

基于双任务一致性的半监督深度学习医学图像分割方法^{*}

罗 港¹, 吕 佳^{1,2}

(1. 重庆师范大学 计算机与信息科学学院; 2. 重庆国家应用数学中心, 重庆 401331)

摘要:【目的】医学图像分割是医学图像分析中的一个重要内容。现有的大部分图像分割算法都是基于监督学习,而实际应用中医学图像标签难以获取,大量标注需依赖领域专家,费时费力。因此,提出一种双任务一致性的半监督医学图像分割模型。【方法】该模型采用一个编码器,两个解码器的网络结构,其中编码器和一个解码器实现图像分割,与另一个解码器实现图像重建。无标签数据通过一致性分割与重建任务得到的两个不同图像背景计算损失并优化网络参数。同时在网络的编码器部分加入注意力模块以更好地获取分割区域的空间位置信息,并使用锐化操作增加无标签数据重建输出的置信度。【结果】在肝脏和细胞数据集上 IOU 分别为 0.953 6 和 0.821 0,相较于 U-Net 提高了 1.5% 和 4.82%;在眼底血管数据集上,SP 值为 0.983 0,与第二名相比提升了 0.18%。【结论】本文模型与有监督方法和半监督方法相比,在医学图像分割的有效性和泛化性上有一定的性能提升,能有效解决数据集数量少、小病灶分割难度大的问题。

关键词:医学图像分割;半监督学习;双任务一致性;坐标注意力;锐化函数

中图分类号:TP391

文献标志码:A

文章编号:1672-6693(2022)06-0099-11

随着医疗科技的快速发展,医学影像成为医生了解和分析病情的重要参考信息之一。医学图像分割是将医学图像中具有特定含义的区域分割出来,并提取相关特征,为临床诊疗和病理学研究提供可靠的依据,辅助医生做出更为准确的诊断^[1]。

近几年来,医学人工智能(Medical artificial intelligence)受到了国家层面的高度重视与战略扶持^[2]。基于深度学习的图像分割算法展现出了极大的优势与潜力,学者们也因此将深度学习应用于医学图像分割并进行了深入的研究。Ronnebrge 等人^[3]提出一种端到端的 U 型网络结构(U-Net),采用跳跃连接结合了浅层和深层的空间信息,在一定程度上提高了分割精度。Gu 等人^[4]提出上下文编码器网络(Context encoder network, CE-Net)以捕获更多的高级语义信息和保存空间信息,并大幅减少网络的参数量。Xia 等人^[5]提出多尺度扩张卷积网络(Multi-scale dilated convolution network, MD-Net)采用不同大小的扩张卷积形成特征金字塔来提取图像中的语义信息,并使用双线性插值和重复卷积减小网络的参数数量,取得了更好的效果。之后 Ding 等人^[6]同时使用多尺度策略和注意力机制,抑制不相关的部分或背景以突显深层特征的空间信息,进一步提取图像特征的细节信息。

然而,用于医学图像分割的神经网络往往需要大量的原始图像以及相对应的标签图像。而标签图像一般需由专家手动标注生成,对专家的经验要求非常高,因此这是一个极其耗时耗力的过程^[7]。针对上述问题,半监督深度学习的医学图像分割方法应运而生。在半监督医学图像分割中,只有小部分训练图像被认定为具有完整的像素级标签,还有大量的无标签图像,可以用来提高模型的准确性和泛化能力。常见的半监督深度学习医学图像分割方法包括:迁移学习、数据增强、先验知识学习、自训练方法、协同训练方法和一致性正则化等^[8]。

最近,研究者们提出了多种基于半监督学习的深度学习医学图像分割方法。Peng 等人^[9]将协同训练的思想应用于医学图像的半监督分割。他们在有标签训练数据的不同子集上训练多个网络模型,使用一组公共的无标签图像来相互交换信息,并使用有标签数据集和无标签数据集生成的对抗样本来加强模型之间的多样性。Hang 等人^[10]提出一种结构感知熵正则化 Mean Teacher 模型,将熵最小化原理引入到学生网络,对无标签图像产生高

^{*} 收稿日期:2022-10-05 网络出版时间:2022-12-08 16:39:17

资助项目:重庆市教育委员会重点项目(KJZD-K202200511);重庆市科技局技术预见与制度创新项目(2022TFII-OFX0265);重庆师范大学研究生科研创新项目(No. YKC21043)

第一作者简介:罗港,男,研究方向为医学图像处理,E-mail:luogang77qq@163.com;通信作者:吕佳,女,教授,博士,E-mail:lvjia@cqnu.edu.cn

网络出版地址:https://kns.cnki.net/kcms/detail//50.1165.N.20221207.1730.014.html

置信的预测,并设计一个局部结构一致性损失,以鼓励两个网络预测的一致性。Fang 等人^[11]提出差分最小化网络(Difference minimization network, DM-Net),使用一个编码器,两个解码器分支的网络结构,并使两个解码器生产的软掩膜之间的差异最小化。同时采用对抗学习训练策略学习一个鉴别器,以鼓励模型生成更精确的分割预测。Wu 等人^[12]提出一种相互一致性网络,由一个编码器和两个稍微不同的解码器组成。两个解码器的预测误差转化为无监督损失,以促进两个解码器的一致性,使得模型更好地捕获无标签图像的广义特征。Chen 等人^[13]提出一种基于多任务组合的半监督学习(Multi-task attention-based semi-supervised learning, MASSL),将一个自编码器和 U-Net 网络相结合。这里自编码器部分不是重建原始输入,重建的是图像的前景和背景,一旦前景和背景重建效果不好,就会导致分割网络效果变差。另外从第一轮开始就交替训练两个网络,使得在训练前几轮由注意力机制创建的重建标签不准确,自编码器训练不稳定导致分割结果变差。综上所述,如何合理使用无标签数据,避免无标签数据被当作噪声,提高模型的分割性能和泛化性是半监督深度学习医学图像分割的关键。

因此,本文提出一种基于双任务一致性的半监督深度学习医学图像分割方法,首先使用一致性策略以更好利用无标签数据的双任务输出,促使模型学习更多分割区域的特征信息,提升分割精度。其次,提出一种交替训练策略,采用两个不同的损失交替优化同一个网络,进一步提升了分割性能。最后引入坐标注意力模块和锐化操作,使网络更容易获取分割区域的位置信息并提升无标签数据边缘轮廓细节的置信度。

