

基于改进Unet与动态阈值可变FCMSPCNN的医学图像分割

邸敬,马帅,王国栋,廉敬

兰州交通大学电子与信息工程学院,甘肃兰州730070

【摘要】针对深度学习的医学图像分割模型训练时间长和精度不精的问题,提出结合动态阈值可变FCMSPCNN的多尺度上下文编解码结构和注意力机制的CoA Unet(Context Attention Unet)分割方法。首先,使用动态阈值可变的FCMSPCNN预分割出目标矩形区域并使用掩码遮盖背景部分;然后,更深层卷积块加入快捷连接交叉融合不同层次的特征,并通过注意力门突出对目标特征的学习;最后,在编解码器最底层加入改进的多尺度上下文提取器可以更好地提取目标特征信息。模型分别在LiTs和DRIVE数据集上进行验证,肝脏分割指标Miou、Aver_HD、Aver_Dice分别为0.8905、6.3699、0.9477,视网膜血管分割指标分别为0.5892、9.2559、0.7409。实验表明,预处理能缩短4.3%~20.33%的训练时间并提升2%~6%分割精度,与其他5种分割方法相比,CoA Unet能取得更好的整体分割性能。

【关键词】CoA Unet;脉冲耦合神经网络;注意力机制;肝脏分割;视网膜血管分割

【中图分类号】R318;TP391.9

【文献标志码】A

【文章编号】1005-202X(2023)03-0328-08

Medical image segmentation using improved Unet combined with dynamic threshold changed FCMSPCNN

DI Jing, MA Shuai, WANG Guodong, LIAN Jing

School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Abstract: Aiming at the long training time and low precision of medical image segmentation model based on deep learning, a segmentation method combining the multi-scale context encoding and decoding structure of dynamic threshold changed FCMSPCNN with context attention Unet (CoA Unet) is proposed. Dynamic threshold changed FCMSPCNN is used to pre-segment the target rectangle and mask the background. The deeper convolution block with shortcut connection integrates features at different levels, and highlights the learning of target features through the attention gate. Furthermore, an improved multi-scale context extractor is added to the bottom layer of the codec to better extract target feature information. The model is verified on LiTs and DRIVE datasets, respectively. The Miou, Aver_HD and Aver_Dice of the proposed method are 0.8905, 6.3699, 0.9477 for liver segmentation, and 0.5892, 9.2559, 0.7409 for retinal vascular segmentation. The experiment reveals that the preprocessing can shorten the training time by 4.30%-20.33% and improve the segmentation accuracy by 2%-6%. Compared with other 5 segmentation methods, CoA Unet can achieve better overall segmentation performance.

Keywords: CoA Unet; pulse coupled neural network; attention mechanism; liver segmentation; retinal vascular segmentation

前言

检测和分割病理区域是常见医学图像分析任务之一,其难点在于病灶通常占据整幅图像较少部分,这种不平衡加大了分割的难度。随着卷积神经网络

(CNN)成功应用于2D和3D图像自动分割,深度学习学习方法成为分割任务的主流,并诞生了各种强大的深度架构。早期深度学习的改进大多集中于拓展网络深度或宽度,构造大规模网络,例如 AlexNet^[1]、VGG^[2]、ResNet^[3]和 DenseNet^[4],网络大多具有数百层。为了促进深度网络的训练,并解决网络深度过深引发的梯度消失和特征退化等问题,dropout、批量归一化、快捷连接等解决方案相继提出。Inception系列和 Wide Residual Networks(WRN)^[5]实验结果表明网络宽度的增加对性能的提升是有益的,因为模型可以重复利用更多特征。随着深度学习的发展,特征融合成为研究的重点,ResNet^[3]和 DenseNet^[4]采用

【收稿日期】2022-11-19

【基金项目】国家自然科学基金(62061023, 61941109);甘肃省科技计划资助项目(22JR5RA360);甘肃杰出青年基金(21JR7RA345)

【作者简介】邸敬,硕士,副教授,主要研究方向:数字图像处理、移动通信关键技术、无线宽带技术,E-mail: 46891771@qq.com

残差和密集连接的模式并取得巨大成功。但ResNet中每一层仅被后续一层重复利用,并且许多层是冗余的,显著降低了学习效率。为克服限制DenseNet提出密集连接重用,尽管性能出色,但导致更高的开销和较慢的收敛速度。因此,ADNet提出基于层注意的自适应密集网络,能够自适应地确定分层前特征的重用状态^[6]。最近关于注意力机制的研究十分火热,如SENet^[7]和Attention gated networks^[8-9]等,注意力机制不仅有助于更好地捕捉显著的对象、位置和通道,还可以改善任务目标的表示。脉冲耦合神经网络(Pulse Coupled Neural Network, PCNN)^[10]是Johnson等于20世纪90年代提出的一种基于猫的视觉原理构建的简化神经网络模型,被称为第三代人工神经网络的新型神经网络模型,已成功应用于图像分割、检测等领域,并取得很好的效果。由于PCNN参数设置过于复杂,MSPCNN(Modified Simplified Pulse Coupled Neural Network, Modified Simplified PCNN)^[11]对PCNN的参数进行简化调整,使MSPCNN具有更高的分割准确率和更低的计算复杂度。FC-MSPCNN(Fire-Controlled MSPCNN)^[12]在MSPCNN基础上提供一种参数设置方法来控制有效脉冲周期内的放电和被放电神经元,解决了随机性和神经元放电的不可预测性,但仍存在的问题是神经元点火之后动态阈值快速下降,神经元不激发的间隔时间较短,不具有较好的生物可解释性。

