

基于全局-局部信息融合 LPV-Net 和 3D-EDA 的心脏半监督分割算法研究

胡慧称¹ 刘瑞霞² 刘照阳² 郭振华³

¹ 齐鲁工业大学(山东省科学院)数学与人工智能学部 济南 250353

² 齐鲁工业大学(山东省科学院)山东省人工智能研究院 济南 250014

³ 浪潮电子信息产业股份有限公司 济南 250101

(15092704322@163.com)

摘要 心脏分割算法提供精准把握心脏结构的信息,辅助医生进行诊断、制定治疗计划以及进行手术前的评估工作,提高临床治疗的效果并减轻并发症。然而,心脏分割在应用中存在诸多问题。手动分割方法不仅耗时费力,并且具有很强的主观意识。全监督的心脏分割方法虽已取得一定成果,但标注数据的依赖性过高。现有的半监督分割方法在处理复杂的心脏结构和病理变化时表现仍不够理想,难以在实际临床环境中稳定应用。为解决这一问题,提出了心脏磁共振成像(MRI)分割方法,结合 Linformer and Performer merge V-Net(LP-V-Net)和 3D Enhanced Discriminator with Attention(3D-EDA)技术,实现了全局-局部信息的有效整合。LPV-Net 模块由 LinPerBlock 和改进的 V-Net 联袂打造,旨在规范模型训练过程、实现全局与局部信息的有机整合,有效提高分割效果的准确性与鲁棒性。同时引入新鉴别器 3D-EDA 规范未标记数据,并加入关键模块 CARE-Layer,集成自定义注意力模块以增强对特征重要信息的捕捉能力,辅助网络可提高主网络分割指标性能。在左心房数据集上进行综合实验,并将所提方法与 MC-Net, V-Net 等其他先进的半监督方法进行比较,发现该方法在基准数据集上的表现尤为优异。特别是在使用有限标签数据进行训练时,该方法仍然展示出卓越的性能;当仅使用 10% 和 20% 的标记数据进行训练时,该方法的 Dice 系数分别达到 88.50% 和 90.39%。

关键词: 半监督分割;全局和局部信息;融合;LPV-Net;3D-EDA;左心房;CARELayer

中图分类号 TP391

Research on Heart Semi-supervised Segmentation Algorithm Based on Global-local Information Fusion LPV-Net and 3D-EDA

HU Huichen¹, LIU Ruixia², LIU Zhaoyang² and GUO Zhenhua³

¹ Department of Mathematics and Artificial Intelligence, Qilu University of Technology(Shandong Academy of Sciences), Jinan 250353, China

² Shandong Artificial Intelligence Institute, Qilu University of Technology(Shandong Academy of Sciences), Jinan 250014, China

³ Inspur Electronic Information Industry Co., Ltd., Jinan 250101, China

Abstract Heart disease is one of the main causes of death worldwide, which seriously threatens human life and health. As a non-invasive medical imaging technology, cardiac magnetic resonance imaging(MRI) is widely used in clinical diagnosis, helping doctors accurately and efficiently diagnose and treat heart diseases. However, cardiac MRI segmentation faces great challenges in practical applications, manual segmentation methods are time-consuming and subjective, while existing fully supervised and semi-supervised segmentation methods are not effective in dealing with complex cardiac structures and pathological changes, limited by the scarcity of data sets. This study aims to solve the challenges of cardiac MRI segmentation by proposing a 3D left atrial semi-supervised segmentation framework based on global-local information fusion to address the time-consuming and subjective problems of manual segmentation. Although the current fully supervised heart segmentation methods are effective, they are limited by the scarcity of data sets. The semi-supervised methods come into being, but they are still limited by the small amount of data, especially when dealing with complex cardiac structure and pathological changes. To solve this problem, this study propose a new cardiac MRI segmentation method that combines Linformer and Performer merge V-Net(LP-V-Net) and 3D Enhanced Discriminator with Attention(3D-EDA) technologies to achieve an effective fusion of global-local information. The LPV-Net module, created by LinPerBlock and improved V-Net, aims to standardize the training process of the model and achieve effective fusion of global and local information, thus improving the accuracy and robustness of segmentation. In addition, we also introduced a new discriminator 3D-EDA for the specification of unlabeled data. The most critical module in the model is CARELayer, which integrates a custom attention module to enhance the ability of the model to capture important information in the feature, and the auxiliary seg-

基金项目:山东省自然科学基金创新发展联合基金项目(ZR2023LZH009)

This work was supported by the Shandong Provincial Natural Science Foundation Innovation and Development Joint Fund Project (ZR2023LZH009).

通信作者:刘瑞霞(liurx@sdas.org)

mentation network improves the segmentation performance. By conducting a comprehensive experiment on the left atrial dataset, comparing the proposed method with several advanced semi-supervised methods. The experimental results show that the proposed method performs well on the baseline dataset, especially when training with limited label data. For example, when training with only 10% and 20% labeled data, the Dice coefficients of 88.50% and 90.39% were obtained.