1 本文方法

1.1 双任务一致性模型结构

本文的半监督学习方法如图 1 所示。与原始 U-Net 结构相比,本文采用单编码器、双解码器的网络结构,两个解码器共享同一个编码器实现双任务输出。其中,编码器和解码器(S)组成分割网络实现分割任务,分别输出有标签和无标签数据的分割预测;编码器和解码器(R)组成重建网络实现重建任务,输出无标签数据的重建背景和前景。并对无标签数据的分割预测使用 Hadamard 积连接两个任务以实现一致性,使得网络可以获取无标签数据中的特征信息。本文在编码器的第 2 层与第 3 层之间,加入坐标注意力(Coordinate attention, CA)模块^[14],使得网络提取更多待分割区域的空间位置信息,并在重建网络的输出后使用锐化函数减少无标签数据的熵,提高边缘轮廓的置信度。

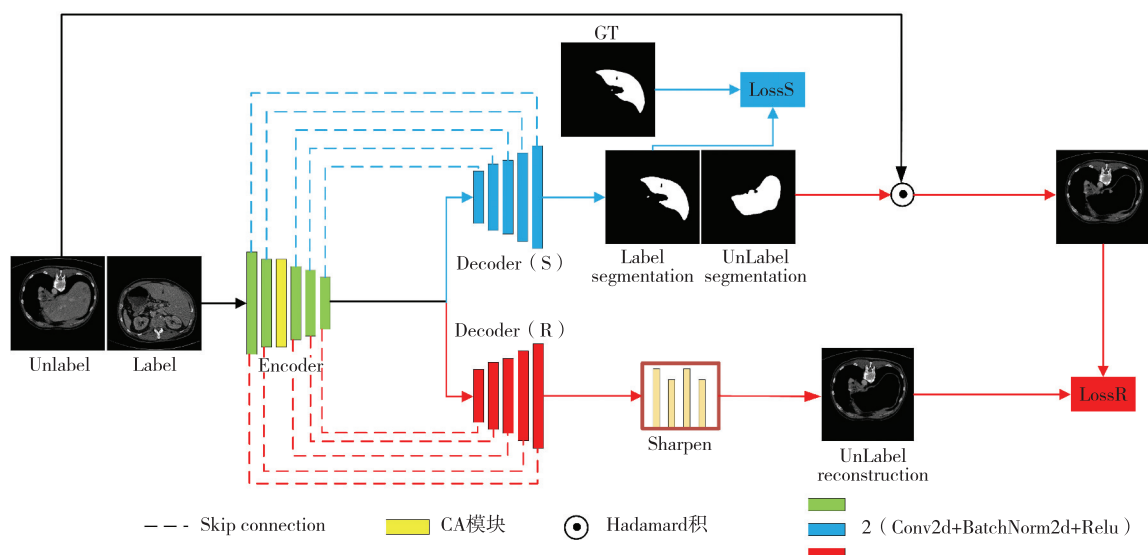


图 1 网络模型结构

Fig. 1 Network model structure

本文的双任务一致性模型在进行每一轮训练时,会交替输入有标签数据和无标签数据。当网络的输入为有标签数据时,有标签图像经分割网络会得到图像的分割预测,并与真实标签计算得到有监督损失,且只更新编码器和解码器(S)两个部分的梯度信息。当网络输入为无标签图像时,无标签图像会同时经过编码器和两个解码器得到分割预测以及图像重建的前景与背景。分割预测与原始图像做 Hadamard 积得到图像的背景,解码器

(R)输出的背景经过锐化操作后,通过一致性分割预测得到的背景和重建的背景得到无监督损失,且只更新编码器和解码器(R)两个部分的梯度信息。

1.2 坐标注意力

有实验表明,通道注意力会给模型带来比较显著的性能提升,但是通道注意力通常会忽略注意力图的位置信息^[15]。因此,本文在网络中添加了 CA 模块(图 2)。

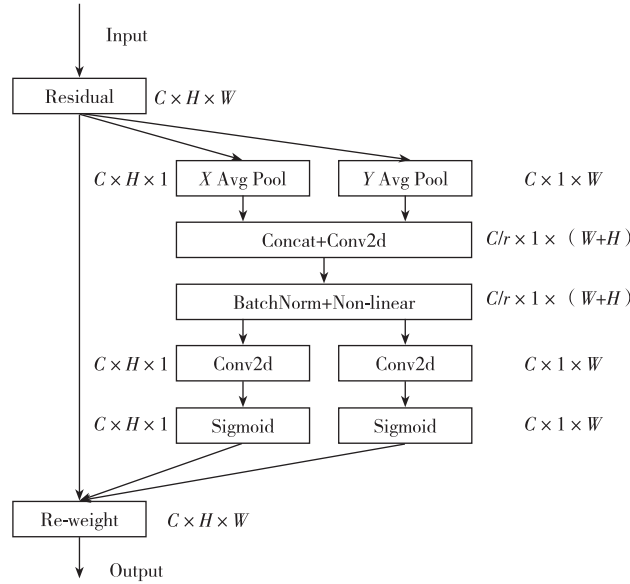


图 2 CA 模块结构

Fig. 2 CA block structure

图 2 展示了 CA 模块的网络结构。CA 模块首先分别沿水平和垂直方向对每个通道进行编码,使得注意力模块可以沿着一个方向捕捉长期依赖,并沿着另一个空间方向保留精确的位置信息,这样有助于网络更准确地定位感兴趣的对象。其次连接生成的 2 个特征图,并在 1×1 卷积函数中进行变换。最后再沿着水平和垂直方向将空间维度拆分成两个独立的张量。两个注意力特征图中的每个元素都反映了相应行和列中是否存在感兴趣的对象,这种编码过程允许协调注意力,能更准确地定位感兴趣对象的确切位置。

