

DOI:10.19651/j.cnki.emt.2416877

融合公平因子的半监督学习医学图像分割模型^{*}武丽^{1,2} 丁琴² 葛彩成²

(1.无锡学院电子信息工程学院 无锡 214000; 2.南京信息工程大学电子与信息工程学院 南京 210000)

摘要:为解决部分稀缺医学图像分割任务中目标语义类别分布不平衡导致的模型泛化能力受限问题,本文提出一种半监督学习的医学图像分割模型 CDCL-SSLNet。通过 UNet 和 Swin-UNet 两种不同属性的分割子模型进行交叉学习,实现特征优势互补。引入类分布公平因子和类学习公平因子对损失函数合理加权,动态地指导模型学习语义类别不平衡数据,有效减小学习偏差,进而提高模型泛化能力。实验选取 Synapse 多器官分割数据集中 5%、10% 的数据模拟标签数据对模型进行训练。CDCL-SSLNet 在仅有 5% 和 10% 的标签数据参与训练的情况下,其 Dice 系数分别达到了 65.71% 和 77.49%,HD95 则分别为 28.97 和 22.07,这两项指标的性能提升幅度均达到了 17%。结果表明 CDCL-SSLNet 能够兼顾大目标及微小目标的精准分割,有效解决了稀缺数据中类分布不平衡导致的模型泛化能力不足的问题,有效提升了医学图像分割的效率与准确性。

关键词: 半监督学习;图像分割;U-Net;公平因子;Transformer

中图分类号: TN391.4 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Semi-supervised learning medical image segmentation model
fused with equity factorsWu Li^{1,2} Ding Qin² Ge Caicheng²

(1. School of Electronic Information Engineering, Wuxi University, Wuxi 214000, China; 2. School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210000, China)

Abstract: In order to solve the problem of limited model generalization ability caused by imbalanced distribution of target semantic categories in some scarce medical image segmentation tasks, this paper proposes a semi-supervised learning medical image segmentation model CDCL-SSLNet, which achieves feature complementarity through cross-learning of two segmentation submodels with different attributes, namely, UNet and Swin-UNet. The introduction of class distribution fairness factor and class learning fairness factor reasonably weights the loss function, dynamically guides the model to learn the unbalanced data of semantic categories, effectively reduces the learning bias, and then improves the model generalization ability. In the experiment, 5% and 10% of the data in Synapse multi-organ segmentation dataset are selected to simulate labeled data to train the model. When only 5% and 10% of the label data participated in the training, the Dice coefficients of CDCL-SSLNet reached 65.71% and 77.49%, respectively, and the performance of HD95 was 28.97 and 22.07, respectively, and the performance of these two indicators was improved by 17%. The results show that CDCL-SSLNet is able to take into account the accurate segmentation of large and small targets, effectively solves the problem of insufficient model generalization ability caused by the imbalance of class distribution in scarce data, and effectively improves the efficiency and accuracy of medical image segmentation.

Keywords: semi-supervised learning; image segmentation; U-Net; fairness factor; Transformer

0 引言

医学图像分割是医学图像处理与分析的关键技术,其根据医学图像内的特征相似性将图像划分为若干个区域,

实现对医学组织结构的精确分离。在临床治疗中,使用相关影像学指标对医学图像进行定量测量和分析,有助于医生诊断、随访或修订治疗方案。

早期的医学影像分割系统主要基于传统的图像分割算

收稿日期:2024-09-12

* 基金项目:国家青年自然科学基金(62106111)、2022年第二批产学研合作协同育人项目(220903806051637)资助

法搭建,如基于边缘检测的分割算法、基于阈值的分割算法和基于区域的分割算法。但医学图像通常具有对比度低、组织纹理复杂、边界区域模糊等特点,极大地限制了此类图像分割算法的效果和应用场景^[1]。在实际问题中,图像的标注需要有专业知识且经验丰富的医生来手动标注,但是存在隐私权以及人为手动勾画的局限性和医疗信息系统集成不足等情况,标注过程往往会面临人为忽略重要细节和边缘不清晰等问题,同时也会浪费大量人力财力。近些年来,深度学习技术在自然语言处理中崭露头角,并广泛应用于医学图像各类分析任务如组织分割、癌症区域检测、癌症分类与识别、细胞检测等^[2]。尽管目前基于全监督学习的模型在图像分割中占主导地位,但这些网络模型的训练依赖海量标注像素级数据,高质量的医学图像数据标注费时费力,标注数据资源稀缺成为深度学习医学图像分割算法性能提升的最大瓶颈^[3]。为克服这一困难,尽管有诸多工作围绕分割模型结构优化展开,如对经典的分割模型 U-Net^[4]进行改进,包括引入注意力机制、构造损失函数等方法^[5]。另外,结合 Transformer^[6]的分割模型,如 TransUNet^[7]、Swin-UNet^[8]等一定程度上提升了医学图像分割精度,但是这些分割模型均为全监督模型,没有充分利用未标注数据资源。在现实环境下,供给模型训练的数据可能属于稀缺类医学图像,这会导致分割精度上限较低,模型难以投入实际应用。

