1

医学影像的多尺度自导向关注

分割

Ashish Sinha和Jose Dolz

*尽管卷积神经网络(CNNs)推动了医学图像分割的发展，但标准模型仍存在一些缺陷。***首先，多尺度方法的使用，即。，导致信息的冗余使用，类似的低层特征在多个尺度上被多次提取。其次，长期的特征依赖没有被有效地建模，导致与每个语义类关联的特征的非最优区分表示。在本文中，我们试图克服这些限制，提出的架构，通过捕获丰富的上下文依赖基于引导自我注意机制的使用。该方法能够将局部特征与其相应的全局依赖性结合起来，并以自适应的方式突出相互依赖的信道映射。此外，不同模块之间的额外损失导致注意机制忽略不相关的信息，通过强调相关的特征关联，将注意力集中在图像中更具区别性的区域。我们在三个不同的数据集:腹部器官、心血管结构和脑肿瘤的语义分割的背景下评估提出的模型。一系列的消融实验证明了这些注意力模块在该架构中的重要性。此外，与其他先进的分割网络相比，我们的模型具有更好的分割性能，提高了预测的准确性，同时减少了标准偏差。这证明了我们的方法的效率，以产生精确和可靠的自动分割的医学图像。我们的代码在以下网址公开:https://github.com/sinAshish/Multi- Scale-Attention**

*索引术语-卷积神经网络，深度学习，医学图像分割，深度关注，自我关注*

我的介绍。

医学图像的语义分割是许多疾病诊断、治疗和跟踪的关键步骤。尽管自动化的任务在过去得到了广泛的研究，手工注释仍然通常用于临床实践，这是一个耗时的过程，并且容易出现观察者之间和观察者内部的变化过程。因此，对精确和可靠的自动分割方法有很高的要求，这些方法可以提高临床场景中的工作流程效率，减轻放射科医生和其他医学专家的工作量。

最近，卷积神经网络(CNNs)在广泛的视觉识别任务中取得了最先进的性能，由于其强大的非线性特征提取能力而变得非常流行。这些深度模型在医学图像分割中占主导地位

A. Sinha就职于印度鲁尔基理工学院。电子邮件:asinha@mt.iitr.ac.in。

J. Dolz就职于加拿大蒙特利尔高等技术学院。电子邮件:jose.dolz@etsmtl.ca。

手稿收到XXX;修改后的XXX。

并在广泛的应用领域取得了优异的表现，包括脑[2]或心脏[3]成像，成为这些问题的事实上的解决方案。在这种情况下，全卷积神经网络[4]或编解码器架构[5]、[6]通常是标准选择。这些架构通常由一个收缩路径和一个扩展路径组成，前者将输入图像折叠成一组高级特征，后者使用高级特征在单个[4]或多个上采样步骤[5]、[6]重建像素级分割掩模。然而，尽管这些多尺度的方法具有很强的表示能力，但却导致了信息流的冗余使用，例如，相似的低级特征在网络的不同层次被多次提取。此外，学习到的特征表示在像素识别方面的鉴别能力可能不足以完成一些具有挑战性的任务，如医学图像分割。

最近在提高特征表示的辨别能力方面的工作包括使用多尺度的上下文融合[7]，[8]，[9]，[10]。赵等人[8]提出了一种金字塔网络，该网络通过聚合多个扩展的卷积块生成的feature map来利用不同尺度的全局信息。上下文多尺度信息的聚合也可以通过[11]池化操作实现。尽管这些策略可能有助于在不同尺度下捕获目标，但所有图像区域的上下文依赖都是同质的、非自适应的，忽略了局部表示和不同类别的上下文依赖之间的差异。而且，这些多上下文表示是手工设计的，缺乏对多上下文表示建模的灵活性。这使得整个图像中的长程目标关系不能在这些方法中得到充分利用，这在许多医学图像分割问题中是至关重要的。

另外，为了有效地整合局部和全局特征，在深度神经网络中，注意力机制也得到了广泛的研究，包括人体姿态估计[12]、情感识别[13]、文本检测[14]、目标检测[15]和分类[16]。与标准的多尺度特征融合方法不同，该方法将整个图像压缩为静态表示，attention允许网络集中在最相关的特征上，而无需额外的监督，避免使用多个相似的特征图，并突出突出对给定任务有用的显著特征。语义分割网络也受益于注意力模块，这导致了像素级识别任务[17]、[18]、[19]、[20]、[21]、[22]的增强模型。例如，Chen等人[17]提出了一种对提取的多尺度特征权重的注意机制

arXiv: 1906.02849 v3 [cs。CV] 2020年2月14日

2

在不同尺度下的自然场景分割。该方法比传统的平均和最大池技术改进了多尺度特征预测的分割性能。

尽管人们对将注意力机制整合到自然场景的图像分割网络中越来越感兴趣，但在医学图像中[23]、[24]、[25]、[26]、[27]的应用仍然很少，局限于简单的注意力模型。因此，在这项工作中，我们探索更复杂的注意机制，可以提高标准深度网络的性能，用于医学图像分割任务。具体地说，我们提出了一种用于医学图像分割的多尺度引导注意力网络。首先，多尺度方法以包含不同语义的不同分辨率生成堆栈。低级堆栈关注本地外观，而高级堆栈将编码全局表示。这种多尺度策略鼓励在不同分辨率下生成的注意力地图编码不同的语义信息。然后，在每个尺度上，一堆注意模块将逐渐去除噪声区域，并强调那些与目标语义描述更相关的区域。每个注意模块包含两个独立的自注意机制，分别关注建模位置和通道特征依赖。这两部分允许建模更广泛和更丰富的上下文表示，并改进通道映射之间的依赖关系，从而增强特性表示。我们在三个不同的分割任务中验证了我们的方法:腹部器官，心血管结构和脑瘤。结果表明，该架构成功地建立了基于局部特征的丰富上下文依赖关系，提高了分割性能。

2相关工作

*A.医学图像分割*

虽然在过去的[28]中医学图像分割已经得到了广泛的研究，但不可否认的是CNNs正在推动着这一领域的发展，在很多应用中都有着出色的表现。大多数可用的医学图像分割架构的灵感来自于著名的全卷积神经网络(FCN)[4]或UNet[5]。在FCN中，将标准分类CNNS的全连通层替换为卷积层，在向前一步实现稠密像素预测。为了恢复输入图像的原始分辨率，预测采用单步上采样的方法。此外，为了提高网络的预测能力，还利用中间特征映射将跳转连接包含在网络中。另一方面，UNet包含了使用卷积层、池层和上行采样层的组合创建的压缩和扩展路径。跳过连接用于连接收缩和扩展路径层的特征。在[30]、[31]、[32]、[33]、[34]、[35]、[36]、[37]、[38]、[39]等广泛的应用中，这些网络被提出了许多扩展以解决像素级分割问题。

*b .深刻的关注*

注意机制旨在强调在局部特征中捕获的重要局部区域，过滤无关的infor-

通过全局特征传递信息，改进了远程依赖的建模。因此，这些模块已经成为需要捕获全局依赖关系的模型的基本部分。这些注意模块的集成在许多视觉问题中被证明是非常成功的，如图像字幕[40]、图像问答[41]和分类[42]。[43]，[44]，[45]最近引起了研究者的注意，因为它在保持计算和统计效率的情况下，表现出了良好的远程依赖建模能力。在这些模块中，每个位置的响应是通过在一个嵌入空间中关注所有位置并取其加权平均来计算的。对于图像视觉问题，[18]、[19]采用自注意来建模局部特征与其相应的全局依赖关系。例如，[18]中提出的PSANet (point-wise spatial attention network, PSANet)通过自适应注意图将feature map中的每个位置与所有其他位置连接起来，使得远程上下文信息能够灵活、动态地聚合。

最近的研究表明，单步产生的注意力特征可能仍然包含来自与给定类无关的区域的噪声，导致次优结果[41]，[46]。为了克服这一问题，一些作品研究了在视觉回答[41]或零镜头学习[46]的背景下，渐进式多重注意层的使用。该策略逐步过滤不需要的噪声，并强调与类语义表示高度相关的区域。就我们所知，堆叠注意力模块的应用在语义分割中仍未被探索。

