

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2023)04-1198-14

论文引用格式: 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(04):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(04):1198-1211)[DOI:10.11834/jig.211019]

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

苏逸欣, 肖志勇*

江南大学人工智能与计算机学院, 无锡 214122

摘要: **目的** 针对人体组织器官及病灶区域的3维图像分割是计算机辅助医疗诊断的重要前提,是医学影像3维可视化的重要技术基础。深度学习方法在医学图像分割任务中的成功通常取决于大量有标注数据。半监督学习利用未标注数据容易获取的优点,在模型训练过程中使用少量标注数据和大量未标注数据进行学习,缓解了数据标注昂贵耗时的问题,在医学图像分割中受到了广泛关注。为更好地利用无标注数据,提升医学图像分割效果,提出一种新的一致性正则方法用于半监督3维医学图像分割。**方法** 模型以V-Net为基础架构,通过扩展网络结构,在均带有分割任务及回归任务属性的双任务主副解码器之间添加了用于正则化约束的交叉损失,构建了具有形状感知的基于双任务的交叉一致性正则网络SACC-Net(shape-aware cross-consistency regular network based on dual tasks),实现将数据层面和模型层面的扰动融合进多任务机制的一致性正则方法,使模型能够更好地利用未标注数据的有效先验信息,并且具有更好的泛化性能。**结果** 在MICCAI 2018(Medical Image Computing and Computer Assisted Intervention Society)心房分割挑战赛公布的数据集中的3维左心房核磁共振成像上验证本文算法,在仅使用训练集中10%的有标注数据的实验组中,Dice系数、Jaccard指数、HD(Hausdorff distance)距离和平均对称表面距离分别为88.01%、78.89%、8.19和2.09。在另一组仅使用20%的有标注数据的实验中Dice系数、Jaccard指数、HD距离和平均对称表面距离分别达到90.11%、82.11%、6.57和1.78。**结论** 本文提出的半监督分割模型性能显著,综合了数据、模型和任务层面的一致性正则约束,与其他半监督方法相比分割效果更好且具有更佳的泛化性能。

关键词: 医学影像;半监督学习;3维图像分割;一致性正则;核磁共振图像(MRI)

Semi-supervised dual-task cross-consistency constraint network for 3D cardiac MRI segmentation

Su Yixin, Xiao Zhiyong*

School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China

Abstract: Objective Three-dimensional (3D) image segmentation of human tissues, organs and lesion areas is projected for computer-aided diagnosis and medical images-related 3D visualization. Thanks to the emerging deep learning technique, fully-supervised network models have been developing intensively in relevant to medical image segmentation tasks. However, it is challenged for a large amount of annotated data and 3D image segmentation data-labeled is costly and ineffi-

收稿日期:2021-10-28;修回日期:2022-05-19;预印本日期:2022-05-26

*通信作者:肖志勇 zhiyong.xiao@jiangnan.edu.cn

基金项目:江苏省优秀青年基金项目(BK20190079)

Supported by: Natural Science Foundation of Jiangsu Province for Excellent Young Scholars (BK20190079)

cient. Semi-supervised learning is focused on a small size of data-labeled and sufficient data-unlabeled in terms of easy acquisition of unlabeled data, which can alleviate the cost and time-consuming problem of data labeling. Our research is focused on new consistency regular method for semi-supervised 3D medical image segmentation model. To improve the medical image segmentation effect, our model can use unlabeled data through the fusion of different consistency methods.

Method The network model is demonstrated on the V-Net, which can remove the residual structure of encoding and decoding. To get efficient features of unlabeled data, the proposed shape-aware cross-consistency regular network is introduced via the V-Net network structure extension on the basis of an encoder and two independent decoders-involved dual tasks (shape-aware cross-consistency regular network based on dual tasks (SACC-Net)), which are divided into a main decoder a and an auxiliary decoder b. The output of encoder-shared is transmitted to the two decoders after noise disturbance. At the same time, the two decoders can output the prediction results after each iteration. To increase the generalization and anti-noise ability of the model, it can minimize the difference between the two parts of the results during the training process. Additionally, the proportion of labeled samples in the training samples is extremely small because the feature distributions between the pre-processed medical image samples are relatively similar. To improve the learning ability of the model to segmented samples further, geometric prior information constraints are melted into the segmentation target. A shape-aware regression layer is added at the end of each decoder as well. During the training phase, each decoder can output two parts of the prediction results at the same time. That is, the total output of each iteration-after network consists of four parts. It can be used to decode the segmentation map SA and the signed distance map output by the decoder A, and the segmentation map SB and the signed distance map output by the decoder B through the dual-task consistency of each decoding part. To enhance the model's ability and learn the effective features of segmentation targets to a greater extent, constraints and cross constraints can be used to realize a consistent regular method that combines data-level and model-level disturbances with multi-task mechanisms, and make better use of unlabeled data. **Result** Our algorithm is validated on the MRI data set published in Atrium Segmentation Challenge held by MICCAI (Medical Image Computing and Computer Assisted Intervention Society) in 2018. The experiment is divided into two test groups based on the amounts proportion of labeled data. In the training set, 10% annotated data is used only in the experimental group, the Dice coefficient, Jaccard index, HD (Hausdorff distance) distance, and average symmetric surface distance is reached to 88.01%, 78.89%, 8.19, and 2.09 of each. In the other group, 20% annotated data of the experiments are used only. The median Dice coefficient, Jaccard index, HD distance and average symmetric surface distance can be reached to 90.14%, 82.11%, 6.57, and 1.78 each as well. Furthermore, In respect of the shape perception method using the level set function for regression tasks, the Dice evaluation index can be improved by 0.69 and 0.60 in comparison with shape-aware semi-supervised net (SASS Net) in 10% and 20% of the marked training results. Each improvement is reached to 1.44% and 0.72% in terms of comparative results of dual-task consistency (DTC) trained with 10% and 20% labeled data. **Conclusion** The semi-supervised segmentation model (SACC-Net) is illustrated for the criteria optimization for both of the region and boundary-based segmentation, which can incorporate the level-consistency of its data, model and task. The constrained method has its potential segmentation effect and generalization performance for semi-supervised methods.

Key words: medical imaging; semi-supervised learning; three-dimensional image segmentation; consistency regularization; magnetic resonance imaging (MRI)

0 引言

语义分割是医学影像分析中的一项基本任务,广泛应用于医学图像中细胞、组织或器官等感兴趣区域的自动划分。核磁共振成像(magnetic resonance imaging, MRI)技术是临床诊断的重要辅助手段之一,以对人体没有电离辐射损伤的非入侵成像

方式提供具有良好图像质量和良好的软组织对比度的图像,MRI已成为精确识别患者心脏结构和病因、指导诊断和生成治疗决策的标准(Karamitsos等, 2009)。3维(three-dimensional, 3D)核磁共振图像中左心房(left atrium, LA)的自动分割可以使医生直观地看到被分割心房的整体形状及大小,还可以通过观察具有空间关系的多幅2维切片来判断病症,对于促进心房颤动的治疗具有重要意义。有了大量

有标注数据并且随着更强大的计算硬件的发展和临床数据库的增长,越来越多的研究人员使用深度学习方法来促进发展更稳健的心脏影像分割方法(Suinesiaputra等,2011),深度学习极大地促进了左心房的分割效果(Xiong等,2019)。此外,在MICCAI 2018 (Medical Image Computing and Computer Assisted Intervention Society)心房分割挑战赛发布的17种方法中,有15种使用深度学习从MRI中分割LA,获得了高精度的结果,优于传统的分割方法(Pop等,2019)。这一成果与2013年的心房分割挑战赛形成了鲜明对比,当时只有一种方法使用了深度学习算法(Tobon-Gomez等,2015)。因此,随着深度学习的应用成果日益增长,心房分割和更广泛的临床影像学发展也随之发生转变,越来越多地转向基于深度学习的方法。