Keywords Semi-supervised segmentation, Global and local information, Fuse, LPV-Net, 3D-EDA, Left atrium, CARELayer

1 引言

心房肌病是一组影响心房肌结构和功能的疾病,包括肥厚性心房病变、心房扩张等,对其的研究对于理解和治疗心房肌病是至关重要的^[1],尤其是心房结构的改变,进而导致心房功能受损,增加心房颤动等心律失常的发生风险^[2]。心脏磁共振成像(MRI)和心脏计算机断层扫描(CT)也可以用于评估心房肌结构和功能的变化^[3]。深入研究心房肌病的心房结构及其变化,在磁共振成像中可自动、准确地分割左心房,对心房肌病的诊断和治疗具有重要意义。

近年来,深度学习作为人工智能领域的重要支柱,发展迅速,尤其是在医学方面备受瞩目。深度学习^[4]技术以解决实际医疗问题为目标取得了显著进展,为医学影像分割和临床诊断提供了强大的工具。在医学影像分析中,基于深度学习的方法被广泛应用于左心房的分割^[5-9],这一领域的研究为医学影像分析和临床诊断带来了革命性的变革。图像分割技术在临床应用中扮演着关键角色,它为后续分析奠定了基础,可用于分析人体器官的结构和状态,从而为疾病诊断提供重要依据。近年来,随着深度学习技术的高速发展,基于深度学习的医学图像分割方法已成为主流研究方向,并取得了显著的成果。然而,在医学图像分割领域,现有的深度学习方法在提升性能的同时往往需要大量高质量的标注数据集。通过对数据集进行训练,网络能够学习到有效的特征表示,实现较好的分割效果。

在医学影像领域,获得带标签的数据集是一项具有挑战性的任务,特别是对于3D图像,其解剖结构分割需要经验丰富的医学专业人士进行操作。然而,这样的数据集获取过程耗时耗力且成本高昂^[10],特别是对于左心房在磁共振图像中的分割。自动医学图像分割技术的发展为应对这一挑战提供了新的可能性。随着深度学习技术的快速发展,全监督神经网络在医学图像分割方面取得了令人满意的成果。然而,由于医学图像数据标记的高成本,限制了全监督方法的应用范围。近年来,半监督学习框架充分利用了有限的标记数据和大量的未标记数据,通过迭代训练模型来提高分割的准确性,在医学图像分割领域备受关注。半监督学习框架为医学影像专家提供了利用现有有限标记数据和丰富的未标记数据的新途径,从而改善左心房分割效果。此外,半监督学习能够减少人工标记的工作量,并提高模型的泛化能力,为医学影像分析的自动化和精确性带来新希望。该发展将进一步推动临床诊断和治疗的可靠性和效率。

在半监督学习中,如何有效地利用大量无标签数据是非常重要的,通常都是通过聚类分析^[11]、生成对抗网络^[12-13]、自编码器等手段来处理无标签数据。双任务一致性^[14-15]半监督框架使用双任务深度网络共同预测目标的像素级分割和可感知集合的水平集表示,通过可以区分的任务转换层将级别集表示形式转换为近似的分割图,在使用无标签数据的水平

集产生的分割图和预测的分割图之间引入双任务一致性正则化。该算法能够同时处理多个任务,从而提高分割的准确性。其基于数据扰动、模型扰动或者是数据模型同时扰动的情况来实现,该做法的主要目的是提高模型的泛化能力,增强模型对噪声和干扰的鲁棒性^[16]。还有基于多任务层面的一致性约束,通过在多个任务之间引入一致性的约束来实现半监督学习,主要方法有共享表示、任务关联、正则化项、协同训练、约束优化以及多任务的损失函数,这些方法都可以帮助模型在处理多个任务时学习到有效且一致的表示,从而提高模型的分割性能,最终达到趋近于带标签的全监督学习的分割效果。Li等^[17]提出了形状感知半监督分割策略,该策略旨在解决在像素级图像注释中常被忽略的几何约束问题,通过强制实施形状约束来提高分割的准确性。Wu等^[18]采用了MC-Net,结合了单个编码器和多个不同编码器的结构,同时应用相互一致性约束。这一策略旨在最大程度地确保模型的输出在不同编码器下生成相似结果。Chen等^[19]提出了CPS方法,利用带有伪标签的数据扩展训练集,并借助两个网络来确保输入图像的预测结果之间高度相似,通过引入符号距离图(SDM)来进一步强化形状约束^[20]以提高性能,利用注意力机制学习标记数据和未标记数据之间的成对关系,以进一步缓解有限标记数据造成的过拟合。这些方法在解决像素级图像分割任务时,通过引入形状约束、多编码器结构或伪标签数据来提高分割模型的性能和稳健性。此外,一些其他领域的模型也开始被应用于医学图像分割领域,如unet、FCN、transform、变分自编码器等,成功的例子如Attention U-Net,它基于U-Net结构,包括编解码器部分,还可以自动学习并聚焦于与当前解码器层特征相关的编码器层特征^[21]。最近的研究显示,基于一致性正则化的半监督分割方法(如MC-Net, CPS方法)在处理有限标签数据时表现优异。然而,这些方法在长程依赖建模和局部复杂特征捕获上仍有改进空间。

