

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2025)02-0575-14

论文引用格式: Zhao X M, Shi P L, Wang D D, Fu Y Y, Zhang S Q and Fang J X. 2025. Distribution-aware mean teacher networks for semi-supervised medical image segmentation. Journal of Image and Graphics, 30(02):0575-0588(赵小明, 石培炼, 王丹丹, 付有瑶, 张石清, 方江雄. 2025. 分布感知均值教师网络的半监督医学影像分割. 中国图象图形学报, 30(02):0575-0588)[DOI:10.11834/jig.240208]

分布感知均值教师网络的半监督医学影像分割

赵小明^{1,2*}, 石培炼¹, 王丹丹², 付有瑶², 张石清², 方江雄²

1. 杭州电子科技大学计算机学院, 杭州 310018; 2. 台州学院智能信息处理研究所, 台州 318000

摘要: 目的 半监督方法旨在通过将未标记数据与标记数据的训练相结合, 能够减少对标记数据的依赖, 并取得较好的医学图像分割结果。然而, 现有的半监督方法通常没有关注到标记数据和未标记数据之间的分布差异所带来的不利影响, 尤其是在标记数据比例较低时, 可能会严重影响模型的分割性能。**方法** 提出一种采用分布感知的均值教师网络 (distribution-aware mean teacher network, DAMTN) 用于半监督医学影像分割。该方法利用标记数据和未标记数据的分布信息来指导模型的学习, 以便在训练阶段使模型对标记和未标记数据的分割结果的分布尽可能相似。该方法采用老师—学生的网络架构, 并嵌入了不同的注意力模块, 以及分布感知 (distribution-aware, DA) 模块、完整性监督 (integrity supervision, IS) 模块和不确定性最小化 (uncertainty minimization, UM) 模块。**结果** 在 MICCAI 2018 (Medical Image Computing and Computer Assisted Intervention Society) 左心房分割挑战 LA (left atrium) 数据集和胰腺 CT (computed tomography) 数据集上的实验结果表明, 该方法使用左心房 10% 标记数据时, 获得的 Dice 系数、Jaccard 指数、HD (Hausdorff distance) 和 ASD (average surface distance) 分别为 88.55%、79.62%、9.07 和 3.08, 与基于不确定性的协同均值教师 (uncertainty-guided collaborative mean teacher, UCMT) 相比, Dice 系数和 Jaccard 指数分别提高了 0.42% 和 0.44%; 而使用左心房 20% 标记数据时, 获得的 Dice 系数、Jaccard 指数、HD 距离和 ASD 距离分别为 90.55%、82.82%、5.78 和 1.77, 与 UCMT 相比, Dice 系数和 Jaccard 指数分别提高了 0.14% 和 0.28%。该方法使用胰腺 CT 10% 标记数据时, 获得的 Dice 系数、Jaccard 指数、HD 和 ASD 分别为 70.20%、56.36%、15.64 和 3.57, 与基于不确定性的互补一致性学习 (uncertainty-guided mutual consistency learning, UG-MCL) 相比, Dice 系数和 Jaccard 指数分别提高了 0.94% 和 1.06%; 而使用胰腺 CT 20% 标记数据时, 获得的 Dice 系数、Jaccard 指数、HD 距离和 ASD 距离分别为 77.89%、64.92%、7.97 和 1.65, 与 UG-MCL 相比, Dice 系数和 Jaccard 指数分别提高了 2.77% 和 3.34%。在 ACDC (automated cardiac diagnosis challenge) 数据集上的实验结果也表明了该方法的优越性。**结论** 提出的方法利用标记数据与未标记数据的分布差异信息, 有效提升了半监督医学影像分割性能。尤其在使用较低数量的标记数据时, 该方法的分割性能明显优于其他使用的半监督方法。

关键词: 分布感知 (DA); 均值教师; 半监督; 医学影像分割; 注意力

Distribution-aware mean teacher networks for semi-supervised medical image segmentation

Zhao Xiaoming^{1,2*}, Shi Peilian¹, Wang Dandan², Fu Youyao², Zhang Shiqing², Fang Jiangxiong²

1. School of Computer Science, Hangzhou Dianzi University, Hangzhou 310018, China;

2. Instituted of Intelligent Information Processing, Taizhou University, Taizhou 318000, China

收稿日期: 2024-04-11; 修回日期: 2024-06-27; 预印本日期: 2024-07-03

* 通信作者: 赵小明 tzxyxm@163.com

基金项目: 国家自然科学基金项目 (62276180, 61976149)

Supported by: National Natural Science Foundation of China (62276180, 61976149)

Abstract: Objective Medical image segmentation has great potential for application in clinical diagnosis. Although supervised medical image segmentation models can achieve good segmentation results, they rely heavily on pixel-level annotated training data. However, acquiring labeled data is costly and can only be performed by experts. In particular, in the case of 3D medical datasets, the annotation process is more intricate and time-consuming than in 2D datasets. Semi-supervised learning is an effective solution that combines labeled and unlabeled data for improved segmentation results. However, existing semi-supervised methods do not address the performance impact caused by the distribution gap between labeled and unlabeled data, especially when the proportion of labeled data is low. **Method** A distribution-aware mean teacher network (DAMTN) is proposed for semi-supervised medical image segmentation to reduce the performance impact caused by the distribution gap between labeled and unlabeled data. This method utilizes the distribution information of both labeled and unlabeled data to guide the learning of the model, aiming to ensure that the segmentation results of labeled and unlabeled data have similar distributions during the training phase. The DAMTN adopts a teacher-student architecture, consisting of a teacher model and a student model. In addition, both the teacher model and the student model's network architectures are based on the V-Net design, which includes an encoder and three decoders. The residual connections between the encoder and decoder have been removed. The decoders are differentiated by embedding different attention modules, which are used to introduce perturbations at the model level. These attention methods, including cross-sampling mutual attention (CMA), position attention (PA), and channel attention (CA), are employed to process high-level features. The CMA is used for inter-sample feature interaction and alignment, the PA is employed to handle spatial position information in the feature maps, the CA is utilized to address channel correlations in the feature maps, and the DAMTN comprises three key modules: the distribution-aware (DA) module, integrity supervision (IS), and uncertainty minimization (UM) module. These three modules compute the losses for the outputs of the student model, and the outputs of the teacher model are used to calculate the consistency loss with the corresponding outputs of the student model. The DA module encourages the model to learn features when the distribution information from labeled and unlabeled data is similar. Each decoder branch is equipped with a dual normalization block consisting of two normalization layers (B_0 and B_1) to obtain the distribution information of labeled and unlabeled data. B_0 handles the unlabeled data, whereas B_1 handles the labeled data. The IS module supervises the predicted mean of the branches and encourages the model to focus on challenging regions, further reducing the differences between branches. The UM module aims to balance the uncertainty of different attention branches, thereby making the predictions of all branches more consistent and improving the model's confidence. The student model incorporates an IS module and a UM module during the training phase to improve the segmentation results, thereby ensuring the consistency and accuracy of the outputs from the three decoders. During the testing phase, the student model takes the average output of the three decoders as its final output. In addition, the parameters of the student model are updated by the optimizer, whereas the parameters of the teacher model are updated using the exponential moving average. **Result** Experiments were conducted on the Medical Image Computing and Computer Assisted Intervention Society 2018 (MICCAI 2018)—Left Atrial Segmentation Challenge (LA) dataset and the Pancreas CT dataset. In the experiment with 10% labeled data for left atrium segmentation, the Dice coefficient, Jaccard index, Hausdorff distance (HD), and average surface distance (ASD) were 88.55%, 79.62%, 9.07, and 3.08, respectively. In the experiment with 20% labeled data for the left atrium, the Dice coefficient, Jaccard index, HD, and ASD were 90.55%, 82.82%, 5.78, and 1.77, respectively. For the LA dataset, when trained with 10% labeled data, compared with those of the uncertainty-guided collaborative mean teacher (UCMT), the Dice coefficient and Jaccard index improved by 0.42% and 0.44%, respectively. For the LA dataset, when trained with 20% labeled data, compared with those of the UCMT, the Dice coefficient and Jaccard index improved by 0.14% and 0.28%, respectively. In the experiment with 10% labeled data for Pancreas CT, the Dice coefficient, Jaccard index, HD, and ASD were 70.20%, 56.36%, 15.64, and 3.57, respectively. In the experiment with 20% labeled data for Pancreas CT, the Dice coefficient, Jaccard index, HD, and ASD were 77.89%, 64.92%, 7.97, and 1.65, respectively. For the Pancreas dataset, when trained with 10% labeled data, compared with uncertainty-guided mutual consistency learning (UG-MCL), the Dice coefficient and Jaccard index improved by 0.94% and 1.06%, respectively. Additionally, compared with those of the mutual consistency network (MC-Net), the Dice coefficient and Jaccard index improved by 3.70% and 4.00%, respectively. For the Pancreas dataset, when trained with 20% labeled data, com-

pared with UG-MCL, the Dice coefficient and Jaccard index improved by 2.77% and 3.34%, respectively. Additionally, compared with those of MC-Net, the Dice coefficient and Jaccard index improved by 0.73% and 0.61%, respectively. Additionally, the experimental results on the automated cardiac diagnosis challenge (ACDC) dataset also demonstrate the superiority of this method. **Conclusion** The model proposed in this study is based on the teacher-student framework, incorporates attention mechanisms, and leverages the distribution information of both labeled and unlabeled data to constrain the student model. It effectively addresses the distribution gap between labeled and unlabeled data, reducing its impact on performance and improving segmentation results. In particular, when labeled data are scarce, this model outperforms other semi-supervised segmentation methods in terms of segmentation performance.