然而,对于医学成像领域中的3D数据,由经验丰富的专家以逐层的方式从3D医学图像中描绘可靠的标注既昂贵又烦琐耗时。由于未标注数据通常是大量且丰富的,本文重点研究利用少量的有标注3D MRI数据和丰富的无标注3D MRI数据对左心房进行分割的半监督学习方法。近年来,半监督图像分割方面的工作主要集中在将未标注数据利用到卷积神经网络的训练中。一类方法主要考虑半监督分割的一般设置,其中大多数采用对抗性学习(adversarial learning)或一致性损失(consistency loss)作为正则化方式,以利用未标注的数据进行模型学习。对抗式学习的方法(Zhang等,2017;Nie等,2018;Zheng等,2019)强制未标记与标记图像的分割分布接近。Nie等人(2018)利用对抗性学习,以高置信度选择未标记图像的区域,以训练分割网络。而一致性损失方法(Tarvainen和Valpola,2017;Li等,2018;Bortsova等,2019;Yu等,2019)利用均值教师方法(mean teacher, MT)(Tarvainen和Valpola,2017)中的师生网络模型进行设计,并要求其输出在输入图像的随机扰动或变换下保持一致。Yu等人(2019)介绍了一种基于MT框架的教师网络利用不确定性指导学生网络学习的方法。尽管这些方法的结果很值得展望,但它们缺乏对语义对象的几何先验信息的建模,常常导致分割对象的覆盖率和边界预测较差。

还有一种半监督方法试图通过在模型学习中结合对感兴趣对象的解剖先验知识来解决上述缺点(Zheng等,2019;He等,2019)。例如,Zheng等人

(2019)介绍了基于深度学习的图集先验信息模型(deep atlas prior, DAP),该模型在损失函数设计中对形状概率进行了编码。He等人(2019)提出了一种自动编码器来学习未标记数据集的先验解剖特征。然而,这种先验知识通常假定输入图像正确对齐,这在实际中对于姿势或形状变化较大的对象很难实现,因此本文方法选择Li等人(2020)的水平集回归方法对目标分割物获取具有动态形状感知的隐式先验信息。

此外,在以往的半监督图像分割方法中,主要以数据层面或模型层面扰动的一致性正则以及近期出现的任务层面的任务间一致性正则来使网络模型更好地利用无标注数据的特征。本文旨在将数据模型层面的扰动结合任务间的一致性正则来增强网络模型的分割准确度以及泛化能力。为了实现这一点,本文设计了一个多解码器结合多任务结构的深度神经网络,该网络中输入数据通过共享编码器后经过数据扰动进入主解码器和辅助解码器,通过分割任务和回归任务分别得到像素级分割图和带符号距离图(signed distance map, SDM)(Luo等,2021a),SDM中为每个像素分配一个值,指示其到目标对象最近边界的有符号距离,使模型中增加了一种隐式形状先验,而后对所有预测结果使用交叉一致性正则化来优化分割网络。

在左心房分割挑战数据集(Xiong等,2021)上对本文方法进行评估,并与现有半监督方法进行广泛比较。实验结果表明,本文方法利用未标注的数据实现了性能提升,可以对分割目标进行高质量的全局形状分割。

1 相关工作

1.1 用于图像分割的深度学习神经网络

深度学习神经网络中各层的神经元数量、类型和连接定义了网络体系结构。其中,卷积神经网络(convolutional neural network, CNN)模型是图像分析中应用最广泛的模型之一。CNN是一种特定的神经网络架构,其隐藏层由一个或多个卷积层组成。卷积层作为输入图像的特征提取器,每个神经元从前一层的多个相邻神经元接收其输入,共享权值,并保留最相关的空间信息。通常在连续卷积层之间插入池化层,用于降低每个生成特征图的维数并保留相关信息以

减少计算成本,同时也扩大后续卷积层的视野。

对于专门用于图像分割的CNN架构,Long等人(2015)首先提出了用于语义分割的全卷积网络(fully convolutional network, FCN)体系结构,随后人们对其进行了调整和进一步扩展,特别是提出了U-Net(Ronneberger等,2015)。这是一种用于生物组织图像分割的U形体系结构,如图1(a)所示,包含一个特征提取的编码器部分和一个通过上卷积进行像素级预测的解码器部分。U-Net架构通过使用下采样特征图和上采样特征图之间的跳跃连接,允许特

征在网络的编码部分和解码部分之间传递,防止奇点并且实现更高的精度(Drozdal等,2016;Mao等,2016;Orhan和Pitkow,2018;何康辉和肖志勇,2021)。而后,为了学习到3D医学影像中切片间的关系获取更加有效的信息,Çiçek等人(2016)和Milletari等人(2016)进一步发展了U形结构,提出广泛用于3D图像分割的基础架构3D U-Net和V-Net,这两种网络模型使用3D图像作为输入,更准确地分割出解剖对象的空间结构。本文方法使用V-Net作为基础网络框架,如图1(b)所示。

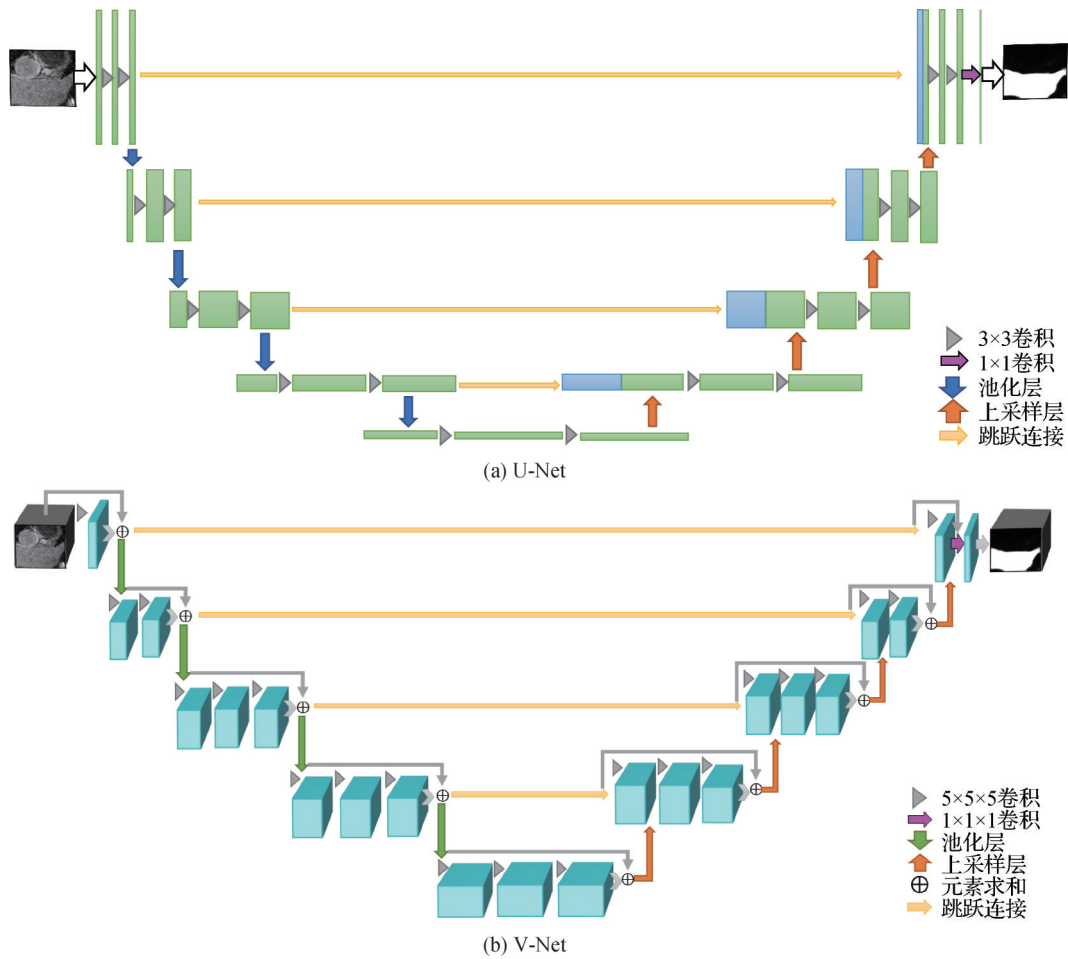


图1 U-Net和V-Net的原网络模型结构图

Fig. 1 The structure diagram of the original network model of U-Net and V-Net ((a) U-Net; (b) V-Net)

1.2 基于半监督的医学图像分割

半监督学习(semi-supervised learning, SSL)广泛应用于各种计算机视觉任务中。目前深度学习方法在医疗实践的各个领域蓬勃发展,半监督学习在有标注数据匮乏的情况下,在医学图像分割中取得了良好效果。为了利用未标注数据进行半监督医学图像分割,一种简单直观的方法是给未标注数据生成伪标