2 本文贡献

本文提出了一种基于全局-局部信息融合LPV-Net和3D-EDA的3D左心房半监督分割框架,LPV-Net通过将LinformerBlock和PerformerBlock结合在一个模块中,同时集成到V-Net中,利用两种不同的注意力机制来捕获数据中不同的依赖关系,从而提高模型的表现。医学图像通常具有高分辨率和复杂的结构,需要模型能够捕获长距离的依赖关系,结合Linformer^[22]和Performer^[23]可以充分利用Linformer处理长距离的依赖关系,同时利用Performer的表现力来捕捉更复杂的局部结构。Linformer和Performer结合后,可以解决Performer在处理长序列时的计算复杂度问题,同时保持较高的表现力,在医学图像分割任务中提高模型的效率和性能。医学图像中常常存在不同尺度的特征,结合Linformer和Performer可以使模型更好地适应不同尺度的特征,从而提高分割的准确性和鲁棒性。V-Net中编码器和

解码器的输入和输出皆是 3D 张量,在集成 Linformer 和 Performer 的特征提取块中,接收的输入张量和输出张量也应该是三维的,以保持与 V-Net 模型的输入和输出一致^[24]。同时,作为特征提取的一部分,通过该提取块的处理,编码器可以更好地提取到全局和局部的特征信息,从而提高模型对输入数据的理解能力,并为解码器提供更具信息丰富性的特征表示。LPV-Net 作为新提出的 GLB-Net 的分割网络,对输入图像进行分割,并根据标签数据来计算监督损失。对于一致性损失,损失的权重随着训练的进行而逐渐增加。最后,借助对抗学习的思想设置了一个名为 3D-EDA 的判别器模型用于分类。模型中有一个最为关键的模块 CARELayer,其集成了自定义的注意力模块,目的是增强模型对特征中重要信息的捕捉能力,用无标签数据来生成无监督损失。这些损失组合在一起构成了模型的整体损失函数,通过反向传播算法和优化器来进行训练,逐步调整模型参数以最小化总损失。结合 Linformer 和 Performer 可以在医学图像分割任务中充分发挥两者的优势,提高模型的性能和效率,同时适应不同尺度和复杂度的医学图像数据。通过整合相关的边缘信息,一定程度地改善了边缘模糊、运动伪影等问题。通过对全局-局部特征融合方式处理的同一数据的不同预测结果进行一致性约束,使得每个预测结果相互之间可以学习到更多的特征。此外,利用标签数据和无标签数据之间的一致性,以及深度学习中的分割网络,设计一种简单而高效的、基于全局-局部信息

融合 LPV-Net 和 3D-EDA 的半监督医学图像分割框架,在 LA 数据集上对提出的方法进行了多次评估。本文的贡献可以概括为以下几点:

1)提出了一种用于半监督分割的 GLB-Net 模型,其核心思想是通过全局-局部信息融合来生成分割图,并与水平集图一起与真实值保持一致性,从而提高预测效果;

2)设计了一个基于自定义的 CARELayer 注意力模块的判别器 3D-EDA,增强模型对特征中重要信息的捕捉能力,以便利用大量的无标签数据;

3)在同一 A100 环境下,将采用多种其他半监督分割方法进行实验,并对它们的结果进行详尽比较,以确保公平性和科学性。所提半监督方法在左心房(LA)分割方面,可利用大量的无标签数据,展现出了其有效性。

3 网络结构

现存的半监督方法很多,学者们也都在为利用无标签数据提高图像整体的分割能力而付出诸多辛勤工作。在模型介绍中,使用 $x \in X$ 表示输入图像,给定一个数据集 $F = \{F^l, F^u\}$,其中有标签的数据集 $F^l = \{x_i^l, y_i^l\}_{i=1}^M$ 包含 M 个样本,无标签的数据集 $F^u = \{x_i^u, y_i^u\}_{i=M+1}^{M+N}$ 包含 N 个样本。 $x_i^l, x_i^u \in R^{H \times W \times D}$,分别是宽度 W ,高度 H 以及深度 D 。本文提出的 GLB-Net 半监督分割模型包括基于全局-局部信息融合的 LPV-Net 和 3D-EDA 判别器网络,网络结构如图 1 所示。

