广泛应用。多尺度注意力U-Net (Multi-scale and Attention U-Net, MA-Unet)^[11]将注意力模块和多尺度特征融合结合起来,通过完善编码部分的特征提取提升网络性能,但依旧在一定程度上增加了该网络的计算开销。

U-Net网络的主要特点是轻量级网络^[12],而现有改进的U-Net大都是通过加宽加深网络来获得更好的分割性能,这些网络对内存的高要求和高的计算量严重限制了深度学习在医学图像分割领域的

应用。如何在保持令人满意的分割性能的同时减少所需计算资源成为研究者新的关注点。基于此,本文提出了一种新的多尺度偶数卷积注意力U-Net (Multiscale Even Convolution Attention U-Net, MECAU-Net)网络^[13]。该网络通过利用偶数卷积降低分割所需参数以减少额外计算开销^[14],再利用卷积注意力模块 (Convolutional Block Attention Module, CBAM)^[15]和新增的4×4卷积层来取得更好的分割性能,其创新性如下:

(1)提出了一种新的多尺度偶数卷积方法。针对对现有的改进U-Net医学图像分割网络产生参数爆炸的问题,为了更好地平衡计算开销和分割性能,MECAU-Net网络将关注点放在了卷积运算上。卷积运算是整个分割网络最常用的运算之一,也是能否合理的降低卷积计算量的关键因素。MECAU-Net的编码部分将3×3卷积换成2×2偶数卷积,通过降低卷积核的大小,将卷积运算计算量尽可能降低。同时借鉴多尺度的思想,MECAU-Net还在编码端建立了一条与2×2偶数卷积平行的4×4偶数卷积通路,将得到的信息直接传递给编码网络的主体2×2偶数卷积部分,提供在更大感受野内提取特征信息,以保证在获取更全面的图像信息的同时减少额外计算开销。同时还采用对称填充解决偶数卷积提取信息过程中产生的偏移问题。

(2)卷积过程嵌入了一种注意力模块。针对大多数的分割网络对图片的所有像素进行无差别处理,而整张图片只有较小区域包含有用信息,计算产生极大浪费的问题,为了更有效的分割出医学图像中的目标区域,MECAU-Net网络在2×2偶数卷积模块后嵌入了卷积注意力模块,通过增加注意力模块来关注有用的特征信息,在提取更丰富的信息的同时几乎不增加额外开销。该模块先通过通道注意力和空间注意力这两个模块提取出有用的特征信息,再和通过偶数卷积块提取出的特征信息融合,传递到下一卷积层继续提取更深层的特征信息。

2 相关工作

2.1 U-Net网络

Ronneberger等在全卷积神经网络^[16]的基础上,不再使用全连接层,避免产生过多参数,提出了专注于解决医学图像分割的U-Net网络。该网络分为

编码部分、解码部分和为编解码之间信息传递提供通路的跳跃连接三部分。编码部分的功能是通过卷积和池化等一系列操作提取图像的特征信息,该部分包含四个子模块,每个子模块包含两个3×3的卷积层,每个子模块之后是通过最大池化实现的下采样层,依次采集到更深层次的语义信息;解码部分同样包含四个子模块,每个子模块由上采样和解码实现,解码时通过上采样恢复分辨率,直到与输入图像的分辨率一致;跳跃连接将特征提取的每一层编码得到的语义信息及时传递给对应的解码端,完整的保留了前三层编码获取到的图像特征。

2.2 偶数卷积核

现有的卷积神经网络主要通过深度卷积、扩展通道和复杂的拓扑结构来提高性能,这加剧了训练过程的复杂度。此外,在这些网络中,3×3卷积核占主导地位,而偶数大小的卷积核很少被采用。Wu等^[13]提出了采用2×2偶数卷积可以进一步减少网络参数。但偶数卷积核具有不对称的感受野,导致在生成的特征图中产生像素偏移。当堆叠多个卷积时,位置偏移会累积,从而严重的侵蚀空间信息。为了解决这一问题,Wu等^[13]提出使用偶数大小的卷积核的同时在特征图的每一层使用对称填充,不仅很好的消除了移位问题,而且扩展了偶数卷积核的感受野,从而在降低参数的同时取得相当甚至更好的性能。通过偶数卷积和对称填充,提取出的特征信息相当于3×3卷积的感受野,利用更小的计算开销获取相同大小感受野内的信息,达到在获得相当甚至更好的性能的同时,减少卷积的参数量。