签。伪标签通常是通过迭代生成的,其中模型通过学习自身对未标注数据的预测来迭代提高伪标签的质量(Bai等,2017)。Chen等人(2021)使用锐化函数或固定阈值将概率映射转换为伪标签,在伪标签的监督下,模型可以学习并产生低熵的预测结果。Rizve等人(2021)将概率和不确定性阈值结合起来,为SSL选择最准确的伪标签。尽管带有伪标签的半监督学习已

展示出良好的性能,但模型生成的标签仍然会有噪声,并对后续分割模型产生不利影响。使用对抗性学习来利用未标注的数据也是一种流行的半监督医学图像分割方法。Zhang等人(2017)提出一种新的基于深度对抗网络的医学图像分割模型,利用此方法激励未注释图像的分割与注释图像的分割达到相似。同时,Miyato等人(2019)和Mittal等人(2021)进一步采用对抗性训练,应用一致性约束来增强SSL。此外,通过使用指数移动平均(exponential moving average, EMA)操作,在平均教师模型MT(Tarvainen和Valpola, 2017)中讨论了模型参数的一致性。Yu等人(2019)扩展了MT模型,使用不确定性来指导其学生神经网络进行半监督左心房分割。之后,Li等人(2020)引入了一种具有形状感知的半监督分割策略,利用未标注数据对分割输出实施几何形状约束,实现了隐式先验信息的融入。进一步考虑到不同任务的一致性,

Luo等人(2021a)将多任务框架应用于分割和水平集函数回归,从而在任务级别加强了一致性表达。

1.3 一致性正则化

一致性正则化(Bachman等, 2014; Sajjadi等, 2016; Bortsova等, 2019)利用未标记数据时依据的假设是,对于类似的输入,模型应该产生一致的预测,图2(a)是在单分割模型下对同一个无标注数据进行数据扰动后产生的预测进行一致性正则化,图2(b)是对同一个无标注数据在双模型中产生的预测进行一致性正则化。一致性正则化使训练后的模型在输入扰动或转换过的数据后,可以输出相同的分类预测或分割预测。这类方法通过最小化深度神经网络对传递扰动或数据扰动后得到的预测之间的差异,以获得具有更好泛化能力的模型。例如,Ouali等人(2020)对中间特征图进行了若干数据增强操作,并约束模型输出不变的分割预测图。

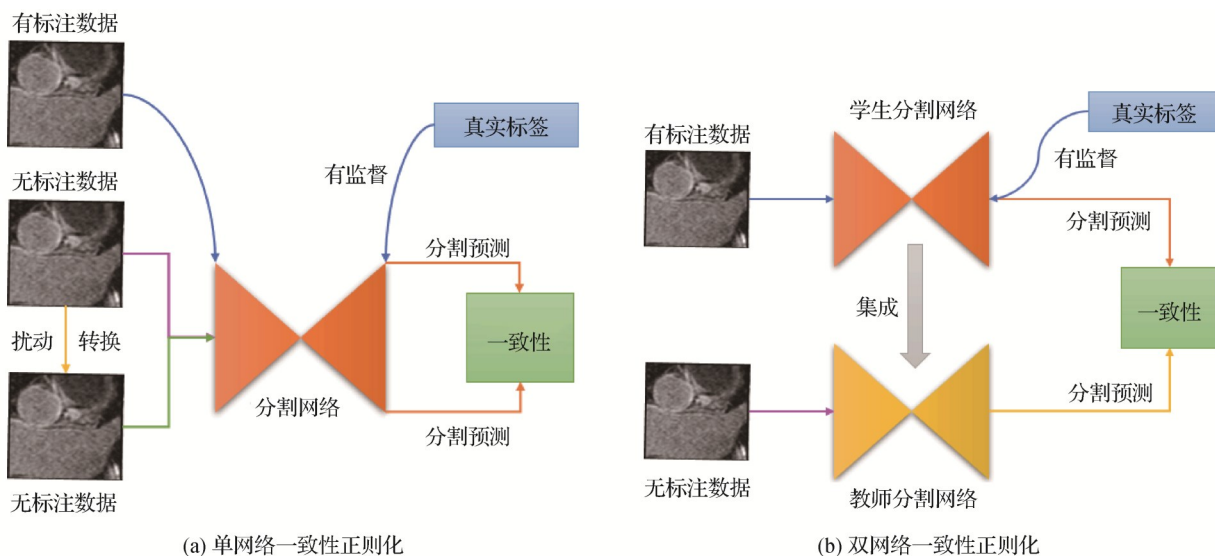


图2 一致性正则化基本结构

Fig. 2 Basic structure of consistency regularization

((a) single network consistency regularization; (b) double network consistency regularization)

一致性正则化方法中大多使用图2(b)的结构。例如,对于不同噪声扰动下有相同输入的教师和学生模型,Yu等人(2019)进一步将不确定性估计引入MT模型。Li等人(2020)将几何变换一致性损失引入MT模型中,应用于病灶、眼底视网膜和肝脏肿瘤的半监督分割。在MT框架中,Zhou等人(2020)对无标注数据的利用是通过教师和学生网络在小扰动下的预测在特征和语义水平上保持一致,并将模型应用于宫颈细胞的半监督分割。Fotedar等人

(2020)明确强调了MT框架在扰动下不同样本间关系的一致性。Bortsova等人(2019)没有使用自训练策略(即MT范式),而是通过一种暹罗网络,评估和学习一个输入样本在通过不同的变换方式后生成的一对输出样本之间的相似度,以此来强制产生有标注和无标注数据之间转化的一致性,并在胸部X射线图分割上实现了最先进的性能。Laradji等人(2021)对CT(computed tomography)图像中的COVID-19(corona virus disease 2019)的弱监督分割也采用了类

似的思想。此外, Sohn等人(2020)利用训练样本在弱扰动和强扰动下的一致性来促进模型的训练。Wang等人(2022)则利用特征空间中的语义方向实现语义数据增强, 从而对模型进行一致性约束。对于半监督医学图像分割, Peng等人(2020)进一步采用基于交互信息的聚类损失, 明确加强了未标注图像和随机扰动后的未标注图像中邻近像素之间的预测一致性。

2 本文方法

2.1 网络模型概述

本文设计的用于半监督左心房分割的具有形状感知的基于双任务的交叉一致性正则网络(shape-aware cross-consistency regular network based on dual tasks, SACC-Net)的总体结构如图3所示。原始的V-Net由一个编码器部分和一个解码器部分组成, 深层特征图经过上采样后, 与从编码器部分通过跳跃连接而来的浅层特征图相结合, 以弥补编码器中的池化操作带来的特征损失, 从而使神经网络可以加速学习到分割目标的特征。本文方法使用的网络模型以UA-MT(uncertainty-aware mean teacher)模型(Yu等, 2019)去除编码部分和解码部分残差结构的V-Net为主干。为了更好地利用无标注数据的有价值的特

征, SACC-Net包含一个编码器和两个相互独立的解码器, 分为主解码器A和辅助解码器B, 对共享编码器的输出分别添加噪声扰动 δ 和 δ' 后分别传入解码器A和解码器B, 两个解码器在每次迭代后同时输出预测结果。在训练过程中, 最终使两部分结果获得最小化差异, 以增加模型的泛化性以及抗噪能力。

另外, 虽然训练样本中有标签样本的占比极小, 但由于通过预处理后的医学影像样本之间的特征分布较为相似, 因此考虑通过添加针对分割目标的几何先验信息约束来进一步提升模型对分割样本的学习能力。在每个解码器的最后分别多添加了一层形状感知回归层, 在训练阶段使每个解码器同时输出两个任务即分割任务(任务1)及回归任务(任务2)的预测结果, 即网络经过每次迭代的总体输出共包含4部分, 分为主解码器A输出的分割图SA、带符号距离图SDM_A, 以及辅助解码器B输出的分割图SB、带符号距离图SDM_B, 通过对每个解码部分分别进行双任务一致性的自我约束和解码器间的交叉约束来更大程度地增强模型对分割目标的有效特征学习能力。