1.3 锐化函数

在半监督学习中,一个常见的假设是分类器的决策边界不应通过边缘数据分布的高密度区域,实现方法是要求分类器对未标记数据输出低熵预测。由于本文方法中的重建网络是无监督学习方式,通常重建的背景和前景边缘轮廓细节不清晰,置信度低且包含较高的信息熵。为了使重建的背景和前景具有低熵和高置信度,本文通过对未标记数据的目标分布使用锐化函数^[11]来隐式实现熵最小化,降低未标记数据的预测熵,锐化函数如下:

$$S(\hat{y}_{h,w,c}, T) = \frac{(\hat{y}_{h,w,c})^{1/T}}{\sum_{i=1}^K (\hat{y}_{h,w,c})^{1/T}}, \forall h \in [1, H], w \in [1, W], T \in (0, 1). \quad (1)$$

其中: $\hat{y}$ 表示重建网络输出, K 表示网络输出的通道数, T 是一个超参数。 H 、 W 分别表示图片的高度和宽度。

1.4 损失函数

在图像分割问题中通常采用 Dice 损失作为损失函数,但是 Dice 损失对高度不平衡的分割任务效果差^[16]。分割中一旦有部分像素预测错误,就会导致 Dice 损失大幅度变动,导致训练不稳定。因此针对这一问题,本文方法中的损失函数采取加权 Dice 和 BCE 构成的联合损失见下式:

$$L_S = \lambda L_{\text{Dice}} + (1 - \lambda) L_{\text{BCE}}, \quad (2)$$

其中: L_S 是分割网络的损失, λ 是超参数, L_{Dice} 和 L_{BCE} 分别定义为:

$$L_{\text{Dice}} = 1 - 2 \times \frac{|\hat{y} \cap y|}{|\hat{y}| + |y|}, \quad (3)$$

$$L_{\text{BCE}} = -\frac{1}{N} \sum_i^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]。 \quad (4)$$

其中: N 是总样本数, y 表示真实标签, $\hat{y}$ 表示分割网络的预测值。

在重建网络部分,通常采用均方误差(Mean square error, MSE)作为损失函数,定义为:

$$L_R = [x \odot (1 - \hat{y}) - \tilde{y}]^2 \quad (5)$$

其中: L_R 是重建网络的损失, x 是输入图像, $\tilde{y}$ 和 $\hat{y}$ 是重建的背景和分割路径的预测, $\odot$ 是 Hadamard 积。

1.5 交替训练策略

文献[13]采用交替训练策略对网络的分割部分和重建部分交替训练优化网络。当每轮网络训练时,会从大量无标签数据中随机抽取部分无标签数据输入网络。然而在网络前几轮,若分割网络性能不好,则由 Hadamard 积得到的合成分割标签不准确。这就使得重建网络的权重向错误的方向优化,最后导致实验结果变差。

为了解决上述问题,本文提出了一种新的学习策略。为了使得重建网络向好的方向优化,先给定一个超参数 k ,在训练的前 k 轮,只对分割网络进行训练,这样就使得重建网络的标签更具有真实性, k 轮之后收敛到交替算法。在进行交替训练时,当输入的是有标签数据时,网络使用 L_S 对分割网络部分进行权值优化,而解码器(R)的权值保持不变。当输入的是无标签数据时,网络使用 L_R 对重建网络部分进行权值优化,而解码器(S)的权值保持不变。定义为:

$$L(x) = \begin{cases} L_S(x), & x = x_L \\ L_R(x), & x = x_U \end{cases}。 \quad (6)$$

其中: x_L 和 x_U 分别表示有标签数据和无标签数据。

2 实验结果与分析

2.1 实验数据集

为了验证本文模型在医学图像分割中的有效性,本文选择肝脏数据集 LiTS^[17]、眼底血管数据集 DRIVE^[18]和细胞数据集 DSB2018^[19]对分割模型性能进行测试。其中肝脏数据集包含 400 张尺寸为 512×512 的二维样本图像,细胞数据集包含 670 张尺寸为 256×256 的二维样本图像,眼部数据集包含 40 张尺寸为 565×584 的视网膜 RGB 图像,由官方划分为 20 幅训练集图像和 20 幅测试集图像,数据包含 2 名专家手工分割结果。本文使用旋转 180° 和水平镜像操作对肝脏、细胞以及眼底血管训练图像进行数据增强,将图像数据量扩充到原始的 4 倍。

实验时,将有标签数据和无标签数据的比例设置为 2 : 8。在肝脏数据集中随机抽取 200 张有标签图像,其中 10% 用于测试,剩余 90% 用于训练,并做十折交叉验证,另外随机选择 720 张图像作为无标签数据。在细胞数据集中随机抽取 400 张有标签图像,其中 10% 用于测试,剩余 90% 用于训练,并作十折交叉验证,另外随机选择 1 440 张图像作为无标签数据。眼底血管数据集将原本的 20 张训练数据作为有标签数据,20 张测试数据作为测试集,其余图像作为无标签数据。然后将 20 张有标签数据随机裁剪为 2 000 张尺寸为 48×48 的图像,60 张无标签数据随机裁剪为 8 000 张尺寸为 48×48 的图像。数据集的具体信息如表 1 所示。

表 1 实验数据集

Tab. 1 Experimental dataset

数据集	原始数据	增强方式	增强后数据	训练有标签数据	训练无标签数据	测试数据
肝脏	400	旋转、水平镜像	1 600	180	720	20
细胞	670	旋转、水平镜像	2 680	360	1 440	40
眼底血管	20/20	旋转、水平镜像	80/20	2 000	8 000	20

2.2 评价指标

本文使用了 Dice 相似系数(Dice similarity coefficient, DSC)、重叠度(Intersection over union, IOU)、准确率(Accuracy, ACC)、敏感性(Sensitivity, SE)、F1-score、特异性(Specificity, SP)以及计算接受者工作特性曲线(Receiver operating characteristic curve, ROC)下的面积(Area under curve, AUC)作为分割结果的客观评价指标。 x_{DSC} , x_{IOU} 计算公式为:

$$x_{\text{DSC}}=2\times\frac{|\hat{y}\cap y|}{|\hat{y}|+|y|},$$

(7)

$$x_{\text{IOU}}=\frac{|\hat{y}\cap y|}{|\hat{y}\cup y|}。$$

(8)