半监督学习的主要思想是利用标注数据和未标注数据所提供的信息,在标注数据较少的情况下,集合大量未标注的数据来改善学习性能^[9]。所以诸多研究围绕半监督学习展开,这些工作专注于利用未标注数据对模型进行训练,使得模型能够充分学习现有数据语义信息,并使之具有更强鲁棒性,如对抗学习、一致性正则化、伪标签等方法^[10]。因此,使用半监督学习方法既可以避免大量的数据标注,又可以达到与全监督学习方法相媲美的图像分割精度^[11]。交叉伪监督(cross pseudo supervision, CPS)^[12]方法是基于一致性正则化方法的代表。这种方法巧妙地结合了伪标签和一致性正则化的思想,通过在网络训练过程中引入伪标签作为监督信息,并要求模型在不同扰动下保持预测的一致性,从而实现了在未标记数据的有效利用和分割性能的提升。Luo 等^[13]吸取了伪标签法和一致性正则化法的优点,创新地提出了利用不同范式的模型来构建交叉学习策略,从而实现了高精度的图像分割。Yu 等^[14]在估计不确定性指导下,计算一致性损失时过滤掉不可靠的预测,只保留可靠的预测(低不确定性)。但是在这些半监督学习方法中,大多没有关注到数据内在语义类别分布情况,特别是在多类语义分割任务中,各类别的语义信息分布不均匀,模型倾向学习分布广的语义信息而削弱对分布较稀少的语义信息学习,如常见的多器官分割任务,其中肝脏分布和胆囊分布差异大,大多模型对胆囊的分割效果并不理想。

一些针对数据语义信息中类分布不平衡的半监督算法

研究在医学图像分割任务中取得新进展,如 Guo 等^[15]提出的具有自适应学习类不平衡半监督学习方法 Adash,利用设定的类平衡阈值自适应地选择伪标签来训练模型,但这种方法是基于显性的训练策略。Wang 等^[16]提出一种针对类分布不平衡的基线模型 SimiS,通过将伪标签补充到标记集合中的非频繁类中,有效地降低了训练过程中标记数据的不平衡率。但是该方法可能高度依赖于网络的初始化。不同的初始化可能导致不同的原型生成和一致性正则化效果,进而影响最终的分割性能。Lin 等^[17]提出的校准标签分布(calibrating label distribution, CLD)半监督框架,通过引入类感知加权损失、概率感知随机裁剪和双不确定性感知采样监督等方法,以实现高效的半监督学习。但上述方法在降低类不平衡率时仅考虑类像素占比,忽视了类特征的提取难度,从而降低特征提取质量,对分割性能产生消极影响。并且半监督方法虽然解决了标注数据资源不足的问题,但其分割精度难以达到全监督模型的水平。

因此,本文提出了一种基于类分布和类学习的半监督学习模型(class distribution and class learning-based semi-supervised learning network, CDCL-SSLNet),针对性解决半监督模型分割精度不足以及数据类分布不平衡问题。CDCL-SSLNet 包括 UNet 和 Swin-UNet 两个子网络,分别捕获局部特征与长距离依赖特征。CDCL-SSLNet 通过一致性正则化生成准确的伪标签,用于监督神经网络优化,提升模型性能。CDCL-SSLNet 使用类分布公平因子和学习公平因子加权的损失函数对两个子模型进行分别优化,分别构造公平的类分布损失权重与类学习性能提升指标,从类像素占比与类特征提取难度两个维度降低类不平衡率。

本文的主要工作如下:

1) 构建了全新的半监督医学图像分割模型 CDCL-SSLNet,利用 UNet 和 Swin-UNet 两种子网络提取局部特征与长距离依赖特征。

2) 利用隐式一致性正则化产生更稳定和准确的伪标签,构造交叉伪监督引导 CDCL-SSLNet 优化。

3) 提出类分布公平因子和类学习公平因子对损失函数进行加权,基于类像素占比与类特征提取难度综合判断类不平衡率,有效地减小模型的学习偏差,提高模型整体分割性能。

1 模型介绍

1.1 CDCL-SSLNet 介绍

CDCL-SSLNet 模型框架如图 1 所示,其包含 U-Net 和 Swin-UNet 两个异质化子模型,子模型无需预先训练。其中,U-Net 作为卷积神经网络(CNN)的典型代表,U-Net 通过其编码器-解码器,结构有效地提取了局部特征。编码器逐步下采样输入特征图,可以有效地捕获局部信息;解码器则通过跳跃连接(skip connections)恢复细节信息,使得低层次的空间信息和高层次的语义信息能够结合,从而保证

分割结果的精准性 U-Net 尤其在医学图像分割中表现突出,因为它能够很好地处理小尺寸的特征和局部细节。;而 Swin-UNet 作为 Transformer 的代表,它结合了 Swin Transformer 的结构,则能够捕获上下文中的长距离依赖信息。Swin Transformer 的多头自注意力机制可以在图像中建立全局上下文关系,使得 Swin-UNet 能够在图像的不同区域之间建立更加全局的关联,而不仅仅依赖于局部信息。通过层次化结构,Swin-UNet 可以在不同尺度下捕获