*C.深度关注医学图像分割*

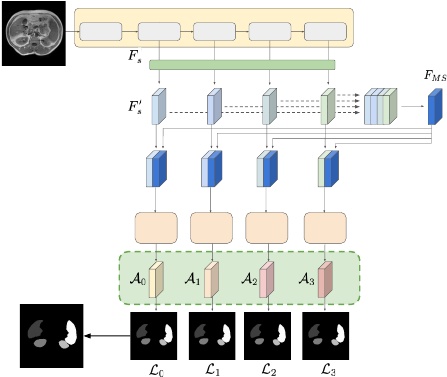
尽管注意机制在许多视觉问题上变得越来越流行，但关于具有注意的医学图像分割的文献仍然很少，只有简单的注意模块[23]、[24]、[25]、[26]、[27]。Wang等人[23]采用了多分辨率的注意模块，将局部深度注意特征(DAF)与全局上下文相结合，用于超声图像的前列腺分割。为了建立长期依赖关系的模型，将局部和全局特征结合在一个简单的注意力模块中，该模块包含三个卷积层，然后是一个softmax函数来创建注意力地图。将类似的注意模块集成到UNet集成的层次聚集框架中，用于左心房分割[24]，该注意模块由两个卷积层和softmax组成。最近，附加的注意门模块被集成到UNet解码路径的跳过连接中，目的是更好地模拟来自编码器[25]的互补信息。

3方法

*答:概述*

医学成像中的目标结构通常在大小、形状和纹理上表现出类内和类间的多样性，特别是在二维图像处理中。传统的神经网络分割具有局部接受域，这导致产生局部特征表示。作为远程

3.



resnet - 101 Conv-1

Res-2

Res-3

Res-4

Res-5

输入

双线性插值

Conv

Conv

Conv

Conv

Conv

引导的注意

引导的注意

引导的注意

引导的注意

分割

关注的功能

图1:本文提出的多尺度引导注意力网络概述。我们使用ResNet-101来提取密集的局部特征。从[Res-2, Res-3, Res-4, Res-5]的输出中获得四张不同大小的feature map。引导注意模块将产生多尺度的注意特征，去除噪声区域，帮助网络强调与语义类更相关的区域。

上下文信息编码不当，局部特征表示可能导致相同标签[19]像素对应的特征之间存在潜在差异。这可能会引入内部类不一致，从而最终影响到识别性能[47]。为了解决这个问题，我们研究了在特征之间建立关联的注意机制。首先，通过使用多尺度策略来获取全局上下文。然后，将学习到的多尺度特征输入到引导注意模块中，引导注意模块由一组空间自注意模块和通道自注意模块组成。空间自注意模块和信道自注意模块有助于自适应地将局部特征与其全局依赖性结合起来，而注意模块的叠加则有助于侧重于相关信息的逐步滤除噪声。建议的框架的概述如图1所示。

*B.多比例尺注意图*

早在[48]深度学习时代之前，多尺度特征就已经被认为在计算机视觉问题中很有用。在深度分割网络的背景下，多尺度特征融合[17]、[49]、[50]表现出惊人的性能。受这些工作的启发，我们利用在多个尺度上学习的特征，这有助于编码全局和局部的上下文。具体来说，我们遵循最近[23]中提出的多尺度战略。在此设置中，多尺度特征表示为Fs，其中s表示体系结构中的层(图1)。由于每层s的特征分辨率不同，因此采用双线性插值将其上采样到一个共同的分辨率，导致特征图F放大s0。然后,Fs0 从所有的刻度被连接起来

形成一个张量，卷积来创建一个常用的多尺度feature map, FMS = conv([F00, F10, F20, F30])。因此，FMS从浅层编码低层次的细节信息，以及从更深层次学习到的高级语义。然后将该多尺度特征图与不同尺度的各特征图结合，输入到引导的注意模块中，生成注意特征As:



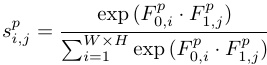
(1)

其中AttMod代表每个引导注意模块(Section III-D)。为多比例尺特征图FMS 在每一个单独的层，补充低层信息和高级语义从FMS 是联合编码的，从而产生更强大的表示。在接下来的章节中，我们将详细说明注意力特征是如何实现的s 得到了。

*C.空间和通道自我注意模块*

如前所述，传统的深度分割模型中的接受域被简化为一个局部邻近区域。这限制了建模范围更广、内容更丰富的上下文表示的能力。另一方面，通道映射可以看作是特定于类的响应，其中不同的语义响应相互关联。因此，增强特定语义的特征表示的另一种策略是改进[51]通道映射之间的依赖关系。为了解决标准cnn的这些限制，我们采用了最近在[19]中提出的位置和频道注意模块，如图2所示。

*位置注意模块(PAM):表示到注意模块的输入feature map F∈RC x W x H，其中C、W、H分别表示通道维度、宽度维度和高度维度。*在上层分支中，F经过一个卷积块，得到一个特征映射F0p∈ RC 0乘以W乘以H，其中C0等于C/81。然后,F0 将pis重塑为形状(W×H)×C的feature map0。在第二个分支中，输入特征图F进行同样的操作，然后转置，得到F1p∈ RC 0W××(H)。将两张图相乘，对结果矩阵应用softmax，生成空间注意图Sp∈R(W×H)×(W×H):



(2)

spi,j评估i的影响th 第j个位置。输入F被送入第三个分支的另一个卷积块，得到F2p∈RC乘以(W乘以H)，与F具有相同的形状2p被重塑为F2p∈RC乘以(W乘以H)。然后将其乘以一个空间注意图的排列版本，其输出被重塑为一个RC \* (W \* H)。位置注意模块对应的注意特征图，即FP AM ，可表示为:

1我们用上标p表示特征图属于位置注意模块。同样，我们将使用上标c作为通道注意模块的特征。

4

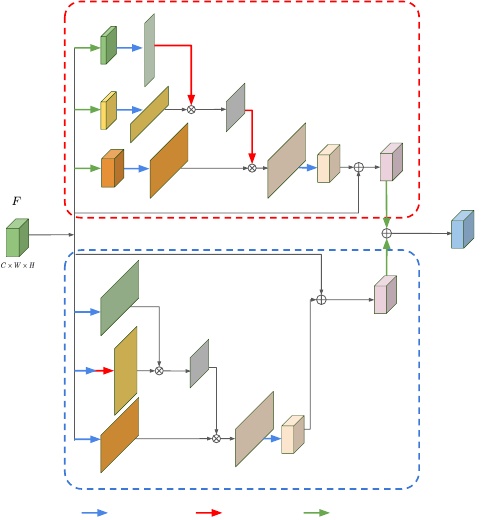
(3)

FP AM,j =

+ Fj

i = 1

在[19]中，表示的是“青”p 初始化为0，逐渐学会了对空间注意图给予更多的重视。因此，位置注意模块在空间注意图的引导下，有选择地将全局上下文聚合到学习到的特征上。



(WxH) xC / 8

(C / 8) xWxH

注意模块的位置

(WxH) x (WxH)

(C / 8) x (WxH)

Softmax

CxWxH

CxWxH

CxWxH

残雪(WxH)

残雪(WxH)

CxWxH

(WxH)我

Self-Attention特性

CxWxH

科学家

Softmax

残雪(WxH)

残雪(WxH)

残雪(WxH)

CxWxH

频道关注模块

重塑

交换

卷积

图2:[19]启发的位置和通道注意模块细节。

*通道注意模块(CAM):通道注意模块的管道如图2底部所示。*输入F∈RC×W ×H 在CAM的前两个支路进行重塑，在第二个支路进行置换，分别得到F0c∈R(W×H)×C和F1c∈RC×(W×H)。然后，我们在F之间进行矩阵乘法0c 和F1c，并得到通道注意图Sc ∈RC×C 为:

exp (F0,ic·F1,jc) sci,j =PC

i=1 exp (F0,i c·F1,j c)

(4)

i的影响在哪里th j上的频道th 由sci给出，j。然后乘以输入F的转置形式，也就是。，其结果被重塑为RC \* (W \* H)。得到最终通道注意图为:

(5)

FCAM j =λc

C

X

i = 1

scij F2,jc

+ Fj

在λc 控制通道注意图在输入特征图f上的重要性。类似于样例p,λc 初始设置为0，逐渐学习。这个公式将所有通道的功能的加权版本聚合为原始功能，突出显示类依赖的功能

映射和增加类之间的特征区分能力。在两个注意模块的最后，新生成的特征被送入卷积层，然后执行元素智慧和操作来生成位置通道注意特征。

*d .指导注意力*

受最近在图像分类[46]中堆叠注意模块的工作启发，我们提出通过顺序细化模块增加注意特征的递进细化。直觉上，这种顺序细化将逐步加权不同局部区域的重要性，同时掩盖不相关的噪声。特别是，给定引导注意模块输入处以s为比例的特征图F，通过连接F生成MS 和Fs0-，通过多步细化生成注意特征(图3)。第一步，位置和信道注意模块使用F生成自注意特征。同时，我们整合了一个编码器-解码器网络，该网络将输入特征F压缩成潜在空间[46]中的压缩表示。其目的是将类信息嵌入到后续的引导注意模块中，通过迫使编解码器的潜在表示接近，可以表示为:

M−1

kEi(FAi−1)−Ei+1(FAi)k22

LG =

X

(6)

我

在Ei(·)为第i个编解码器网络F的编码表示Ai 表示第i个双重注意模块后生成的注意特征，M为迭代次数。请注意,FAi−1 为语义引导注意模块F输入处的特征，即将第一编解码器重构的特征图(n = 0)与第一注意模块生成的自注意特征进行矩阵乘法组合，生成FSA。另外，为了确保重构的特征对应于位置-通道注意模块输入处的特征，编码器的输出被迫接近其输入:

LRec =

M X

kFi −Fˆik22

(7)

我

Fˆi 重构的特征图，即Di(Ei(F)第i个编解码网络的。。

由于引导注意模块是在多个尺度上使用的，所以所有模块的组合引导损失为:

LGT otal =

年代

X

s = 0

汉莎天厨

(8)

同样，重建损失的总和为:S LRecT otal = X LsRec

s = 0

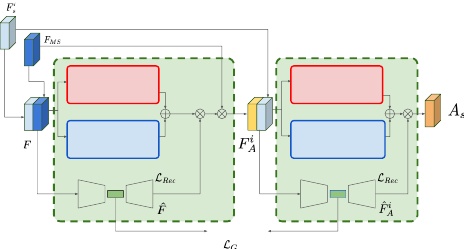
(9)

W×H X

spij F2,摩根大通

5

语义引导注意模块



注意模块的位置

注意模块的位置

频道的关注

模块

Encoder-decoder 0

频道的关注

模块

Encoder-decoder 1

引导损失

1优化步骤

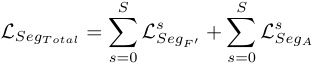
2优化步骤

图3:具有2个细化步骤的语义引导注意模块示意图。对于每个规模，这个模块提供了一组注意的特征，即。,一个s。

在LRec1 和LRec2 为导向注意模块的第一和第二块中编码器-解码器结构的重建损失。

*大肠深处监督*

虽然注意模块不需要辅助的目标函数，但是我们发现在每个尺度[52]上使用额外的监督提高了所提模型的分割性能，这与文献[17]、[23]、[25]的类似工作是一致的。



(10)

其中第一个术语是指分割结果在原始特征Fs0 第二项评价由注意特征提供的分割结果。在所有的情况下，网络预测和ground truth标签之间的多类交叉熵都被用作分割损失。最终优化的目标函数为:



(11)

其中，在主损失函数中，对每个词的重要性进行了控制。

四、实验

*答:实验设置*

*数据集:我们使用三个公共分割基准来评估我们的方法。*首先，来自联合健康腹部器官分割(CHAOS)的腹部MRI数据集挑战了[53]、[54]、[55]。特别地，我们关注的是MRI (T1-DUAL in phase)上腹部器官(脾、肝和肾)的分割，其中包括来自20个被试的扫描，用于训练并给出相应的ground truth注释，以及来自20个被试的扫描，用于没有注释的测试。扫描的分辨率为每片256×256像素，在26到50片之间。由于在数据集中没有提供测试标签，所以我们使用了训练数据集来进行实验，将它分成13、2和5个主题的子集，用于训练、验证和

测试。我们选择不同的受试者重复这个过程三次，并报告三次的平均结果。然后，我们评估了我们在HVSMR 2016挑战赛[56]提供的3D心血管MRI先天性心脏病全心和大血管分割任务上的方法。在这种情况下，心肌和血库是特别针对的对象。训练集包括10个MRI轴位扫描和相应的手动扫描。不同受试者的图像尺寸不同，平均为390×390×165体素体积。通过使用5倍交叉验证策略，我们报告了训练数据的结果，每一倍包含6次训练扫描，2次验证扫描和2次测试扫描。为了增加数据的可变性，我们随机旋转、翻转和镜像图像，但不增加数据集的大小。对于第三个任务，我们使用了医学分割十项挑战中提供的大脑分割数据集2。特别地，这个数据集包含了BRATS ' 16和BRATS ' 17挑战[57]、[58]、[59]的多模态MRI数据(FLAIR、T1w、T1gd、T2w)。该数据集的重点是脑胶质瘤坏死区(TC)和活性区(ET)以及水肿区(ED)的分割。我们使用了484个扫描，分为训练(388个扫描)，验证(48个扫描)和测试(48个扫描)。与前面的任务类似，我们随机旋转、翻转和镜像图像，但不增加数据集的大小。

*网络架构:所提网络的多尺度策略是基于[23]最近的工作，在我们的实验中被认为是较低的基线。*首先，我们对提出的不同模块进行烧蚀研究，以评估每个选择对分割性能的影响。前两个网络——即。(PAM)和(CAM) -分别通过将注意模块替换为空间自注意模块或通道自注意模块来扩展基线(图2)。然后，将两个模块同时结合，形成提出的DANet模型。在下一个模型中-即。由双重注意模块生成的注意特征在一个多步骤的过程中被细化，其中包括第二个双重注意模块。最后，我们提出的架构(被称为拟议的(MS-Dual- guided))扩展了拟议的(MS-Dual)模型，加入了语义指导(图3)。我们还评估了除了注意力模块以外的不同元素对拟议的多尺度架构的影响。首先，我们去掉了模型上的深度监督(eq. 10的第一项)。其次，我们没有使用编码器-解码器结构来重建每个双重注意模块的输入特征，而是将其去除，并用每个注意模块的输入和输出之间的平均误差平方损失替换eq. 7。这种模型称为w/out编码器-解码器(dist)。最后，我们也研究了没有编码器-解码器的影响，即。，没有指导，在细化步骤，这是称为w/out编码器-解码器)。此外，我们评估了有多个改进步骤n的影响，n = 1、2、3和5。

此外，我们还比较了pro- plus的性能

2 http://medicaldecathlon.com

6

提出网络与其他最先进的整合注意力的建筑:注意力UNet [25]， DANet[19]和金字塔注意力网络(PAN)[20]。

*培训和实现细节:我们使用Adam optimizer和mini-batch size为8，以及使用*1 和β2 分别设置为0.9和0.99。虽然大多数网络在前250个epoch内聚合，但我们发现PAN[20]和DANet[19]需要大约400个epoch才能获得最佳结果。初始学习率被设置为0.001,50个epoch后乘以0.5，验证集没有改进。作为分割目标函数，我们在所有网络的所有类别中使用每个像素的交叉熵误差。此外，如第三节所介绍的，我们在所提出的架构中使用eq.(11)中的目标函数，根据经验将次于次于次于设置为1,0.25和0.1。我们使用大小为256×256的二维轴向图像作为网络的输入。实验是在一台装有Titan v的服务器上进行的。我们的模型代码在https://github.com/sinashish/multiscale-attention上公开。

*评估:ground truth和CNN分割的相似性是通过使用几个比较指标来评估的。*首先，我们利用广泛使用的骰子相似系数(DSC)来比较体积的重叠。此外，我们还基于体积相似度(VS)来评估分割性能。此外，为了测量分割轮廓的敏感性，我们考虑了平均表面距离(MSD)的使用。这些指标的制定在补充材料中有详细说明。由于每一次扫描的片间距离和x-y间距没有提供，我们报告这些结果在体素。

*b .结果*

*针对所提出的attention模块的烧蚀研究:为了验证不同组件对分割性能的个人贡献，我们在不同的设置下进行了烧蚀实验。*表一报告了不同注意模块的结果。与基线相比，我们观察到，通过在基线架构的每个尺度上集成空间(PAM)或注意模块(CAM)，平均而言，在重叠和体积相似性方面性能提高了2-3%，在表面距离方面提高了12-18%。另一方面，让两个模块并行。，提出的(DANet) -在DSC方面带来稍好一点的结果，但在使用表面距离度量时实现较低的性能。然而，尽管在MSD上的平均性能较低，提出的DANet模型仍然在4个结构中的3个结构中取得了比channel attention模型更好的结果。这一趋势在DSC度量上得到了重复，DANet在肝脏和左右肾脏这3个相同的结构上超过了提出的CAM架构。这表明，尽管空间注意和通道注意都能提高绩效，但通道注意模块的贡献大于空间注意模块的贡献。如果由提议的DANet模型生成的特征在第二步中被细化-网络称为提议的(MS-Dual) -