2.2 算法原理

2.2.1 基于模型扰动和数据扰动的一致性正则

在上述整体半监督分割架构SACC-Net中, 设计了基于V-Net结构的双解码器结合双任务预测的方

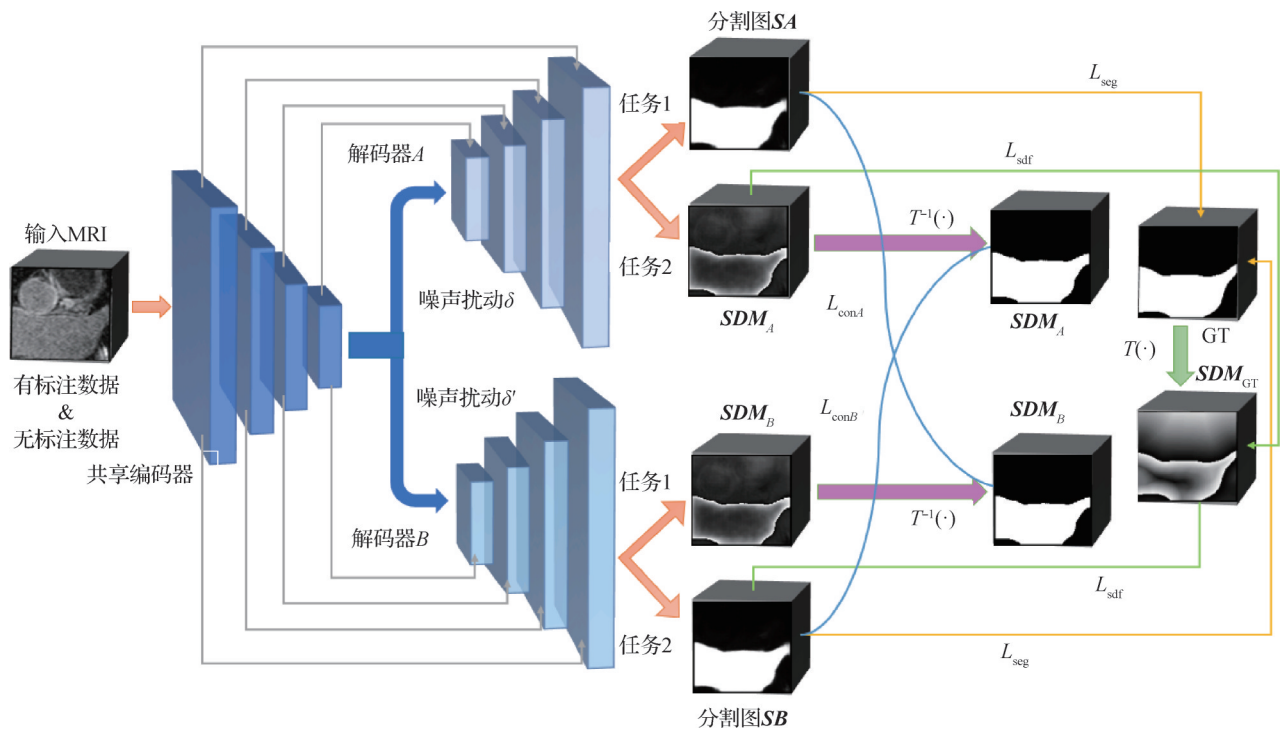


图3 SACC-Net整体网络结构

Fig. 3 SACC-Net overall network structure

法。在本文方法中,训练模型输入的数据增强后的3D MRI样本 $X \in \mathbf{R}^{H \times W \times D}$ 中包括 N 个有标注样本和 M 个无标注样本,其中 $N \ll M$ 。定义有标注数据为 $D^l = \{X_n, Y_n, Z_n\}_{n=1}^N$, 无标注数据为 $D^u = \{X_m\}_{m=N+1}^{N+M}$, 共享编码器连同添加了噪声扰动的主解码器 A 和辅助解码器 B 一起执行 N 次随机正向传播,如图3所示,当输入有标注样本 $X_n \in \mathbf{R}^{H \times W \times D}$ 时, $Y_n \in \{0, 1\}^{H \times W \times D}$ 是样本的真实分割标签, $Z_n \in \mathbf{R}^{H \times W \times D}$ 是从样本的真实分割标签 (ground truth, GT) Y_n 通过水平集函数 $T(\cdot)$ 转换而来的带符号距离图 SDM_{GT} 。

在训练过程中,主解码器 A 和辅助解码器 B 通过学习添加了不同扰动的训练样本来相互独立地进行反向传播。此外,由于网络模型中的两个解码器还添加了 dropout 层来进行模型层面的扰动,并且由于神经网络模型参数的优化过程本身也具有随机性,从而会使两个解码器分别进行反向传播时均可以对用来提取特征的共享编码器里的参数进行优化,使编码器在特征提取能力增强的同时提升了模型整体对扰动的抵抗能力和泛化能力,实现了数据层面与模型层面的扰动相结合。

2.2.2 基于先验知识的任务间一致性正则

本节介绍将双任务一致性融合进多解码器的训练架构。如上述整体框架图(图3)所示,每个解码器由末尾的两个任务部件组成,即在训练过程中可同时输出像素级分割概率图和传统回归任务中通过水平集函数表示的分割目标几何形状图。本文利用先验知识时,使用的水平集函数为物体的3D轮廓的一种表示方式,可使用定义的函数公式将输入的3D图像上的3维空间点 x 映射到1维空间,从而产生输入图像所对应的带符号距离图 SDM 。由于分割结果可以由像素级的分割标签图和水平集函数(Luo等,2021a)捕捉的全局几何轮廓来表示,因此这两个预测任务应当具有一致性。为了将任务间一致性约束融合进半监督分割框架,更好地利用无标注数据,将这种双任务一致性策略应用在双解码器之间,通过最小化两个解码器间的相互交叉一致性损失,进一步通过几何先验知识增强模型的分割性能。

为了建立主解码器和辅助解码器之间的几何约束一致性,使用转换式 $T^{-1}(\cdot)$ 将回归任务生成的带符号距离图 SDM_A 和 SDM_B 分别转换为对应的像素

级的分割概率图 SDM_{segA} 和 SDM_{segB} , 由平滑 Heaviside 函数实现。本文方法中应用的将分割标签图转化为带符号距离图的转换层 $T(\cdot)$ 为

$$T(x) = \begin{cases} -\inf_{y \in \partial D} \|x - y\|_2 & x \in D_i \\ 0 & x \in \partial D \\ +\inf_{y \in \partial D} \|x - y\|_2 & x \in D_o \end{cases} \quad (1)$$

式中, $\|x - y\|$ 表示两个不同体素点 x 和 y 之间的欧氏距离, \inf 表示计算取距离值的最大下界, ∂D 是0水平集即0距离,代表分割目标的边界轮廓, D_i 表示分割目标内部, D_o 表示分割目标外部。因此在一般情况下,转换式的值在像素点属于分割目标内部时取负值,反之取正值。由于输入的图像样本具有不同的视野和不同的分割目标体积,进一步在输出层使用 tanh 激活函数将 $T(\cdot)$ 归一化,使每个输入样本得到的值在 $[-1, 1]$ 。通过转换式 $T(\cdot)$ 即可将训练样本中有标注数据的 GT 转化为对应的带符号距离图 SDM_{GT} 。此外,由于转换式的函数表达有不可微分的性质,所以无法直接用转换式的逆向算式将带符号距离图 SDM_A 和 SDM_B 直接转化为分割概率图。对此,使用 Luo 等人(2021a)方法中的平滑近似方式进行逆转换 $T^{-1}(\cdot)$, 具体为

$$T^{-1}(z) = \frac{1}{1 + e^{-k \cdot z}}, \quad z \in SDM \quad (2)$$

式中, z 是体素 x 在带符号距离图 SDM 中的值, k 应选择尽可能大的变换因子,以实现平滑的近似变化效果。通过以上转换式和逆转换式,模型中的双任务输出得以进行任务间一致性的优化操作。