重要的全局信息,并将这些特征传递给解码阶段。U-Net 依赖于 CNN 的局部特征提取能力,擅长捕获精细的边缘和局部模式,而 Swin-UNet 的 Transformer 机制能够在全局范围内捕获长距离依赖信息。训练过程中,这两种特征互为补充,U-Net 保持了图像的精细局部信息,而 Swin-UNet 提供了全局的上下文依赖,从而提高了模型的整体性能和分割精度。在训练过程中,这两种特征相互补充,从而提升了模型的性能。

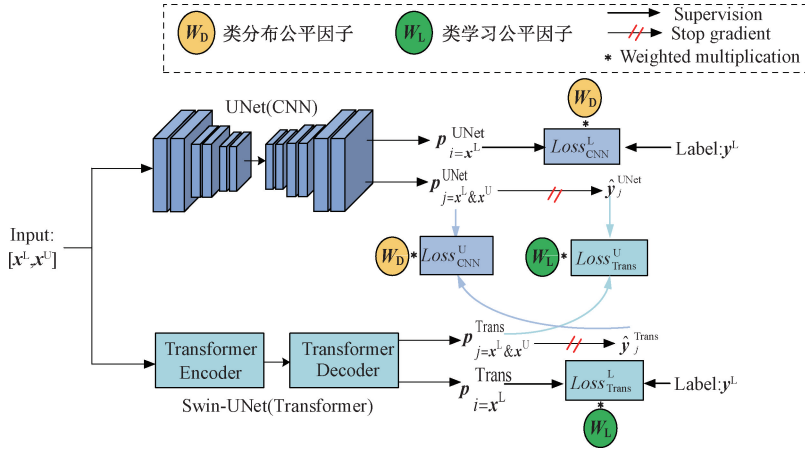


图 1 CDCL-SSLNet 整体框架

Fig. 1 The overall CDCL-SSLNet framework

假设整个数据集由 N_L 个标记样本 $\{(x_i^l, y_i)\}_{i=1}^{N_L}$ 和 N_U 个未标记样本 $\{(x_i^u)\}_{i=1}^{N_U}$ 组成,且 $N_L \ll N_U$, 其中 $x_i \in R^{D \times H \times W}$ 是输入数据, $y_i \in R^{K \times D \times H \times W}$ 是具有 K 个类(包括背景)的标签数据。CDCL-SSLNet 中的两个子模型通过使用两种不同的动态加权策略来最小化目标函数,从而引导模型参数优化。两个子模型的目标函数分别如下式(1)和(2)所示。

$$\overline{Loss}_S = \frac{1}{N_L} \frac{1}{K} \sum_{i=0}^{N_L} [W_i^D Loss_S(p_i^{UNet}, y_i) + W_i^L Loss_S(p_i^{Trans}, y_i)] \quad (1)$$

$$\overline{Loss}_U = \frac{1}{N_L + N_U} \frac{1}{K} \sum_{j=0}^{N_L + N_U} [W_j^D Loss_U(p_j^{UNet}, \hat{y}_j^{Trans}) + W_j^L Loss_U(p_j^{Trans}, \hat{y}_j^{UNet})] \quad (2)$$

其中, $p_i^{(*)}$ 是第 i 个样本输出概率图, $\hat{y}_i^{(*)} = \operatorname{argmax}\{p_{i,k}^{(*)}\}_{k=0}^K$ 是第 i 个样本的伪标签, $Loss_S(x, y) = Loss_{CE}(x, y)$ 是监督标记数据输出的监督交叉熵损失函数。 $Loss_U(x, y) = \frac{1}{2} [Loss_{CE}(x, y) + Loss_{Dice}(x, y)]$ 是无监督损失函数,两个子模型输入相同的数据 x_i 得到两个输出,通过计算输出后的交叉熵损失和 Dice 损失来衡量子模型的预测一致性。这里值得注意的是无监督损失部分,为了避免两个子模型出现预测一致性较高但预测均错误的情况,模型引入少量标签数据对子模型提供额外约束,因

此全部数据集 ($N_L + N_U$) 都参与到无监督损失计算更合理。 $W_i^D, W_i^L, W_j^D, W_j^L$ 即损失加权公平因子,它们能够动态地调整损失比重,从而确保模型能够更加合理地学习分割目标,2.2 和 2.3 节将具体介绍两类公平因子。最后得到总损失函数,如式(3)所示。

$$Loss_{total} = \overline{Loss}_S + \lambda \overline{Loss}_U \quad (3)$$

其中, λ 是可调参数,根据经验,设置 λ_{max} , 并使用与时

间相关的高斯斜升函数 $\lambda(t) = \lambda_{max} \times e^{-5(1 - \frac{t}{t_{max}})^2}$ 来逐步调整无监督损失参与的比重, t 是当前训练时间, t_{max} 是训练时间的总数。模型训练完毕后,将一张待分割的图像切片输入参数固定的模型中,得到两个子模型的预测结果,通过计算它们的均值来获得最终的分割结果。这种方法有效地融合了两种子模型各自的优势,从而提升了整体分割性能。

1.2 类分布公平因子

如图 2 所示的多器官分割图像中各个分割目标像素分布情况,其中各个器官的像素分布并不均匀,如肝部的组成像素分布占分割目标总像素的 41.63%,而胆囊像素仅占 0.41%,二者比约为 102:1。这对于模型而言,在同步的学习过程中会对肝部投入更多关注,但对胆囊的关注投入较少,因此类分布不均衡会导致模型学习偏差。