平均结果在基于容量和距离的度量上分别进一步提高了0.7%和10%。最后，提出的语义引导丢失(MS- Dual-Guided)的引入带来了性能的额外提升，在三个指标中产生了最佳值:86.75%(DSC)、93.85%(VS)和0.66体素(MSD)。这些结果表明，在DSC、VS和MSD中，与[23]中的基线相比，分别提高了4.5%、4%和26%，显示了提出的注意网络相对于单个注意组件的效率。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 方法 | **DSC (%)** | | **VS (%)** | | **默沙东公司(压)** | |
| 基线(DAF) [23]) | 82.48 | (±6.06) | 89.68 | (±4.48) | 0.92 | (±0.33) |
| 提出(PAM) | 84.46 | (±6.68) | 91.84 | (±4.77) | 0.80 | (±0.43) |
| 提出(凸轮) | 85.08 | (±5.62) | 92.18 | (±5.07) | 0.74 | (±0.32) |
| 提出(DANet) | 85.52 | (±5.86) | 92.07 | (±5.23) | 0.77 | (±0.41) |
| 提出(MS-Dual) | **86.17** | **(±5.78)** | **92.74** | **(±4.76)** | **0.67** | **(±0.30)** |
| 提出(MS-Dual-Guided) | **86.75** | **(±5.05)** | **93.85** | **(±3.50)** | **0.66** | **(±0.27)** |

表一:混沌数据集上不同注意模块的消融研究。数值显示了三次实验的平均结果。最佳和次最佳结果分别用红色和蓝色粗体表示。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 拟议(MS-Dual和MS-Dual- guided) | | | |  |  |
| 模型 | **DSC (%)** | | **VS (%)** | | **默沙东公司(压)** | |
| *1优化步骤*  MS-Dual (No guidance) 85.75  MS-Dual-Guided 86.34  没有深入监督  w/out解码器(dist) 85.92 | | (±5.08)(±5.17)(±4.86)(±5.17) | 92.72 93.47 91.39 92.94 | (±3.65)(±3.78)(±3.55)(±4.04) | 0.71 0.68 0.75 0.76 | (±0.28)(±0.29)(±0.17)(±0.34) |
| *2优化步骤*  MS-Dual(无指导)86.17  MS-Dual-Guided 86.75  没有深入监督  w/out解码器(dist) 86.67 | | (±5.78)(±5.05)(±5.52)(±4.98) | 92.74 93.85 91.80 93.67 | (±4.76)(±3.50)(±3.66)(±3.38) | **0.67 0.66 0.75 0.77** | **(±0.30)(±0.27)(±0.16)(±0.31)** |
| *3优化步骤*  MS-Dual(无引导)86.26  MS-Dual-Guided 86.14  没有深入监督  w/out解码器(dist) 85.88 | | (±5.71)(±5.89)(±5.72)(±4.78) | 93.62 93.50 90.95 93.23 | (±4.72)(±3.98)(±4.31)(±3.71) | 0.71 0.67 0.80 0.79 | (±0.34)(±0.36)(±0.17)(±0.39) |
| *5优化步骤*  MS-Dual(无引导)86.33  MS-Dual-Guided 86.30  没有深入监督  w/out解码器(dist) 86.16 | | (±4.98)(±5.05)(±5.78)(±4.23) | **93.74 93.16 91.03 92.98** | **(±3.91)(±4.11)(±3.66)(±2.93)** | 0.71 0.68 0.87 0.80 | (±0.31)(±0.22)(±0.34)(±0.31) |

表二:MS-Dual和MS-Dual- guided架构上不同元素的消融研究在混沌数据集上评估。数值显示了三次实验的平均结果。最好的结果用红色粗体表示，而蓝色用于突出显示第二好的性能。

表二报告了改进步骤以及若干要素对MS-Dual和MS-Dual- guided模型的影响。首先，我们可以观察到增加细化步骤的数量通常不会提高方法的性能。事实上，只有两个注意力引导模块往往能获得最佳效果。我们认为，逐步细化特征地图可能会产生对注意区域的过度聚焦，导致强烈挖掘注意特征。这有一个不利的影响

7

注意特征可能集中在最具辨别能力的区域，而不是覆盖对象的整个范围。此外，我们观察到，提出的包括引导注意的模型优于所有的变体，特别是在基于距离的度量。此外，我们提供了复杂性方面的比较，其结果在补充材料表VIII中进行了描述。

*与最先进的比较:由几个最先进的分割网络得到的实验结果报告在表III。*在第一个数据集(top)中，与在医学图像分割上下文中提出的其他网络相比，即。、UNet[5]、Attention UNet[25]和DAF[23]——我们的网络分别实现了5.6%、4.3%和2.0%(按DSC计算)、4.9%、4.2%和2.1%(按VS计算)和25%、26%和6%(按MSD计算)的平均改善。这种性能上的差异可以解释为[23]和[25]中集成的注意力模块比我们架构中提出的要简单得多。另一方面，一般计算机视觉任务中的注意力模块受到更多的关注，导致策略更加精细，通常可以取得更好的分割效果。在这些架构中，以ResNet101为骨干的PAN模型[20](和我们的一样)在针对自然场景提出的分割网络中取得了最好的结果。尽管有这些竞争性的结果，提出的模型在DSC、VS和MSD上仍然比PAN架构的性能高出2.4%、1.9%和12%。由于PAN[20]也采用了多尺度的架构，这些差异表明双重自注意和引导细化模块的使用实际上可以提高分割网络的性能。同样，本文提出的模型在第二和第三个数据集(中间和底部)上也优于其他网络，表明它可以广泛应用于一般医学图像的分割。在补充材料的表V, VI和VII中给出了这两个数据集的每门课的分数。除了这些值之外，我们还描述了所有模型在混沌中用于评估的15名受试者的DSC、VS和MSD值的分布(补充材料中图7)。

*定性评价:为了可视化不同注意模块的影响，图4显示了对三个混沌主题的分割结果。*尽管在表III中报告了一些类似的结果，定性结果描述了有趣的发现。首先，我们可以观察到UNet通常在某些器官下节段，很容易混淆。例如，在第二行，它混淆了小肠子和脾脏，而脾脏甚至没有出现在那一层。整合注意力可以克服这些局限性，通过将注意力集中在相关区域来提高分割性能。这可以从其他网络获得的结果中观察到，这些网络在一定程度上减少了假阳性的数量。然而，它会产生更平滑的分段，从而导致丢失细粒度的细节。一个有趣的结果是最后一行的分割，除了提出的网络外，所有的模型都被混淆了来分割左肾。虽然DANet和PAN模型混淆了左肾和右肾，但DAF无法检测到该区域与肾脏相关的任何相关区域。此外，联合国新闻网和

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **混乱** |  |  |  |  |
| 模型 |  | **DSC** |  | **VS** |  | **默沙东公司** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 81.14  83.89  82.70  84.34  82.48  **84.77**  **86.75** | (±7.88)  (±9.54)  (±6.51)  (±6.17)  (±6.06)  **(±5.27)**  **(±5.05)** | 89.01  91.42  90.32  **91.93**  89.68  91.79  **93.85** | (±4.82)  (±4.52)  (±5.27)  **(±4.71)**  (±4.48)  (±3.53)  **(±3.50)** | 0.91  0.78  0.86  0.78  0.92  **0.72**  **0.66** | (±0.49)  (±0.23)  (±0.29)  (±0.31)  (±0.33)  **(±0.24)**  **(±0.27)** |
|  |  | **HSVM** |  |  |  |  |
|  |  | **DSC** |  | **VS** |  | **默沙东公司** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 79.80  **82.55**  80.97  82.26  81.78  81.58  **83.20** | (±6.72)  **(±5.91)**  (±7.76)  (±5.08)  (±5.71)  (±6.84)  **(±4.93)** | 93.41  **94.65**  93.76  94.33  94.31  **94.61**  94.45 | (±6.44)  **(±4.45)**  (±5.85)  (±3.69)  (±3.21)  **(±4.17)**  (±2.39) | 1.68  1.27  1.62  1.24  1.48  **1.25**  **1.19** | (±1.28)  (±0.46)  (±1.19)  (±0.38)  (±0.50)  **(±0.42)**  **(±0.37)** |
|  | **小鬼的18** | |  |  |  |  |
|  |  | **DSC** |  | **VS** |  | **默沙东公司** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 73.65  **79.09**  74.12  76.89  76.78  78.61  **80.37** | (±12.39)  **(±10.89)**  (±12.76)  (±11.53)  (±11.77)  (±10.58)  **(±10.74)** | 87.72  **93.32**  89.85  91.76  90.58  92.66  **93.08** | (±8.70)  **(±6.99)**  (±9.93)  (±8.11)  (±9.03)  (±6.86)  **(±7.20)** | 1.65  **0.95**  1.42  1.17  1.21  1.02  **0.90** | (±0.57)  **(±0.33)**  (±0.52)  (±0.47)  (±0.46)  (±0.40)  **(±0.36)** |