2.3 损失函数

如图3所示,为了实现上述方法中描述的数据模型和任务间的一致性优化,本文使用的损失函数包含两部分,即作用在有标注样本上的有监督训练损失 L_{seg} 、 L_{sdf} 和应用于包含无标注数据的全部训练样本的无监督训练损失 L_{conA} 、 L_{conB} 。其中,对于有标注的训练样本 $D^l = \{X_n, Y_n, Z_n\}_{n=1}^N$, 定义3D样本中的体素 $(x, y) \in (X, Y)$, 训练过程中主解码器 A 和辅助解码器 B 输出的两个像素级预测概率图通过 sigmoid 函数生成分割图 SA 和分割图 SB , 选用 Dice 损失函数进行分割图和真实标签之间的损失计算,具体计算为

$$L_{\text{seg}}(x, y) = \frac{1}{2} \left(\sum_{x_i, y_i \in D'} \left(1 - \frac{2 \times \sum_{x_j \in x_i, y_j \in y_i} f_{A1}(x_j) \times y_j}{2 \times \sum_{x_j \in x_i, y_j \in y_i} f_{A1}(x_j) + \sum_{y_j \in y_i} y_j} \right) + \sum_{x_i, y_i \in D'} \left(1 - \frac{2 \times \sum_{x_j \in x_i, y_j \in y_i} f_{B1}(x_j) \times y_j}{2 \times \sum_{x_j \in x_i, y_j \in y_i} f_{B1}(x_j) + \sum_{y_j \in y_i} y_j} \right) \right) \quad (3)$$

式中, $\sum_{x_j \in x_i, y_j \in y_i}$ 和 $\sum_{x_i, y_i \in D'}$ 分别表示3D图像样本中的体素和训练数据集中的有标注样本, $f_{A1}(\cdot)$ 和 $f_{B1}(\cdot)$ 表示

解码器输出的分割图 SA 和分割图 SB 。两个解码器在输出分割图的同时, 输出的带符号距离图 SDM_A 和 SDM_B 分别与真实标签通过上述逆转变式 $T(\cdot)$ 生成的 SDM_{CT} 进行损失计算, 具体计算为

$$L_{\text{sdf}}(x, y) = \sum_{x_i, y_i \in D'} \left\| f_{A2}(x_i) - T(y_i) \right\|^2 + \sum_{x_i, y_i \in D'} \left\| f_{B2}(x_i) - T(y_i) \right\|^2 \quad (4)$$

式中, $f_{A2}(\cdot)$ 和 $f_{B2}(\cdot)$ 表示解码器输出的 SDM_A 和 SDM_B 。从而有监督部分的总体损失函数定义为

$$L_{\text{sup}} = L_{\text{seg}} + \beta L_{\text{sdf}} \quad (5)$$

式中, β 是平衡两个损失项的加权系数。为了更好地利用大量的无标注样本, 本文将任务一致性与模型一致性相融合, 设计了一种在两个解码器各自的两项任务输出之间用一种一致性损失函数进行交叉约束, 面向全体训练集数据, 训练优化过程中分别对主解码器 A 的分割任务输出的概率图 A 和辅助解码器 B 的回归任务输出的距离图之间, 以及辅助解码器 B 的分割任务输出的概率图 B 和主解码器 A 的回归任务输出的距离图之间进行一致性优化操作, 具体计算为

$$L_{\text{cross}}(x) = \sum_{x_i \in D} \left\| f_{A1}(x_i) - T^{-1}(f_{B2}(x_i)) \right\|^2 + \sum_{x_i \in D} \left\| f_{B1}(x_i) - T^{-1}(f_{A2}(x_i)) \right\|^2 \quad (6)$$

式中, $\sum_{x_i \in D}$ 表示训练集中的所有样本。由上可得出

本文设计的半监督图像分割模型的总体损失函数为

$$L = L_{\text{sup}} + \alpha L_{\text{cross}} \quad (7)$$

式中, α 为用来控制有监督训练损失与无监督训练损失之间平衡的超参数, 从而稳定整体训练的损失权重。本文使用与时间相关的高斯 warm-up 函数

(Tarvainen 和 Valpola, 2017; Yu 等, 2019) 的退火策略来使其缓慢增加。

3 实验和讨论

3.1 数据集

为了评估本文方法, 将算法应用于 MICCAI (Medical Image Computing and Computer Assisted Intervention) 在 2018 年发布的心房分割挑战赛 (Xiong 等, 2021) 中提供的 3D 左心房 MRI 公开数据集。该数据集是一种广泛用于半监督医学图像分割定量分析的数据集 (Yu 等, 2019; Li 等, 2020; Luo 等, 2021a), 由 100 幅 3D 左心房 MRI 样本和专家标注的左心房分割标签组成, 分辨率为 $0.625 \text{ mm} \times 0.625 \text{ mm} \times 0.625 \text{ mm}$, 各向同性分辨率为 $1.0 \text{ mm} \times 1.0 \text{ mm} \times 1.0 \text{ mm}$ 。实验时, 将其中 80 个样本用做训练集, 20 个样本用做测试集 (Yu 等, 2019), 并采用相同的数据预处理方法, 为了提高分割精度, 将所有 MRI 裁剪成体素尺寸为 $112 \times 112 \times 80$ 的图像, 丢弃了无用混杂背景。

在本文工作中, 使用了两种比例对训练集进行有标注和无标注样本划分, 分别训练和验证所提方法的性能, 即分别设定 80 个训练样本中的 10% 和 20% 的样本作为两组中的有标注样本, 这是典型的半监督学习实验设置。另外, 由于半监督学习中对于训练过程中少量有标注数据的选取对模型的表现结果会有一些的偏差性影响, 由此基于 Yu 等人 (2019) 方法中设定的数据划分比例, 从数据集总体 100 个样本中基于 5 折交叉验证方法随机划分为 5 组数据分别进行训练和测试, 将所得分割结果平均后用于模型评估, 以此减少结果的偶然性, 增加模型的鲁棒性。

3.2 实验细节及参数设置

实验使用 NVIDIA GTX1080Ti 单个显卡进行模型训练, 本文所有网络模型均在 PyTorch 平台中实现。使用 V-Net (Milletari 等, 2016) 作为文中所有实验的网络主干, 并在原始 V-Net 中添加解码器 B , 将解码器 A 和解码器 B 的末尾添加回归层在生成分割图的同时实现形状感知任务, 模型中的解码器 B 的噪声扰动设置为 $[-0.2, 0.2]$ 。该框架采用 SGD (stochastic gradient descent) 优化方法, 训练时迭代 6 000 次, 初始学习率设置为 0.01, 每 2 500 次迭代后

衰减为原来数值的1/10, batch大小为4, 由2个有标注样本和2个无标注样本组成。

在所有的实验中, 将 β 设置为0.3, 将 α 设置为与时间相关的高斯升温函数 $\alpha(t) = 0.001 \times e^{-s\left(1 - \frac{t}{t_{\max}}\right)^2}$, 其中 t 表示迭代的次数, t_{\max} 表示最大迭代次数。与Luo等人(2021a)的方法一样, 实验中, k 的值设置为1500, 为了避免训练过程中的过拟合现象, 使用了标准的数据增强方法, 包括将图像沿轴面随机翻转和旋转 90° 、 180° 和 270° 。在测试过程中, 使用滑动窗口策略获得未加噪声扰动的解码器 A 的最终分割结果用于模型评估, stride设置为步幅为 $18 \times 18 \times 4$ 。此外, 采用非极大值抑制(non-maximum suppression, NMS)作为后处理, 用以去除孤立的无关区域。

3.3 评价指标

为了评估本文模型的性能, 选用了4个相互补充的评价指标, 包括Dice系数(Dice coefficient, DSC)、Jaccard指数(Jaccard)、95%的Hausdorff距离(95% Hausdorff distance, 95HD)和平均对称表面距离(average surface distance, ASD)。

Dice和Jaccard是基于区域的度量指标, 具体定义为

$$DSC = \frac{2 \times TP}{2 \times TP + FN + FP} \quad (8)$$

$$Jac = \frac{TP}{FP + FN + TP} \quad (9)$$

式中, TP (true positive)为真阳性, 即左心房的像素点中正确分类为左心房的像素点数目; FP (false positive)为假阳性, 即分类为左心房部分却实际为背景点的像素点数目; TN (true negative)为真阴性, 表

示正确分类的背景点的像素点数目; FN (false negative)为假阴性, 表示实际属于左心房却分类为背景的像素点数目。

95HD和ASD用于评估预测结果与真实标签之间的边界误差。具体定义为

$$95HD = \max_{t_2} \left(\max_{t_1} \left(\sqrt{T^2 - P^2} \right) \right) \times 0.95 \quad (10)$$

$$\begin{cases} B_{\text{seg}} = \left\{ \forall t1 \in P, \min \left(\sqrt{t1^2 + t2^2} \right) \mid \exists t2 \in T \right\} \\ B_{\text{gt}} = \left\{ \forall t2 \in T, \min \left(\sqrt{t1^2 + t2^2} \right) \mid \exists t1 \in P \right\} \\ ASD = \text{mean} \left\{ B_{\text{seg}}, B_{\text{gt}} \right\} \end{cases} \quad (11)$$