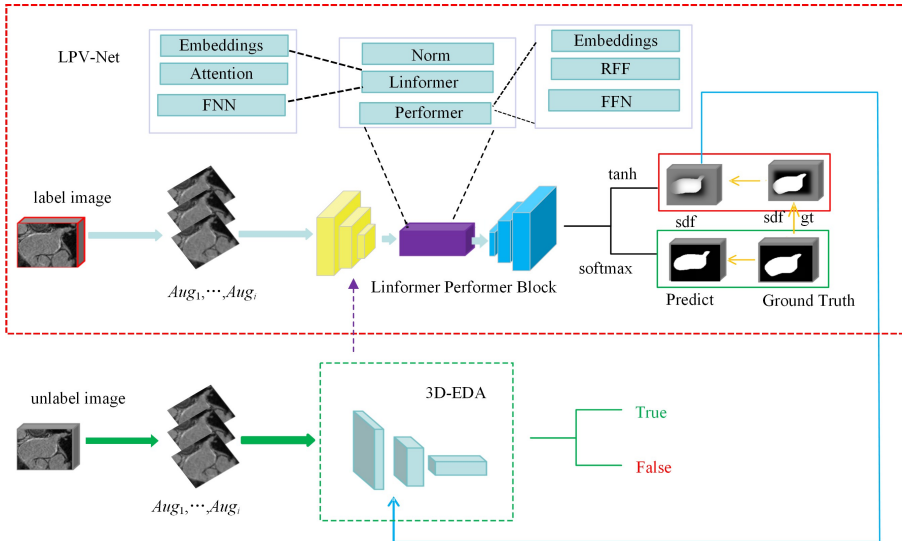


图 1 GLB-Net 模型

Fig. 1 GLB-Net model

3.1 GLB-Net 模型

GLB-Net 结合了全局-局部信息融合 LPV-Net 和 3D-EDA,通过结合两种不同的注意力机制来捕获医学图像中的不同依赖关系,提高模型性能。此外,该框架还采用一个名为 3D-EDA 的判别器模型,利用对抗学习的思想来增强模型的性能,特别是对无标签数据的处理。最终,该框架的整体损失函数由监督损失和无监督损失组成,通过反向传播和优化器进行训练,逐步调整模型参数来最小化总损失,使之具有良好的性能,可实现对左心房的精准分割。

3.2 LPV-Net

本节主要介绍 LPV-Net 的结构和工作原理。传统的 transformer 是一种用于序列建模的架构,在处理 3D 医学图

像分割任务时存在计算复杂度高、数据效率低、空间信息捕获不足、超参数调整繁琐等问题。为了解决这些问题,提出了 LPformer,与传统 Transformer 相比,结合 Linformer 和 Performer 在医学图像分割中的优势,能更高效地处理长距离依赖关系、降低计算复杂度以及适应不同尺度的特征。Linformer 具有较低的计算复杂度和适用于大规模数据的优势,而 Performer 则具有更高的表现力、较低的内存占用和适用于多种任务的特点。这两种方法的结合可以充分发挥它们各自的优势,在医学图像分割任务中提高模型的性能和效率。结合了 Linformer 和 Performer 的关键模块,用于对输入特征进行编码和解码。该模块包含多个 Transformer Encoder 和 Decoder 层,其中每个层都包含 Linformer 和 Performer 结构,

以捕获全局和局部的特征依赖关系。将三维的医学图像作为输入,其表示为一个三维张量,具有 depth,height,width 的维度,经过一系列的卷积和下采样操作后,图像的空间维度会逐渐减小,通道数则会逐渐增加。在经过所有卷积和下采样模块后,特征图的维度通常会变成(batch_size,channels,depth,height,width),具体来说,LinformerPerformerBlock 特征图的维度应该是(batch_size,n_filters * 16,depth/16,height/16,width/16),这也作为该块的有效输入。U-Net 和 V-Net 在几个方面略有差异。U-Net 是一种经典的编码器-解码器结构,用于处理二维图像分割任务,通过对称的编码器和解码器以及跳跃连接实现特征融合。相比之下,V-Net 专门用于处理三维数据,如医学图像中的体积数据,采用了类似残差连接的结构,并且使用 3D 卷积操作。此外,U-Net 主要使用交叉熵等损失函数进行监督学习,而 V-Net 在某些情况下使用特定于体积数据的损失函数,如形状距离函数^[25],其中预处理输入都是空间分辨率为 $H \times W \times D$ 的 3D 图像。LinPerBlock 模块通过结合 Linformer 的高效全局依赖建模能力和 Performer 的强局部捕获能力,使得网络在长距离依赖与局部复杂结构建模中表现出色。

3.3 3D-EDA

基于先前的知识,提出了一个新的判别器 3D-EDA,通过无标签数据来进行正则化。3D-EDA 由多个 3D 卷积层、线性分类器以及 CARELayer 模块组成。CARELayer 模块作为一个通道注意力机制,可以增强神经网络对输入特征的重要性的感知,将每个通道的重要性权重作用在输入特征上,使网络可学习有效的特征信息,从而促使预测的边缘检测图在无标签数据和带标签数据上保持一致性。3D-EDA 接收两个具有相同空间尺寸的三维数据的输入 map 和 image,最终得到二分类输出优化无监督损失,使得分割网络可以充分利用无标签数据。在训练初期,判别器主要依赖于有标签数据以建立初步特征分布,随着训练的进行,无标签数据的权重逐渐增加。CARELayer 通过分配不同通道的重要性权重,优化特征提取过程,显著提高了模型在无标签数据上的鲁棒性。