2.3 注意力模块

现有的大多数改进U-Net网络都是通过加深网络来获取更多特征信息,网络性能虽得到了一定的提升,但很大一部分计算量用来分析图像的背景部分,而医学图像的背景部分占比较大,且一般都不包含有用信息,如何有针对性的利用有限的计算量提取医学图像的有用信息是很重要的。注意力模块是一种能让模型对重要信息重点关注并充分吸收学习的技术^[17],它可以帮助网络快速锁定待处理部分,减少不必要的计算损耗,是一种非常有用的降低网络计算量的方法。大量实验证明注意力模块的加入提升了图像分割网络的分割性能^[18]。压缩激励网络 (Squeeze-and-Excitation Networks, SENet)^[19]以考虑

特征通道之间的关系为出发点,提出了SE注意力模块,使用全局上下文对通道权值进行重新标定,调整特征通道之间的依赖关系,但SE注意力模块没有充分利用全局上下文信息。CBAM是一种轻量级注意力模块^[15],它将注意力分为通道注意力和空间注意力两个过程,在给定一个中间特征图后,沿着通道和空间两个维度依次推断出注意力权重,然后与原特征图相乘来对特征进行自适应调整。和其他注意力模块相比,CBAM可以更好的嵌入到卷积模块中,并且不会给原有网络带来过多计算开销。

3 多尺度偶数卷积注意力U-Net

3.1 动机

医学图像分割较之普通图像分割更困难,因其可供训练网络模型的图片很难获取,所以人们在构建医学图像分割网络时,都会格外注意特征获取部分^[20]。U-Net在编码、解码过程中只采用 3×3 卷积和反卷积来提取或恢复图像特征,其改进分割网络也都在沿用 3×3 卷积的基础上利用更大尺度的 5×5 卷积核获取到更大感受野内的内容,这些网络的性能虽都有一定的提升,但是大的卷积核也导致了计算速度更慢,参数量更多。U-Net分割网络虽然取得了不错的分割结果,但是之后的改进网络都在以计算量为代价提升分割精度,违背了U-Net的轻量级网络的本质。为此,本文提出以利用较少的计算量实现更好的分割结果为目的的网络,采用CBAM注意力模块锁定图像的特征信息,用 2×2 偶数卷积替换 3×3 卷积降低参数量,再通过新增一层 4×4 偶数卷积核,代替多尺度特征融合在更大感受野内采集特征信息,再将采集到的特征信息传递给网络主体部分。

3.2 多尺度偶数卷积U-Net

卷积操作在整个计算过程中所占计算开销最大,所以考虑从卷积的角度来降低整个网络的计算量。现有改进U-Net网络仍是以 3×3 的卷积作为网络主体进行特征提取。在多尺度编码部分,也都选择使用更大尺度的 5×5 等奇数卷积核获取大感受野内的信息。大卷积核的感受野更大,提取特征的性能更好,但是其计算速度更慢,参数量更多。此外,多尺度特征融合通常可以给网络带来更好的性能,但也通常会引起的参数爆炸。为了解决这两个问

题,提出了如图1所示的多尺度偶数卷积U-Net模块。采用更小的卷积核降低计算量非常有帮助,但 1×1 的卷积核几乎没办法提取出像素的位置信息,为了降低计算损耗,本文采用 2×2 偶数卷积核对输入特征图中的信息进行提取。Szegedy等人^[21]在Inception-v3中提出采用两个连续 3×3 卷积组成的小网络来代替一个 5×5 卷积,在保持感受野的同时降低参数量。借鉴类似思想,提出采用两个 2×2 卷积组合代替一个 3×3 卷积,以此来降低卷积部分的参数量。此外,为了避免像素偏移问题,提出的网络延用了Wu等^[13]提出的对称填充的方法,在消除了偶数卷积带来的像素偏移的同时扩大感受野。