其他评价指标见表 2。

表 2 评价指标

Tab.2 Evaluation index

ACC	SE	F1	SP
$\frac{x_{\text{TP}}+x_{\text{TN}}}{x_{\text{TP}}+x_{\text{FP}}+x_{\text{FN}}+x_{\text{TN}}}$	$\frac{x_{\text{TP}}}{x_{\text{TP}}+x_{\text{FN}}}$	$\frac{2x_{\text{TP}}}{2x_{\text{TP}}+x_{\text{FP}}+x_{\text{FN}}}$	$\frac{x_{\text{TN}}}{x_{\text{FP}}+x_{\text{TN}}}$

其中: $\hat{y}$ 和 y 分别表示分割预测和专家标签。真阳性(True positive, x_{TP})表示正确分类前景的像素个数,假阴性(False negative, x_{FN})表示错误分类前景的像素个数,假阳性(False positive, x_{FP})表示错误分类背景的像素个数,真阴性(True negative, x_{TN})表示正确分类背景的像素个数。

2.3 实验环境及参数设置

模型的训练在配备了 NVIDIA RTX3060Ti GPU 和 AMD 5600X 处理器的个人计算机上进行,使用 Pytorch 框架实现。批数量(batchsize)设置为 4,用带有默认参数的 Adam 优化器来训练模型,并使用 Cosine warmup 学习率策略辅助学习,初始学习率设置为 0.001,训练迭代次数为 100 轮, λ 根据不同数据集设置, T 设置为 0.5, k 设置为 10,有标签数据集与无标签数据集的比例设置为 2:8。

2.4 实验结果及分析

2.4.1 肝脏分割 在肝脏数据集中,由于分割目标较大,将 λ 设置为 1,此时(2)式中的 BCE 构成的损失项权值为 0。本文将肝脏图像分辨率缩小为 256×256 大小,使用 200 张有标签数据和 720 张无标签数据,其中 10%的有标签数据作为测试,其余图片用于训练,并做十折交叉验证来评估本文方法的分割性能。表 3 展示了不同算法在肝脏数据集上的 DSC、IOU、F1 和 SE。

表 3 肝脏数据集上不同算法对比

Tab.3 Comparison of different algorithms on liver dataset

算法	评价指标			
	DSC	IOU	F1	SE
U-Net	0.967 6	0.937 9	0.961 9	0.979 7
ATU-Net	0.973 8	0.949 6	0.966 6	0.988 1
U-Net++	0.974 5	0.950 9	0.968 2	0.984 5
CE-Net	0.975 5	0.952 5	0.969 1	0.985 9
ResCEAttUnet	0.975 9	0.953 2	0.968 8	0.986 4
Mean Teacher	0.975 8	0.954 8	0.968 6	0.986 6
MASSL	0.975 7	0.953 2	0.968 1	0.986 1
本文算法	0.976 1	0.953 6	0.968 5	0.989 8

注:加黑的数字表示该算法下最优的性能,下同

本文方法的 DSC 为 0.975 8,IOU 为 0.952 9,与 U-Net 相比分别提升了 0.82%和 1.5%,本文方法体现了一定的优越性。ATU-Net^[20]在 U-Net 的每一个上采样层都加入一个注意力模块,能提取到更多的空间信息。U-Net++^[21]重新设计了跳跃路径,并采用密集跳跃连接,将更多的语义信息带到网络的深层中,提高了分割效果。Mean Teacher^[22]采用半监督学习的方式,通过最小一致化两个有略微差异的网络的分割预测,提高了模型的分割性能。由于肝脏的分割目标比较大,特征位置信息比较明显,CA 模块与锐化函数的作用变小,故与其他算法相比本文方法提升较小。