Key words: distribution-aware(DA); mean teacher; semi-supervised; medical image segmentation; attention

0 引言

作为常见的3D医学影像,电子计算机断层扫描(computed tomography, CT)和核磁共振成像(magnetic resonance imaging, MRI)技术,能够为医生提供全面、准确的诊断信息,从而进行后续的治疗工作。CT和MRI作为重要的临床诊断辅助手段,在医学领域发挥着关键作用。CT利用X射线和计算机处理技术生成横断面的图像,能够提供详细的解剖结构信息。MRI则利用磁场和无线电波产生高分辨率的图像,可以提供更多的组织对比度。

医学影像分割在CT和MRI中具有重要意义。医学影像分割是将图像中的结构或组织区域进行分离和识别的过程。在3D医学影像中,医学影像分割可以帮助医生准确定位病变区域,提取感兴趣的解剖结构或病变区域。这将大大减少临床诊断时医生的工作量,对于临床诊断和治疗规划至关重要。

随着深度学习方法的研究,越来越多的研究人员开始搭建各种深度学习模型用于3D医学影像分割任务。这些模型利用神经网络的强大学习能力和特征提取能力,进行自动学习并提取医学影像中的关键特征,从而得到准确的分割结果。典型的深度学习模型通常以卷积神经网络(convolutional neural network, CNN)架构为基础,采用编码器(encoder)-解码器(decoder)的结构(Ronneberger等,2015)。近年来,出现了各种基于U-Net结构的改进模型,如UNet++(Zhou等,2018)、Attention UNet(Oktay等,2018)、UNet3+(Huang等,2020)和U2Net(Qin等,2020)。UNet++基于U-Net(Ronneberger等,2015)增加了更多的跳跃连接和多尺度特征融合。Attention UNet则是在U-Net的编码器和解码器中引入了注意

力机制。UNet3+主要是更多跨尺度的密集连接和特殊的损失函数设计。U2Net是由U-Net单元组成的U-Net,是一个两级嵌套的U形结构。有学者提出将CNN与Transformer相结合的分割网络模型(赵小明等,2024)。相较于2D的医学图像而言,3D医学图像多出了一个维度,处理难度相应增加。目前,处理3D图像的U-Net模型主要有3D UNet(Çiçek等,2016)和V-Net(Milletari等,2016)。3D图像作为这些模型的输入,模型内部的3D卷积能够捕获图像的3D空间信息。同样,在经过编码器、解码器及跳跃连接操作之后得到像素级的分割结果。虽然这些监督学习方法往往能够实现很好的分割效果,但是这通常建立在大量的标记数据上。对于3D数据集而言,获取大量的标记数据是昂贵、耗时的,因为这些标记数据必须由专家以逐层的方式从3D数据中进行标记形成。

针对上述问题,半监督学习(semi-supervised learning, SSL)方法可能提供了一种解决方案,因为其可以有效减小对有标记数据的依赖。在半监督医学图像分割中,一些代表性方法,如一致性正则化、符号距离图(signed distance maps, SDM)、扰动一致性、对比学习和不确定性引导等得到广泛采用。一致性正则化包括交叉伪标签监督的方法(Wu等,2021, 2022a; Chung等,2023; Gao等,2023),生成伪标签来监督不同分支的预测。SDM用于捕捉图像的形状特征(Li等,2020; Luo等,2021; Zhou等,2022a; Shi等,2023; 苏逸欣和肖志勇,2023),并对输出施加几何形状约束。扰动一致性涉及输入数据扰动的技术(Yu等,2019; Sohn等,2020),生成不一致的输入,并约束模型产生一致的输出。对比学习方法则是将样本之间的相同特征进行聚合,并对不同的特征进行分离(Bachman等,2019; Chen等,2020)。

不确定性引导则根据模型生成的不确定性图,来促使模型产生高置信度的伪标签(Shen等,2023;Wang等,2021)。此外,对抗性学习方法(Hung等,2018;Li等,2020),旨在使标记数据的预测分布与未标记数据的预测分布一致,可以看做是一致性正则化的特殊形式。另外,还有采用熵最小化的正则化方法(Zhong等,2023)。而注意力机制则是通过自适应地调整特征图中的权重,提高对重要结构和病变区域的关注度,改善分割结果的准确性、鲁棒性和计算效率。常用的注意力方法包括:通道注意力(channel attention, CA)(Zhong等,2023)、空间注意力(spatial attention, SA)(Zhong等,2023)和自注意力(self-attention)(Gao等,2023)等。Zhong等人(2023)在多个解码器中分别引入了CA和SA模块,从而使不同的解码器分别具有对关键通道的特征表示能力和对特征图空间位置的关注,提升了模型对医学图像重要的通道和空间位置的敏感性。Gao等人(2023)在解码器中利用self-attention构建了样本内的不同空间位置之间的关系,从而提高模型的性能。随着SAM(segment anything model)(Kirillov等,2023)的出现,半监督医学图像分割领域也引入了SAM辅助模型进行训练。例如,SemiSAM(Zhang等,2024)利用领域知识训练的分割模型为SAM的定位和生成输入提示提供了信息,SAM的伪标签被用做额外的监督,以协助半监督框架的学习过程。然而,这些方法未能有效缓解标记和未标记数据样本之间的分布差异问题。换句话说,标记数据和未标记数据的分布是不同的,模型此时一般不能直接将标记数据学到的知识转移到未标记数据上。当标记数据比例较低时,这种分布差异尤为明显。

为了避免直接将标记数据的知识应用于未标记数据,一种可能的方案是分别提取标记和未标记数据的分布信息,并强制使它们之间保持一致。通过减小两个集合之间的分布差距,可以有效地减少其对模型性能的不利影响,特别是在标记数据比例较低的情况下。

为此,本文提出一种采用分布感知的均值教师网络(distribution-aware mean teacher network, DAMTN)的半监督医学图像分割方法。该方法利用标记数据和未标记数据的分布信息来指导模型的学习,以便在训练阶段使模型对标记和未标记数据的分割结果的分布尽可能相似。DAMTN是一种采用老师

(Teacher)—学生(Student)模型搭建的网络,包含3个模块:分布感知(distribution-aware, DA)模块、完整性监督(integrity supervision, IS)模块和不确定性最小化(uncertainty minimization, UM)模块。DA模块计算学生模型输出的分布损失;IS模块对分支的平均预测进行监督,以确保标记数据的分割没有明显的偏差;UM模块旨在平衡不同注意力分支的不确定性。在LA(left atrium)、Pancreas CT和ACDC(automated cardiac diagnosis challenge)数据集上的实验结果表明了该方法的有效性。

本文的主要工作为:1)提出一种分布感知均值教师网络用于半监督医学图像分割,该网络鼓励模型学习整个训练集的分布特征;2)引入的分布感知模块能够有效建立标记和未标记数据之间的分布一致性;3)在LA、Pancreas CT和ACDC数据集上的实验结果表明,本文方法取得的半监督图像分割性能优于其他使用的方法。

1 DAMTN模型设计

1.1 模型概述

图1给出了用于3D医学影像分割的DAMTN的系统结构。Teacher和Student模型采用共享的编码器和3个不同的解码器分支的设计。解码器分支之间的关键区别在于使用不同的注意力嵌入来生成统计差异,以形成模型方面的扰动。其中,编码器和解码器部分来自于V-Net(Yu等,2019)的设计。