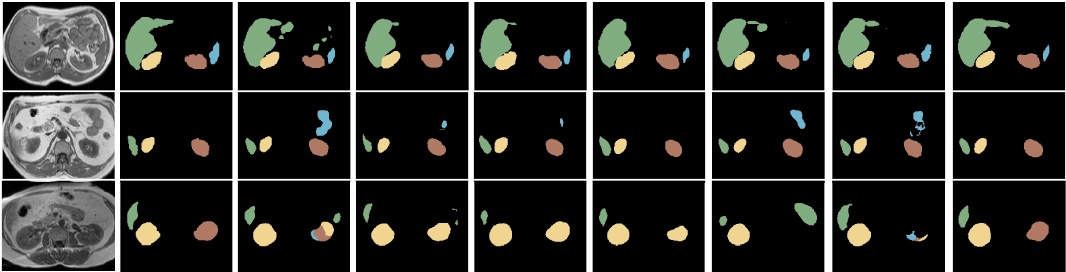
表三:与其他最先进的架构在四个分析数据集的比较。最佳和次最佳结果分别用红色和蓝色粗体表示。

带有注意模型的UNet生成了左肾的三个部分，即左肾。左、右肾和脾脏，这在解剖学上不可信。与所有这些模型不同的是，所提出的架构没有受到模糊区域的干扰，一些错误分类的结构现在被正确分类。

在BRATS’18的图像中，我们可以观察到类似的结果(图5)。特别的是，我们可以看到，我们提出的网络比其他架构获得了更细的细节。例如，对水肿的小分枝(绿色部分)可以更好地由所提议的模型(第二行)捕获。同样，我们的方法分割坏死区域(红色部分)更接近于ground truth，特别是当区域形状复杂时(第一行)。这些视觉结果表明，我们的方法可以成功地恢复更细的分割细节，同时避免在模糊区域分心。空间信息和通道图之间的选择性集成，以及引导注意模块有助于捕获上下文信息。实验表明，该多尺度引导注意力模型能够有效地对互补信息进行编码，实现医学图像的精确分割。

*特征图的视觉检查:仅通过消融研究和定量评估显示性能差异可能不足以完全理解新模型的好处和行为。*虽然如结果所示，提出的模块有助于性能改进，但研究不同模块是否按预期工作是很有趣的。为此，我们从空间和通道两方面分析了一些参与的特征图

8

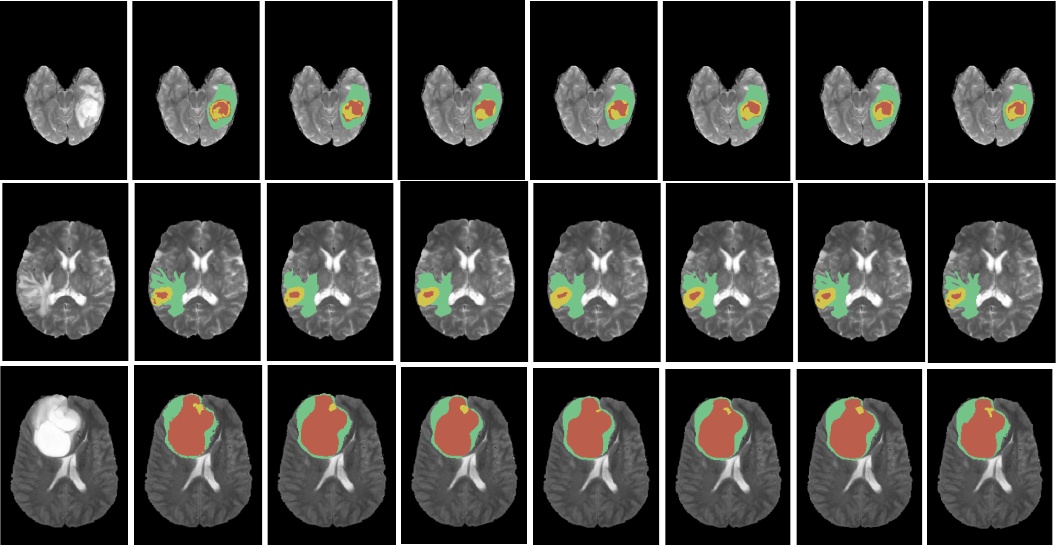


输入地面盘盘注意

求婚了

Image Truth (ResNet34) (ResNet101) UNet

图4:三个主题在CHAOS Challenge数据集上的结果。所提出的多尺度引导注意力网络在质量上优于其他集成注意力模块的网络。



锅(ResNet-34)

锅(resnet - 101)

输入核磁共振

GT

UNet

DANet

UNet关注

提出了

图5:结果在三个科目上的BRATS挑战数据集。在这些图中，描绘了以下肿瘤结构:水肿(绿色)，增强核(黄色)和坏死或肿瘤核(红色)。

注意模块(图6)。我们发现，经过第二个引导注意模块后，特定语义类的反应更加明显，即:PAM 2和CAM 2的专注特征。虽然空间注意和通道注意可以在引导模块的第一步(第二列和第三列)突出特定的类语义，但语义映射中仍然突出了一些非目标区域。此外，最高值也在整个图像上分布得更广。相反，所提出的引导注意模块生成的特征(第四和第五栏)更好地聚焦于感兴趣结构的特定区域。特别的是，可以看到存在一些特征图，突出的区域集中在一个单一的器官上，避免了可能导致某些区域分类错误的模糊区域。

诉的结论

在这项工作中，我们引入了一种新的注意力结构用于医学图像分割任务。该模型采用多尺度策略结合不同层次的语义信息和自我注意模块逐步聚合相关的上下文特征。最后，通过引导细化模块滤除图像中的噪声区域，帮助网络聚焦于图像中与类相关的区域。为了验证我们的方法，我们在腹部器官、心血管结构和脑瘤三个不同的分割任务上进行了实验。我们提供了大量的实验来评估所提议的体系结构的各个组件的影响。此外，我们将我们的模型与现有的整合注意力的方法进行了比较

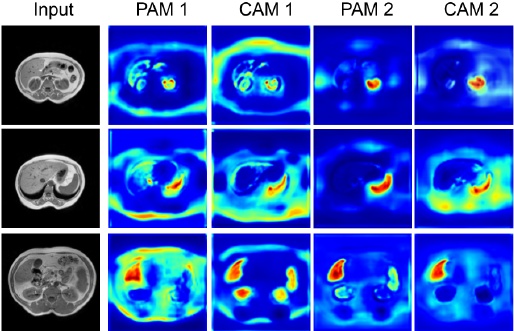


图6:通道图可视化结果。对于每一行，我们显示一个输入图像，以及图3的导视模块处的空间注意模块(PAM)和通道注意模块(CAM)的输出对应的通道映射。

最近提出用于自然场景[19]、[20]和医学图像[5]、[23]、[25]的分割。实验结果表明，所提出的模型在定量和定性上都优于以往的所有方法，这可能是由于增强了基于局部特征的丰富上下文依赖的建模能力。这证明了我们的方法的效率，提供精确和可靠的医学图像自动分割。

致谢

我们要感谢英伟达的慷慨捐赠的Titan V GPU用于这项工作。

参考文献

[1] G. Litjens等，“医学图像分析中的深度学习研究”，《医学图像分析》，第42卷，第60-88页，2017年。

[2] J. Dolz等，“超密网络:用于多模态图像分割的超密连接CNN”，《IEEE医学成像学报》，2018年。

深度学习技术用于自动MRI心脏多结构分割和诊断:问题解决了吗?IEEE医学影像汇刊，第37卷，第2期。11，第2514-2525页，2018。

王建民，“全卷积网络在语义分割中的应用”，《计算机视觉与模式识别》，2015年，页331 - 3440。

王建民，“卷积网络在生物医学影像分割中的应用”，国立中山大学医学研究所硕士论文。施普林格，2015,pp 234-241。

“精细网络:高分辨率语义分割的多路径精细网络”，《IEEE计算机视觉与模式识别会议论文集》，2017年，页1925-1934。

[7] L.-C。陈等，“deep plab:基于深度卷积网络、atrous卷积和完全连接CRFs的语义图像分割”，《IEEE模式分析与机器智能学报》，第40卷，第1期。4, 2018年第834-848页。

赵、史、齐、王、贾，“金字塔场景解析网络”，《IEEE计算机视觉与模式识别会议论文集》，2017年，第2881-2890页。

[9] L.-C。陈等，“可分离卷积的编码解码器用于语义图像分割”，《欧洲计算机视觉会议论文集》，2018，第801-818页。

[10] F. Yu和V. Koltun，“扩大卷积的多尺度上下文聚合”，ICLR, 2016。

9

b . W. Liu, A. Rabinovich, A. C. Berg， < Parsenet: Looking wider to see better >， arXiv预印本:1506.04579,2015。