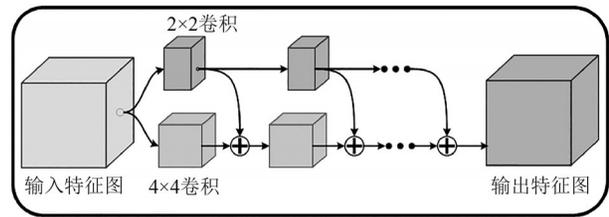


图1 多尺度偶数卷积U-Net模块

Fig. 1 Multi-scale even convolution U-Net module

感受野的大小直接影响到图像特征信息和图像空间信息的提取,为了增强图像信息的提取能力,常用多尺度特征融合的方式平衡特征信息和位置信息两者之间的关系。多尺度特征融合通常可以给网络带来更好的性能,常见的多尺度特征融合是在 3×3 卷积的网络中引入 5×5 卷积,以此来获得更大感受野内的图像信息^[22],但是多尺度特征融合会在网络中引入过多的参数和复杂度。考虑到参数增长的问题,在分割网络的主体部分以外的编码部分新增了一层 4×4 偶数卷积编码网络,利用额外的分支单独对图像信息进行提取,相比于常见的 3×3 和 5×5 组合形式的多尺度特征融合,其计算开销又得到了降低。也就是说多尺度偶数卷积核使得网络在额外计算开销增加较少的情况下,有效提升了分割性能。首先,采用更大尺度的 4×4 偶数卷积核,对图像信息进行提取,为了避免偶数卷积带来的像素偏移问题,同样在每一层对待分割的特征图进行对称填充,再将获取到的信息通过拼接的方式传递给对应的主体网络进行下一步池化。通过这种方式获取更全面的图像信息也避免了多尺度特征融合带来的参数爆炸问题。

3.3 注意力 U-Net

为了提升网络的分割精度,研究者采用各种方法设计了新的分割网络,常见的方法有残差连接^[7]、密集连接^[23]等,大多数的改进都是通过增加卷积块内的操作对图片的每一个部分信息进行进一步的提取,这种特征提取方式通常是将全局内的信息无差别的进行再一次的提取,虽然包含了一部分有用的信息,但多余的无用信息也产生了大量额外的计算开销。注意力模块通过增加某一位置的权重,有针对性的提取目标位置的图像特征,不仅提供了大量的有用信息,也节省了不必要的计算开销。

CBAM 是一个轻量级的通用模块,非常适合在特征提取的过程中使用,不会占用太多的额外开销,所以提出将 CBAM 嵌入到编码部分的每一层卷积中进行更深层次的特征提取。提出的网络将 2×2 卷积后的特征图作为输入,通过 CBAM 对每一层卷积提取出的特征图进行进一步的处理。经过处理后的编码层会得到更精确的特征信息,每一编码层都将卷积和两路注意力提取出的结果加权,得出本层最终提取出的特征,有利于后续更好的提取图片中的特征信息。

3.4 MECAU-Net 网络

U-Net 以其独特的 U 形和通过跳跃连接融合编解码两侧信息的结构在医学图像分割领域取得了很好的性能^[24],受到了研究者的广泛关注,对基于

U-Net 的医学图像分割网络进行改进,一般都是以提高计算复杂度为代价获得更好的分割性能。为了平衡计算量和网络性能之间的关系,本文对 U-Net 进行改进,提出了一种新的基于 MECAU-Net 网络的医学图像分割方法,主要对 U-Net 的编码部分进行改进,其结构图如图 2 所示。该网络从上到下共五层,网络主体沿用 U-Net 网络的 U 形结构和跳跃连接,在编码部分采用更小的 2×2 的偶数卷积核、嵌入 CBAM 注意力模块以及在主体外新增一层 4×4 偶数卷积的特征提取,在稍微增加计算成本的情况下,取得了更好的分割效果。

MECAU-Net 网络的具体改进如下:

(1) 在网络主体部分采用两个大小为 2×2 的偶数卷积核代替 3×3 卷积核,提取图像中的信息,卷积核变小,卷积部分需要的计算开销降低。同时,借鉴多尺度的思想,在编码端建立了一条与 2×2 偶数卷积平行的 4×4 偶数卷积通路,将得到的信息直接传递给编码网络的主体部分,在更大感受野内提取特征信息,融合不同感受野内的特征信息,以保证在获取更全面的图像信息的同时减少额外计算开销。此外,为了解决偶数卷积核带来的像素偏移问题,借鉴 Wu 等^[13]的思想,在卷积之前对图像特征进行对称填充,在消除偏移的同时扩展了偶数卷积核的感受野,从而达到降低参数的同时取得相当甚至更好性能的目的。