3.4 学习策略

该模型整体的损失是由生成器损失、判别器损失组成,其中,生成器通过最小化监督损失和一致性损失来提高生成图像的质量和准确性,同时判别器通过最小化对抗损失来提高对生成图像的辨别能力。Loss^[24] 的计算式如下:

$$L_{\text{sup}} = L_{\text{dice}} + \beta L_{\text{mse}} \quad (1)$$

其中, L_{dice} ^[26]和 L_{mse} ^[27]分别代表一个分割损失函数和一个形状损失函数; β 是调节因子,用于调节分割损失和形状损失在总体损失中的权重。

$$L_{\text{dice}} = 1 - \frac{2 \times \sum_i p_i \times q_i}{\sum_i p_i + \sum_i q_i} \quad (2)$$

$$L_{\text{mse}} = \frac{1}{N} \sum_{i=1}^N (p_i - q_i)^2 \quad (3)$$

其中, p_i 是生成器输出的第 i 个像素的预测概率, q_i 是标签的第 i 个像素的真实值, N 是像素总数。

$$L_{\text{con}} = \sigma \times \text{sigmoid} \left(\frac{\text{iter}}{\text{Rampup}} \right) \quad (4)$$

其中, σ 是一致性损失的系数,iter 是当前迭代次数,Rampup

是一致性损失的梯度上升时间^[28]。

$$L_{\text{adv}} = - \sum_{i=1}^{\mu} (t_i \cdot \log(\text{softmax}(Dout))_i) \quad (5)$$

其中, $Dout$ 是判别器对生成器输出的预测结果, t_i 是判别器的目标标签, μ 是类别数。

$$L_{\text{total}} = L_{\text{sup}} + L_{\text{con}} \times L_{\text{adv}} \quad (6)$$

损失函数共同构成了训练过程中的目标函数,通过优化这些损失函数^[29],可以有效地训练生成器和判别器,以实现所需的任务目标。

4 实验与分析

在心房分割挑战数据集上评估并验证所提方法,左心房的数据集经过前期的处理后用于整个实验。数据集中包含了 100 个 3D 钆增强的 MR 成像扫描和 LA 分割标签,用于实验的训练和验证。框架是在 A100 GPU 和 PyTorch 中实现的,提出的网络是采用改进的 V-Net 作为一个基本骨干网络,分别使用了 10%,20% 和 100% 的带标签训练数据,随机裁剪 $112 \times 112 \times 80$ 大小的立方体,并采用数据增强技术来避免过拟合现象。

4.1 评价指标

所提模型由分割网络 LPV-Net 和鉴别器 3D-EDA 组成,包含两个带标签的数据和两个无标签的数据,以 epoch 为单位,每 1000 个 epoch 保存一次。正如此前的研究,半监督分割的评估是通过使用 4 个性能指标来定量评估的,包括 Dice 系数、Jaccard 指数、95% Hausdorff 距离 (95HD) 和平均表面距离 (ASD),能够全面评价模型在分割任务上的表现。

$$Dice = \frac{2|E \cap D|}{|E| + |D|} \quad (7)$$

$$Jaccard = \frac{|E \cap D|}{|E \cup D|} \quad (8)$$

其中, E 是模型预测的分割结果, D 是实际的分割标签, $E \cap D$ 是预测分割区域与实际分割区域的交集大小, $E \cup D$ 是预测分割区域和实际分割区域的并集大小。

$$HD = \max(hd(C_D, C_E), hd(C_E, C_D)) \quad (9)$$

$$hd(C_D, C_E) = \max_{x \in C_D} (\min_{y \in C_E} (||x, y||)) \quad (10)$$

$$hd(C_E, C_D) = \max_{x \in C_E} (\min_{y \in C_D} (||x, y||)) \quad (11)$$

95 HD 指在累积分布函数中达到 95% 的分布所对应的距离值,用于衡量模型预测结果与实际分割区域之间的边界误差。

$$ASD = \frac{1}{|C_E| + |C_D|} \times (\sum_{x \in C_E} d(x, C_E) + \sum_{x \in C_D} d(y, C_D)) \quad (12)$$

$$d(x, A) = \min_{y \in A} ||x, y|| \quad (13)$$

ASD 是衡量模型预测结果与实际分割区域之间的平均距离误差。

4.2 结果与分析

以上是在左心房数据集上本文模型与 7 种最先进的半监督方法的对比,其中包括基于一致性的(MC-Net,DTC)、基于不确定性的(UA-MT,DUWM)、基于约束的(SASSNet)和基于低熵的(SS-Net)、基于多尺度的(URPC),分别采用了全监督的模型训练结果、10%带标签数据以及 20%带标签数据,