为了克服模型学习偏差,本工作提出类分布公平因子作为损失函数的加权项,促使模型更多地关注像素分布较少的类如主动脉、胆囊等,使得模型根据数据特点给予不

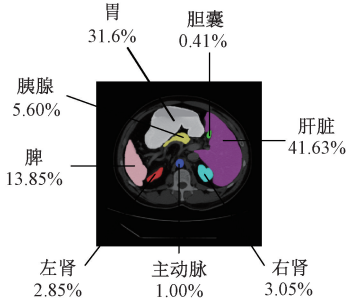


图 2 多器官分割目标像素占比图

Fig. 2 Proportion of target pixels for multi-organ segmentation

同类别公平的学习关注度,降低数据类不平衡率。具体来说,通过统计模型生成的伪标签中每个类的像素数量 N_k , $k = 0, 1, 2 \dots$ 代表分割目标类别,得到第 k 类的分布情况 $P_k = \frac{N_k}{\max\{N_i\}_{i=0}^K}$, 其中 $k = 0, 1, \dots, K$, 接着构造第 k 类的加权系数即第 k 类分布公平因子,使得分布较广(较小)的类得到更小(大)的权重,如式(4)所示。

$$\omega_k^c = e^{P_k - \max\{P_i\}_{i=0}^K} \quad (4)$$

以上仅是在迭代后得到的各个类分布公平因子 $W_i^c = [\omega_1^c, \omega_2^c, \dots, \omega_K^c]$, 要使得模型能够动态地且平稳地实现权重更新,这里使用指数移动平均方法对类分布公平因子 W_i^c 进行更新,如式(5)所示,其中 α 是动量参数,实验中设置为 0.99。

$$W_i^c = \alpha W_{i-1}^c + (1 - \alpha) W_i^c \quad (5)$$

1.3 类学习公平因子

对损失施加类分布公平因子仅是从数据层面减轻模型学习偏见,对于医学图像分割任务而言,不同场景下的器官或病灶像素可能占据相似的比重,仅对损失施加类分布公平因子不足以应对组织病灶较为复杂的情况。而且有些器官或病灶的像素虽然占比较大,但边缘不规则、形态多变,特征提取难度高,众多数据也难以让模型捕捉到其中的规律性,因此盲目地对损失施加类分布公平因子可能会适得其反。为了合理充分地消除模型学习偏见,要从模型本身的学习状况着手。根据模型在学习类的过程中的表现出的难易程度,本文引入了类学习公平因子对损失进行加权。类学习公平因子综合考虑到模型在学习某个类的过程中的学习速度及学习质量,判断类特征提取难度,降低类不均衡率。

首先使用群体稳定性指标(population stability index, PSI)来衡量第 t 次迭代后每个类的学习速度,如式(6)所示。

$$\begin{cases} v_{k,t} \downarrow = \sum_{i=\tau}^t I(\Delta) \ln\left(\frac{D_{k,t}}{D_{k,t-1}}\right), & \Delta \leq 0 \\ v_{k,t} \uparrow = \sum_{i=\tau}^t I(\Delta) \ln\left(\frac{D_{k,t}}{D_{k,t-1}}\right), & \Delta > 0 \end{cases} \quad (6)$$

其中, $D_{k,t}$ 是指模型第 t 次迭代后在第 k 类目标上取得的 Dice 值, $I(\Delta)$ 是指示函数,而 $\Delta = D_{k,t} - D_{k,t-1}$ 代表

着在每一次迭代中 Dice 值的变化情况,如果 $\Delta \leq 0$, 则表示本次迭代学习在退步或止步;如果 $\Delta > 0$, 则表示本次迭代学习在进步。 τ 指累积迭代的次数,根据经验,实验中将它设置为 50。因此,模型在整体 τ 次的迭代学习过程中,将退步或止步的迭代过程单独累积计算 PSI 值得到 $v_{k,t} \downarrow$, 将进步的迭代过程单独累积计算 PSI 值得到 $v_{k,t} \uparrow$, 那么模型的整体学习速度可以定义成式(7):

$$v_{k,t} = \frac{v_{k,t} \uparrow + \epsilon}{v_{k,t} \downarrow + \epsilon} \quad (7)$$

ϵ 是一个较小的值作为平滑项,实验中设置为 0.001。学习速度越快的类意味着 $v_{k,t}$ 越大,为了降低学习速度以适应学习慢的类,损失函数中相应的权重应设置的越小,经过充分的迭代学习,所有类的学习速度将趋于一致。

另外,可以通过 Dice 值即 $D_{k,t}$ 的高低来衡量学习质量,那么第 k 类学习公平因子最终可以定义为式(8):

$$\omega_k^D = -\log(D_{k,t} \cdot (v_{k,t})^\beta \cdot \beta) \quad (8)$$

其中, β 作为调节参数可以减少离群值影响,根据经验,实验中将它设置为 0.33 可以起到最佳效果,最后可得各个类学习公平因子为 $W_i^D = [\omega_1^D, \omega_2^D, \dots, \omega_K^D]$, 其仍采用指数移动平均方法更新。CDCL-SSLNet 整体训练步骤可定义为算法 1。