本文提出一种动态阈值可变FCMSPCNN(Dynamic Threshold Changed FCMSPCNN, DTC-FCMSPCNN)和CoA Unet相结合的分割方法。DTC-FCMSPCNN通过调整动态阈值衰减因子和添加可变阈值参数,使模型具有调节动态阈值衰减速度的能力,具有更好的生物可解释性,并在识别出分割目标后使用掩码遮盖背景部分,突出任务目标从而缩短训练时间。本文提出基于Unet^[13]的分割模型CoA Unet,在更深层的卷积块中加入快捷连接,交叉融合不同层次的特征信息,使用注意力门突出对任务目标的学习,解码器最底层加入改进的多尺度上下文提取器模块,通过串并联 3×3 卷积构成不同尺度的感受野,生成多尺度高层特征图,保留更多空间信息,将提出的模型在LiTs和DRIVE数据集上进行验证,实验结果表明,使用DTC-FCMSPCNN对数据集预处理可以显著缩短训练时间并提升分割精度,CoA Unet取得了更好的整体分割效果。

1 本文方法

实验流程如图1所示。首先通过旋转、平移、剪切变换等方式进行数据扩充,其次使用DTC-

FCMSPCNN结合膨胀腐蚀算法剔除与任务不相关区域并使用掩码遮盖处理,将目标所在矩形区域分割出来,最后使用训练好的CoA Unet进行目标分割。

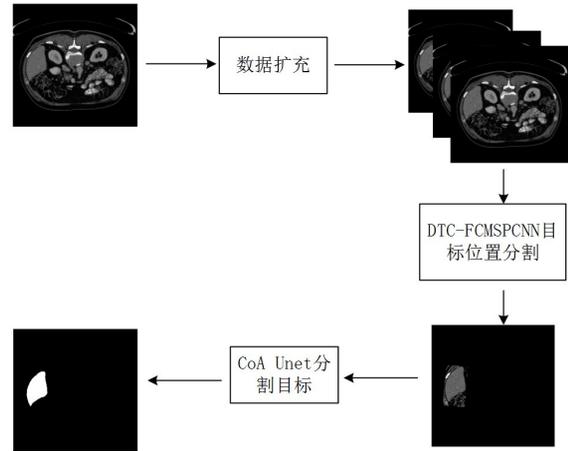


图1 分割算法流程图

Figure 1 Flowchart of segmentation algorithm

1.1 改进的DTC-FCMSPCNN

PCNN系列模型相比基于深度学习的图像分割模型具有不需要训练就可以得到较好分割效果这一显著优点。DTC-FCMSPCNN在实验基础上对FC-MSPCNN参数设置进行简单调整,使模型具有调节动态阈值衰减速度能力,具有更好的生物可解释性。标准FC-MSPCNN模型结构如图2所示。图中, α 、 β 、 V 、 R_n 为自适应参数;突触权重矩阵 W_{ijk} 表示中心神经元与周围神经元的关系,通过该参数控制邻域神经元和中心神经元之间作用力的大小; S_{ij} 表示图像的归一化大津阈值。

DTC-FCMSPCNN数学公式概括为:

$$U_{ij}[n] = e^{-\alpha} U_{ij}[n-1] + S_{ij} (1 + \beta \sum_{kl} W_{ijk} Y_{kl}[n-1]) \quad (1)$$

$$E_{ij}[n] = \alpha E_{ij}[n-1] + BQ + (V + R) Y_{ij}[n] \quad (2)$$

$$Y_{ij}[n] = \begin{cases} 1, & \text{if } U_{ij}[n] > E_{ij}[n-1] \\ 0, & \text{else} \end{cases} \quad (3)$$

通过实验研究对FC-MSPCNN中的 W_{ijk} 进行改进,使用基于高斯分布的参数形式替换原矩阵,如式(4)所示:

$$W_{ijk} = \begin{bmatrix} 0.6546 & 0.8454 & 0.6546 \\ 0.8454 & 0 & 0.8454 \\ 0.6546 & 0.8454 & 0.6546 \end{bmatrix} \quad (4)$$

在该模型中衰减因子参数 α 的值参考传统PCNN模型的设置方法,其表达式如下:

$$\alpha = S' \quad (5)$$

其中, S' 表示整幅图像的大津阈值,为简化计算将动态阈值中的幅度参数 V 设置为:

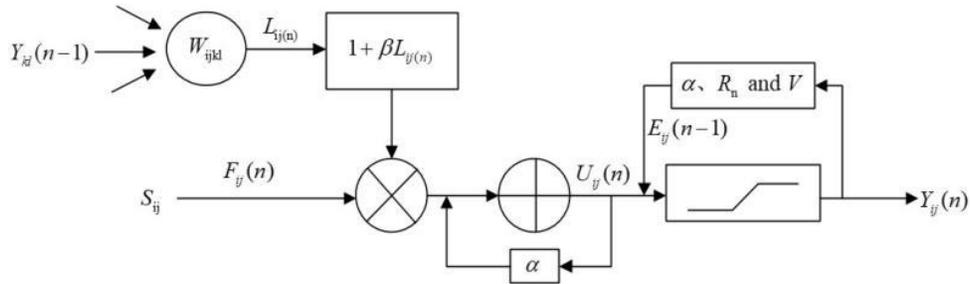


图2 FC-MSPCNN模型

Figure 2 FC-MSPCNN model

$$V = \alpha \tag{6}$$

同时,将参数R设置为:

$$R = \alpha e^{-\alpha} \tag{7}$$

因为传统PCNN模型动态阈值的衰减因子数值较大,这表明神经元在点火之后,动态阈值快速下降,神经元的不应期间隔时间较短,不具有较好的生物可解释性,为了解决这个问题,动态阈值衰减因子A的大小重新调整为:

$$A = e^{-\frac{\alpha}{s}} \tag{8}$$

此外加入了可变阈值参数B和Q,这两个参数可根据具体情况来调节神经元动态阈值的大小,这两个参数的表达式如下:

$$B = \alpha \tag{9}$$

$$Q = e^{-2\alpha} \tag{10}$$

1.2 改进的CoA Unet模型

改进模型基于医学图像分割任务中流行的Unet,这种模型能够适应医学图像训练数据少的特点,并且将高低层信息相融合以实现病灶目标的精确分割。经典Unet由编解码模块和跳跃连接组成,编码模块实现信息压缩并提取局部信息,解码模块融合上下文信息并实现信息恢复,跳跃连接将高分辨率局部信息与低分辨率全局信息相融合。为实现对病灶目标的多尺度感知和保留更多语义信息,提出CoA Unet模型,如图3所示。

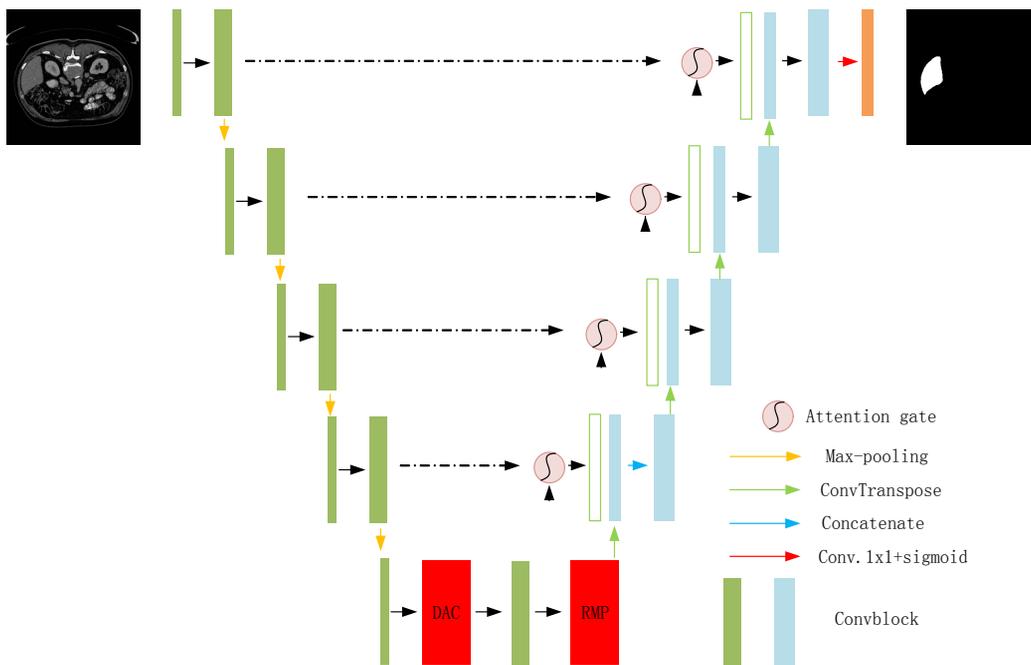


图3 CoA Unet模型

Figure 3 CoA Unet model

1.2.1 交叉连接卷积块 Unet中的Convblock由级联的双层卷积、归一化和非线性映射构成,在此基础上

将双层卷积加深为3层卷积,并且融入ResNet跨层连接的思想,在相邻卷积层之间加入1x1卷积的快捷连

接 (Shortcut Connection)。更深的卷积层数有利于抽象特征提取, 快捷链接的加入有效防止了网络加深引起的梯度爆炸或消失, 并且充分保留了图像信息, 实验发现在编码阶段加入快捷连接能够有效提升模型性能。设计细节如图 4 所示。