式中, T 表示专家手动分割的边界, P 表示模型预测出的分割区域的边界, $t1$ 、 $t2$ 均为边界像素点, 其中, $t1$ 属于 P , $t2$ 属于 T 。

3.4 结果与分析

3.4.1 消融实验与分析

首先, 验证在双任务输出情况下, 增加带有数据扰动的辅助解码器后的效果, 实验结果如表1所示。表1中第1、2行是含有单个解码器 A 和其双任务输出的对比模型(D_A+DT_A), 第3、4行是含有主解码器 A 和辅助解码器 B 且两个解码器均为双任务输出, 但是没有解码器 A 和 B 之间的交叉一致性约束的对比模型(D_A+D_B+DT_A&B(无交叉评估))。训练过程中的可视化结果如图4, 图中第1、2行分别为左心房MRI训练样本的2维切片及其对应的真实标签GT, 样本经过共享编码器并进行噪声扰动后分别通过含有dropout解码器 A 和 B , 同时得到分割任务的分割图(第3、5行)以及回归任务的带符号距离

表1 16(20%)个标签情况下的消融实验测试结果

Table 1 Test results of ablation experiments with 16 (20%) labels

对比方法	主干网络	Dice/%	Jaccard/%	95HD/体素	ASD/体素
D_A+DT_A	3D U-Net	87.59	78.08	10.1	1.96
D_A+DT_A	V-Net	88.69	79.99	8.24	2.22
D_A+D_B+DT_A&B(无交叉评估)	3D U-Net	87.89	78.55	9.38	1.83
D_A+D_B+DT_A&B(无交叉评估)	V-Net	89.60	81.30	7.35	1.97
D_A+D_B+DT_A(无交叉评估)	3D U-Net	88.35	79.27	9.25	1.88
D_A+D_B+DT_A(无交叉评估)	V-Net	89.47	81.12	8.85	2.00
D_A+D_B+DT_A&B(有交叉评估)(本文)	3D U-Net	87.84	78.46	9.63	1.82
D_A+D_B+DT_A&B(有交叉评估)(本文)	V-Net	90.14	82.17	6.62	1.83

注: 加粗字体表示各列最优结果。

图(第4、6行)。可以看出,不同的数据扰动和模型扰动会使模型在同一样本同一时间下得到带有差异的预测结果。由评估数据可以证明,在解码器拥有双任务的情况下,增加带有数据扰动的辅助解码器可以使模型具有更好的性能,Dice提升了0.91%,并且在两个随机优化的解码器的作用下,模型扰动使得模型更具有泛化性。

其次,验证本文方法利用的将模型扰动和任务

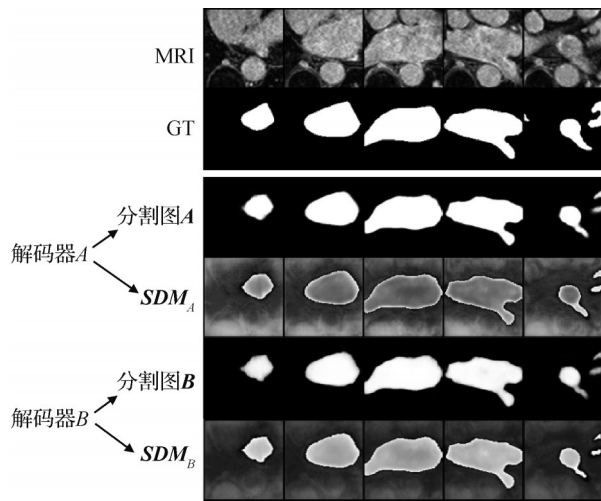


图4 主编码器A和辅助编码器B在LA数据集上的2D分割结果图

Fig. 4 2D segmentation results of the main encoder A and auxiliary encoder B on the LA dataset

扰动相结合的交叉一致性损失的有效性,结果如表1所示。表1最后两行是本文模型,即具有解码器A和B,且解码器均为双任务解码器(D_A+D_B+DT_A&B(有交叉评估))。第4行对比模型与之不同之处是缺少了交叉一致性损失函数(D_A+D_B+DT_A&B(无交叉评估)),第6行对比模型与之不同之处是仅解码器A是双任务解码器,解码器B只有分割任务而无回归任务(D_A+D_B+DT_A(无交叉评估))。从实验结果来看,在模型和任务扰动相融合的情况下,在解码器间进行交叉的任务一致性优化可以更大程度地提升半监督分割模型的效果,与无模型扰动且无交叉的任务一致性优化的模型相比,Dice提升了1.45%。

3.4.2 对比实验

在两组不同的数据划分设定下,将本文方法与最近的一些半监督分割方法进行对比,结果如表2和表3所示。为了验证方法的鲁棒性以及减小数据带来的偏差性,选择了两种通常半监督方法评估所设置的有标注样本在训练集中的占比,在LA数据集中随机划分出5组分别包含80个训练样本且分别设置8个(10%)有标注样本和16个(20%)有标注样本的实验组,根据5折交叉验证方法得到均值结果,实验均在相同的训练和测试样本上进行。

表2 本文模型在10%有标注的训练结果上与现有半监督模型的定量对比

Table 2 The quantitative comparison between this model and the existing semi-supervised model on the 10% labeled training results

对比方法	有标注	无标注	Dice/%	Jaccard/%	95HD/mm	ASD/mm
V-Net(Milletari等,2016)	8(10%)	0	79.99	68.12	21.11	5.48
V-Net(Milletari等,2016)	80(全部)	0	91.14	83.82	5.75	1.52
3D U-Net(Çiçek等,2016)	8(10%)	0	62.12	58.34	36.56	5.21
3D U-Net(Çiçek等,2016)	80(全部)	0	88.21	79.13	12.9	2.33
DAP(Zheng等,2019)	8(10%)	72	81.89	71.23	15.81	3.8
UA-MT(Yu等,2019)	8(10%)	72	84.25	73.48	13.84	3.36
SASS Net(Li等,2020)	8(10%)	72	87.32	77.72	9.62	2.55
LG-ER-MT(Hang等,2020)	8(10%)	72	85.54	75.12	13.29	3.77
DUWM(Wang等,2020)	8(10%)	72	85.91	75.75	12.67	3.31
DTC(Luo等,2021a)	8(10%)	72	86.57	76.55	14.47	3.74
PPNet(Luo等,2021b)	8(10%)	72	85.08	74.88	17.18	3.13
CVRL(You等,2022)	8(10%)	72	87.72	78.29	9.34	2.23
SACC-Net(本文)	8(10%)	72	88.01	78.89	8.19	2.09

注:加粗字体表示各列最优结果。

表3 本文模型在20%有标注的训练结果上与现有半监督模型的定量对比

Table 3 The quantitative comparison between this model and the existing semi-supervised model on the 20% labeled training results

对比方法	有标注	无标注	Dice/%	Jaccard/%	95HD/mm	ASD/mm
V-Net(Milletari等,2016)	16(20%)	0	86.03	70.06	14.26	3.51
V-Net(Milletari等,2016)	80(全部)	0	91.14	83.82	5.75	1.52
3D U-Net(Çiçek等,2016)	16(20%)	0	66.74	61.09	37.42	4.43
3D U-Net(Çiçek等,2016)	80(全部)	0	88.21	79.13	12.9	2.33
DAP(Zheng等,2019)	16(20%)	64	87.89	78.72	9.29	2.74
UA-MT(Yu等,2019)	16(20%)	64	88.88	80.21	7.32	2.26
CCT(Ouali等,2020)	16(20%)	64	88.83	80.06	8.44	2.49
SASS Net(Li等,2020)	16(20%)	64	89.54	81.24	8.24	2.20
LG-ER-MT(Hang等,2020)	16(20%)	64	89.62	81.31	7.16	2.06
DUWM(Wang等,2020)	16(20%)	64	89.65	81.35	7.04	2.03
DTC(Luo等,2021a)	16(20%)	64	89.42	80.98	7.32	2.10
CVRL(You等,2022)	16(20%)	64	89.87	81.65	6.96	1.72
PPNet(Luo等,2021b)	16(20%)	64	89.58	81.71	10.24	2.19
MT-HU-HS(Li等,2021)	16(20%)	64	90.04	81.98	6.93	2.18
SACC-Net(本文)	16(20%)	64	90.14	82.17	6.62	1.83