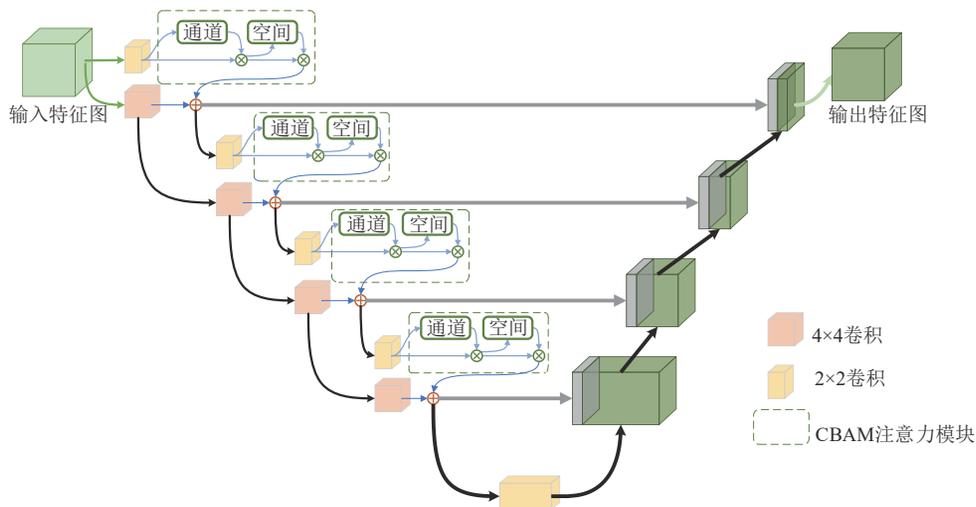


图2 MECAU-Net网络

Fig. 2 The overall architecture of MECAU-Net

(2)为了更有效的分割出医学图像中的目标区域,在 2×2 偶数卷积模块后嵌入了CBAM注意力模块,沿通道和空间两个不同维度依次推断注意力图,然后将注意力图和特征图融合后得到最后的结果。通过增加CBAM注意力模块来关注有用的特征信息,在提取更丰富的信息的同时几乎不增加额外开销。

4 实验结果及分析

在实验中选择DRIVE^[25]和CHAOS^[26]数据集作为分割对象。DRIVE是眼底血管分割数据集,包含40张图片,其中7张图片显示早期轻度糖尿病视网膜病变的迹象。该数据集中有很多待分割的细小血管,所以对DRIVE数据集来说最重要的是如何在图像中提取尽可能多的细节信息。CHAOS是医学图像分割中很有代表性的数据集,该数据集中的图片以肝肾脾的横截面为主,在分割过程中的侧重点是如何将边缘信息准确的分割出来。本次实验选取DRIVE和CHAOS作为实验数据集,通过对比各种网络分割性能,验证了提出的MECAU-Net网络在医学图像分割的优越性。

为了更清晰、更直观地定量评价提出的MECAU-Net网络的性能,在仿真实验中选取像素准确率(Pixel Accuracy, PA)、交并比(Intersection over Union, IOU)和平均交并比(Mean Intersection over Union, MIoU)作为性能好坏的评价指标^[27],选取参数量(单位为MB)和预测时间(单位为秒(s))作为计算量的评价指标。PA是预测类别正确的像素数占总像素数的比例,IOU是某一类别预测值与真实值的交集和并集之比,MIoU是对每一类预测值与真实值的交集和并集之比求和再求平均。PA、IOU和MIoU这三个评价指标的取值都在0到1之间,在区间内的值越大表明网络的分割性能越好。

由于图像大小并不完全相同,为了方便后续分割,先将图像大小调成统一的值。此外,为了缓解过拟合问题,在图像处理过程中加入图像增强步骤,先将图像横向和纵向翻转 90° ,再用弹性变换对数据集中的图像进行处理,扩充数据集。为了进一步避免训练过程中产生的过拟合问题,并合理评估MECAU-Net网络的分割性能,采用10倍交叉验证来优化整个网络。在训练过程中将验证集的比例

设置为0.1,抽取数据集中的10%作为测试集。此外,为了达到更好的分割结果,在对测试集进行分割前也做了图像增强。

4.1 消融实验

将U-Net网络中的 3×3 卷积替换成 2×2 偶数卷积,称这种方法为偶数卷积U-Net;在偶数卷积U-Net网络的基础上再添加 4×4 偶数卷积,称为多尺度偶数卷积U-Net;在U-Net网络编码部分的卷积中添加CBAM模块,称此网络为注意力U-Net。为了验证提出的偶数卷积、多尺度和CBAM注意力模块的有效性,对DRIVE数据集进行消融实验,其结果如表1所示。