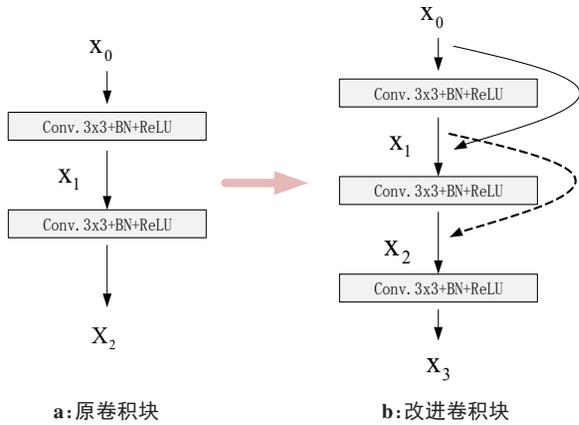


图 4 卷积块结构对比

Figure 4 Convblock comparison

1.2.2 上下文提取器模块 Unet 中连续的跨步卷积和池化运算会导致图像空间信息的丢失, 针对这一问题 Context Encoder Network (CE-Net)^[14] 提出上下文提取器模块进行解决。上下文提取器模块由并联的空洞卷积模块 DAC (Dense Atrous Convolution) 和不同尺度的池化模块 RMP (Residual Multi-Kernel Pooling) 构成, 希望提取更高层次、多尺度的信息。但是在实验过程中发现 DAC 模块的使用会导致分割结果填充不完全和评价指标 Miou 全 0 等问题, 因此对 DAC 模块进行改进。改进后的 DAC 模块不使用空洞卷积, 通过串联 3x3 卷积构成不同尺度的感受野进行代替, 并且前置 1x1 卷积有效降低运算量, 结构细节如图 5 所示。这种结构的优点在 Inception 结构中早有证明^[15]。RMP 模块用 4 个不同大小的感受野编码全局上下文信息: 2x2、3x3、5x5 和 6x6, 然后将特征输入到 1x1 卷积中, 降低特征映射的维数。最后, 将上采样特征与原始特征串联。

1.2.3 注意力门 为提升对特定目标的学习效果, 突

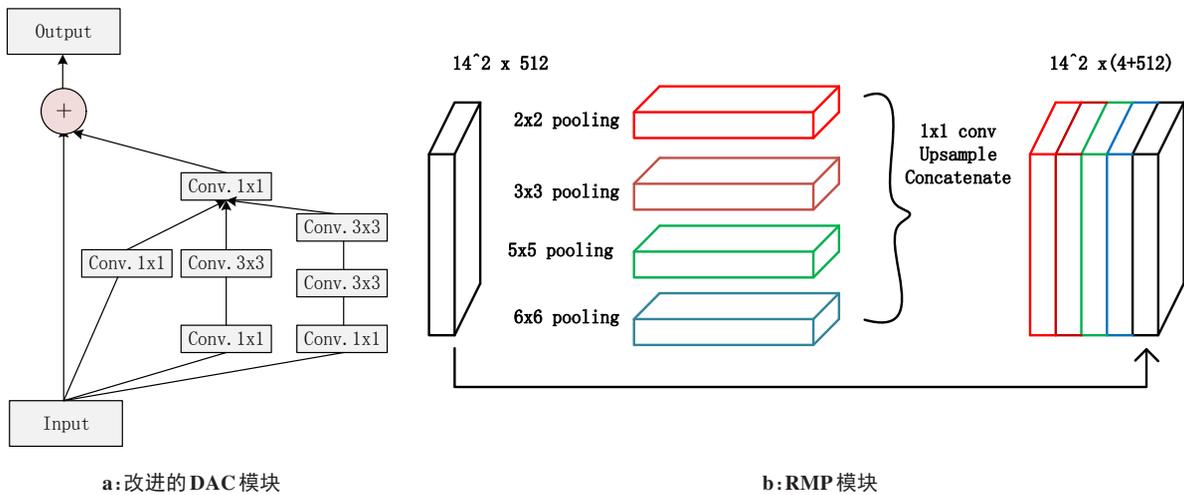


图 5 上下文提取器模块

Figure 5 Context extractor module

出目标显著特征, 隐式学习非任务目标区域, 使用 Attention Gate Unet^[8-9] 提出的 Attention Gate (AG) 模块代替 Unet 架构中的跳跃连接, 该架构的计算费用相对较低, 同时也可以提升模型的精度与准确性。AG 模块结构如图 6 所示。图中, $\mathbf{x}^l = \{\mathbf{x}_i^l\}_{i=1}^n$ 设为所选层 $l \in \{1, \dots, L\}$ 的激活映射, 其中每个层 \mathbf{x}_i^l 表示长度的像素级特征向量 F_l (即通道数), 对于每个层 \mathbf{x}_i^l , AG 计算系数 $\alpha^l = \{\alpha_i^l\}_{i=1}^n$, 其中 $\alpha_i^l \in [0, 1]$ 。AG 的输出为 $\hat{\mathbf{x}}^l = \{\alpha_i^l \mathbf{x}_i^l\}_{i=1}^n$, 其中每个特征向量按相应的注意系数进行缩放。