结果如表 1 所列。值得注意的是,在带标签训练数据的情况下,V-Net 模型平均 Dice 系数达到了 89.01%,Jaccard 系数达到 82.13%;所提框架在左心房数据集上 20%带标签数据的

Dice 为 90.39%,Jaccard 为 82.54%,以及 ASD 为 1.77 mm 方面,均优于其他方法的上限,但 95HD 效果低于 V-Net,具体数据如表 2 所列。

表 1 10%带标签数据集与其他方法的对比

Table 1 Compares other methods with a 10% labeled data set

Method	# scan used		Metrics			
	Labeled	Unlabeled	Dice/%	Jaccard/%	95HD/voxel	ASD/voxel
V-Net ^[24]	80	0	89.01	82.13	5.42	1.85
UA-MT ^[15]	8	72	85.65	75.19	19.10	5.27
SASSNet ^[17]	8	72	86.29	75.34	11.21	3.34
MC-Net ^[18]	8	72	88.29	79.16	8.63	2.30
URPC ^[30]	8	72	83.99	73.01	23.14	2.52
SS-Net ^[31]	8	72	85.92	75.59	18.63	2.76
DTC ^[16]	8	72	86.49	76.56	15.01	3.98
DUWM ^[32]	8	72	85.32	76.25	15.25	2.82
Ours	8	72	88.50	79.47	8.66	2.27

表 2 20%带标签数据集与其他方法的对比

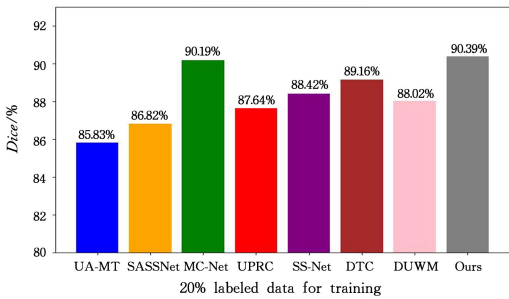
Table 2 Compares other methods with a 20% labeled data set

Method	# scan used		Metrics			
	Labeled	Unlabeled	Dice/%	Jaccard/%	95HD/voxel	ASD/voxel
V-Net ^[24]	80	0	89.01	82.13	5.42	1.85
UA-MT ^[15]	16	64	85.83	75.68	15.69	3.71
SASSNet ^[17]	16	64	86.82	76.73	10.35	3.25
MC-Net ^[18]	16	64	90.19	82.28	7.08	1.82
URPC ^[30]	16	64	87.64	78.59	12.67	2.14
SS-Net ^[31]	16	64	88.42	79.28	14.58	3.95
DTC ^[16]	16	64	89.16	80.17	7.30	2.08
DUWM ^[32]	16	64	88.02	81.04	8.18	2.26
Ours	16	64	90.39	82.54	6.84	1.77

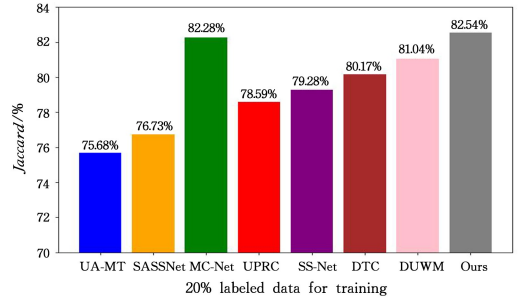
为了公平比较,在这些方法中将提出的改进 V-Net 看作是标准基线网络模型。从表 2 中可以看出,与其他方法相比,所提方法比先进的半监督方法实现了更好的性能表现,除了 HD 效果稍差。这证实了基于全局-局部信息融合 LPV-Net 和 3D-EDA 的 3D 左心房半监督分割框架具有完整的能力,可以从无标签数据中提取到丰富的信息。

表 2 的结果表明,本文方法在使用 10%和 20%带标签数据时的性能与全监督方法 V-Net 相当接近,尤其是 Dice 系数,略低于完全有标签情况下的性能上限,展示了其强大的鲁棒性。