根据图1所示,整个网络的输入数据包括有标记数据和无标记数据。输入数据会分别进入Teacher模型和Student模型,且在进入Teacher模型之前进行加噪处理。Teacher模型和Student模型则是使用相同的分割主干,包括1个共享编码器和3个不同的解码器分支。分割主干产生3个解码器的输出。特别地,Teacher模型的参数更新是通过指数移动平均(exponential moving average, EMA)实现的。Teacher模型对应的输出与Student模型对应的输出进行一致性损失。完整性监督和不确定性最小化模块则是对Student模型的输出的平均值进行一致性正则化,分布感知模块是对Student模型的输出进行分布一致性约束。所提出模型的整体损失计算式为

$$L = L_{\text{sup}} + \lambda_{\text{um}} L_{\text{um}} + \lambda_{\text{dis}} L_{\text{dis}} + \lambda_{\text{con}} L_{\text{con}} \quad (1)$$

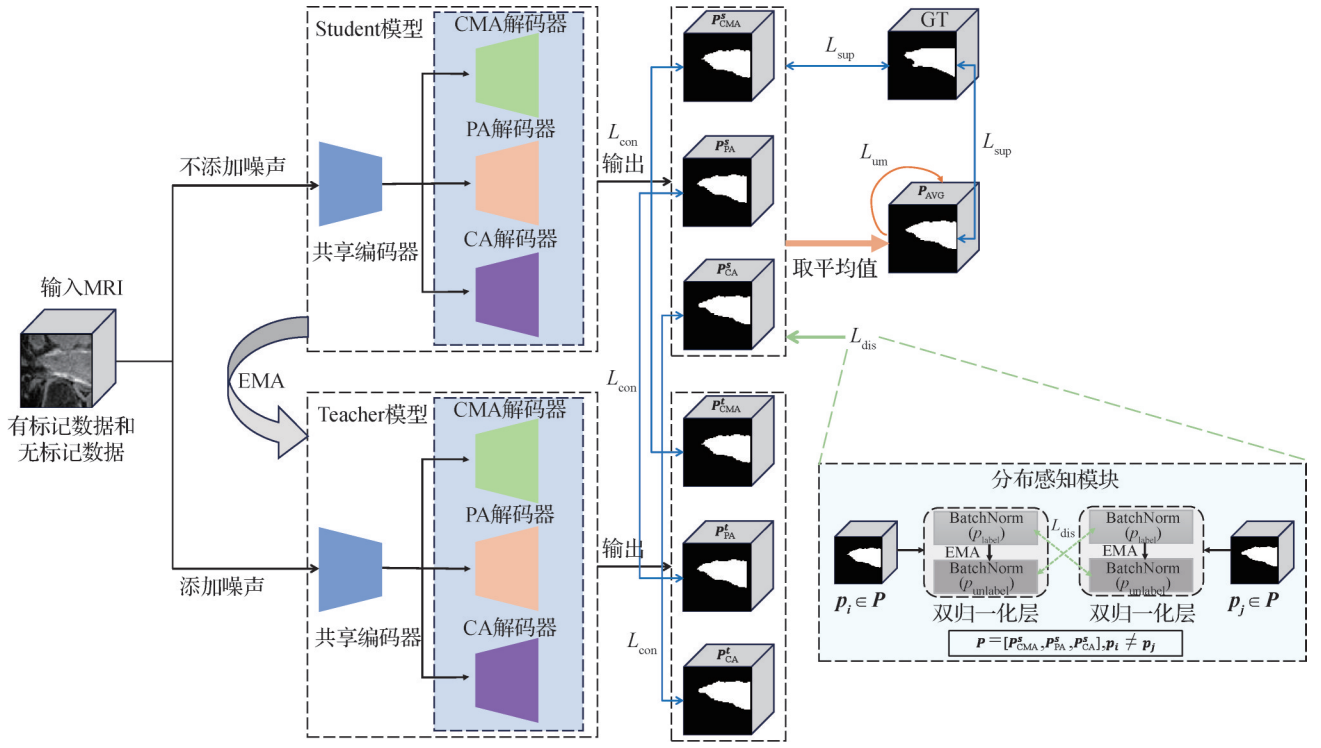


图1 DAMTN网络架构示意图

Fig. 1 DAMTN architecture illustration

式中, L_{sup} 代表3个解码器分支的监督损失和它们平均值的监督损失, L_{um} 表示不确定性最小化损失, L_{dis} 表示分布感知模块内的分布损失, L_{con} 表示 Teacher 和 Student 模型之间的一致性损失。 λ_{um} 、 λ_{dis} 和 λ_{con} 是用来调整 L_{um} 、 L_{dis} 和 L_{con} 的权重值。

1.2 分割主干

为了增强半监督医学图像的分割性能,拟在分割骨干中引入注意力模块。具体而言,分割骨干由共享的编码器和3个解码器分支组成。这些分支分别利用交叉样本互注意力(cross-sample mutual attention, CMA) (Gao 等, 2023) 模块、位置注意力(position attention, PA) (Fu 等, 2019) 模块和通道注意力(channel attention, CA) (Woo 等, 2018) 模块。解码器中嵌入的注意力模块用于处理高层特征。

1.2.1 CMA分支

CMA 分支由两个连续的 Transformer 编码器层组成,分别表示为 E_1 和 E_2 。对于高层特征图 $F_{in} \in \mathbf{R}^{b \times c \times h \times w \times d}$, 其中 b 是批量大小, c 是通道数, 以下是具体的操作:

E_1 在每个样本的空间维度上进行操作,通过执行16个头的样本内多头自注意力,捕捉同一样本中不同像素位置之间的交互作用。这一步骤关注的是

单个样本内的关系。另一方面, E_2 在批次维度上进行操作,通过执行4个头的样本间多头自注意力,捕捉不同样本之间相同空间位置的信息。这一步骤允许在相同的空间位置上对不同样本的信息进行聚合。

具体来说, E_1 的输入维度是 $\mathbf{R}^{(h \times w \times d) \times b \times c}$, 而 E_2 的输入维度是 $\mathbf{R}^{b \times (h \times w \times d) \times c}$ 。Gao 等人 (2023) 指出, CMA 注意力可以在一组样本之间建立跨样本的相关性,有效地将标记的知识转移到未标记的数据上。

1.2.2 PA分支

PA 分支采用了类似于 Transformer 中使用的位置注意力的计算方法,用于强调图像中的关键位置信息。PA 的具体细节为

$$out = V(\Phi(Conv2(F')), \delta(\Phi(Conv1(F')^T, Conv1(F'))))^T \quad (2)$$

$$F_{PA} = Unsqueeze(F' + \gamma \times out) \quad (3)$$

式中, 张量 $F' \in \mathbf{R}^{b \times c \times (h \times w) \times d}$ 是通过将 F_{in} 进行重新改变形状得到的四维张量。Conv1 和 Conv2 都是具有输入通道大小为 c 和卷积核大小为 1×1 的二维卷积层。Conv1 的输出通道大小为 $c/8$, 而 Conv2 的输出通道大小为 c 。在 Conv1 和 Conv2 的操作之后, 输出图像的维度被重新转换为 $\mathbf{R}^{b \times c/8 \times (h \times w \times d)}$ 和

$\mathbf{R}^{b \times c \times (h \times w \times d)}$ 。 Φ 表示在PyTorch中实现的批次矩阵乘法。 δ 表示对最后一个维度采用softmax操作。 V 用于将维度恢复为 $\mathbf{R}^{b \times c \times (h \times w) \times d}$ 。 γ 是一个可学习的参数。*Unsqueeze*用于扩展维度为 $\mathbf{R}^{b \times c \times h \times w \times d}$ 。

1.2.3 CA分支

CA分支利用空间注意力来处理高层特征并突出重要的通道。具体的计算为

$$\mathbf{F}_{CA} = \mathbf{F}_{in} \times \Omega(MLP(Pool_{avg}(\mathbf{F}_{in})) + MLP(Pool_{max}(\mathbf{F}_{in}))) \quad (4)$$

式中, $\mathbf{F}_{CA}$ 表示通过通道注意力增强的特征图, $\mathbf{F}_{in}$ 表示前面提到的高级特征图。 $Pool_{avg}$ 和 $Pool_{max}$ 分别表示在 $h \times w \times d$ 维度上的平均池化和最大池化操作。 MLP 代表多层感知机(multi-layer perceptron), Ω 表示sigmoid激活函数。