注意系数 α_i^l 的计算方法如下: 在标准的 CNN 架构中, 为了捕获一个足够大的接受域, 从而获得语义上下文信息, 特征图逐渐被降采样。粗空间网格级特征识别目标对象的位置, 并在全局尺度上对它们的关系进行建模。假设 $\mathbf{g} \in R^{f_g}$ 是这样的全局特征向量, 并为 AG 提供信息, 以消除 \mathbf{x}_i^l 中与任务无关的特征内容的歧义。其想法是共同考虑每个 \mathbf{x}_i^l 和 \mathbf{g} , 并共同关注每个尺度 l 与被最小化的目标最相关的特征。数学公式如下:

$$q_{att,i}^l = \psi^T (\sigma_1 (W_x^T \mathbf{x}_i^l + W_g^T \mathbf{g} + \mathbf{b}_{xg})) + b_\psi \quad (11)$$

$$\alpha^l = \sigma_2 (q_{att}^l (\mathbf{x}^l, \mathbf{g}; \Theta_{att})) \quad (12)$$

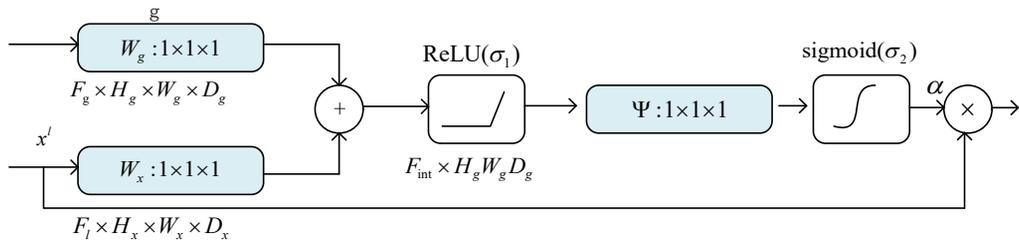


图6 AG模块

Figure 6 AG module

其中, Θ_{att} 包含线性变换 $W_x \in R^{F_l \times F_{int}}$, $W_g \in R^{F_g \times F_{int}}$, $\Psi \in R^{F_{int} \times 1}$, 偏置项 $b_\psi \in R$, $b_{xg} \in R^{F_{int}}$ 。线性转换的计算使用 $1 \times 1 \times 1$ 卷积。AG块用于突出前景并减少背景影响,其运算过程主要分为3个步骤:(a)输入 g, x' 经过通道方向的 $1 \times 1 \times 1$ 卷积进行线性变换相加后通过 ReLU层(σ_1);(b)采用 $1 \times 1 \times 1$ 卷积作为附加线性变换,并通过 Sigmoid 激活函数(σ_2)生成注意力图;(c)注意力图与来自输入的 x' 对应元素相乘生成最终的注意力图。

2 结果与分析

2.1 数据集和评估指标

实验使用公共数据集 LiTs 的 420 张图片子集,训练

集和测试集分别为 400 张和 20 张。将训练集和测试集图片数据扩充到 800 张和 40 张,然后使用 DTC-FCMSPCNN 模型将测试集目标矩形区域分割出来,其余背景部分使用掩码进行遮盖处理。为增加模型可信度,进一步从 DRIVE 视网膜血管分割数据集获取训练和测试图片对分割模型进行测试,其中训练和测试各 20 张图片。由于适用性问题,DRIVE 数据集未使用 DTC-FCMSPCNN 预处理。数据集示例如图 7 所示。实验采用常见的评估指标:交并比 IoU、Dice 系数(Dice)、Hausdorff 距离(HD)、准确率(Scc)、灵敏度(Sens)和特异度(Spec),并且由于测试集含有多张图片,所以各评估指标均取平均值,符号表示为 Miou、Aver_Dice、Aver_HD、Aver_Scc、Aver_Sens 和 Aver_Spec。

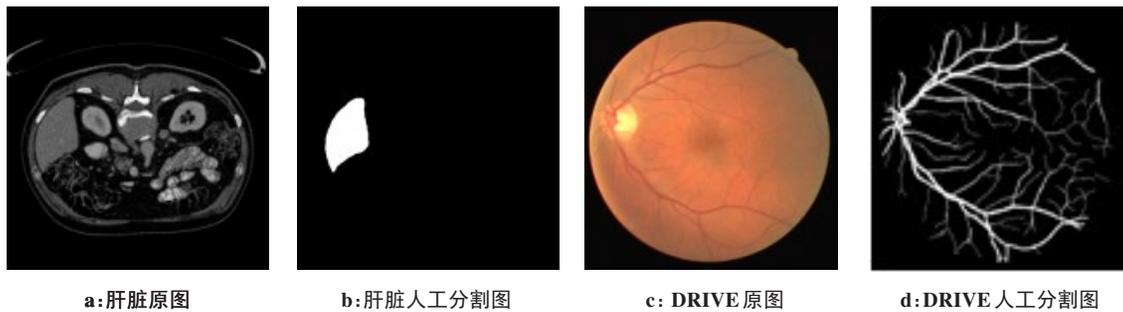


图7 数据集图像示例

Figure 7 Data set image examples

2.2 实施细节

实验框架使用 Pytorch 库,并在 NVIDIA RTX3080 GPU 进行。在 DTC-FCMSPCNN 模型目标矩形区域分割阶段,由于肝脏图片包含多种人体器官,而且灰度值接近,在输入前通过式(13)对图片灰度值进行变换,配合对模型输出使用内核尺寸为 3 的椭圆形膨胀腐蚀运算可以消除大部分不相关区域,准确分割目标区域。式中,90 这一数值的设置通过多组实验得出,若做出更改需要配合膨胀腐蚀算法一同更改。