表1 消融实验
Tab. 1 The ablation experimental results

网络	IOU	MIoU	PA	参数量/MB
U-Net	0.748	0.740	0.9209	5.43
偶数卷积U-Net	0.752	0.748	0.9275	4.87
多尺度偶数卷积U-Net	0.771	0.769	0.9470	5.61
注意力U-Net	0.759	0.772	0.9346	5.52

如表1所示,偶数卷积U-Net分割图像的IOU、MIoU、PA和参数量分别为0.752、0.748、0.9275和5.43 MB,虽然分割性能提升并不明显,但由于使用了 2×2 偶数卷积,有效降低了U-Net的参数量,降低后的参数量只有4.87 MB;同时采用 2×2 和 4×4 偶数卷积构成的多尺度偶数卷积U-Net分割图像的IOU、MIoU、PA和参数量分别为0.771、0.769、0.9470和5.61 MB,相比U-Net和偶数卷积U-Net分割性能都有了显著提升。虽然多尺度会不可避免的增加网络的参数量,但由于使用更小的偶数卷积核,网络的参数量维持在5.61 MB。与U-Net相比,参数量几乎不变的情况下,注意力U-Net的分割图像的IOU、MIoU和PA都有一定程度的提升,改进的编码部分提高了网络的性能,这是因为增加了CBAM注意力模块后,在分割过程中通过通道和空间注意力模块,更有针对性地提取出有用的特征信息,进而提高了网络的性能,由于CBAM本身是轻量的注意力机制,所以加入了CBAM注意力机制的U-Net网络参数只增长到5.52 MB,与U-Net相比,只增加了0.09 MB。总的来说,多尺度偶数卷积U-

Net和注意力U-Net相比,在分割性能方面多尺度偶数卷积U-Net更占优势,但其所需的计算开销也会略大于注意力U-Net的计算开销。

4.2 对比实验

为了验证提出的MECAU-Net网络的有效性和优越性,在DRIVE和CHAOS数据集上进行分割实验并与U-Net、V-Net、MA-Net、MDU-Net、MultiResU-Net和CE-Net进行对比,同时,为保证实验结果的公平性,本文在相同的实验环境下分别运行这六个对比网络,其视觉效果如图3和图4所示。

图3是各种网络在DRIVE数据集上的分割效果图,第一行和第三行图片分别为在两张示例图片上各个网络的分割结果图,第二行和第四行分别给

出了对应的红色框内的细节部分,放大的细节部分都是两张图片中包含细节信息最多的位置。由细节部分可以看出U-Net、V-Net、MA-Net和MDU-Net都只分割出了主体部分较粗的血管,U-Net的分割图中几乎没有分割出细节部分的血管,MultiResU-Net和CE-Net对细节部分的处理结果稍好一点,但提出的MECAU-Net成功的分割出了最多的细节信息,分割效果最好。图4是各种网络在CHAOS数据集上的分割效果,第一行和第三行为完整的分割结果图,第二行和第四行分别给出了放大的红色框内部分的分割细节图,可以看到U-Net、V-Net、MA-Net、MDU-Net和MultiResUNet在分割过程中都发生了不同程度的误判,CE-Net的分割结果相对精确,