1.3 完整性监督和不确定性最小化

1.3.1 完整性监督

由于3个不同解码器的预测学习了不同的图像特征,具有互补性。在测试阶段,拟取Student模型的3个预测图均值作为模型的输出。在测试过程中,模型将生成3个不同分支的预测图,并选择具有较高预测概率的像素作为输出结果。因此,在训练阶段引入完整性监督,具体为

$$P_{AVG} = \frac{P_{CMA}^s + P_{PA}^s + P_{CA}^s}{3} \quad (5)$$

式中, P_{AVG} 代表Student模型的输出取平均后得到的结果。 P_{CMA}^s 是CMA分支的输出, P_{PA}^s 是PA分支的输出, P_{CA}^s 是CA分支的输出。对于 P_{AVG} 的监督损失可以表示为

$$L_{sup}^{AVG} = CE(P_{AVG}, Label) + Dice(\delta(P_{AVG}), Label) \quad (6)$$

式中, δ 表示softmax函数, $CE(\cdot)$ 代表交叉熵损失, $Dice(\cdot)$ 表示Dice损失, $Label$ 表示标签数据的真实标签。类似地,对于其他3个分支,监督损失可以表示为

$$L_{sup}^{CMA} = CE(P_{CMA}^s, Label) + Dice(\delta(P_{CMA}^s), Label) \quad (7)$$

$$L_{sup}^{PA} = CE(P_{PA}^s, Label) + Dice(\delta(P_{PA}^s), Label) \quad (8)$$

$$L_{sup}^{CA} = CE(P_{CA}^s, Label) + Dice(\delta(P_{CA}^s), Label) \quad (9)$$

最后,整体的监督损失可以表示为

$$L_{sup} = \frac{1}{3}(\lambda(L_{sup}^{CMA} + L_{sup}^{PA} + L_{sup}^{CA}) + \lambda_{AVG}L_{sup}^{AVG}) \quad (10)$$

式中,参数 λ 和 λ_{AVG} 是加权因子,其中 $\lambda = 0.5$, $\lambda_{AVG} = 0.06$ 。考虑到已经在不同的分支中应用了监督损失,在训练过程中,这些分支倾向于在主要区域的分

割上达成一致,而在更具挑战性的区域出现差异。因此,增加了 λ_{AVG} 参数,以鼓励模型在训练的后期更加关注具有挑战性的区域,进一步减少分支之间的差异。

1.3.2 不确定性最小化

在整个模型设计中,3个解码器上采用了不同形式的注意力,导致它们的输出存在统计差异,即在解码器分支之间的某些区域的分割结果中存在显著的变化。为了尽可能地减小模型的不确定性,确保3个分支的预测一致性,拟采用Zhong等人(2023)和Vu等人(2019)方法中的不确定性最小化方法。通过最小化不确定性,有效地调整未标记图像的预测结果,增加模型对其预测的置信度。

具体来说,先对3个分支得到的预测图进行softmax操作,再将它们相加,然后除以3得到预测的标签,表示为 $\bar{P}$ 。不确定性最小化意味着希望预测标签 $\bar{P}$ 中每个像素的预测概率要么为0,要么为1,具体表述为

$$L_{um} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=0}^C \bar{P}_i^c \log(\bar{P}_i^c) \quad (11)$$

式中, $\bar{P} = \delta(P_{AVG})$ 。 C 和 N 分别表示类别数和像素数。 $\bar{P}_i^c$ 表示像素 i 上类别 c 的平均概率。当最小化 L_{um} (不确定性最小化损失)时,给定像素的所有分支的预测趋向于一致,即要么全部为0,要么全部为1。

1.4 分布感知模块

在标记数据比例有限且大量未标记数据存在的场景中,如果没有适当的指导,模型可能会容易拟合标记数据,并忽视未标记数据中的潜在信息。现有方法通常采用锐化操作来帮助模型学习标记数据和未标记数据的模糊区域特征,从而减少不确定性。这些方法通常假设模型在训练期间能够很好地对未标记数据进行分割。而在训练过程中,模型对未标记数据的分割性能是随机的,可能产生错误并使模型的性能不稳定。

此外,在半监督医学图像分割的背景下,标记数据和未标记数据通常具有一定程度的分布相似性。为此,引入Dual-BN块(Zhou等,2022b),表示为 B_0 和 B_1 ,每个解码器分支都由两个BatchNorm层组成。 B_0 用于处理来自未标记数据的预测标签,而 B_1 用于处理来自标记数据的预测标签。当预测标签经过BatchNorm层时,这些层捕获了预测标签的均值和方差信息,即它们的分布信息。因此,它们计算的分

布损失为

$$mse_{ij}^{11} = MSE(B_1^i(P_{label}^i), B_1^j(P_{label}^j)) \quad (12)$$

$$mse_{ij}^{00} = MSE(B_0^i(P_{unlabel}^i), B_0^j(P_{unlabel}^j)) \quad (13)$$

$$mse_{ij}^{10} = MSE(B_1^i(P_{label}^i), B_0^j(P_{unlabel}^j)) \quad (14)$$

$$mse_{ii}^{10} = MSE(B_1^i(P_{label}^i), B_0^i(P_{unlabel}^i)) \quad (15)$$

$$L_{dis} = \frac{1}{3} \sum_{i=1}^3 \sum_{j=1, j \neq i}^3 (mse_{ij}^{11} + mse_{ij}^{00} + mse_{ij}^{10} + mse_{ii}^{10}) \quad (16)$$

式中, P_{label}^i 表示第 i 个分支对标记数据的预测标签映射, $P_{unlabel}^i$ 表示第 i 个分支对未标记数据的预测标签映射。 P_{label}^i 和 $P_{unlabel}^i$ 遵循相同的模式。 B_1^i 是与第 i 个分支处理标记数据相关联的 BatchNorm 层, B_0^i 是与第 i 个分支处理未标记数据相关联的 BatchNorm 层。通过 BatchNorm 层后, 获得了预测标签图的分布信息。 $MSE(\cdot)$ 表示计算分布信息的均方误差 (mean squared error, MSE) 损失。

为了利用来自标记数据的分布信息来指导未标记数据的学习, DA 模块使用 EMA (Tarvainen 和 Valpola, 2017) 分别更新 3 个 B_0^i 层的参数, 具体为

$$e_0^t = \alpha e_0^{t-1} + (1 - \alpha) e_1^t \quad (17)$$

式中, e_k^t ($k = 0, 1$) 表示相应 BatchNorm 层的参数, t 表示训练迭代次数, α 是控制 EMA 更新的参数, 具体为 $\alpha = \min(1 - 1/(t + 1), 0.99)$ 。

1.5 一致性损失

输入 Teacher 模型的图像是经过加噪处理的, 而 Student 模型的输入图像则没有添加噪声。此外, 用于 Teacher 模型的参数更新的 EMA 可以参考式 (17)。一致性损失是利用 Teacher 模型的输出对学生 Student 模型的输出进行均方误差损失的计算, 以约束学生模型的输出, 具体为

$$L_{con} = MSE(P_{CMA}^s, P_{CMA}^t) + MSE(P_{PA}^s, P_{PA}^t) + MSE(P_{CA}^s, P_{CA}^t) \quad (18)$$

2 实验和结果分析

2.1 数据集

1) LA 数据集。来自 2018 年 MICCAI (Medical Image Computing and Computer Assisted Intervention Society) 分割挑战赛的 LA 基准数据集 (Yu 等, 2019)。它提供了 100 个 3D 钆增强成像 MRI 扫描和 LA 分割掩膜。这些扫描结果各向同性分辨率为 $0.625 \text{ mm} \times 0.625 \text{ mm} \times 0.625 \text{ mm}$ 。根据 Yu 等人 (2019) 的设置, 该数据集分为两组: 80 次扫描用于

训练和 20 次扫描用于评估。LA 图像的尺寸处理为 $112 \times 112 \times 80$ 。

2) Pancreas CT 数据集。该数据集包含 82 例腹部三维增强 CT 扫描患者。这是美国国立卫生研究院临床中心对 53 名男性和 27 名女性受试者进行 82 次腹部对比增强 3D CT 扫描, 17 名受试者是肾切除术前扫描的健康肾脏供体, 剩下的 65 名患者由放射科医生从既没有腹部病变也没有胰腺癌病变的患者中选择。根据 Wu 等人 (2022a) 的设置, 分配 62 次扫描用于训练, 20 次扫描用于评估。Pancreas CT 图像的尺寸处理为 $96 \times 96 \times 96$ 。