注:加粗字体表示各列最优结果。

本文使用现有的用于3D图像分割的主流网络主干框架3D U-Net和V-Net分别做了相应的消融对比实验(见表1)以及使用不同比例的标注数据的基准对比实验(见表2和表3)。表1的消融实验数据表明,以通过卷积操作实现下采样的V-Net网络框架为主体可以有更好的特征提取能力,从而使得效果更佳。表2和表3的基准网络实验数据证明,在不添加任何半监督方法的情况下,V-Net对少量有标注样本分割的表现也优于3D U-Net,因此本文中对比实验的主干网络框架均使用V-Net。

可以看出,与其他方法相比,本文方法在所有评估指标上都取得了最好的准确性,尤其是在基于边界的评判标准ASD和95HD有更加显著的提升。与仅使用10%和20%有标注样本且不使用无标注样本的基准方法V-Net相比,本文方法在使用相同数量的有标注样本的情况下增加无标注样本后在Dice评判标准上分别提升了8.02%(10%有标注)和4.11%(20%有标注),证明本文方法使无标注医学影像得到了非常有效的利用。

此外,在与同样使用水平集函数进行回归任务的形状感知方法的对比中,本文方法的Dice评价指

标,在10%和20%有标注的训练结果上,相比SASS Net(shape-aware semi-supervised net)分别提升了0.69%和0.60%,相比DTC(dual-task consistency)分别提升了1.44%和0.72%。这两组对比实验中,在其他3个评价指标上,本文方法相比而言效果也是最好。说明本文设计的数据模型扰动和任务间扰动相融合的交叉一致性优化方法将几何先验约束的优势更大程度地运用到了半监督医学影像分割中,使大量无标注医学影像样本的有效特征发挥了更加显著的作用。

本文方法与几种较先进分割方法的3D和2D分割结果的可视化对比如图5所示,从左到右分别为LA数据集中提供的真实标签、UA-MT(Yu等,2019)、SASS Net(Li等,2020)、DTC(Luo等,2021a)以及本文的SACC-Net。可以直观地看到,在测试集中同一个相对其他样本而言较为模糊因而较难分割的测试样本下,通过本文方法分割出的效果与专家标注的真实标签重叠性更高,预测出了更完整的左心房形状和更多细节,尤其是在分割目标的边界处,这进一步证明了本文提出的3D半监督医学影像分割方法的有效性。

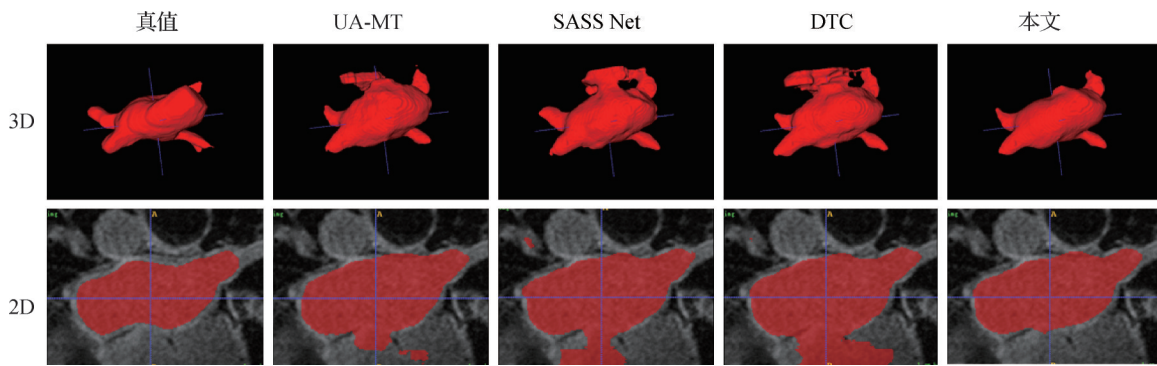


图5 半监督模型在LA数据集上的3D和2D分割结果对比图

Fig. 5 Comparison of the 3D and 2D segmentation results of the semi-supervised model on the LA dataset

4 结论

本文提出了一种具有形状感知的基于双任务的交叉一致性正则分割网络 SACC-Net, 应用于3维左心房核磁共振影像分割。本文方法使用双解码器双任务网络, 利用交叉损失函数对分割任务得到的分割预测图和回归任务得到的带符号距离预测图进行一致性约束, 使分割模型更具形状感知性, 实现了数据层面、模型层面和任务层面的一致性结合。

与其他先进的半监督方法相比, 本文的半监督3D MRI分割方法在经过多折交叉验证后得到的平均结果优于其他应用于本数据集的半监督分割方法中的最优结果, 实验数据偏差性更小, 且模型的分割精确度和性能鲁棒性更具优势。

如何进一步充分利用未标注数据, 以及如何使用更多有价值的领域先验知识来使模型具有更好的泛化性是一个极其值得探究的问题。因此今后将考虑将本文算法应用于不同的医学图像数据集, 并尝试将更多的解剖先验知识应用于半监督分割模型中, 期望进一步改善半监督医学图像分割的性能。

参考文献 (References)

- Bachman P, Alsharif O and Precup D. 2014. Learning with pseudo-ensembles//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press: 3365-3373
- Bai W J, Oktay O, Sinclair M, Suzuki H, Rajchl M, Tarroni G, Glocker B, King A, Matthews P M and Rueckert D. 2017. Semi-supervised learning for network-based cardiac MR image segmentation//Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada: Springer: 253-260 [DOI: 10.1007/978-3-319-66185-8_29]
- Bortsova G, Dubost F, Hogeweg L, Katramados I and De Bruijne M. 2019. Semi-supervised medical image segmentation via learning consistency under transformations//Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China: Springer: 810-818 [DOI: 10.1007/978-3-030-32226-7_90]
- Chen X K, Yuan Y H, Zeng G and Wang J D. 2021. Semi-supervised semantic segmentation with cross pseudo supervision//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 2613-2622 [DOI: 10.1109/CVPR46437.2021.00264]
- 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]
- Drozdal M, Vorontsov E, Chartrand G, Kadoury S and Pal C. 2016. The importance of skip connections in biomedical image segmentation//Proceedings of the 1st International Workshop on Deep Learning in Medical Image Analysis. Athens, Greece: Springer: 179-187 [DOI: 10.1007/978-3-319-46976-8_19]
- Fotadar G, Tajbakhsh N, Ananth S and Ding X W. 2020. Extreme consistency: overcoming annotation scarcity and domain shifts//Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention. Lima, Peru: Springer: 699-709 [DOI: 10.1007/978-3-030-59710-8_68]
- Hang W L, Feng W, Liang S, Yu L Q, Wang Q, Choi K S and Qin J. 2020. Local and global structure-aware entropy regularized mean teacher model for 3D left atrium segmentation//Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention. Lima, Peru: Springer: 562-571 [DOI: 10.1007/978-3-030-59710-8_55]