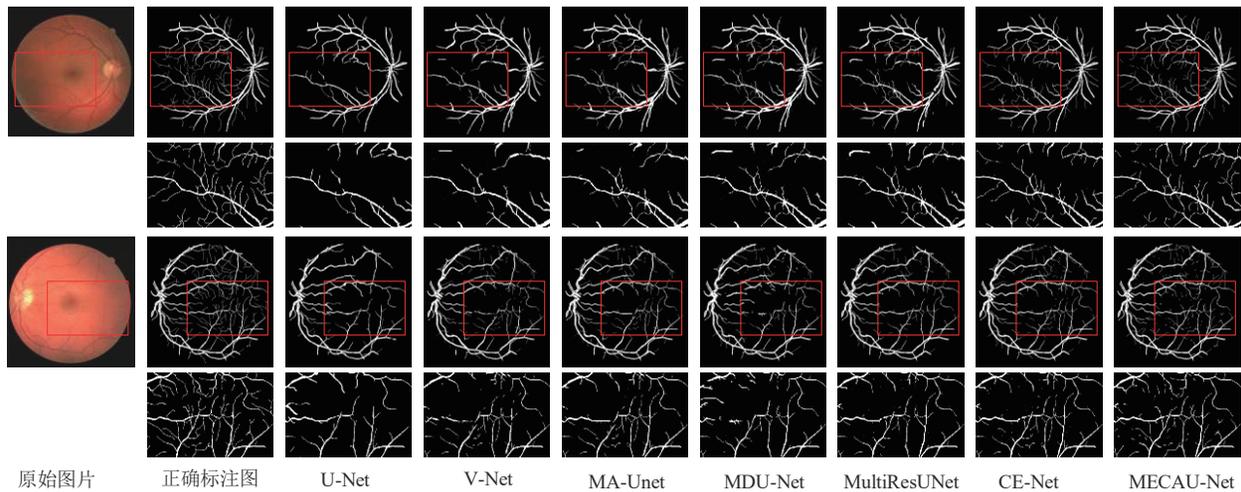


图3 各种网络在DRIVE数据集上的分割效果

Fig. 3 The visual segmentation results on the DRIVE dataset

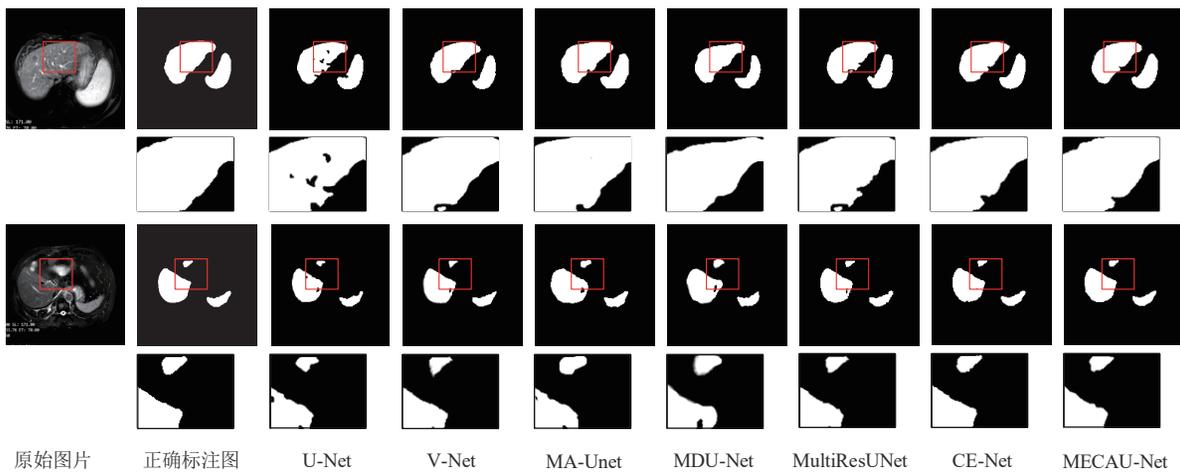


图4 各种网络在CHAOS数据集上的分割效果

Fig. 4 The visual segmentation results on the CHAOS dataset

但是提出的MECAU-Net网络对边缘部分的处理优于CE-Net,取得了最好的分割性能。

为了定量评价提出的MECAU-Net网络的性能,分别对DRIVE和CHAOS数据集进行分割实验,其实验结果分别如表2和3所示。

表2 各种网络在DRIVE数据集上的实验结果

Tab. 2 The experimental results of different networks on the

DRIVE dataset					
网络	IOU	MIoU	PA	参数量/MB	预测时间/s
U-Net	0.748	0.740	0.9209	5.43	17
V-Net	0.766	0.755	0.9283	8.93	23
MA-UNet	0.767	0.762	0.9312	10.57	42
MDU-Net	0.766	0.767	0.9330	12.33	61
MultiResUNet	0.767	0.763	0.9392	12.45	68
CE-Net	0.764	0.778	0.9393	15.24	85
MECAU-Net	0.777	0.786	0.9491	5.89	19