3) ACDC 数据集。该数据集是在法国第戎大学医院 (University Hospital of Dijon) 临床实践中收集的 (Bernard 等, 2018)。ACDC 数据集包含来自 100 名患者的心脏 MRI 样本, 用于训练。此外, 参照 Wu 等人 (2022a) 的工作, 本文在患者级别使用固定数据分割进行实验, 其中新的训练集、验证集和测试集分别包含 70、10 和 20 个患者的数据。ACDC 图像的尺寸处理为 256×256 像素。

LA 和 Pancreas CT 数据集是 3D 分割数据集, 训练数据是整个 3D 图像; 而 ACDC 数据集是将其 3D 图像切割成 2D 分片进行训练, 即为 ACDC 数据集设置了相应的 2D 模型进行训练。

2.2 实验设置

实验使用 NVIDIA RTX A6000 GPU 在 PyTorch 平台中实现所有网络模型进行训练。采用随机梯度下降 (stochastic gradient descent, SGD) 优化器更新模型参数, 权重衰减为 0.0001, 动量为 0.9。初始学习率设置为 0.01, 每 2500 次迭代除以 10 次。总共训练了 6000 次迭代, 批处理大小为 4, 包含相同数量的有标记和未标记图像。

使用一个随时间变化的高斯上升函数 $\lambda(t) = \beta \times e^{(-5((1 - t/t_{\max}))^2)}$ ($\beta = 0.1$) 来控制监督损失和其他损失之间的相对权重, 其中 t 表示当前的训练步骤, $t_{\max}$ 为最大的训练步骤。这种设计确保了最初的总体损失由监督损失项控制, 避免了网络过早地学习不确定的特征。式 (1) 中的 λ_{um} 和 λ_{con} 直接使用这个高斯上升函数, 而 $\lambda_{dis} = 45 \times \lambda(t)$ 。

2.3 评价指标

评价指标包括 Dice 系数 (Dice coefficient, DSC)、Jaccard 指数 (Jaccard)、95% 的 Hausdorff 距离 (95%

Hausdorff distance, 95HD)和平均表面距离(average surface distance, ASD)。

95HD 和 ASD 用于评估预测结果与真实标签之间的边界误差。

2.4 与其他半监督方法的比较

比较的方法有 UA-MT (uncertainty-aware mean teacher) (Yu 等, 2019)、SASSNet (shape-aware semi-supervised network) (Li 等, 2020)、DTC (dual-task consistency) (Luo 等, 2021)、MC-Net (mutual consistency network) (Wu 等, 2021)、SS-Net (smoothness and inter-class separation network) (Wu 等, 2022b)、UCMT (uncertainty-guided collaborative mean teacher) (Shen 等, 2023)和 UG-MCL (uncertainty-guided mutual consistency learning) (Zhang 等, 2023)。

2.4.1 在 LA 数据集上的比较

表 1 给出了不同的半监督方法在 LA 数据集上的比较结果。从表 1 可以看出,在 LA 数据集上,当只有 20% 的标记数据时,相较于 UCMT (Shen 等,

2023), Dice 系数和 Jaccard 指数分别提高了 0.14% 和 0.28%。在 LA 数据集上,当使用 10% 标记数据进行训练时,与 UCMT 相比, Dice 系数和 Jaccard 指数分别提高了 0.42% 和 0.44%。这说明本文设计的分布感知均值教师网络有效地利用了医学图像的分布信息,提升了半监督医学图像的分割性能,尤其是在低比例下的标记数据时。同时,该模型的 ASD 指标并不是最佳的,这是由于分布信息对于边界信息的不敏感所导致。

图 2 为本文和其他不同的半监督方法在 LA 数据集上的可视化结果。由图 2 的结果表明,与这些方法相比,本文方法取得了最好的分割效果。

2.4.2 在 Pancreas CT 数据集上的比较

表 2 给出了不同的半监督方法在 Pancreas CT 数据集上的比较结果。从表 2 可以看出,在 Pancreas 数据集上,当使用 20% 标记数据进行训练时,与 UG-MCL (Zhang 等, 2023) 相比, Dice 系数和 Jaccard 指数分别提高了 2.77% 和 3.34%。此外,与 MC-Net (Wu

表 1 DAMTN 和不同的半监督分割方法在 LA 数据集上的定量比较

Table 1 Quantitative comparison between DAMTN and different SSL segmentation methods on the LA dataset

方法	有标记数据	无标记数据	Dice/%	Jaccard/%	95HD/体素	ASD/体素
V-Net (Milletari 等, 2016)	8 (10%)	0	79.87	67.60	26.65	7.94
V-Net (Milletari 等, 2016)	16 (20%)	0	85.94	75.99	16.70	4.80
V-Net (Milletari 等, 2016)	80 (全部)	0	90.98	83.61	8.58	2.10
UA-MT (Yu 等, 2019)	8 (10%)	72 (90%)	85.81	75.41	18.25	5.04
SASSNet (Li 等, 2020)	8 (10%)	72 (90%)	85.71	75.35	14.74	4.00
DTC (Luo 等, 2021)	8 (10%)	72 (90%)	84.55	73.91	13.80	3.69
MC-Net (Wu 等, 2021)	8 (10%)	72 (90%)	86.87	77.98	12.67	2.18
SS-Net (Wu 等, 2022b)	8 (10%)	72 (90%)	86.56	76.61	12.76	3.02
UCMT (Shen 等, 2023)	8 (10%)	72 (90%)	88.13	79.18	9.14	3.06
DAMTN (本文)	8 (10%)	72 (90%)	88.55	79.62	9.07	3.08
UA-MT (Yu 等, 2019)	16 (20%)	64 (80%)	88.18	79.09	9.66	2.62
SASSNet (Li 等, 2020)	16 (20%)	64 (80%)	88.11	79.08	12.31	3.27
DTC (Luo 等, 2021)	16 (20%)	64 (80%)	87.79	78.52	10.29	2.50
MC-Net (Wu 等, 2021)	16 (20%)	64 (80%)	90.12	82.12	8.07	1.99
SS-Net (Wu 等, 2022b)	16 (20%)	64 (80%)	88.19	79.21	8.12	2.20
UCMT (Shen 等, 2023)	16 (20%)	64 (80%)	90.41	82.54	6.31	1.70
DAMTN (本文)	16 (20%)	64 (80%)	90.55	82.82	5.78	1.77

注:加粗字体表示各组各列最优结果。

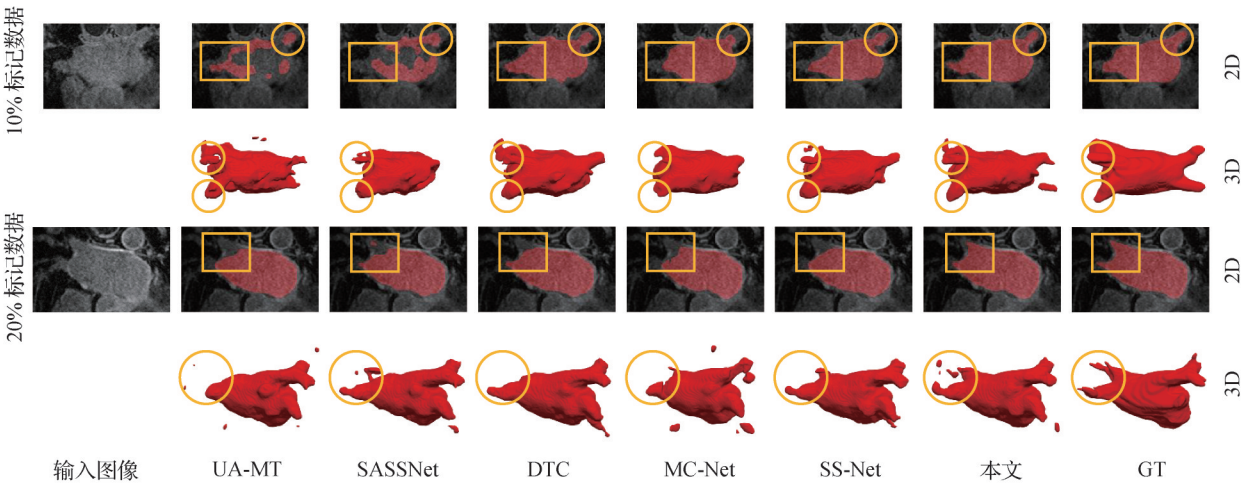


图2 不同的半监督方法在LA数据集上的可视化结果

Fig. 2 Visualization results of different semi-supervised methods on the LA dataset