王志伟，“基于多上下文的姿态估计方法”，《计算机视觉与模式识别》，《计算机视觉与模式识别IEEE会议论文集》，2017，页1831-1840。

“群体层面情绪认知的注意模型”，载于《2018年多模态互动国际会议论文集》。《ACM》，2018，第611-615页。

[14] . Huang等，“基于金字塔注意力网络的场景文本检测的掩模R-CNN”，《2019年IEEE计算机视觉应用冬季会议》。IEEE, 2019，第764-772页。

陈淑君、陈欣君、王淑君、胡欣君，“显著目标检测的逆向注意力”，《欧洲计算机视觉会议论文集》，2018年，第234-250页。

“引导注意力推理网络”，载于《IEEE计算机视觉与模式识别会议论文集》，2018年，第9215-9223页。

[17] L.-C。陈等，“关注尺度:尺度感知的语义图像分割”，《IEEE计算机视觉与模式识别会议论文集》，2016，第3640-3649页。

王志强，“空间注意力网络在视觉分析中的应用”，《计算机视觉技术与应用》，北京:中国计算机科学出版社，2003年。

王志军，“基于双注意力的场景分割网络”，《计算机视觉与模式识别》，北京，2003年。

李宏辉，熊平，安俊杰，王良，“金字塔注意网络的语义分割”，载于《视觉科学》，2018年。

余秋雨等，“实时语义分割的双边分割网络”，《计算机视觉欧洲会议论文集》，2018年，页325-341。

张培平，“基于深度门控的大规模街道场景分割”，《模式识别》，第88卷，第702-714页，2019年。

王玉英等，“超声对前列腺分割的深层注意特征”，《中华医学》，2018年。

[24] C. Li等，“基于注意力的左心房三维分割层次聚集网络”，心脏统计地图集与计算模型国际研讨会。施普林格，2018，第255-264页。

王建民，“关注门控网络:学习利用医学影像中的显著区域”，《医学影像分析》，第53卷，第197-207页，2019年。

王丽萍，“基于注意力的半监督深度网络在医学图像分割中的应用”，国立中山大学医学研究所硕士论文。施普林格，2018，第370-378页。

王建民，“全卷积网络中的并发空间和通道挤挤与激发”，国立中山大学医学影像计算与计算机辅助干预研究所硕士论文。施普林格，2018，第421-429页。

[28] T. Heimann和h。“用于三维医学图像分割的统计形状模型:综述”，《医学图像分析》，第13卷，第1期。4，第543-563页，2009。

刘建民，“核磁共振成像中脑皮层下结构的分割算法:一项调查”，中华医学杂志，第36卷，第1期。4, 200-212页，2015年。

[30] T. Fechter等，“三维全卷积神经网络与随机游走在CT中的食管分割”，《医学物理》，第44卷，第2期。12，第6341-6352页，2017。

李志伟，“H-DenseUNet:用于肝脏和肿瘤CT卷分割的混合紧密连接UNet”，《IEEE医学影像学报》，第37卷，第1期。12, 2018年第2663-2674页。

王建民，“基于几何感知的U-Net的深度Q学习驱动CT胰腺分割”，中华医学会医学影像学杂志，2019年。

黄志平，“三维全卷积网络在MRI中用于皮层下分割的研究”，《神经影像》，第170卷，第456-470页，2018年。

“先进的全自动小脑分离技术与磁共振成像技术的比较”，《神经影像》，2018年。

[35] C. Zotti, Z. Luo, A. Lalande, p.m。“具有形状的卷积神经网络在心脏mri分割中的应用”，《IEEE生物医学与健康信息学杂志》，第23卷，第1期。2018年，第1119 - 1128页。

王建民，“膀胱肿瘤结构的多区域分割研究”，国立中山大学医学研究所硕士论文。12, 2018年第5482-5493页。

金，郭，杜益。何，A. P. Harrison, J. Xiao, c.k。曾、吕丽萍，“双流链式三维深度网络融合在PET/CT中准确分割食管肿瘤体积”，医学影像计算与计算机辅助介入国际会议。施普林格，2019，第182-191页。

王建民，“基于稀疏可变形卷积算法的三维多器官分割方法”，《医学图像分析》，第54卷，第1-9页，2019年。

王元义，周元义，沈文伟，朴志强，“基于器官关注网络的腹部多器官分割方法”，医学影像分析，第55卷，第88-102页，2019年。

王建民，“图像字幕的关注领域”，载于《计算机视觉与电子工程师学会国际会议论文集》，2017，第1242-1250页。

杨振荣等，“用于图像问题回答的叠加注意力网络”，《IEEE计算机视觉与模式识别会议论文集》，2016年，第21-29页。

王永平，“剩余注意力网络在图像分类中的应用”，《IEEE计算机视觉与模式识别会议论文集》，2017，第3156-3164页。

“自然语言推理的可分解注意模型”，载于《自然语言推理》，2016年。

“你所需要的就是注意力”，《神经信息处理系统进展》，2017，第5998-6008页。

王永平，“非局部神经网络”，《IEEE计算机视觉与模式识别会议论文集》，2018年，第7794-7803页。

“用于细粒度零射击学习的叠加语义引导注意力模型”，《神经信息处理系统进展》，2018年，第5995-6004页。

“基于全局卷积网络的大核问题-改进语义分割”，《IEEE计算机视觉与模式识别会议论文集》，2017，第4353-4361页。

“轮廓检测与分层图像分割”，国立中山大学计算机科学与工程研究所硕士论文，第33卷，第1期。5，第898-916页，2010。

王建民，“基于超列的目标分割与精细定位”，《计算机视觉与模式识别IEEE会议论文集》，2015，页447-456。

“具有放大特征的前馈语义分割”，《IEEE计算机视觉与模式识别会议论文集》，2015年，第376 - 3385页。

陈丽玲，“卷积网络中用于图像字幕的空间和信道注意”，《IEEE计算机视觉和模式识别会议论文集》，2017，第5659-5667页。

[52]彭译葶。李、谢、张、图，“深度监督的网络”，《人工智能与统计》，2015年，第562-570页。

“在腹部器官直接体积绘制的传递函数规范中探索基于brushlet的三维纹理”，《IEEE可视化与计算机图形学学报》，第21卷，第1期。2，第174-187页，2014。

王建民，“利用多层层次分类方法从MR影像中分割腹部器官”，国立中山大学工程与建筑学院学报，第30卷，第1期。3，第533-546页，2015年。

王建民，“应用多层神经网络技术实现腹部器官的分割”，国立中山大学医学研究所硕士论文。3, 830 - 8522014页。

刘建民，“心脏全心分割技术在先天性心脏病中的应用”，国立中山大学医学研究所硕士论文。施普林格，2015，第80-88页。

王建民，“多模态脑肿瘤影像分割基准研究”，国立中山大学医学影像研究所硕士论文，第3期。1993-2024, 2014年10页。

“利用专家分割标签和放射学特征进行癌症基因组图谱收集”，《科学数据》，第4卷，第170117页，2017年。

“对脑肿瘤分割的最佳机器学习算法的识别，进展。

10

评估和总体生存预测在brats挑战，" arXiv预印arXiv:1811.02629, 2018。

**补充材料**

*评价指标:配方*

在本部分中，我们给出了用于评估所提议的体系结构的度量标准的正式定义。

*a)骰子相似系数(DSC):给定a和B两卷，其DSC可定义为:*



(12)

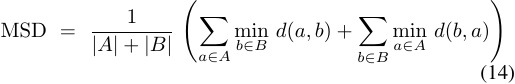
在这个度量中，接近1的值表示高度重叠，而接近0的值表示根本没有重叠。

*b)体积相似度(VS):此外，我们还基于体积相似度来评估分割性能，其公式为:*



(13)

*c)平均表面距离(MSD):定义轮廓A与轮廓B之间的MSD:*





(15)

其中，d(a, b)为a曲面上a点到b曲面的距离，由欧几里得范数的极小值给出:



(16)

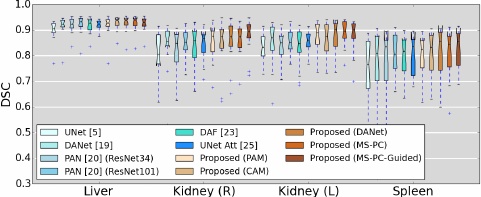
*额外的结果*

表IV、V和VI报告了消融研究的实验结果的扩展版本，并与其他最先进的网络进行了比较。在这些表格中，对单个器官的个别结果也包括在内，以便读者更广泛地了解不同方法的性能。我们可以观察到，所提议的架构始终优于其他模型，在所有评估指标的几乎所有器官中排名第一或第二。唯一的例外是肝脏分割在体积相似方面的结果，所有的模型都得到了几乎相同的结果。