- He K H and Xiao Z Y. 2021. LRUNet: a lightweight rapid semantic segmentation network for brain tumors. *Journal of Image and Graphics*, 26(9): 2233-2242 (何康辉, 肖志勇. 2021. LRUNet: 轻量级脑肿瘤快速语义分割网络. *中国图象图形学报*, 26(9): 2233-2242) [DOI: 10.11834/jig.200436]
- He Y T, Yang G Y, Chen Y, Kong Y Y, Wu J S, Tang L J, Zhu X M, Dillenseger J L, Shao P F, Zhang S B, Shu H Z, Coatrieux J L and Li S. 2019. DPA-DenseBiasNet: semi-supervised 3D fine renal artery segmentation with dense biased network and deep priori anatomy//*Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention*. Shenzhen, China: Springer: 139-147 [DOI: 10.1007/978-3-030-32226-7_16]
- Karamitsos T D, Francis J M, Myerson S, Selvanayagam J B and Neubauer S. 2009. The role of cardiovascular magnetic resonance imaging in heart failure. *Journal of the American College of Cardiology*, 54(15): 1407-1424 [DOI: 10.1016/j.jacc.2009.04.094]
- Laradji I, Rodriguez P, Mañas O, Lensink K, Law M, Kurzman L, Parker W, Vázquez D and Nowrouzezahrai D. 2021. A weakly supervised consistency-based learning method for covid-19 segmentation in CT images//*Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision*. Waikoloa, USA: IEEE: 2452-2461 [DOI: 10.1109/WACV48630.2021.00250]
- 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]
- Li S M, Zhao Z Y, Xu K X, Zeng Z and Guan C T. 2021. Hierarchical consistency regularized mean teacher for semi-supervised 3D left atrium segmentation//*Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Guadalajara, Mexico: IEEE: 3395-3398 [DOI: 10.1109/EMBC46164.2021.9629941]
- Li X M, Yu L Q, Chen H, Fu C W and Heng P A. 2018. Semi-supervised skin lesion segmentation via transformation consistent self-ensembling model//*Proceedings of 2018 British Machine Vision Conference*. Newcastle, UK: BMVC: #63
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Luo X D, Chen J N, Song T and Wang G T. 2021a. Semi-supervised medical image segmentation through dual-task consistency. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(10): 8801-8809 [DOI: 10.1609/aaai.v35i10.17066]
- Luo X D, Liao W J, Chen J N, Song T, Chen Y N, Zhang S C, Chen N Y, Wang G T and Zhang S T. 2021b. Efficient semi-supervised gross target volume of nasopharyngeal carcinoma segmentation via uncertainty rectified pyramid consistency//*Proceedings of the 24th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Strasbourg, France: Springer: 318-329 [DOI: 10.1007/978-3-030-87196-3_30]
- Mao X J, Shen C H and Yang Y B. 2016. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections//*Proceedings of the 30th International Conference on Neural Information Processing Systems*. Barcelona, Spain: Curran Associates Inc.: 2810-2818
- 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]
- Mittal S, Tatarchenko M and Brox T. 2021. Semi-supervised semantic segmentation with high-and low-level consistency. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(4): 1369-1379 [DOI: 10.1109/TPAMI.2019.2960224]
- Miyato T, Maeda S I, Koyama M and Ishii S. 2019. Virtual adversarial training: a regularization method for supervised and semi-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(8): 1979-1993 [DOI: 10.1109/tpami.2018.2858821]
- Nie D, Gao Y, Wang L and Shen D. 2018. ASDNet: attention based semi-supervised deep networks for medical image segmentation//*Proceedings of 2018 International Conference on Medical Image Computing and Computer-Assisted Intervention*. Granada, Spain: Springer: 370-378 [DOI: 10.1007/978-3-030-00937-3_43]
- Orhan A E and Pitkow X. 2018. Skip connections eliminate singularities//*Proceedings of the 6th International Conference on Learning Representations*. Vancouver, Canada: ICLR
- Ouali Y, Hudelot C and Tami M. 2020. Semi-supervised semantic segmentation with cross-consistency training//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 12671-12681 [DOI: 10.1109/CVPR42600.2020.01269]
- Peng J Z, Pedersoli M and C. Desrosiers. 2020. Mutual information deep regularization for semi-supervised segmentation//*Proceedings of 2020 International Conference on Medical Imaging with Deep Learning*. Montréal, Canada: MIDL: 601-613
- Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young A A, Rhode K and Mansi T. 2019. Statistical atlases and computational models of the heart. Atrial segmentation and LV quantification challenges//*Proceedings of the 9th International Workshop*. Granada, Spain: Springer: #11395 [DOI: 10.1007/978-3-030-12029-0]
- Rizve M N, Duarte K, Rawat Y S and Shah M. 2021. In defense of pseudo-labeling: an uncertainty-aware pseudo-label selection framework for semi-supervised learning [EB/OL]. [2021-10-13]. <https://arxiv.org/pdf/2101.06329.pdf>
- Ronneberger O, Fischer P and Brox T. 2015. U-net: convolutional net-

- works 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]
- Sajjadi M, Javanmardi M and Tasdizen T. 2016. Regularization with stochastic transformations and perturbations for deep semi-supervised learning//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain: Curran Associates Inc.: 1171-1179
- 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 [EB/OL]. [2021-10-13]. <https://arxiv.org/pdf/2001.07685.pdf>
- Suinesiaputra A, Cowan B R, Finn J P, Fonseca C G, Kadish A H, Lee D C, Medrano-Gracia P, Warfield S K, Tao W C and Young A A. 2011. Left ventricular segmentation challenge from cardiac MRI: a collation study//Proceedings of the 2nd International Workshop on Statistical Atlases and Computational Models of the Heart. Toronto, Canada: Springer: 88-97 [DOI: 10.1007/978-3-642-28326-0_9]
- 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
- Tobon-Gomez C, Geers A J, Peters J, Weese J, Pinto K, Karim R, Ammar M, Daoudi A, Margeta J, Sandoval Z, Stender B, Zheng Y F, Zuluaga M A, Betancur J, Ayache N, Chikh M A, Dillenseger J L, Kelm B M, Mahmoudi S, Ourselin S, Schlaefler A, Schaeffter T, Razavi R and Rhode K S. 2015. Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. *IEEE Transactions on Medical Imaging*, 34(7): 1460-1473 [DOI: 10.1109/TMI.2015.2398818]
- Wang Y L, Huang G, Song S J, Pan X R, Xia Y T and Wu C. 2022. Regularizing deep networks with semantic data augmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(7): 3733-3748 [DOI: 10.1109/TPAMI.2021.3052951]
- Wang Y X, Zhang Y, Tian J, Zhong C, Shi Z C, Zhang Y and He Z Q. 2020. Double-uncertainty weighted method for semi-supervised learning//Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention. Lima, Peru: Springer: 542-551 [DOI: 10.1007/978-3-030-59710-8_53]
- Xiong Z H, Fedorov V V, Fu X H, Cheng E, Macleod R and Zhao J C. 2019. Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network. *IEEE Transactions on Medical Imaging*, 38(2): 515-524 [DOI: 10.1109/TMI.2018.2866845]
- Xiong Z H, Xia Q, Hu Z Q, Huang N, Bian C, Zheng Y F, Vesal S, Ravikumar N, Maier A, Yang X, Heng P A, Ni D, Li C Z, Tong Q Q, Si W X, Puybureau E, Khoudli Y, Géraud T, Chen C, Bai W J, Rueckert D, Xu L C, Zhuang X H, Luo X Z, Jia S M, Sermesant M, Liu Y S, Wang K Q, Borra D, Masci A, Corsi C, De Vente C, Veta M, Karim R, Preethe C J, Engelhardt S, Qiao M Y, Wang Y Y, Tao Q, Nuñez-García M, Camara O, Savioli N, Lamata P and Zhao J C. 2021. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. *Medical Image Analysis*, 67: #101832 [DOI: 10.1016/j.media.2020.101832]
- You C Y, Zhao R H, Staib L H and Duncan J S. 2022. Momentum contrastive voxel-wise representation learning for semi-supervised volumetric medical image segmentation//Proceedings of the 25th International Conference on Medical Image Computing and Computer-Assisted Intervention. Singapore, Singapore: Springer: 639-652 [DOI: 10.1007/978-3-031-16440-8_61]
- 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 segmentation//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 Z, Yang L, Chen J X, Fredericksen M, Hughes D P and Chen D Z. 2017. Deep adversarial networks for biomedical image segmentation utilizing unannotated images//Proceedings of the 20th International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec City, Canada: Springer: 408-416 [DOI: 10.1007/978-3-319-66179-7_47]
- Zheng H, Lin L F, Hu H J, Zhang Q W, Chen Q Q, Iwamoto Y, Han X H, Chen Y W, Tong R F and Wu J. 2019. Semi-supervised segmentation of liver using adversarial learning with deep atlas prior//Proceedings of the 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China: Springer: 148-156 [DOI: 10.1007/978-3-030-32226-7_17]
- Zhou Y N, Chen H, Lin H J and Heng P A. 2020. Deep semi-supervised knowledge distillation for overlapping cervical cell instance segmentation//Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention. Lima, Peru: Springer: 521-531 [DOI: 10.1007/978-3-030-59710-8_51]

作者简介

苏逸欣,女,硕士研究生,主要研究方向为3维医学图像处理及分割、深度学习。E-mail: 1318387726@qq.com

肖志勇,通信作者,男,副教授,主要研究方向为人工智能、机器视觉、医学影像分析。

E-mail: zhiyong.xiao@jiangnan.edu.cn