算法 1 CDCL-SSLNet 训练流程

Algorithm 1 CDCL-SSLNet training process

输入: D_L : 标注训练数据, D_U : 无标注训练数据
 定义: M_{UNet} : 子模型 1, $M_{Swin-UNet}$: 子模型 2
 T : 训练迭代次数, $\lambda(t)$: 高斯斜升函数
 输出: 训练完毕的子模型 M_{UNet} 、 $M_{Swin-UNet}$
 while ($t < T$) do
 随机初始化类分布公平因子 W_D 和类学习公平因子 W_L
 小批次 $x \subset D_L$ 分别输入 M_{UNet} 、 $M_{Swin-UNet}$ 训练
 监督损失 $Loss_S = W_D \cdot Loss_S^{UNet}(x, y) + W_L \cdot Loss_S^{Swin-UNet}(x, y)$
 小批次 $x \subset (D_U \& D_L)$ 分别输入 M_{UNet} 、 $M_{Swin-UNet}$
 获得伪标签 $\hat{y}_{UNet} = M_{UNet}(x)$, $\hat{y}_{Swin-UNet} = M_{Swin-UNet}(x)$
 统计伪标签像素分布,更新类分布公平因子损失加权项 W_D
 统计伪标签中各类学习速度,更新类学习公平因子损失加权项 W_L
 小批次 $x \subset (D_U \& D_L \& \hat{y}_{UNet})$ 输入 $M_{Swin-UNet}$
 生成伪标签 $\hat{y}_{Swin-UNet} = M_{Swin-UNet}(x)$
 小批次 $x \subset (D_U \& D_L \& \hat{y}_{UNet})$ 输入 M_{UNet}
 生成伪标签 $\hat{y}_{UNet} = M_{UNet}(x)$
 无监督损失 $Loss_S = W_D \cdot Loss_S^{UNet}(x, \hat{y}_{Swin-UNet}) + W_L \cdot Loss_S^{Swin-UNet}(x, \hat{y}_{UNet})$

总损失 $Loss = Loss_L + \lambda(t) \cdot Loss_U$

梯度计算 $G \leftarrow Loss \cdot \nabla$

更新子模型 $M_{UNet}, M_{Swin-UNet}$ 参数

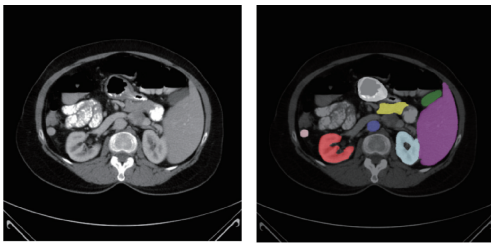
endwhile

2 实验

2.1 实验准备工作

1) 数据集

本文使用 MICCAI2015 多图谱腹部标记挑战赛中的 Synapse 多器官分割 CT 数据集进行实验。该数据集收集 50 例腹部 CT 扫描数据,对脾、右肾、左肾、胆囊、食道、肝、胃、主动脉、下腔静脉、门静脉和脾静、胰腺、右肾上腺、左肾上腺等 13 个器官进行了标注。为了公平对比实验结果,本文遵循了 TransUNet 方法介绍中对该数据集的处理方式,即从中选取 30 例患者的腹部 CT 扫描数据,共 3 779 个轴向对比增强腹部 CT 切片,并对其中的主动脉、胆囊、脾、左肾、右肾、肝、胰、胃等 8 类器官进行分割。图 3(a)所示为原图,图 3(b)为标签图,两幅图展示了该数据集中原图和标注后的图片区别。



(a) 原图 (b) 标签
(a) Original drawing (b) Labels

图3 Synapse 多器官分割数据集

Fig. 3 Synapse multi-organ segmentation dataset

2) 实验配置

对 Synapse 数据集按照 7:2:1 的比例分为训练集、测试集和验证集,同时将每个切片调整至 224×224 像素,强度重新缩放至 $[0, 1]$,接着利用随机裁剪、随机旋转、随机翻转等操作扩充训练样本至 5 290 张。这样既可以充分释放模型性能,又能避免模型过拟合。为了探究半监督学习模型标签数据在训练过程中的比例与分割性能之间的关系,特别设计了两组实验,将训练集中的标签数据占比分别设置为 5% 和 10% 进行对比分析。

实验硬件环境:CPU 为 Intel Core i7-12700H, GPU 为 NVIDIA GeForce RTX 3090; 软件环境: Python3.8, Pytorch1.12.1; 操作系统为 Windows 10。模型使用 SGD 优化器训练,学习率为 0.01, 动量 0.9, 权重衰减为 1×10^{-4} , Batch size 设为 24, 训练迭代次数为 14 k。