图 3 对比了不同方法对肝脏图像的分割结果。从对比图片可知,本文方法在较大的分割区域上的表现与其

他算法相比差别不大。从第 3 行图片中可以看出,本文方法准确性要比其他算法更好;从第 4 行的图片中可以看出本文方法对异常值敏感,错误的识别了分割区域。

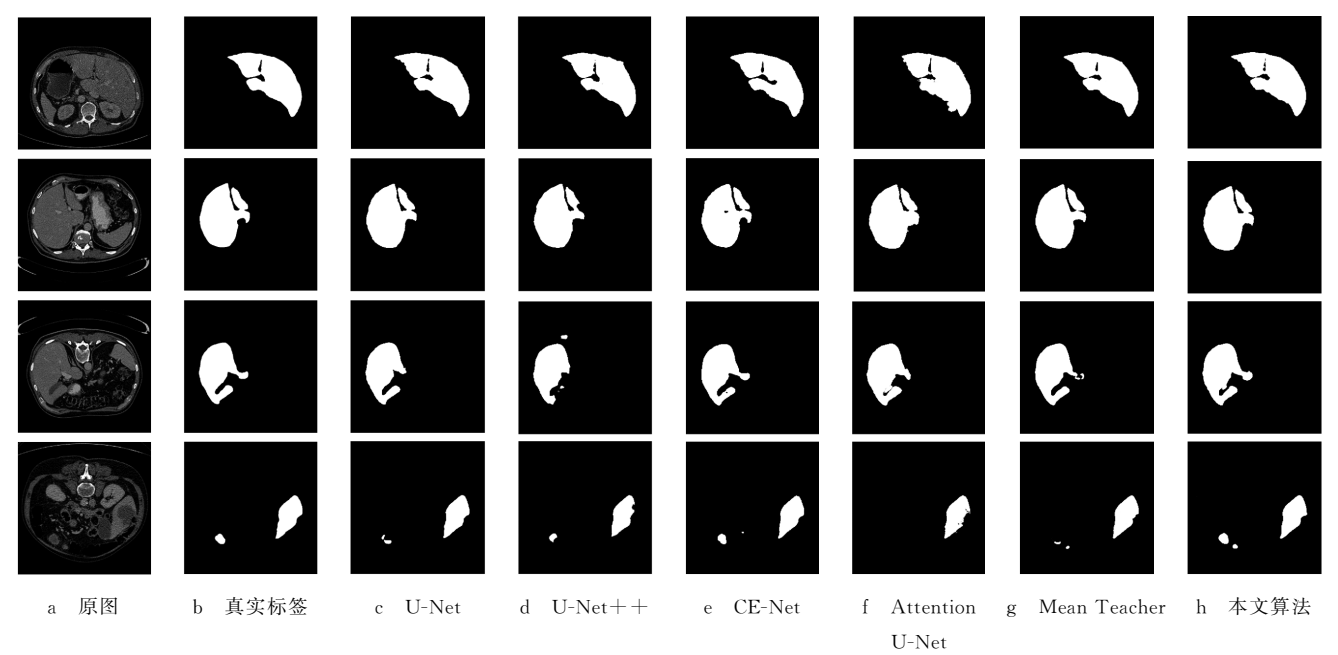


图 3 肝脏数据分割结果

Fig. 3 Visual segmentation of liver dataset

2.4.2 细胞分割 在细胞数据集中,由于分割区域较小,将 λ 设置为 0.6,此时(2)式中的 BCE 构成的损失项权值为 0.4,Dice 构成的损失项权值为 0.6。表 4 为不同算法在细胞数据集的 DSC、IOU、F1 和 SE。

表 4 细胞数据集上不同算法对比

Tab. 4 Comparison of different algorithms on cell dataset

算法	评价指标			
	DSC	IOU	F1	SE
U-Net ^[3]	0.867 2	0.772 8	0.880 2	0.844 2
ATU-Net ^[20]	0.871 4	0.779 0	0.876 9	0.849 8
U-Net++ ^[21]	0.891 5	0.807 3	0.902 6	0.888 7
CE-Net ^[4]	0.880 8	0.790 4	0.892 8	0.891 4
X-Net ^[24]	—	0.825 0	0.903 6	—
Mean Teacher ^[23]	0.888 2	0.807 2	0.889 9	0.877 9
MASSL ^[13]	0.877 4	0.805 1	0.881 0	0.866 4
本文算法	0.900 0	0.829 7	0.899 2	0.892 4

本文方法在 DSC、IOU、F1 和 SE 上取得了较好的性能。X-Net 使用双编解码器的网络结构,通过卷积和 Transformer 两种编码器同时提取局部和全局特征,得到了更好的分割效果。由于细胞数据集的背景相较于肝脏更加复杂,本文方法使用 CA 模块为网络提供更多的空间位置信息,并且锐化函数降低了未标记数据的预测熵,分割性能比 MASSL 更好,特别地,IOU 高出 2.46%。分割结果如图 4 所示,从对比图可以看出,本文方法可以有效地分割出不同大小的细胞且噪声较少、边缘轮廓清晰。

2.4.3 眼底血管分割 在眼底血管数据集中,由于眼底血管分割区域细小,将 λ 设置为 0.2,此时(2)式中的 BCE 构成的损失项权值为 0.8,Dice 构成的损失项权值为 0.2。表 5 为不同算法在眼底血管数据集的 F1、SE、SP、ACC 和 AUC。

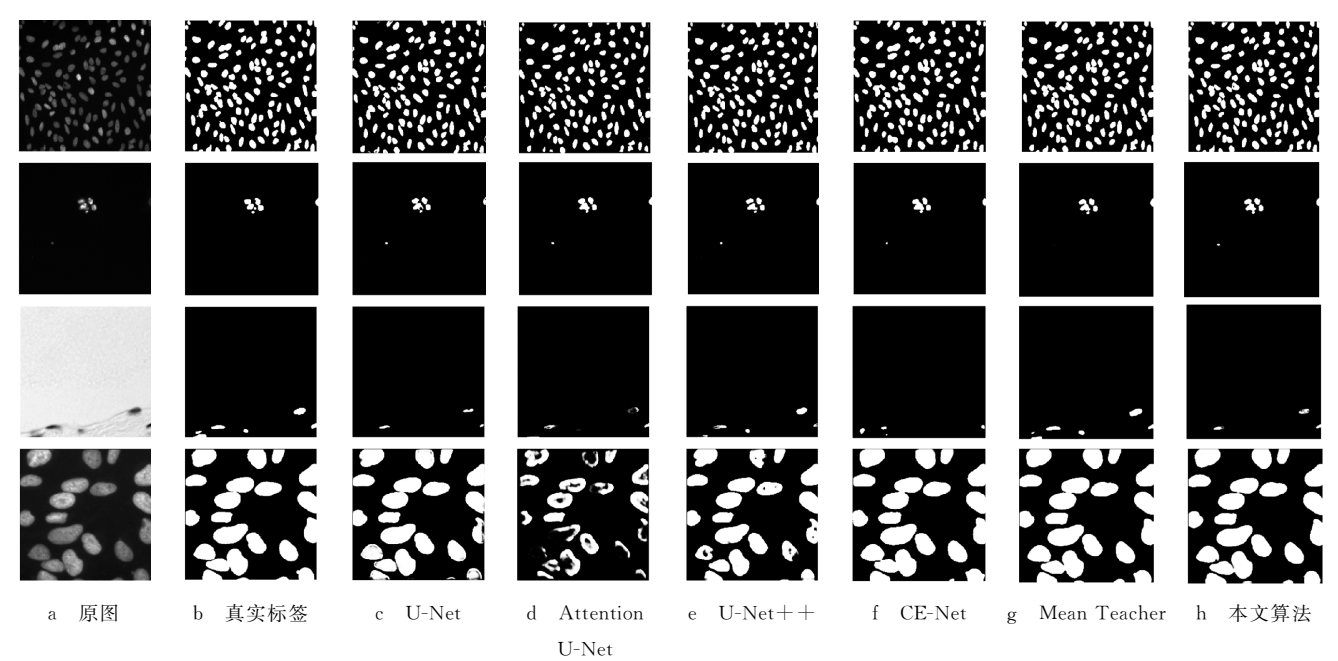


图 4 细胞分割可视化

Fig. 4 Visual segmentation of cell

表 5 眼底血管数据集上不同算法对比

Tab. 5 Comparison of different algorithms on retinal vessels dataset

算法	评价指标				
	F1	SE	SP	ACC	AUC
U-Net	0.820 6	0.828 8	0.972 1	0.953 9	0.973 5
ATU-Net	0.820 9	0.830 9	0.971 8	0.953 9	0.974 8
U-Net++	0.823 0	0.831 0	0.972 5	0.954 5	0.972 5
CcNet	—	0.762 5	0.980 9	0.952 8	0.967 8
HA-Net	0.829 3	0.799 1	0.981 3	0.958 1	0.982 3
RE-GAN	0.834 5	0.829 4	0.981 2	0.956 3	0.983 0
PPMU-Net	—	0.781 4	0.981 0	0.955 6	0.978 0
本文算法	0.824 0	0.833 0	0.973 9	0.956 9	0.979 5