表3 各种网络在CHAOS数据集上的实验结果

Tab. 3 The experimental results of different networks on the

CHAOS dataset					
网络	IOU	MIoU	PA	参数量/MB	预测时间/s
U-Net	0.743	0.741	0.9172	5.79	22
V-Net	0.758	0.747	0.9230	9.45	37
MA-UNet	0.760	0.753	0.9307	10.98	58
MDU-Net	0.763	0.762	0.9346	13.10	74
MultiResUNet	0.778	0.776	0.9378	13.52	78
CE-Net	0.773	0.768	0.9308	17.85	90
MECAU-Net	0.779	0.78	0.9489	6.24	25

从表2可以看出,对DRIVE数据集,U-Net网络的IOU、MIoU和PA分别为0.748、0.740和0.9202,参数量为5.43 MB,预测时间为17 s,提出的MECAU-Net网络的IOU、MIoU和PA分别为0.777、0.786和0.9491,参数量为5.89 MB,预测时间需要19 s,和U-Net相比,提出的MECAU-Net的IOU、MIoU和PA分别提升了2.89%、4.6%和2.9%,但是参数量只增长了0.46 MB,所需的预测时间也只增加了2 s。V-Net、MA-Net、MDU-Net、MultiResUNet和CE-Net的参数量分别为8.93 MB、10.57 MB、12.33 MB、12.45 MB、15.24 MB,所需预测时间也呈递增趋势,他们的参数量都远超过U-Net网络和提出的

MECAU-Net网络,并且其IOU、MIoU和PA与V-Net、MA-Net、MDU-Net、MultiResUNet相比都有了明显增长,IOU、MIoU和PA分别最多提升了1.3%、3.1%和2.1%。

由表3可以看出,对CHAOS数据集,U-Net的参数量为5.79 MB,预测时间为22 s,IOU、MIoU和PA分别为0.743、0.741和0.9172;V-Net、MA-Net、MDU-Net、MultiResUNet和CE-Net的IOU、MIoU、PA相较于U-Net都有提升,但MECAU-Net网络的IOU、MIoU和PA都最高。此外,提出的MECAU-Net的参数量相比于CE-Net低了11.61 MB,预测时间低了65 s,但是达到了更好的分割性能。综上所述,提出的MECAU-Net网络相对于U-Net在稍微增加计算成本的情况下,分割性能取得了较大的提升,并且比其他对比算法取得了更好的分割性能的同时还降低了参数量。

5 结论

本文提出了一种以更低复杂度得到更好分割性能的MECAU-Net医学图像分割网络,首先为了降低参数量引入两个 2×2 偶数卷积代替尺度更大的 3×3 卷积,并且在卷积的填充部分利用对称填充解决偏移问题。同时为了进一步提高分割精度,借鉴多尺度思想,在整体网络使用 2×2 偶数卷积的基础上,增加了 4×4 偶数卷积的分支,作为单独通路向提出的MECAU-Net网络主体传输在更大感受野内提取到的特征信息;此外,在编码部分的卷积层嵌入CBAM注意力模块,进一步提取通道和空间两个方向上的特征图信息,与卷积提取出的特征信息进行融合,得到加权后的特征图,方便更精确的提取出后续的特征信息。最后,为了进一步验证提出的MECAU-Net网络的分割性能,将其和其他几个U-Net的改进图像分割网络在医学图像数据集上进行分割实验。与其他改进的U-Net分割网络相比,提出的MECAU-Net网络用更少的计算量得到了更好的分割结果。

参考文献

- [1] 田萱,王亮,丁琪.基于深度学习的图像语义分割方法综述[J].软件学报,2019,30(2):440-468.
TIAN Xuan, WANG Liang, DING Qi. Review of image semantic segmentation based on deep learning[J]. Jour-