$$y = \ln(90 + x) \tag{13}$$

DRIVE 数据集训练模型设置 batch size 大小为

1, Adam 优化器使用默认参数,损失函数对比了最近流行的 3 种,分别为二进制交叉熵损失函数(Binary Cross-Entropy, BCE)、Dice Loss^[16] 和 Focal Tversky Loss^[17]。Focal Tversky Loss 引入 3 个额外参数 α, β, γ , 原文建议参数设置为 0.7、0.3 和 4/3,但原文应用场景与本文不同,因此通过多组实验对参数细节进行研究,Epoch 次数为 32,如表 1 所示。在确定 Focal Tversky Loss 参数 $\alpha = 0.6, \beta = 0.4, \gamma = 3$ 后,对 3 种损失函数在各自数据集上的表现进行对比,如表 2 所示。最终确定在 LiTs 数据集肝脏分割实验中使用二进制交叉熵损失函数 BCE,在 DRIVE 数据集视网膜血管分割实验中使用 Focal Tversky Loss。

表1 Focal Tversky Loss 参数设置
Table 1 Focal Tversky Loss parameter setting

α, β, γ	Miou	Aver_HD	Aver_Dice	Aver_Acc	Aver_Sens	Aver_Spec
0.1,0.1,1	0.5185	9.5847	0.6635	0.9325	0.7507	0.9512
0.1,0.1,2	0.5293	9.4693	0.6741	0.9338	0.7669	0.9510
0.1,0.1,3	0.5360	9.0079	0.6732	0.9444	0.6511	0.9737
0.7,0.3,4/3	0.5324	9.5144	0.6768	0.9332	0.7788	0.9490
0.7,0.3,4	0.5310	9.4838	0.6757	0.9325	0.7855	0.9476
0.7,0.3,3	0.5389	9.4365	0.6821	0.9475	0.5836	0.9836
0.1,0.3,3	0.5357	8.7858	0.6722	0.9475	0.6214	0.9801
0.3,0.7,3	0.5565	8.7662	0.7043	0.9537	0.6759	0.9820
0.6,0.4,3	0.5892	9.2559	0.7409	0.9456	0.8871	0.9516
0.5,0.5,3	0.5363	9.2473	0.6790	0.9377	0.7488	0.9570

表2 损失函数表现对比
Table 2 Performance comparison among loss functions

损失函数	肝脏分割实验			视网膜血管分割实验		
	Miou	Aver_HD	Aver_Dice	Miou	Aver_HD	Aver_Dice
BCE	0.8905	6.3699	0.9477	0.5425	9.4081	0.6852
DiceLoss	0.8313	6.9137	0.9300	0.5389	9.4365	0.6821
Focal Tversky Loss	0.8793	6.7600	0.9356	0.5892	9.2559	0.7409

2.3 对比实验

实验将是否使用DTC-FCMSPCNN对肝脏图像预处理作为变量进行控制,并选择5种先进的分割算法作为对比实验。其中,Unet++^[18]提出编码器和解码器子网络通过一系列嵌套的密集跳过路径连接减少特征映射之间的语义差距,该方法本质是一种深度监督网络;Unet3+^[19]与Unet++相比,重新设计了编码器和解码器之间的相互连接以及解码器之间的内部连接,该方法能够从全尺度捕获细粒度细节和粗粒度语义;Recurrent residual U-Net(R2U-Net)^[20]为卷积块添加了递归结构和残差连接使网络有更好的分割表现;Attention gated networks(Attention-Unet)^[8-9]提出使用AG模块代替跳跃连接的设计,该方法能突出对任务有用部分的学习,并隐式学习非目标区域;CE-Net^[14]使用ResNet-34作为编码模块,并提出上下文提取器模块,该方法从多尺度捕获图像高层语义信息,保留更多的空间信息用于图像分割。

表3给出了肝脏分割的定量结果。由表3可知,除Attention-Unet外其他模型使用DTC-FCMSPCNN预先分割肝脏目标矩形区域都能实现更好的分割效果,其中R2U-Net的提升最为明显,这得益于DTC-FCMSPCNN对于目标区域的准确定位,配合掩码遮盖非目标区域使分割模型能够快速学习到有用特

征,加快模型收敛速度。而Attention-Unet表现更差与模型结构有关,非任务区域的学习提供其所需的空间和纹理信息,对其模型训练同样重要。

图8展示了6种模型在同一张测试图像上的分割效果,横纵坐标单位均为像素,表示图像大小。这张图像的分割难度较大,肝脏周围器官灰度值接近,不易区分,而且图像构成复杂,可以看到所有模型的分割都存在瑕疵,但本文提出模型目标分割相对完整,分割效果优于其他模型。