3) 评价指标

为了能够客观地评估医学图像分割效果,本文采用

Dice 系数和 Hausdorff 距离^[18] 作为评估依据。Dice 系数和 95% Hausdorff 距离(hd95)是常用于评估图像分割算法性能的两个指标。它们分别从重叠度和边缘距离的角度衡量预测结果与真实标签的相似性。对这两者进行对比分析的方法通常属于图像分割性能评价范畴。Dice 系数是一种用于衡量两个样本相似度的指标,通常用于图像分割和生物信息学。其计算公式为:

$$Dice = \frac{2(A \cap B)}{|A| + |B|} \quad (9)$$

其中, A 和 B 是两个集合, $|A|$ 和 $|B|$ 是它们的大小。Dice 系数也称重叠指数,取值范围为 $[0, 1]$, Dice 值越大越接近 1 表示相似度越高则模型分割效果越好。

Hausdorff 距离(Hausdorff distance, HD)是描述点集间相似程度的一种度量方式。HD 能够更好的衡量分割病灶区域边界准确度,其值越小,分割效果越好。其公式为:

$$d_H = \max \{ \max_{a \in A} \min_{b \in B} d(a, b), \max_{b \in B} \min_{a \in A} d(a, b) \} \quad (10)$$

其中, A 和 B 是两个点集, A 是分割结果的边界点集, B 是真实标注边界的点集, $d(a, b)$ 是点 a 和点 b 之间的距离。本文中评估 HD 使用 HD95, 即 Hausdor 距离最终结果值乘以 95%, 目的是为了消除离群值中极小的子集的影响。

2.2 实验结果分析

为了公正地评估 CDCL-SSLNet 的分割性能,将其与多种先进的半监督分割模型进行了对比,包括 CPS^[19]、C&T^[20]、UA-MT^[21]、SimiS^[22]、CLD^[23] 以及 Adsh^[24]。表 1 展示了各模型的分割结果,可以看出 CDCL-SSLNet 在主动脉、胆囊、左肾、肝、脾等 5 个器官上均取得了最佳的分割效果,整体性能明显优于其他分割模型。特别是在处理一些难以分割的小器官(如胆囊)时,CDCL-SSLNet 展现出了独特的优势。在仅有 5% 和 10% 的标签数据参与训练的情况下,其 Dice 系数分别达到了 65.71% 和 77.49%, HD95 则分别为 28.97 和 22.07。Dice 值越大表示相似度越高,模型分割效果越好,HD 值越小,分割效果越好,这两项指标的性能提升幅度均达到了 17%,充分证明了 CDCL-SSLNet 在数据类分布不均衡的场景下,对于小目标的分割具有显著的优势。

表 2 展示了各模型的参数量、浮点运算数(FLOPs)、推理分割一张切片的时间,其中 UA-MT 模型参数量最少,CLD 的 FLOPs 最小且推理所需时间少,但是纵向对比各模型推理时间差距不大,在此基础上 CDCL-SSLNet 能够获得较好的分割性能,则表明其综合性能最佳。

如图 4 所示,展示了各模型在使用 10% 的 Synapse 多器官分割标签数据进行训练后的分割效果图,CPS、C&T 和 UA-MT 模型在左肾的分割中出现明显漏割,并且在胃部的分割中出现了过度分割的问题。当面对小目标,如主动脉、胆囊和胰腺时,这些模型仅能分割出主要区域,对边

表 1 各模型在 Synapse 多器官分割数据集上的 Dice 系数、HD95 和 Dice(%) / HD95 分割结果
Table 1 Dice coefficients, HD95 and Dice (%) / HD95 of each model on the Synapse multi-organ segmentation dataset split the result