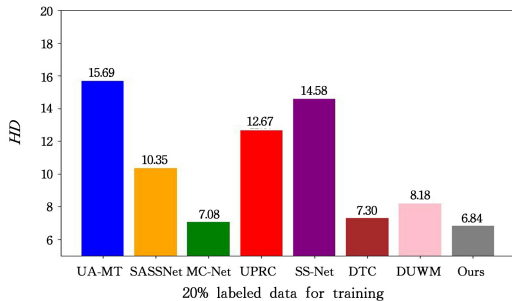
图 2 给出了不同半监督方法在左心房数据集上使用 20%带标签数据的具体情况。



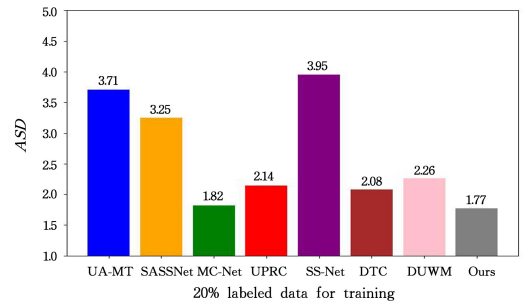
(a) Dice



(b) Jaccard



(c) 95HD



(d) ASD

图 2 不同模型的各种系数对比

Fig. 2 Comparison of various coefficient of different models

4.3 可视化分析

图 3 给出了不同模型的实验结果。图 3 中,所提框架在

分割性能方面表现更加出色,在捕捉相对复杂的特征时具有很好的效果,并且可以更清楚地划分边缘区域。

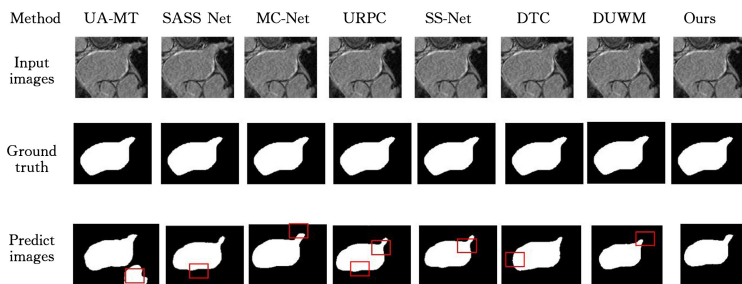


图 3 不同模型的实验结果

Fig. 3 Experimental results of different models

4.4 消融实验

为了验证所提方案的有效性,评估了原始 V-Net 以及改变后的 V-Net 的性能。结果如表 3 所列,所提方法的性能提升明显。此外,专门研究了 transform 和 MLP,实验效果都不如所提模块。从中可以看出,模块有助于提高整体的效果,同

时设计一个比较优秀的框架。实验中,仅仅使用带标签数据来进行训练,由于带标签数据和无标签数据之前的孤立,会导致它们之间无法进行互动以及传递相关知识,体现出融合的重要性。另一方面,该模型通过全局-局部信息成功融合,从而使得整体效果变好,证明了模型的有效性。

表 3 LA 数据集上的消融实验

Table 3 Ablation experiment on LA dataset

Method	# scan used		Metrics			
	Labeled	Unlabeled	Dice/%	Jaccard/%	95HD/voxel	ASD/voxel
V-Net	16	64	79.99	68.12	21.11	5.48
V-Net+transformer	16	64	82.21	70.75	19.64	5.72
V-Net+LPformer	16	64	86.09	76.15	11.35	3.05
VNet+LPformer+MLP	16	64	89.41	80.95	7.77	1.93
V-Net+LPformer+3D-EDA	16	64	90.39	82.54	6.84	1.77

结束语 本文提出了一种新颖的基于全局-局部信息融合的 3D 半监督分割框架,结合了 Linformer 和 Performer 的优势,能够有效捕获医学图像中的长程和局部特征。为了解决无标签数据在医学图像分割中的挑战,引入了 3D-EDA 判别器和 CARELayer 模块,显著提高了模型在无标签数据上的鲁棒性和特征捕捉能力。通过综合实验验证,所提方法在左心房数据集上的表现优于现有的多种半监督分割方法,特别是在使用有限标记数据时仍能展示出较高的分割精度和鲁棒性。本文方法的创新性体现在两方面:首先,通过结合 Linformer 和 Performer,解决了传统 Transformer 架构在长距离依赖建模上的计算复杂度问题;其次,通过引入 CARELayer 和 3D-EDA,增强了对无标签数据的特征捕获能力,进一步提升了半监督学习的效果。未来工作将探索更多类型的医学影像数据集(如脑部或肝脏),以进一步验证所提方法的泛化能力,从而解决迁移学习的挑战。

参考文献

[1] ZHANG Y, JI Y. HINT1 (Histidine Triad Nucleotide-Binding Protein 1) Attenuates Cardiac Hypertrophy Via Suppressing HOXA5 (Homeobox A5) Expression[J]. *Circulation*, 2022, 145(8): e151-e152.

[2] JIANG L S, HAO Z Y, XIE X Y, et al. Left atrial appendage angiography for stroke risk prediction in patients with atrial fibrillation[J]. *Eurointervention*, 2023, 19(8): 695-702.

[3] GASSENMAIER S. Deep learning applications in magnetic resonance imaging: has the future become present? [J]. *Diagnostics*, 2021, 11(12): 2181.

[4] OLYA, HESSAM M. An integrated deep learning and stochastic optimization approach for resource management in team-based healthcare systems[J]. *Expert Systems with Applications*, 2022 (187): 115924.

[5] XIONG Z, FEDOROV V V, FU X, et al. Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(2): 515-524.

[6] XIA Q, YAO Y, HU Z, et al. Automatic 3D atrial segmentation from GE-MRIs using volumetric fully convolutional networks [C]// *International Workshop on Statistical Atlases and Computational Models of the Heart*. 2018: 211-220.