表2 DAMTN和不同的半监督分割方法在Pancreas CT数据集上的定量比较

方法	有标记数据	无标记数据	Dice/%	Jaccard/%	95HD/体素	ASD/体素
V-Net(Milletari 等, 2016)	6(10%)	0	67.78	53.74	21.59	3.86
V-Net(Milletari 等, 2016)	12(20%)	0	71.89	59.45	10.13	1.59
V-Net(Milletari 等, 2016)	62(全部)	0	81.95	69.94	5.75	1.36
UA-MT(Yu 等, 2019)	6(10%)	56(90%)	67.92	53.94	17.66	4.32
SASSNet(Li 等, 2020)	6(10%)	56(90%)	67.08	52.75	18.35	3.35
DTC(Luo 等, 2021)	6(10%)	56(90%)	67.32	53.78	15.76	2.03
MC-Net(Wu 等, 2021)	6(10%)	56(90%)	66.5	52.36	19.95	6.91
SS-Net(Wu 等, 2022b)	6(10%)	56(90%)	67.58	54.52	21.35	3.53
UG-MCL(Zhang 等, 2023)	6(10%)	56(90%)	69.26	55.3	15.85	3.67
DAMTN(本文)	6(10%)	56(90%)	70.2	56.36	15.64	3.57
UA-MT(Yu 等, 2019)	12(20%)	50(80%)	70.52	58.15	17.46	2.67
SASSNet(Li 等, 2020)	12(20%)	50(80%)	76.28	62.77	9.16	2.97
DTC(Luo 等, 2021)	12(20%)	50(80%)	71.56	58.62	15.47	1.48
MC-Net(Wu 等, 2021)	12(20%)	50(80%)	77.16	64.31	10.93	2.47
SS-Net(Wu 等, 2022b)	12(20%)	50(80%)	74.5	62.04	13.88	1.4
UG-MCL(Zhang 等, 2023)	12(20%)	50(80%)	75.12	61.58	11.45	2.68
DAMTN(本文)	12(20%)	50(80%)	77.89	64.92	7.97	1.65

注:加粗字体表示各组各列最优结果。

等,2021)相比,Dice系数和Jaccard指数分别提高了0.73%和0.61%。在Pancreas数据集上,当使用10%标记数据进行训练时,与UG-MCL相比,Dice系数和Jaccard指数分别提高了0.94%和1.06%。此外,与MC-Net相比,Dice系数和Jaccard指数分别提

高了3.70%和4.00%。与LA数据集相比,在Pancreas数据集的10%标记数据时,该模型获得了更好的性能提升。其次,ASD指标也与LA数据集一致,即没有达到最优。这表明,本文提出的方法具有一定的泛化能力,在Pancreas数据集上也能够表现得

很好。

图3为本文和其他不同的半监督方法在Pancreas数据集上的可视化结果,这些方法和在LA数据集上的可视化对比方法相同。从图3的可视化结果可以看出,与这些方法相比,本文方法获得了最好的分割效果。

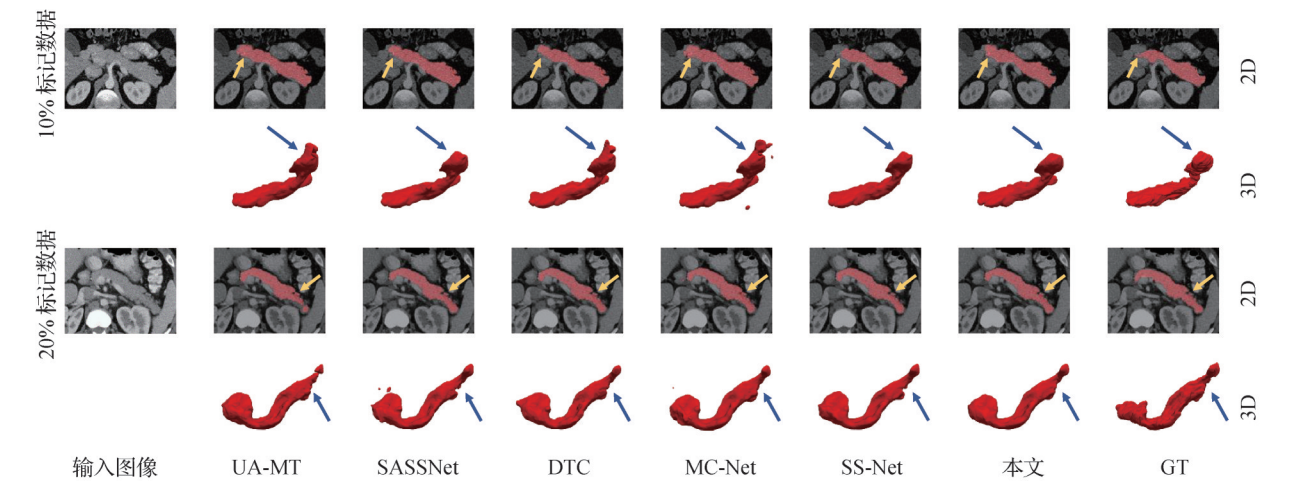


图3 不同的半监督方法在Pancreas CT数据集上的可视化结果
Fig. 3 Visualization results of different semi-supervised methods on the Pancreas CT dataset

2. 4. 3 在ACDC数据集上的比较

表3给出了不同的半监督方法在ACDC数据集上的比较结果。从表3可以看出,在ACDC数据集上,当使用20%标记数据进行训练时,相较于MC-Net(Wu等,2021),Dice系数和Jaccard指数分别提升了0.18%和0.50%。在ACDC数据集上,当使用

表3 DAMTN和不同的半监督分割方法在ACDC数据集上的定量比较

方法	有标记数据	无标记数据	Dice/%	Jaccard/%	95HD/体素	ASD/体素
U-Net(Ronneberger等, 2015)	7(10%)	0	77.34	66.20	9.18	2.45
U-Net(Ronneberger等, 2015)	14(20%)	0	85.15	75.48	6.20	2.12
U-Net(Ronneberger等, 2015)	70(全部)	0	91.65	84.93	1.89	0.56
UA-MT(Yu等, 2019)	7(10%)	63(90%)	81.58	70.48	12.35	3.62
SASSNet(Li等, 2020)	7(10%)	63(90%)	84.14	74.09	5.03	1.40
DTC(Luo等, 2021)	7(10%)	63(90%)	82.71	72.14	11.31	2.99
MC-Net(Wu等, 2021)	7(10%)	63(90%)	86.34	76.82	7.08	2.08
SS-Net(Wu等, 2022b)	7(10%)	63(90%)	84.07	73.84	9.33	2.69
UG-MCL(Zhang等, 2023)	7(10%)	63(90%)	86.24	76.29	6.07	2.04
DAMTN(本文)	7(10%)	63(90%)	86.69	77.15	6.68	2.14
UA-MT(Yu等, 2019)	14(20%)	56(80%)	85.87	76.78	5.06	1.54
SASSNet(Li等, 2020)	14(20%)	56(80%)	87.04	78.13	7.84	2.15
DTC(Luo等, 2021)	14(20%)	56(80%)	86.28	77.03	6.14	2.11
MC-Net(Wu等, 2021)	14(20%)	56(80%)	87.83	79.14	4.94	1.52
SS-Net(Wu等, 2022b)	14(20%)	56(80%)	86.07	76.61	7.26	2.27
UG-MCL(Zhang等, 2023)	14(20%)	56(80%)	87.28	78.59	6.84	2.37
DAMTN(本文)	14(20%)	56(80%)	88.01	79.64	5.85	1.64

注:加粗字体表示各组各列最优结果。

10% 标记数据进行训练时,与 MC-Net 相比,Dice 系数和 Jaccard 指数分别提升了 0.35% 和 0.33%。这表明,本文方法能够有效应用于 2D 数据集上的分割任务,证明了其通用性。此外,在 ACDC 数据集上,HD95 和 ASD 指标均没有达到最优,进一步表明,分布信息对于边界信息的不敏感。