表4为视网膜血管分割的定量结果。由表4可知,本文算法整体表现优于其他算法。算法充分结合Attention-Unet和CE-Net设计优点,在特征信息利用和多尺度感知方面更加优秀,在4个指标方面优于其他架构。虽然Unet3+和R2U-Net在Aver_HD或Aver_Spec方面处于领先,但是更高的特异性代表了更好的背景分割(真阴性),这不如更好的视网膜血管分割(真阳性)来得重要^[21]。本文4个领先指标除Aver_Acc提升较少,其他均有3%~10%的提升。综合表4和表5可以看到,提出算法在部分指标方面略有不足,但整体表现优于其他对比算法。

2.4 烧蚀实验

在肝脏数据集上进行烧蚀实验,以验证提出的改进对模型性能的贡献,如图9所示。实验数据为未

表3 肝脏分割评估指标比较

Table 3 Comparison of evaluation indexes of liver segmentation

算法	DTC-FCMSPCNN 预处理	Miou	Aver_HD	Aver_Dice	时间
Unet++	是	0.8624	6.9524	0.9253	272 min 35 s
	否	0.8479	7.0230	0.9170	291 min 1 s
Unet3+	是	0.8706	6.8676	0.9304	518 min 35 s
	否	0.8653	7.1045	0.9273	575 min 50 s
R2U-Net	是	0.8685	6.6759	0.9291	380 min 10 s
	否	0.5369	10.3186	0.6813	477 min 2 s
Attention-Unet	是	0.7618	9.0318	0.8623	289 min 44 s
	否	0.8113	7.1177	0.8957	302 min 3 s
CE-Net	是	0.8430	5.6288	0.9417	237 min 26 s
	否	0.7915	6.1985	0.9325	250 min 12 s
本文方法	是	0.8905	6.3699	0.9477	246 min 5 s
	否	0.8903	6.6616	0.9414	265 min 35 s