[7] LI L, WENG X, SCHNABEL J A, et al. Joint left atrial segmentation and scar quantification based on a DNN with spatial encoding and shape attention[C]// *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. 2020: 118-127.

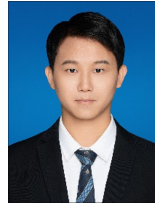
[8] ZHANG H, ZHANG W, SHEN W, et al. Automatic segmentation of the cardiac MR images based on nested fully convolutional dense network with dilated convolution[J]. *Biomedical Signal Processing and Control*, 2021, 68: 102684.

[9] CHEN C, BAI W, RUECKERT D. Multi-task learning for left atrial segmentation on GE-MRI [C]// *Statistical Atlases and Computational Models of the Heart*. 2018: 292-301.

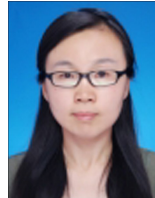
[10] XIONG Z, XIA Q, HU Z, et al. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging[J]. *Medical Image Analysis*, 2021, 67: 101832.

[11] RODRIGUEZ A, ALESSANDRO L. Clustering by fast search

- and find of density peaks[J]. *Science*, 2014, 344(6191): 1492-1496.
- [12] GOODFELLOW I. Generative adversarial networks[J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [13] MIRZA M, SIMON O. Conditional generative adversarial nets [J]. *arXiv*:1411.1784, 2014.
- [14] BORTSOVA G, DUBOST F, HOGEWEG L, et al. Semi-supervised medical image segmentation via learning consistency under transformations[C]// *Medical Image Computing and Computer Assisted Intervention (MICCAD)*. 2019: 810-818.
- [15] YU L, WANG S, LI X, et al. Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation [C] // *Medical Image Computing and Computer Assisted Intervention (MICCAD)*. 2019: 605-613.
- [16] LUO X, CHEN J, SONG T. Semi-supervised medical image segmentation through dual-task consistency[C]// *Proceeding of the AAAI Conference on Artificial Intelligence (AAAI)*. 2021: 8801-8809.
- [17] LI S, ZHANG C, HE X. Shape-aware semi-supervised 3D semantic segmentation for medical images [C] // *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. 2020: 552-561.
- [18] WU Y, XU M, GE Z, et al. Semi-supervised left atrium segmentation with mutual consistency training [C] // *MICCAI 2021*. 2021: 297-306.
- [19] CHEN X, YUAN Y, ZENG G, et al. Semi-supervised semantic segmentation with cross pseudo supervision [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021: 2613-2622.
- [20] MA J, WEI Z, ZHANG Y, et al. How distance transform maps boost segmentation CNNs: an empirical study [C] // *Medical Imaging with Deep Learning*. PMLR, 2020: 479-492.
- [21] OKTAY O. Attention u-net: Learning where to look for the pancreas [J]. *arXiv*:1804.03999, 2018.
- [22] WANG S N, WEI B Z, DNLE S, et al. Linformer: Self-attention with linear complexity [J]. *arXiv*:2006.04768, 2020.
- [23] CHOROMANSKI K. Rethinking attention with performers [J]. *arXiv*:2009.14794, 2020.
- [24] MILLETARI F, NASSIR N, SEYED-AHMAD A. V-net: Fully convolutional neural networks for volumetric medical image segmentation [C] // *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016: 565-571.
- [25] PARK J J, FLORENCE P, STRAUB J, et al. DeepSDF: Learning continuous signed distance functions for shape representation [C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019: 165-174.
- [26] DICE L R. Measures of the amount of ecologic association between species [J]. *Ecology*, 1945, 26(3): 297-302.
- [27] ZHANG K. A comprehensive review on medical image segmentation [J]. *Computers in Biology and Medicine*, 2017, 91: 11-29.
- [28] BERTHELOT D, BEGAN; Boundary equilibrium generative adversarial networks [J]. *International Conference on Learning Representations (ICLR)*. *arXiv*:1703.10717, 2017.
- [29] GOODFELLOW I J, JEAN P A, MEHDI M, et al. Generative adversarial nets [J]. *NeurIPS 2014*, 2014, 2: 2672-2680.
- [30] LUO X, WANG G, LIAO W, et al. Semi-supervised medical image segmentation via uncertainty rectified pyramid consistency [J]. *Medical Image Analysis*, 2022, 80: 102517.
- [31] WU Y, WU Z, WU Q, et al. Exploring smoothness and class-separation for semi-supervised medical image segmentation [C] // *25th International Conference Medical Image Computing and Computer Assisted Intervention (MICCAI 2022)*. 2022: 34-43.
- [32] WANG Y, ZHANG Y, TIAN J, et al. Doubleuncertainty weighted method for semi-supervised learning [C] // *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2020: 542-551.



HU Huichen, born in 1998, postgraduate. His main research interests include medical artificial intelligence and image processing.



LIU Ruixia, born in 1976, Ph.D, professor. Her main research interests include medical artificial intelligence and image processing.