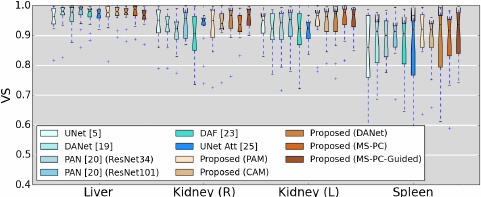
除了报道值表IV和V的补充材料,我们也描绘DSC的分布,VS和MSD值15科目用于评价所有的模型(图7)。在这些情节,我们可以先观察的影响不同的关注模块分割模型的性能。随着我们逐步将所提出的注意力模块纳入基线网络，分割性能得到改善，这体现在分割精度值分布较好且方差较小。这种结果分布的差异在与其他网络比较时更加突出

1

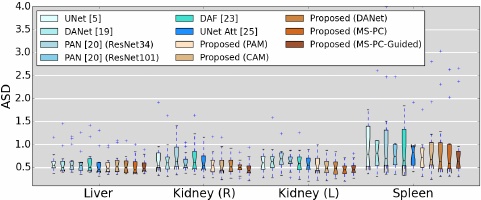
最先进的网络，用蓝色的盒子图表示。我们还可以观察到，这种模式在各个器官和指标之间是恒定的，这表明所提出的注意力网络比目前最先进的结构能够获得更好和更健壮的分割结果。



(a)骰子相似系数(%)



(b)体积相似度(%)



(c)平均表面距离(体素)

图7:这些图描绘了四个被分割器官不同评价指标的分布。蓝色代表由其他先进网络获得的结果，而由我们提出的模型获得的结果显示在褐色的箱线图中。

*1)收敛性:我们还比较了不同架构的收敛性，结果如图8所示。*特别地，每个网络在一个验证折叠上的四个结构上的平均DSC值被显示出来。可以观察到，尽管大多数网络取得的结果在某种程度上可能被认为是“相似的”，但收敛行为是完全不同的。当有三个网络具有相似的收敛曲线时，即在UNet, DANet和DAF -， PAN需要更多的迭代来收敛，最终在将近400个epoch后比这些网络表现得更好。另一方面，我们发现注意UNet和所提出的网络收敛速度最快，分别在epoch 48和73处达到最佳结果。

2

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | **DSC (%)** | |  |  |  |  |
| 方法 |  | **肝** | **肾脏R** | | **肾脏L** | | **脾** | |  | **的意思是** |
| 基线(DAF) [23])  提出(PAM)  提出(凸轮)  提出(DANet)  提出(MS-Dual)  提出(MS-Dual-Guided) | 91.66(±2.99)  91.89(±4.29)  92.58(±2.65)  **92.60(±3.20)**  **92.62(±3.08)**  92.46(±2.82) | | 79.28(±18.68)  85.47(±7.04)  84.52(±9.34)  85.29(±7.96)  **86.29(±5.98)**  **87.96(±6.46)** | | 83.63  86.84  86.38  87.74  **88.82**  **88.01** | (±7.56)  (±6.53)  (±6.27)  (±6.37)  **(±4.84)**  **(±6.16)** | 75.35  73.65  76.84  76.44  **76.96**  **78.61** | (±20.41)  (±22.62)  (±20.56)  (±22.17)  **(±19.87)**  **(±18.69)** | 82.48(±6.06)  84.46(±6.68)  85.08(±5.62)  85.52(±5.86)  **86.17(±5.78)**  **86.75(±5.05)** | |
|  |  |  |  | **体积相似度(VS) (%)** | | | |  |  |  |
|  |  | **肝** | **肾脏R** | | **肾脏L** | | **脾** | |  | **的意思是** |
| 提出(DAF) [23])  提出(PAM)  提出(凸轮)  提出(DANet)  提出(MS-Dual)  提出(MS-Dual-Guided) | 96.69(±3.21)  96.62(±4.62)  **97.25(±2.95)**  97.04(±3.03)  **97.47(±3.07)**  96.44(±3.15) | | 86.75(±16.41)  92.83(±7.43)  93.78(±6.04)  **94.50(±5.96)**  93.30(±4.11)  **96.14(±3.15)** | | 90.29  93.96  93.98  93.43  **95.27**  **94.95** | (±8.39)  (±6.46)  (±5.48)  (±7.03)  **(±4.89)**  **(±4.48)** | **84.98**  83.93  83.72  83.30  84.90  **87.87** | **(±14.42)**  (±20.54)  (±20.97)  (±22.53)  (±16.86)  **(±15.23)** | 89.68(±4.48)  91.84(±4.77)  92.18(±5.07)  92.07(±5.23)  **92.74(±4.76)**  **93.85(±3.50)** | |
|  |  |  |  | **平均表面距离(体素)** | | | | |  |  |
|  |  | **肝** | **肾脏R** | | **肾脏L** | | **脾** | |  | **的意思是** |
| 基线(DAF) [23])  提出(PAM)  提出(凸轮)  提出(DANet)  提出(MS-Dual)  提出(MS-Dual-Guided) | 0.64  0.55  0.58  **0.54**  **0.53**  **0.54** | (±0.29)  (±0.19)  (±0.22)  **(±0.19)**  **(±0.18)**  **(±0.16)** | 0.97  0.56  0.57  0.56  **0.51**  **0.48** | (±1.08)  (±0.23)  (±0.24)  (±0.19)  **(±0.14)**  **(±0.18)** | 0.63  0.55  0.52  0.50  **0.46**  **0.48** | (±0.25)  (±0.21)  (±0.20)  (±0.18)  **(±0.14)**  **(±0.14)** | 1.45  1.54  1.29  1.49  **1.19**  **1.13** | (±2.04)  (±2.40)  (±1.64)  (±2.29)  **(±1.42)**  **(±1.24)** | 0.92  0.80  0.74  0.77  **0.67**  **0.66** | (±0.33)  (±0.43)  (±0.32)  (±0.41)  **(±0.30)**  **(±0.27)** |
|  |  |  |  |  |  |  |  |  |  |  |

表四:混沌数据集(MRI任务中的多器官分割)上不同注意力模块的消融研究。数值显示了三次实验的平均结果。最好的结果用红色粗体表示，而蓝色用于突出显示第二好的性能。

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | **DSC (%)** | |  |  |  |  |
| 方法 |  | **肝** | **肾脏R** | | **肾脏L** | | **脾** | |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 90.94(±4.01)  91.69(±4.07)  91.99(±2.98)  **92.13(±3.51)**  91.66(±2.99)  92.02(±1.93)  **92.46(±2.82)** | | 79.14(±15.23)  83.85(±9.40)  81.51(±9.03)  **85.02(±5.16)**  79.28(±18.68)  84.33(±5.91)  **87.96(±6.46)** | | 82.51  84.49  83.62  85.36  83.63  **85.57**  **88.01** | (±7.48)  (±8.60)  (±6.21)  (±4.87)  (±7.56)  **(±4.09)**  **(±6.16)** | 71.95  75.54  73.70  74.84  75.35  **77.18**  **78.61** | (±21.61)  (±16.08)  (±19.97)  (±21.23)  (±20.41)  **(±15.95)**  **(±18.69)** | 81.14(±7.88)  83.89(±9.54)  82.70(±6.51)  84.34(±6.17)  82.48(±6.06)  **84.77(±5.27)**  **86.75(±5.05)** | |
|  |  |  |  | **体积相似度(VS) (%)** | | | |  |  |  |
|  |  | **肝** | **肾脏R** | | **肾脏L** | | **脾** | |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 95.54(±4.43)  96.90(±4.18)  96.56(±3.55)  **96.99(±3.64)**  96.69(±3.21)  **96.95(±1.89)**  96.44(±3.15) | | 87.68(±5.77)  92.88(±5.12)  90.89(±5.64)  **93.77(±4.63)**  86.75(±16.41)  92.29(±6.41)  **96.14(±3.15)** | | 89.55  91.52  91.83  **92.69**  90.29  91.79  **94.95** | (±4.68)  (±6.73)  (±7.75)  **(±6.88)**  (±8.39)  (±3.53)  **(±4.48)** | 83.28  84.37  81.98  84.24  84.98  **85.94**  **87.87** | (±14.78)  (±16.15)  (±20.67)  (±17.37)  (±14.42)  **(±11.88)**  **(±15.23)** | 89.01(±4.82)  91.42(±4.52)  90.32(±5.27)  **91.93(±4.71)**  89.68(±4.48)  91.74(±3.91)  **93.85(±3.50)** | |
|  |  |  |  | **平均表面距离(体素)** | | | | |  |  |
|  |  | **肝** | **肾脏R** | | **肾脏L** | | **脾** | |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 0.59  0.61  0.62  **0.57**  0.64  **0.57**  **0.54** | (±0.18)  (±0.27)  (±0.25)  **(±0.22)**  (±0.29)  **(±0.25)**  **(±0.16)** | 0.69  0.65  0.75  **0.61**  0.97  **0.61**  **0.48** | (±0.38)  (±0.31)  (±0.31)  **(±0.19)**  (±1.08)  **(±0.23)**  **(±0.18)** | 0.61  0.67  0.69  0.64  0.63  **0.56**  **0.48** | (±0.19)  (±0.30)  (±0.21)  (±0.15)  (±0.25)  **(±0.18)**  **(±0.14)** | 1.76  1.17  1.37  1.30  1.45  **1.15**  **1.13** | (±2.57)  (±0.94)  (±1.43)  (±1.47)  (±2.04)  **(±1.01)**  **(±1.24)** | 0.91  0.78  0.86  0.78  0.92  **0.72**  **0.66** | (±0.49)  (±0.23)  (±0.29)  (±0.31)  (±0.33)  **(±0.24)**  **(±0.27)** |