LD 比例	模型	Dice(%) / HD95									
		Dice ↑ / %	HD95 ↓	主动脉	胆囊	左肾	右肾	肝	胰腺	脾	胃
100%	U-Net	76.85	39.70	<u>89.07</u> 40.89	<u>69.72</u> 38.88	<u>77.77</u> 40.46	<u>68.60</u> 40.76	94.43 38.34	53.98 38.81	86.67 40.46	75.58 39
	Swin-UNet	79.13	21.55	85.47 21.74	66.53 21.69	83.28 22.33	79.61 22.4	94.29 20.63	56.58 22.17	90.66 19.58	76.60 21.85
5%	CPS	62.34	32.51	<u>67.46</u> 32.91	<u>51.49</u> 31.73	<u>67.1</u> 33.45	<u>65.84</u> 32.48	<u>65.2</u> 31.91	50.89 31.74	70.62 31.74	60.12 34.11
	C&T	62.48	31.99	<u>68.02</u> 33.54	<u>50.93</u> 30.86	67.92 31.85	<u>64.21</u> 30.91	<u>66.31</u> 31.55	49.93 33.43	71.5 31.57	61.02 32.2
	UA-MT	62.85	31.11	<u>68.24</u> 30.25	<u>52.71</u> 31.85	<u>64.26</u> 31.94	<u>65.01</u> 30.85	<u>67.19</u> 31.47	51.64 30.72	73.17 29.49	60.58 32.33
	SimiS	63.47	30.09	<u>67.94</u> 29.54	<u>51.66</u> 31.65	<u>64.85</u> 29.47	<u>63.28</u> 30.94	<u>67.59</u> 29.13	51.53 31.6	78.8 28.87	62.11 29.52
	CLD	64.28	29.34	<u>68.1</u> 29.79	<u>52.41</u> 28.88	<u>67.18</u> 30.25	<u>64.22</u> 28.49	<u>69.42</u> 28.67	50.5 28.96	79.7 31.14	62.71 28.55
	Adash	63.57	29.55	<u>69.38</u> 30.66	<u>52.99</u> 29.26	<u>64.51</u> 28.43	<u>63.99</u> 30.67	<u>69.54</u> 29.3	53.70 30.27	72.46 28.69	61.99 29.12
	CDCL-SSLNet (本文)	65.71	28.97	70.03 27.95	56.7 29.32	67.89 30.37	66.29 27.84	70.01 27.56	52.94 28.81	77.58 29.94	64.24 29.97
	CPS	74.28	24.55	<u>79.22</u> 23.24	<u>59.37</u> 23.45	<u>75.94</u> 25.73	<u>74.43</u> 24.37	<u>86.95</u> 25.48	57.18 23.72	85.65 24.82	75.5 25.59
	C&T	75.31	23.73	<u>80.26</u> 22.48	<u>60.28</u> 25.41	<u>77.21</u> 23.45	<u>76.76</u> 23.54	<u>87.96</u> 23.63	56.73 22.71	87.59 23.32	75.69 25.3
	UT-MT	75.88	24.34	<u>79.68</u> 24.22	<u>61.86</u> 24.87	<u>77.98</u> 23.13	<u>78.22</u> 22.48	<u>88.99</u> 24.98	57.27 22.48	88.06 25.38	74.98 25.41
SimiS	76.04	22.97	<u>81.94</u> 23.05	<u>61.93</u> 24.04	<u>78.13</u> 24.28	<u>77.83</u> 22.68	<u>87.86</u> 21.87	56.89 21.28	87.83 23.44	75.91 23.12	
CLD	76.39	23.89	<u>81.71</u> 22.27	<u>62.57</u> 24.81	<u>78.62</u> 25.09	<u>78.37</u> 23.92	<u>89.91</u> 22.81	55.67 23.01	87.93 24.58	76.34 24.63	
Adash	76.46	23.91	<u>82.96</u> 23.44	<u>62.79</u> 24.18	<u>79.15</u> 23.7	79.86 25.3	<u>89.97</u> 25.07	54.98 23.77	87.08 22.34	74.89 23.48	
CDCL-SSLNet (本文)	77.49	22.07	83.81 21.37	63.47 23.81	81.43 21.66	79.69 22.74	90.12 21.00	55.99 21.23	89.45 21.15	75.96 23.61	

缘的像素判断能力较弱,导致边缘分割模糊不清。尽管 SimiS、CLD、Adash 以及 CDCL-SSLNet 都具备处理类别分布不均衡的能力,但它们在分割效果上呈现出显著差异,CDCL-SSLNet 的分割结果几乎与真实标签完全吻合,这充分证明了其半监督学习策略的优越性。

2.3 消融实验及模型探究

1) CDCL-SSLNet 半监督学习框架探究

为了验证 CDCL-SSLNet 模型的有效性,本研究从半

监督策略、类分布公平因子和类学习公平因子 3 个方面进行了深入探究。CDCL-SSLNet 模型采用 UNet 和 Swin-UNet 两个子网络进行交叉学习,这是一种隐式一致性正则化学习策略。该策略通过整合 CNN 学习到的局部信息与 Transformer 学习到的全局信息,实现了精细化学习。为了证明此方法的优势,在 Synapse 多器官分割数据集上进行了实验,实验中设置 10% 的标签数据参与训练。将 CDCL-SSLNet 中的两个子模型均置换成同种模型,实质

表 2 各模型参数量、FLOPs、推理时间对比

Table 2 Comparison of parameters, FLOPs, and inference time of each model

模型	参数量/	FLOPs/	推理时间/
	M	G	ms
CPS	82.50	60.51	327
C&T	96.32	48.84	322
UA-MT	58.44	42.72	325
SimiS	78.60	42.68	320
CLD	74.30	42.42	321
Adash	86.25	44.28	322
CDCL-SSLNet(本文)	99.28	45.94	323