2.4.4 在 LA 数据集上不同标注数据比例下的 Dice 性能

图 4 为不同方法在 LA 数据集上的不同标注数据比例下的 Dice 性能变化曲线图。由图 4 可以看出,随着标注数据比例的增大,所有的方法 Dice 性能均得到增加,至 50% 的标注数据时,Dice 指标几乎停止了增加,而本文提出的模型则一直保持着最佳的性能。

2.5 消融实验

2.5.1 验证模块的有效性

为了探究模型的不同模块的有效性,表 4 给出了在 LA 数据集上的消融实验结果。表 4 的结果表明,在两个不同比例的标记数据下,分别去除 UM、IS 和 DA 模块时,所有指标均变差。这证明了本文设计模块的有效性。尤其是,在 10% 的标记数据时,去

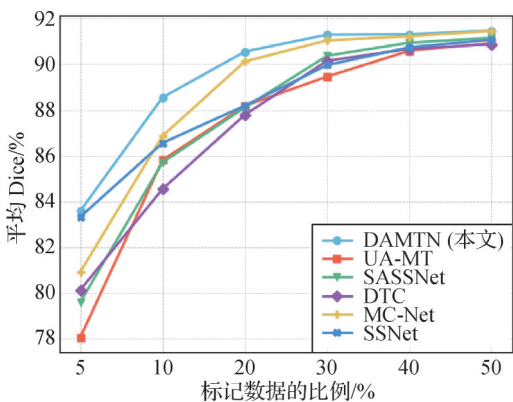


图 4 在 LA 数据集上,不同标注数据比例下的 Dice 性能

Fig. 4 Dice performance under different annotation data ratios on the LA dataset

除 UM 模块时,与其他模块相比,所有指标都发生了较大的变化。这表明在低比例下的标记数据时,解码器之间的不确定性较大,这不利于模型提升分割性能。消去 IS 模块时,所有指标均下降,这表明该模块对解码器进行了有效的一致性正则化。此外,随着标记数据的比例下降,消去 DA 模块,模型性能下降得更多。这表明低比例标记数据时,模型能够有效减少有标记数据和未标记数据之间的分布差距。

表 4 DAMTN 在 LA 数据集上进行的消融实验
Table 4 Ablation experiments of DAMTN on the LA dataset

单独消去的模块	有标记数据	无标记数据	Dice/%	Jaccard/%	95HD/体素	ASD/体素
UM	8(10%)	72(90%)	87.42	77.98	14.46	4.41
IS	8(10%)	72(90%)	88.29	79.38	12.21	3.60
DA	8(10%)	72(90%)	88.39	79.34	14.00	3.71
DAMTN(不消)	8(10%)	72(90%)	88.55	79.62	9.07	3.08
UM	16(20%)	64(80%)	90.39	82.56	6.49	2.17
IS	16(20%)	64(80%)	90.34	82.32	6.62	2.28
DA	16(20%)	64(80%)	90.51	82.76	8.40	2.38
DAMTN(不消)	16(20%)	64(80%)	90.55	82.82	5.78	1.77

注:加粗字体表示各组各列最优结果。

2.5.2 探究分布感知模块的作用

图 5 是在 LA 数据集上有无分布感知模块时,10% 标注数据比例下的特征分布图。图 5 表明,在添加分布感知模块后有标记数据和未标记数据的特征分布更加靠近了,即两者的重叠区域增大了。这表明模型使用分布感知模块能够使得标记数据特征

分布和未标记数据特征分布更加接近。

3 结 论

本文提出一种基于分布感知均值教师网络用于半监督医学图像分割领域,旨在利用标记和无标记

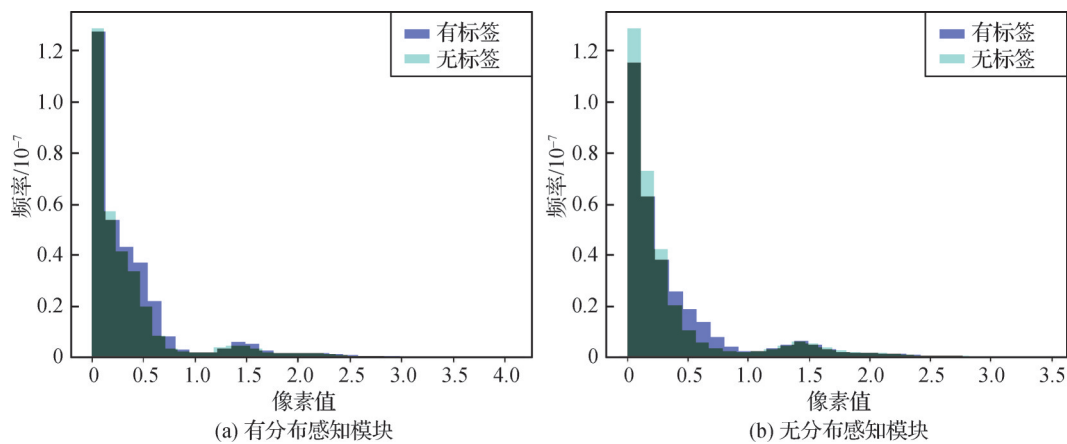


图5 在LA数据集上有无分布感知模块时,10%标注数据比例下特征分布情况

Fig. 5 With and without a distribution-aware module on the LA dataset, the feature distribution at a 10% annotated data ratio
(a) with distribution-aware module; (b) without distribution-aware module)

数据的分布信息来指导模型学习。该模型核心的损失计算由3个模块完成:分布感知(DA)模块、完整性监督(IS)模块和不确定性最小化(UM)模块。分布感知模块旨在有标记和无标记数据的分布信息相似时,鼓励模型学习特征;完整性监督和不确定性最小化模块则主要对模型进行一致性正则化。该模型利用分布感知模块进一步平衡了有标记数据和无标记数据之间的分布差距,进一步提升了模型的性能。在LA、Pancreas CT和ACDC的实验结果表明,所提出的模型性能优于其他使用的方法,验证了利用分布信息提升半监督模型性能的有效性。

未来将考虑将在更低比例的标记数据下,尝试改进分布信息对半监督分割模型的约束方法,以便进一步提升半监督医学图像分割的性能。例如,设置缓存模块,存储训练过程中的样本特征,损失则约束每个批次的样本特征与缓存特征之间的分布。

参考文献(References)

- Bachman P, Hjelm R D and Buchwalter W. 2019. Learning representations by maximizing mutual information across views//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: 15535-15545
- Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng P A, Cetin I, Lekadir K, Camara O, Ballester M A G, Sanroma G, Napel S, Petersen S, Tziritis G, Grinias E, Khened M, Kollerathu V A, Krishnamurthi G, Rohé M M, Pennec X, Sermesant M, Isensee F, Jäger P, Maier-Hein K H, Full P M, Wolf I, Engelhardt S, Baumgartner C F, Koch L M, Wolterink J M, Išgum I, Jang Y, Hong Y, Patravali J, Jain S, Humbert O and Jodoin P M. 2018. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Transactions on Medical Imaging, 37 (11): 2514-2525 [DOI: 10.1109/TMI.2018.2837502]
- Chen T, Kornblith S, Norouzi M and Hinton G. 2020. A simple framework for contrastive learning of visual representations//Proceedings of the 37th International Conference on Machine Learning. Virtual Event: JMLR.org: 1597-1607
- Chung Y, Lim C, Huang C, Marrouche N and Hamm J. 2023. FBA-Net: foreground and background aware contrastive learning for semi-supervised atrium segmentation//Proceedings of the 2nd International Workshop on Medical Image Learning with Limited and Noisy Data. Vancouver, Canada: Springer: 106-116 [DOI: 10.1007/978-3-031-44917-8_10]
- Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O. 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation//Proceedings of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens, Greece: Springer: 424-432 [DOI: 10.1007/978-3-319-46723-8_49]
- Fu J, Liu J, Tian H J, Li Y, Bao Y J, Fang Z W and Lu H Q. 2019. Dual attention network for scene segmentation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3141-3149 [DOI: 10.1109/cvpr.2019.00326]
- Gao S B, Zhang Z J, Ma J C, Li Z H and Zhang S. 2023. Correlation-aware mutual learning for semi-supervised medical image segmentation//Proceedings of the 26th International Conference on Medical Image Computing and Computer Assisted Intervention. Vancouver, Canada: Springer: 98-108 [DOI: 10.1007/978-3-031-43907-0_10]
- Huang H M, Lin L F, Tong R F, Hu H J, Zhang Q W, Iwamoto Y,