从表 5 可以看出,本文方法在 SE 和 ACC 上取得了较好的性能,但是在其他指标上表现一般。原因是血管分割区域细长,CA 模块不能很好地提取到分割区域的位置信息,且受光照度等因素的影像,眼底血管图像比其他医学图像更复杂,重建网络性能减弱,导致分割性能降低。

图 5 展示了不同算法在眼底血管数据集上的分割结果,本文方法在 DRIVE 数据集上的表现相较于 U-Net、ATU-Net 和 U-Net++,出现血管断连的情况更少,能提取更多细节。总体来说,本文方法在 DRIVE 数据集上的表现较好。

2.4.4 消融实验 为对本文方法模型的有效性进行评估,将未使用锐化函数(Sharpen)、坐标注意力(CA)模块的本文方法作为基线(Baseline)网络,再将这两个模块依次融入其中,并在 3 个数据集上进行实验。结果如表 6、表 7 所示。实验结果表明,在肝脏和细胞数据集上使用了 CA 和锐化函数的网络在各个评价指标上均有明显提升。锐化函数增强了边缘轮廓的细节,提高了重建网络的置信度,对分割结果有一定的优化作用。CA 模块保留了图像在不同方向的位置信息,让网络更加关注感兴趣的区域,能更好地检测带分割区域中的小目标。本文方法与 Baseline 相比,在肝脏数据上 DSC 与 IOU 分别提高了 0.98%和 1.76%,在细胞数据集上,DSC 与 IOU 分别提高了 2.06%和 1.98%,在眼底血管数据集上,SE 与 ACC 分别提高了 1.12%和 0.16%。综合实验结果表明,各

个模块对于目标分割区域的性能有一定的贡献度。

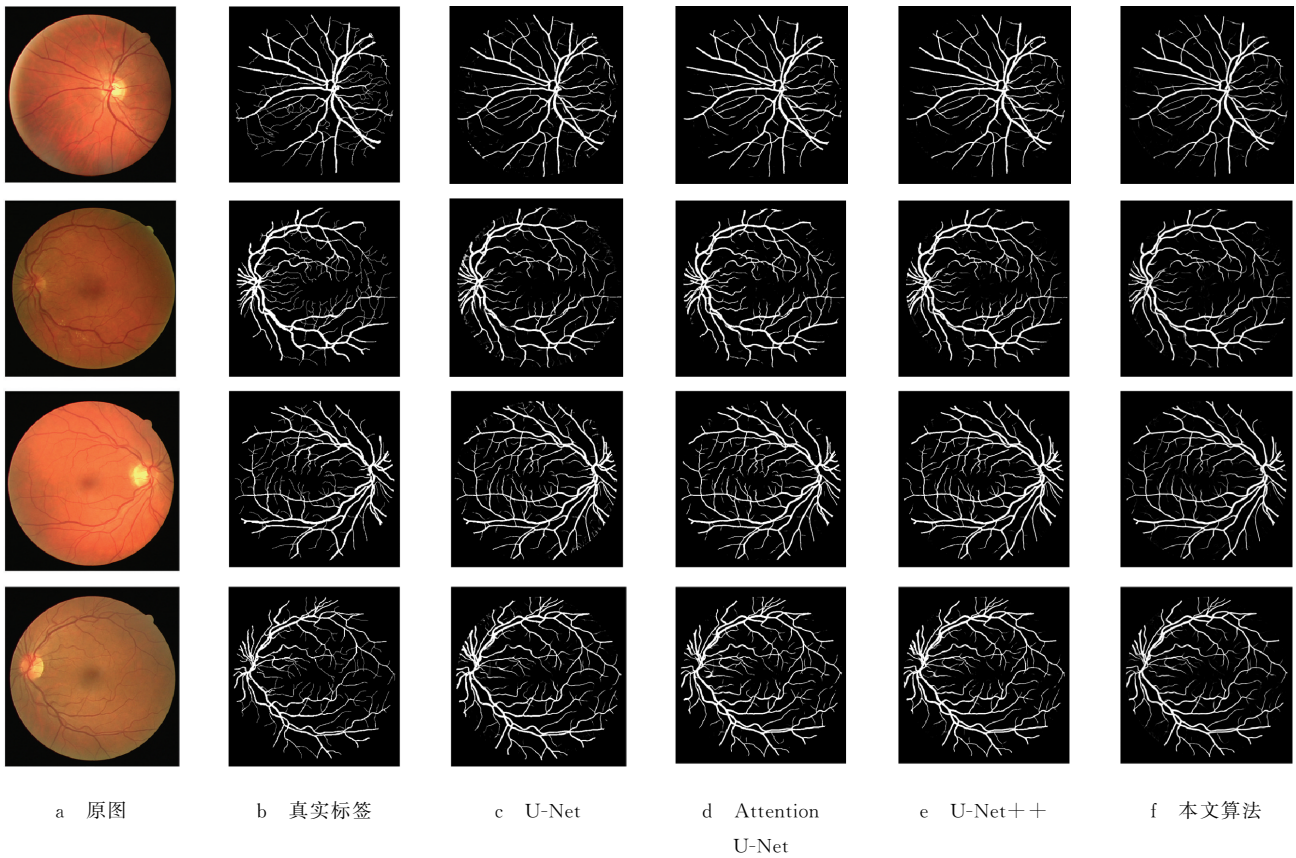


图 5 眼底血管分割可视化

Fig. 5 Visual segmentation of retinal blood vessels

表 6 肝脏与细胞数据集上的消融实验结果

Tab. 6 Results of ablation experiment on liver dataset and cell dataset

肝脏	评价指标			
	DSC	IOU	F1	SE
Baseline	0.965 1	0.934 9	0.958 7	0.987 8
Baseline+CA	0.975 6	0.948 8	0.966 3	0.988 2
Baseline+sharpen	0.970 9	0.944 2	0.964 9	0.980 4
本文算法	0.976 1	0.953 6	0.968 6	0.989 8
细胞	DSC	IOU	F1	SE
Baseline	0.879 4	0.809 9	0.878 9	0.868 0
Baseline+CA	0.891 6	0.815 9	0.894 2	0.874 3
Baseline+sharpen	0.896 7	0.826 6	0.895 9	0.887 1
本文算法	0.900 0	0.829 7	0.899 2	0.892 4