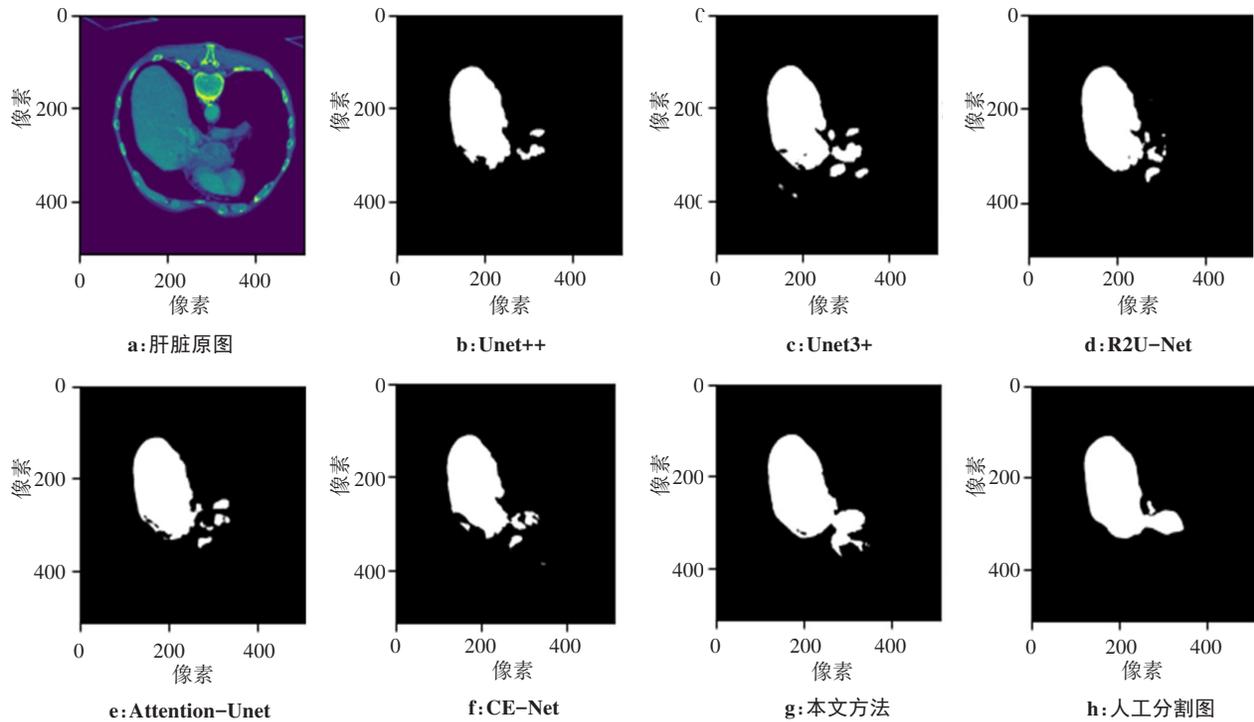


图8 不同模型分割效果图

Figure 8 Segmentation results obtained by different models

表4 视网膜血管分割评估指标比较

Table 4 Comparison of evaluation indexes of retinal vascular segmentation

算法	Miou	Aver_HD	Aver_Dice	Aver_Acc	Aver_Sens	Aver_Spec
Unet++	0.5360	9.0079	0.6732	0.9444	0.6511	0.9737
Unet3+	0.4926	9.0861	0.6345	0.9444	0.5657	0.9812
R2U-Net	0.5494	9.0063	0.6905	0.9409	0.7372	0.9615
Attention-Unet	0.5374	9.3547	0.6807	0.9382	0.7357	0.9587
CE-Net	0.5207	9.5485	0.6683	0.9315	0.7716	0.9478
本文方法	0.5892	9.2559	0.7409	0.9456	0.8871	0.9516

使用DTC-FCMSPCNN预处理得到。类别1表示使用原卷积块,类别2表示跳跃连接替换AG模块,类别3表示原DAC模块替换改进DAC模块,类别4表示删除DAC和RMP模块,类别5表示使用本文所有改进结构,评估指标使用Aver_HD、Miou和Aver_Dice,符号含义与前面相同。通过交叉验证可以看到删除任何一部分的改进都会导致分割表现变差,这是合理的,反映了改进对模型是有益的,甚至是互补的。



图9 烧蚀实验分析

Figure 9 Ablation test

3 结论

本文针对肝脏和视网膜血管分割任务目标较小且复杂多变问题,提出DTC-FCMSPCNN与CoA Unet相结合的分割方法。DTC-FCMSPCNN对肝脏图像预处理以减少训练时间并提高分割精度,基于Unet的改进模型CoA Unet,在更深层卷积块中加入快捷连接促进特征交叉融合,使用注意力门替换跳跃连接突出目标特征,并在编解码器最底层加入改进多尺度上下文提取器模块以保留更多图像信息。模型在两个重要的医学分割任务数据集上表现出很好的分割效果,提出的DTC-FCMSPCNN预处理并结合掩码遮盖非任务区域的方法也可扩展到其他分割或是检测问题上。但目前所做研究未扩展到其他不同类型的公开数据集上,由于医学图像分割任务不同,后续仍需对分割方案进行完善,提高模型的泛化能力和鲁棒性。

【参考文献】

[1] Alex K, Ilya S, Geoffrey EH. ImageNet classification with deep convolutional neural networks[J]. Commun ACM, 2017, 60(6): 84-90.
 [2] Simonyan K, Zisserman A. Very deep convolutional networks for large-

scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.
 [3] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and Pattern Recognition. Las Vegas, 2016: 770-778.
 [4] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 4700-4708.
 [5] Zagoruyko S, Komodakis N. Wide residual networks[J]. arXiv preprint arXiv: 1605.07146, 2016.
 [6] Wang M, Cai H, Huang X, et al. ADNet: adaptively dense convolutional neural networks[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Snowmass, 2020: 1001-1010.
 [7] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[J]. IEEE Trans Pattern Anal, 2020, 42(8): 2011-2023.
 [8] Schlemper J, Oktay O, Schaap M, et al. Attention gated networks: learning to leverage salient regions in medical images[J]. Med Image Anal, 2019, 53: 197-207.
 [9] Zhang S, Fu H, Yan Y, et al. Attention guided network for retinal image segmentation [C]//International conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2019: 797-805.
 [10] Ranganath HS, Kuntimad G, Johnson JL. Pulse coupled neural networks for image processing [C]//Proceedings IEEE Southeastcon'95 Visualize the Future. Raleigh, 1995: 37-43.
 [11] Lian J, Yang Z, Sun W, et al. An image segmentation method of a modified SPCNN based on human visual system in medical images [J]. Neurocomputing, 2019, 333: 292-306.
 [12] Lian J, Yang Z, Sun W, et al. A fire-controlled MSPCNN and its applications for image processing[J]. Neurocomputing, 2021, 422: 150-164.
 [13] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.
 [14] Gu Z, Cheng J, Fu H, et al. CE-net: context encoder network for 2d medical image segmentation[J]. IEEE Trans Med Imaging, 2019, 38(10): 2281-2292.
 [15] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Arizona, 2017.
 [16] Milletari F, Navab N, Ahmadi SA. V-net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV). Stanford, 2016: 565-571.
 [17] Abraham N, Khan NM. A novel focal tversky loss function with improved attention U-net for lesion segmentation[C]//2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). Venice, 2019: 683-687.
 [18] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, et al. Unet++: a nested U-net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3-11.
 [19] Huang H, Lin L, Tong R, 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 (ICASSP). Barcelona, 2020: 1055-1059.
 [20] Alom MZ, Yakopcic C, Hasan M, et al. Recurrent residual U-Net for medical image segmentation[J]. J Med Imaging, 2019, 6(1): 014006.
 [21] Kamran SA, Hossain KF, Tavakkoli A, et al. RV-GAN: segmenting retinal vascular structure in fundus photographs using a novel multi-scale generative adversarial network[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2021: 34-44.

(编辑:陈丽霞)