- nal of Software, 2019, 30(2): 440-468. (in Chinese)
- [2] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015, 2015: 234-241.
- [3] MILLETARI F, NAVAB N, AHMADI S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation [C]//2016 Fourth International Conference on 3D Vision (3DV). Stanford, CA, USA. IEEE, 2016: 565-571.
- [4] ZHOU Z, SIDDIQUEE M M R, TAJBAKSH N, et al. UNet++: A nested U-net architecture for medical image segmentation [J]. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 2018, 11045: 3-11.
- [5] HUANG Huimin, LIN Lanfen, TONG Ruofeng, et al. UNet 3+: A full-scale connected UNet for medical image segmentation [C]//ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona, Spain. IEEE, 2020: 1055-1059.
- [6] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [7] GU Zaiwang, CHENG Jun, FU Huazhu, et al. CE-net: Context encoder network for 2D medical image segmentation [J]. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281-2292.
- [8] IBTEHAZ N, RAHMAN M S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Networks, 2020, 121: 74-87.
- [9] ZHANG Jiawei, JIN Yuzhen, XU Jilan, et al. MDU-net: Multi-scale densely connected U-net for biomedical image segmentation [EB/OL]. 2018: arXiv: 1812.00352 [cs.CV]. <https://arxiv.org/abs/1812.00352>.
- [10] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation [C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar. Stroudsburg, PA, USA: Association for Computational Linguistics, 2014: 1724-1734.
- [11] CAI Yutong, WANG Yong. MA-Unet: An improved version of Unet based on multi-scale and attention mechanism for medical image segmentation [EB/OL]. 2020: arXiv: 2012.10952 [cs.CV]. <https://arxiv.org/abs/2012.10952>.
- [12] 邹亚君, 李翌昕, 马尽文. 基于深度学习的酒标分割研究 [J]. 信号处理, 2019, 35(4): 623-630.
ZOU Yajun, LI Yixin, MA Jinwen. Research on deep learning based wine label segmentation [J]. Journal of Signal Processing, 2019, 35(4): 623-630. (in Chinese)
- [13] WU S, WANG G, TANG P, et al. Convolution with even-sized kernels and symmetric padding [EB/OL]. 2019: arXiv: 1903.08385 [cs.CV]. <https://arxiv.org/abs/1903.08385>.
- [14] WANG Huiyu, KEMBAVI A, FARHADI A, et al. ELASTIC: improving CNNs with dynamic scaling policies [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA. IEEE, 2019: 2253-2262.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19.
- [16] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. IEEE, 2015: 3431-3440.
- [17] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-net: Learning where to look for the pancreas [EB/OL]. 2018: arXiv: 1804.03999 [cs.CV]. <https://arxiv.org/abs/1804.03999>.
- [18] PETIT O, THOME N, RAMBOUR C, et al. U-net transformer: Self and cross attention for medical image segmentation [C]//Machine Learning in Medical Imaging, 2021: 267-276.
- [19] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [20] 刘敏, 方志军, 高永彬. 多阶段冠状动脉CTA三维分割算法 [J]. 信号处理, 2019, 35(11): 1911-1918.
LIU Min, FANG Zhijun, GAO Yongbin. Multi-stage three-dimensional segmentation algorithm for coronary CTA [J]. Journal of Signal Processing, 2019, 35(11): 1911-1918. (in Chinese)
- [21] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. IEEE, 2016: 2818-2826.
- [22] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going

- deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA. IEEE, 2015: 1-9.
- [23] YANG Zhenzhen, XU Pengfei, YANG Yongpeng, et al. A densely connected network based on U-net for medical image segmentation [J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2021, 17(3): 1-14.
- [24] 田娟秀, 刘国才, 谷珊珊, 等. 医学图像分析深度学习研究方法研究与挑战 [J]. 自动化学报, 2018, 44(3): 401-424.
- TIAN Juanxiu, LIU Guocai, GU Shanshan, et al. Deep learning in medical image analysis and its challenges [J]. Acta Automatica Sinica, 2018, 44(3): 401-424. (in Chinese)
- [25] HAFIZ A M, BHAT G M. A survey on instance segmentation: State of the art [J]. International Journal of Multimedia Information Retrieval, 2020, 9(3): 171-189.
- [26] LIU Zhihua, CHEN Long, TONG Lei, et al. Deep learning based brain tumor segmentation: a survey [EB/OL]. 2021: arXiv: 2107.04537 [cs. CV]. <https://arxiv.org/abs/2107.04537>.
- [27] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: Faster and better learning for bounding box regression [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

作者简介



杨真真(通信作者) 女, 1984年生, 山东临沂人。南京邮电大学副教授、博士。研究方向为机器学习、图像处理。
E-mail: yangzz@njupt.edu.cn



孙雪 女, 1999年生, 辽宁朝阳人。南京邮电大学研究生。研究方向为深度学习、图像分割。
E-mail: 546293349@qq.com



邵静 女, 1995年生, 江苏徐州人。南京邮电大学研究生。研究方向为机器学习、图像处理。
E-mail: 1481726137@qq.com



杨永鹏 男, 1984年生, 山东高密人。南京信息职业技术学院讲师、南京邮电大学博士研究生。研究方向为机器学习、图像处理。
E-mail: yangyp@njcit.cn