- Han X H, Chen Y W and Wu J. 2020. UNet 3+: a full-scale connected UNet for medical image segmentation//Proceedings of 2020 International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain; IEEE: 1055-1059 [DOI: 10.1109/icassp40776.2020.9053405]
- Hung W C, Tsai Y H, Liou Y T, Lin Y Y and Yang M H. 2018. Adversarial learning for semi-supervised semantic segmentation [EB/OL]. [2024-04-11]. <http://arxiv.org/pdf/1802.07934.pdf>
- Kirillov A, Mintun E, Ravi N, Mao H Z, Rolland C, Gustafson L, Xiao T T, Whitehead S, Berg A C, Lo W Y, Dollár P and Girshick R. 2023. Segment anything//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 3992-4003 [DOI: 10.1109/ICCV51070.2023.00371]
- Li S L, Zhang C Y and He X M. 2020. Shape-aware semi-supervised 3D semantic segmentation for medical images//Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention. Lima, Peru: Springer: 552-561 [DOI: 10.1007/978-3-030-59710-8_54]
- Luo X D, Chen J N, Song T and Wang G T. 2021. Semi-supervised medical image segmentation through dual-task consistency//Proceedings of the 35th AAAI Conference on Artificial Intelligence. Virtually: AAAI: 8801-8809 [DOI: 10.1609/aaai.v35i10.17066]
- Milletari F, Navab N and Ahmadi S A. 2016. V-Net: fully convolutional neural networks for volumetric medical image segmentation//Proceedings of the 4th International Conference on 3D Vision (3DV). Stanford, USA: IEEE: 565-571 [DOI: 10.1109/3dv.2016.79]
- Oktaç O, Schlemper J, Le Folgoc L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla N Y, Kainz B, Glocker B and Rueckert D. 2018. Attention U-Net: learning where to look for the pancreas [EB/OL]. [2024-04-11]. <https://arxiv.org/pdf/1804.03999.pdf>
- Qin X B, Zhang Z C, Huang C Y, Dehghan M, Zaiane O R and Jagersand M. 2020. U2-Net: going deeper with nested U-structure for salient object detection. Pattern Recognition, 106: #107404 [DOI: 10.1016/j.patcog.2020.107404]
- Ronneberger O, Fischer P and Brox T. 2015. U-net: convolutional networks for biomedical image segmentation//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Shen Z Q, Cao P, Yang H, Liu X L, Yang J Z and Zaiane O R. 2023. Co-training with high-confidence pseudo labels for semi-supervised medical image segmentation//Proceedings of the 32nd International Joint Conference on Artificial Intelligence. Macau, China: Morgan Kaufmann: 4199-4207 [DOI: 10.24963/ijcai.2023/467]
- Shi Y Y, Zhang Y C and Wang S S. 2023. Competitive ensembling teacher-student framework for semi-supervised left atrium MRI segmentation//Proceedings of 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Istanbul, Turkey: IEEE: 2231-2234 [DOI: 10.1109/bibm58861.2023.10385958]
- Sohn K, Berthelot D, Li C L, Zhang Z Z, Carlini N, Cubuk E D, Kurakin A, Zhang H and Raffel C. 2020. FixMatch: simplifying semi-supervised learning with consistency and confidence//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: 596-608
- Su Y X and Xiao Z Y. 2023. Semi-supervised dual-task cross-consistency constraint network for 3D cardiac MRI segmentation. Journal of Image and Graphics, 28(4): 1198-1211 (苏逸欣, 肖志勇. 2023. 面向3维心脏MRI分割的半监督双任务交叉一致性约束网络. 中国图象图形学报, 28(4): 1198-1211) [DOI: 10.11834/jig.211019]
- Tarvainen A and Valpola H. 2017. Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 1195-1204
- Vu T-H, Jain H, Bucher M, Cord M, Pérez P. 2019. Advent: adversarial entropy minimization for domain adaptation in semantic segmentation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 2517-2526 [DOI: 10.1109/cvpr.2019.00262]
- Wang K P, Zhan B, Zu C, Wu X, Zhou J L, Zhou L P and Wang Y. 2021. Triple-uncertainty guided mean teacher model for semi-supervised medical image segmentation//Proceedings of the 24th International Conference on Medical Image Computing and Computer Assisted Intervention. Strasbourg, France: Springer: 450-460 [DOI: 10.1007/978-3-030-87196-3_42]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Wu Y C, Ge Z Y, Zhang D H, Xu M F, Zhang L, Xia Y and Cai J F. 2022a. Mutual consistency learning for semi-supervised medical image segmentation. Medical Image Analysis, 81: #102530 [DOI: 10.1016/j.media.2022.102530]
- Wu Y C, Wu Z H, Wu Q Y, Ge Z Y and Cai J F. 2022b. Exploring smoothness and class-separation for semi-supervised medical image segmentation//Proceedings of the 25th International Conference on Medical Image Computing and Computer Assisted Intervention. Singapore, Singapore: Springer: 34-43 [DOI: 10.1007/978-3-031-16443-9_4]
- Wu Y C, Xu M F, Ge Z Y, Cai J F and Zhang L. 2021. Semi-supervised left atrium segmentation with mutual consistency training//Proceedings of the 24th International Conference on Medical Image Computing and Computer Assisted Intervention. Strasbourg, France: Springer: 297-306 [DOI: 10.1007/978-3-030-87196-3_28]
- Yu L Q, Wang S J, Li X M, Fu C W and Heng P A. 2019. Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium seg-

- mentation//Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention. Shenzhen, China: Springer: 605-613 [DOI: 10.1007/978-3-030-32245-8_67]
- Zhang Y C, Jiao R S, Liao Q C, Li D Y and Zhang J C. 2023. Uncertainty-guided mutual consistency learning for semi-supervised medical image segmentation. *Artificial Intelligence in Medicine*, 138: #102476 [DOI: 10.1016/j.artmed.2022.102476]
- Zhang Y C, Yang J, Liu Y C, Cheng Y and Qi Y. 2024. SemiSAM: enhancing semi-supervised medical image segmentation via SAM-assisted consistency regularization [EB/OL]. [2024-05-18]. <https://arxiv.org/pdf/2312.06316.pdf>
- Zhao X M, Liao Y H, Zhang S Q, Fang J X, He X X, Wang G Y and Lu H S. 2024. Method of classifying benign and malignant breast tumors by DCE-MRI incorporating local and global features. *Journal of Image and Graphics*, 29(1): 256-267 (赵小明, 廖越辉, 张石清, 方江雄, 何遐遐, 汪国余, 卢洪胜. 2024. 融合局部与全局特征的DCE-MRI乳腺肿瘤良恶分类. *中国图象图形学报*, 29(1): 256-267) [DOI: 10.11834/jig.230092]
- Zhong L F, Liao X, Zhang S T and Wang G T. 2023. Semi-supervised pathological image segmentation via cross distillation of multiple attentions//Proceedings of the 26th International Conference on Medical Image Computing and Computer Assisted Intervention. Vancouver, Canada: Springer: 570-579 [DOI: 10.1007/978-3-031-43987-2_55]
- Zhou X G, Li Z Q and Tong T. 2022a. DTSC-Net: semi-supervised 3D biomedical image segmentation through dual-teacher simplified consistency//Proceedings of 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Las Vegas, USA: IEEE: 1429-1434 [DOI: 10.1109/bibm55620.2022.9995371]
- Zhou Z Q, Qi L, Yang X, Ni D and Shi Y H. 2022b. Generalizable cross-modality medical image segmentation via style augmentation and dual normalization//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 20824-20833 [DOI: 10.1109/cvpr52688.2022.02019]
- Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N and Liang J M. 2018. UNet++: a nested U-net architecture for medical image segmentation//Proceedings of the 4th International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada, Spain: Springer: 3-11 [DOI: 10.1007/978-3-030-00889-5_1]

作者简介

赵小明,男,二级教授,主要研究方向为图像处理、机器学习和模式识别。E-mail: tzxyzxm@163.com

石培炼,男,硕士研究生,主要研究方向为医学影像分割。E-mail: 221050067@hdu.edu.cn

王丹丹,女,高级实验师,主要研究方向为机器学习与多媒体信息处理。E-mail: dandanw0707@tzc.edu.cn

付有瑶,男,讲师,主要研究方向为机器学习与图像信号处理。E-mail: fuyouyao828@126.com

张石清,男,教授,主要研究方向为机器学习和模式识别。E-mail: tzcqsq@163.com

方江雄,男,教授,主要研究方向为医学图像分割和深度学习。E-mail: fangchj2002@163.com