为了探究本文方法半监督框架的有效性,对不同占比的有标签数据进行了分割性能比较,使用 U-Net 作为有监督网络,结果见表 8。其中 S 和 SS 分别表示 U-net 和本文的半监督方法,L 表示有标签数据比例,U 表示无标签数据比例。从表 8 可以看出,在有标签数据集和无标签数据集比例为 1 : 9 的情况下,半监督方法性能远高于有监督方法。本文方法在仅有少量有标签数据时,加入无标签数据可以明显提高分割性能。随着标签数据逐渐增加,半监督学习的分割效果提升越来越不明显,且无标签数据有可能会被当作噪声,导致分割效果变差。

表 7 眼底血管数据集上的消融实验结果

Tab. 7 Results of ablation experiment on retinal vessels dataset

眼底血管	评价指标				
	F1	SE	SP	ACC	AUC
Baseline	0.817 8	0.821 8	0.981 2	0.955 3	0.974 6
Baseline+CA	0.818 2	0.825 3	0.979 9	0.955 5	0.976 5
Baseline+sharpen	0.821 9	0.827 8	0.979 8	0.955 9	0.975 5
本文算法	0.824 0	0.833 0	0.973 9	0.956 9	0.979 5

表 8 不同有标签数据比例对比

Tab. 8 Comparison of different label data proportions

S/SS	L/%	U/%	评价指标			
			DSC	IOU	F1	SE
S	10	—	0.941 3	0.902 4	0.940 4	0.978 0
SS	10	90	0.961 6	0.928 4	0.955 7	0.980 9
S	20	—	0.967 6	0.937 9	0.961 9	0.979 7
SS	20	80	0.978 1	0.957 4	0.970 6	0.990 0
S	40	—	0.974 9	0.952 5	0.972 0	0.982 7
SS	40	60	0.978 9	0.959 0	0.971 2	0.991 6
S	50	—	0.983 6	0.967 7	0.976 5	0.995 8
SS	50	50	0.983 4	0.967 4	0.975 3	0.994 4

为了探究 k 值对分割结果产生的影响,使用肝脏数据集比较不同取值下的模型性能,结果见表 9。从表 9 的结果可以看出, k 的取值影响整个模型的性能,当 k 取 5 时网络性能最好。当 k 取 0 时,分割网络分割效果差,由 Hadamard 积得到背景图像质量差,从而影响重建网络的性能,最后导致分割结果变差。当 k 取 10 或者更大时,因为训练前几轮没有使用无标签数据,网络已经收敛到有标签数据的特征上,导致无标签数据作用减小。综合考虑,将 k 的取值设定为 5。

表 9 不同 k 取值下的实验结果

Tab. 9 Experimental results for different values of k

k	评价指标			
	DSC	IOU	F1	SE
0	0.977 0	0.955 3	0.970 1	0.990 4
5	0.978 1	0.957 4	0.970 6	0.990 0
10	0.975 2	0.952 1	0.968 3	0.988 1
15	0.974 2	0.950 1	0.967 1	0.987 5
20	0.976 4	0.953 8	0.969 3	0.989 5

3 结束语

本文从医学图像的标签获取困难以及医学图像分割区域位置不清晰的问题出发,提出一种双任务一致性的半监督深度学习医学图像分割方法。该方法采用一个编码器两个解码器的 U-Net 网络架构,其中编码器与一个解码器做图像分割任务,与另一个解码器做图像重建任务,通过一致化两个任务的输出预测提升模型的分割性能。并采用交替训练的方式,当网络的输入为有标签数据时,只更新分割网络的权重;当网络的输入为无标签数据时,只更新重建网络的权重。在编码器的中间层加入坐标注意力模块,增强网络对分割区域空间位置信息的

感知,提升网络分割性能。在重建网络后加入锐化操作,避免重建网络的输出边缘轮廓细节模糊,提高重建网络的置信度。仿真实验结果表明,本文方法在肝脏、细胞、眼底血管数据集上的分割性能都优于当前主流的网络模型,表现出了良好的泛化性。后续工作将针对如何重建复杂的背景图像问题进行更深入的研究。

参考文献:

- [1] 闫超,孙占全,田恩刚,等. 结合多分支结构与自注意力机制的医学图像分割模型及其应用[J]. 小型微型计算机系统,2021,42(11):2375-2381.
YAN C, SUN Z Q, TIAN E G, et al. Medical image segmentation model combining multi-branch structure and self-attention mechanism and its application[J]. Journal of Chinese Computer Systems, 2021, 42(11): 2375-2381.
- [2] 潘亚玲,王哈琦,陆勇. 人工智能在医学影像 CAD 中的应用[J]. 国际医学放射学杂志,2019,42(1):6-10.
PAN Y L, WANG H Q, LU Y. Application of artificial intelligence in medical image CAD[J]. International Journal of Medical Radiology, 2019, 42(1): 6-10.
- [3] RONNEBERGE O, FISCHER P, BROX T, et al. U-net: convolutional networks for biomedical image segmentation[C]//LNCS 9351: Proceedings of the 18th International Conference on Medical Image Computing and Computer Assisted Interventions. Berlin, Heidelberg: Springer, 2015: 234-241.
- [4] GU Z, CHENG J, FU H, et al. CE-Net: context encoder network for 2D medical image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281-2292.
- [5] XIA H, SUN W, SONG S, et al. Md-net: multi-scale dilated convolution network for CT images segmentation[J]. Neural Processing Letters, 2020, 51(4): 2915-2927.
- [6] 丁才富,杨晨,纪秋浪,等. MCA-Net: 多尺度综合注意力 CNN 在医学图像分割中的应用[EB/OL]. (2021-11-23)[2022-10-05]. 微电子学与计算机. <https://doi.org/10.19304/J. ISSN1000-7180.2021.0950>.
DING C F, YANG C, JI Q L, et al. MCA-net: application of multi-scale comprehensive attention CNN in medical image segmentation[EB/OL]. (2021-11-23)[2022-10-05]. Microelectronics & Computer. <https://doi.org/10.19304/J. ISSN1000-7180.2021.0950>.
- [7] 刘清清. 基于半监督深度学习的肝脏及肿瘤 CT 图像自动分割方法研究[D]. 合肥: 中国科学技术大学, 2021.
LIU Q Q. Automatic segmentation of Liver and tumor CT images based on semi-supervised deep learning[D]. Hefei: University of Science and Technology of China, 2021.
- [8] SENKYIRE I B, LIU Z. Supervised and semi-supervised methods for abdominal organ segmentation: a review[J]. International Journal of Automation and Computing, 2021, 18(6): 887-914.
- [9] PENG J, ESTRADAB G, PEDERSOLI M, et al. Deep co-training for semi-supervised image segmentation[C]//European Conference on Computer Vision. Berlin, Heidelberg: Springer, 2018: 142-159.
- [10] HANG W, FENG W, LIANG S, et al. Local and global structure-aware entropy regularized mean teacher model for 3D left atrium segmentation[C]//Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention. Berlin, Heidelberg: Springer, 2020: 562-571.
- [11] FANGK, LI W J. DMNet: difference minimization network for semi-supervised segmentation in medical images[C]//Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention. Berlin, Heidelberg: Springer, 2020: 532-541.
- [12] WU Y, XU M, GE Z, et al. Semi-supervised left atrium segmentation with mutual consistency training[C]//Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention. Berlin, Heidelberg: Springer, 2021: 297-306.
- [13] CHEN S, BORTSOVA G, JUAREZ G U, et al. Multi-task attention-based semi-supervised learning for medical image segmentation[C]//Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention. Berlin, Heidelberg: Springer, 2019: 457-465.
- [14] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Washington: IEEE Computer Society, 2021: 13713-13722.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2018: 7132-7141.
- [16] MA J. Segmentation loss odyssey[EB/OL]. (2020-03-27)[2022-10-05]. <https://arxiv.org/pdf/2005.13449.pdf>
- [17] BILIC P, CHRIST P F, VORONTSOV E, et al. The liver tumor segmentation benchmark (LiTS)[EB/OL]. (2017-01-13)

- [2022-10-05]. <https://arxiv.org/pdf/1901.04056.pdf>
- [18] STAAL J, ABRAMOFF M D, NIEMEIER M, et al. Ridge-based vessel segmentation in color images of the retina[J]. IEEE Transactions on Medical Imaging, 2004, 23(4): 501-509.
- [19] NAYLOR, PETER, LAE, et al. Segmentation of nuclei in histopathology images by deep regression of the distance map[J]. IEEE Transactions on Medical Imaging, 2019, 38(2): 448-459.
- [20] OKTAV O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: Learning where to look for the pancreas[EB/OL]. (2018-03-20)[2022-10-05]. <https://arxiv.org/pdf/1804.03999.pdf>
- [21] ZHOU Z, SIDDIQUEE M, TAJBAKHS N, et al. U-Net++: a nested U-Net architecture for medical image segmentation [C]//Deep Learning in Medical Image Analysis. Berlin, Heidelberg: Springer, 2018: 3-11.
- [22] BI R, JI C, YANG Z, et al. Residual based attention-Unet combing DAC and RMP modules for automatic liver tumor segmentation in CT[J]. Mathematical Biosciences and Engineering, 2022, 19(5): 4703-4718.
- [23] TAVAINEN A, VALPOLA H. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results[J]. Advances in Neural Information Processing Systems, 2017, 30(2): 26-36.
- [24] LI Y, WANG Z, YIN L, et al. X-Net: a dual encoding-decoding method in medical image segmentation[J]. The Visual Computer, 2021, 37(12): 1-11.
- [25] FENG S, ZHUO Z, PAN D, et al. CcNet: a cross-connected convolutional network for segmenting retinal vessels using multi-scale features[J]. Neurocomputing, 2020, 392: 268-276.
- [26] WANG D, HAYTHA A, POTTENBURGH J, et al. Hard attention net for automatic retinal vessel segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(12): 3384-3396.
- [27] ZHOU Y K, CHEN Z L, SHEN H L, et al. A refined equilibrium generative adversarial network for retinal vessel segmentation [J]. Neurocomputing, 2021, 437: 118-130.
- [28] DU X, WANG J, SUN W. U-Net retinal blood vessel segmentation algorithm based on improved pyramid pooling method and attention mechanism[J]. Physics in Medicine & Biology, 2021, 66(17): 5013-5025.

Semi-Supervised Medical Image Segmentation Method Based on Dual Task Consistency

LUO Gang¹, LÜ Jia^{1,2}

(1. School of Computer and Information Science, Chongqing Normal University;

2. National Center for Applied Mathematics in Chongqing, Chongqing 401331, China)

Abstract: [Purposes] Medical image segmentation is an important content in medical image analysis. Most of the existing image segmentation algorithms are based on supervised learning so far, but the actual application of medical image label is difficult to obtain, numerous labeling need to rely on domain experts, which would cost time and labor. To solve the problems, a dual task consistent semi-supervised medical image segmentation model is proposed. [Methods] The model used a network structure of one encoder and two decoders, where one decoder did image segmentation and the other decoder did image reconstruction. The loss of unlabeled data was calculated and network parameters were optimized through two different image backgrounds obtained by consistent segmentation and reconstruction task. Simultaneously, the attention block is added to the network to better obtain the spatial location information of the segmented region, and sharpening operation is used to increase the confidence of the reconstructed network image. [Findings] The IOU of liver and cell were 0.953 6 and 0.8210 and the SP value of retinal blood vessel was 0.983 0. [Conclusions] Compared with the latest supervised methods in medical image segmentation, the proposed model also has a certain performance improvement in terms of effectiveness and generalization.

Keywords: medical image segmentation; semi-supervised learning; dual task consistency; coordinate attention; sharpening function

(责任编辑 黄 颖)