表V:与其他混沌数据集(MRI任务中的多器官分割)上的最新网络结构的比较。数值显示了三次实验的平均结果。最好的结果用红色粗体表示，而蓝色用于突出显示第二好的性能。

3.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **DSC** |  |  |
| 方法 | **心肌** | | **血池** | |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出了 | 71.77(±9.36)87.84(±4.35)  **75.85(±9.10)89.24(±3.56)**  72.90(±11.93)89.04(±3.69)  74.98(±7.68)89.53(±2.97)  74.08(±8.55)89.48(±3.39)  74.50(±10.13)88.66(±4.25)  **77.10(±6.94)89.30(±3.50)** | | | | 79.80(±6.72)  **82.55(±5.91)**  80.97(±7.76)  82.26(±5.08)  81.78(±5.71)  81.58(±6.84)  **83.20(±4.93)** | |
|  |  |  | **卷相似(VS)** | | |  |
|  | **心肌** | | **血池** | |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出了 | 91.05(±9.75)95.78(±4.04)  91.80(±8.95)97.50(±3.01)  90.58(±10.89)96.93(±3.66)  91.42(±7.59)97.23(±2.36)  91.73(±6.30)96.89(±2.33)  **92.52(±7.66)96.69(±2.20)**  **92.08(±4.39)96.82(±2.76)** | | | | 93.41(±6.44)  **94.65(±4.45)**  93.76(±5.85)  94.33(±3.69)  94.31(±3.21)  **94.61(±4.17)**  94.45(±2.39) | |
|  |  | **平均表面距离(MSD)** | | | | |
|  | **心肌** | | **血池** | |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出了 | 1.82  **1.23**  1.97  1.33  1.41  1.24  **1.15** | (±1.48)  **(±0.51)**  (±1.84)  (±0.53)  (±0.45)  (±0.42)  **(±0.33)** | 1.55  1.32  1.26  **1.15**  1.44  1.25  **1.24** | (±1.08)  (±0.46)  (±0.48)  **(±0.30)**  (±0.46)  (±0.39)  **(±0.43)** | 1.68  1.27  1.62  1.24  1.48  **1.25**  **1.19** | (±1.28)  (±0.46)  (±1.19)  (±0.38)  (±0.50)  **(±0.42)**  **(±0.37)** |

表VI: HVSMR 2016数据集上的建议网络与其他最先进的架构的比较。数值显示了5次实验的平均结果。

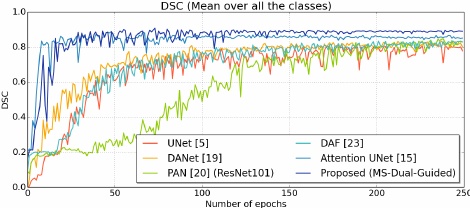


图8:平均验证DSC随时间的演变。

4

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | **DSC (%)** |  |  |  |  |  |
| 方法 |  | **艾德** |  | **等** |  | **TC** |  | **- - - - - -** |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 84.87  **88.24**  85.25  87.07  86.87  87.50  **89.11** | (±6.82)  **(±5.39)**  (±6.64)  (±6.67)  (±5.94)  (±5.66)  **(±4.94)** | 56.38  63.69  55.89  60.77  60.28  **63.74**  **65.25** | (±27.55)  (±22.25)  (±27.76)  (±24.74)  (±24.74)  **(±22.65)**  **(±.2285)** | 79.71(±11.70)  **85.33(±6.92)**  81.23(±8.22)  82.82(±8.76)  83.18(±8.39)  84.59(±7.43)  **86.76(±6.49)** | |  | - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - - | 73.65  **79.09**  74.12  76.89  76.78  78.61  **80.37** | (±12.39)  **(±10.89)**  (±12.76)  (±11.53)  (±11.77)  (±10.58)  **(±10.74)** |
|  |  |  |  | **体积相似度(VS) (%)** | | |  |  |  |  |
|  |  | **艾德** |  | **等** |  | **TC** |  | **- - - - - -** |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 96.36  **99.04**  98.05  **98.68**  97.99  98.14  98.54 | (±4.08)  **(±1.21)**  (±1.98)  **(±2.21)**  (±2.10)  (±1.88)  (±1.76) | 75.81  **83.47**  75.87  80.38  77.86  **82.99**  82.91 | (±27.23)  **(±20.11)**  (±28.17)  (±24.83)  (±24.92)  **(±21.09)**  (±20.17) | 90.99(±11.63)  **97.45(±2.95)**  95.63(±4.41)  96.22(±5.89)  95.88(±5.26)  96.84(±2.87)  **97.78(±2.56)** | |  | - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - - | 87.72(±8.70)  **93.32(±6.99)**  89.85(±9.93)  91.76(±8.11)  90.58(±9.03)  92.66(±6.86)  **93.08(±7.20)** | |
|  |  |  | **平均表面距离(体素)** | | | | | |  |  |
|  |  | **艾德** |  | **等** |  | **TC** |  | **-** |  | **的意思是** |
| UNet [5]  DANet [19]  锅(ResNet34) [20]  锅(ResNet101) [20]  来说,[23]  UNet关注[25]  提出(MS-Dual-Guided) | 0.99  **0.67**  0.86  0.74  0.76  0.69  **0.58** | (±0.33)  **(±0.16)**  (±0.20)  (±0.19)  (±0.17)  (±0.18)  **(±0.14)** | 2.37  **1.43**  2.29  1.79  1.84  1.58  **1.40** | (±1.74)  **(±0.95)**  (±1.87)  (±1.35)  (±1.33)  (±1.12)  **(±1.02)** | 1.56  **0.78**  1.10  0.96  1.02  0.79  **0.71** | (±1.34)  **(±0.25)**  (±0.47)  (±0.48)  (±0.66)  (±0.29)  **(±0.31)** |  | - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - -  - - - - - - | 1.65  **0.95**  1.42  1.17  1.21  1.02  **0.90** | (±0.57)  **(±0.33)**  (±0.52)  (±0.47)  (±0.46)  (±0.40)  **(±0.36)** |

表七:在BRATS 2018数据集(MRI任务中的多器官分割)上，将提出的网络与其他最先进的架构进行比较。数值显示了三次实验的平均结果。最好的结果用红色粗体表示，而蓝色用于突出显示第二好的性能。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **模型的复杂性** | | |  |  |  |  |  |
| **模型** |  |  |  | #参数 |  |  |  |  |
|  |  |  | 1 Iter |  | 2 Iter | 3 Iter | 5 Iter | |
| UNet | 31030853年 | | - |  | - | - |  | - |
| 锅(ResNet34) | 21323991年 | | - |  | - | - |  | - |
| 锅(ResNet101) | 42675415年 | | - |  | - | - |  | - |
| UNet关注 | 34877681年 | | - |  | - | - |  | - |
| DANet (ResNet101) | 68475961年 | | - |  | - | - |  | - |
| 提出(DAF) | 43482179年 | | - |  | - | - |  | - |
| 提出(PAM) | 43486343年 | | - |  | - | - |  | - |
| 提出(凸轮) | 43485543年 | | - |  | - | - |  | - |
| 提出(DANet) | 43980179年 | | - |  | - | - |  | - |
| MS-Dual(没有指导) |  | - | 43485831 44411103 | | | 45337675年 | 47190819年 | |
| MS-Dual-Guided |  | - | 50531399 58499679 | | | 66470539年 | 82412259年 | |
| MS-Dual-Guided(无深Sup) |  | - | 50530099 58498379 | | | 66467939年 | 82407059年 | |
| MS-Dual-Guided(经销) |  | - | 43485831 44411103 | | | 45337675年 | 47190819年 | |

表八:以参数数衡量的评估模型的模型复杂性。