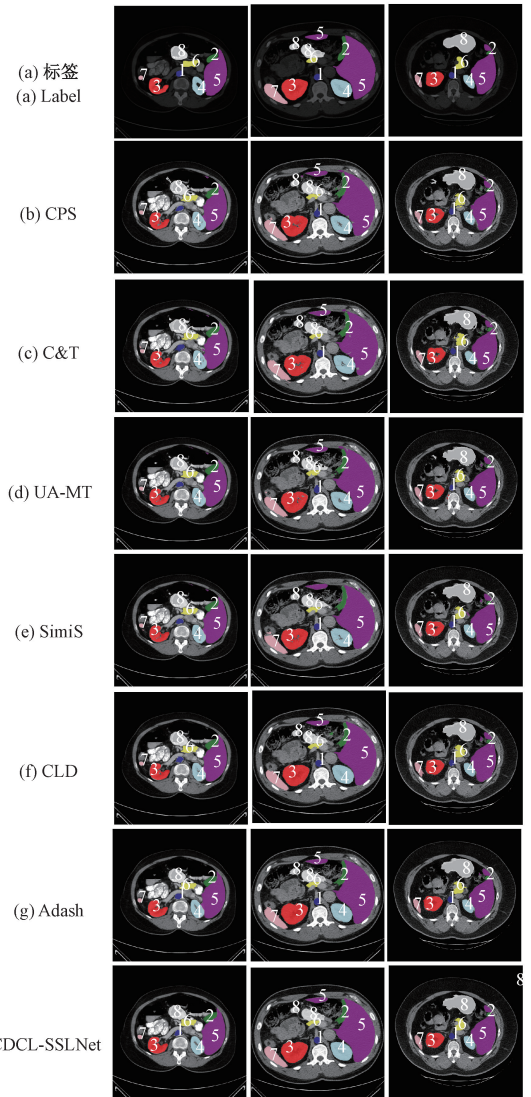


图 4 各模型在 Synapse 多器官分割数据集上的分割结果图

Fig. 4 Segmentation results of each model on the Synapse multi-organ segmentation dataset

表 3 组合模型在 Synapse 多器官数据集上的 Dice 系数、HD95 和 Dice (%) / HD95 分割结果

Table 3 Dice coefficients, HD95 and Dice (%) / HD95 segmentation results of the combined model on the Synapse multi-organ dataset

子模型 1	子模型 2	Dice(%) / HD95									
		Dice ↑ / %	HD95 ↓	主动脉	胆囊	左肾	右肾	肝	胰腺	脾	胃
U-Net	U-Net	63.87	33.01	64.21	62.1	65.87	63.88	67.09	60.99	65.34	61.48
				33.85	31.52	33.38	31.34	33.43	33.99	32.67	33.9
Swin-UNet	Swin-UNet	74.97	34.25	76.39	62.28	76.97	73.04	88.02	69.31	81.80	71.95
				35.21	32.94	34.93	35.06	35.49	33.23	32.96	34.18
U-Net	Swin-UNet	77.49	22.07	83.81	63.47	81.43	79.69	90.12	55.99	89.45	75.96
				21.86	23.01	20.09	21.32	22.86	22.58	21.97	22.88

上就变成了 CPS 方法。因此,参照 CPS 方法设置两个子模型不同的初始权重,以实现两个子模型的学习差异化。

表 3 给出了实验结果,可以看出 UNet 和 Swin-UNet 组合模型获得了最优的分割结果。相比之下,UNet 和 UNet 组合模型的 Dice 系数比标准的 CDCL-SSLNet 低了 17.58%,HD95 高了 49.57%;而 Swin-UNet 和 Swin-UNet 组合模型的 Dice 系数比 CDCL-SSLNet 低了 3.25%,HD95 高了 55.19%。这一方面表明 Swin-UNet 具有较强的全局建模能力,使得其 Dice 系数高于 UNet;另一方面也注意到 Swin-UNet 和 Swin-UNet 组合模型的 HD95 并没有 UNet 和 UNet 组合模型低。这可能是因为 UNet 中的卷积层对局部信息感知较好,特别是对边缘细节的感知。因此,UNet 和 Swin-UNet 组合模型即 CDCL-SSLNet 能够充分利用 CNN 的局部感知能力和 Transformer 的长远程信息感知能力,制造出比显式一致性正则化更稳定和准确的伪标签,为后续分割任务奠定基础。

2) 类分布公平因子与类学习公平因子探究

为了深入探究类分布公平因子(D)与类学习公平因子(L)对 CDCL-SSLNet 性能提升的具体贡献,本研究构建了多种损失加权学习策略,包括 D-D、L-D 和 L-L。同时,借

鉴 SimiS 中提出的对少数类进行标签扩充的方法(S),进一步构建了 D-S、S-D、S-L、L-S 和 S-S 等 5 种标签扩充学习策略。这些策略在 Synapse 多器官分割数据集上进行了实验验证,实验中设置 10% 的标签数据参与训练。值得注意的是,在 CDCL-SSLNet 中,类分布公平因子作用于 U-Net,而类学习公平因子作用于 Swin-UNet。因此,

上述构建的学习策略 D-D、L-D、L-L、D-S、S-D、S-L、L-S、S-S 均具有顺序性,其中“—”前面的部分表示作用于 U-Net,后面的部分表示作用于 Swin-UNet。实验结果如表 4 所示。通过这些实验,本研究旨在全面评估不同学习策略对 CDCL-SSLNet 性能的影响,从而为模型优化提供有力支持。

表 4 组合学习策略在 Synapse 多器官数据集上的 Dice 系数、HD95 和 Dice(%) / HD95 分割结果
Table 4 Dice coefficient, HD95 and Dice(%) / HD95 segmentation results of combinatorial learning strategy on Synapse multi-organ dataset

学习策略			Dice(%) / HD95							
	Dice ↑ / %	HD95 ↓	主动脉	胆囊	左肾	右肾	肝	胰腺	脾	胃
D-D	74.01	29.18	$\frac{78.89}{28.03}$	$\frac{59.97}{27.64}$	$\frac{76.05}{30.31}$	$\frac{75.14}{29.74}$	$\frac{84.32}{27.82}$	$\frac{57.09}{29.61}$	$\frac{85.83}{30.27}$	$\frac{74.79}{30.01}$
L-D	75.73	24.21	$\frac{79.92}{23.71}$	$\frac{61.01}{25.31}$	$\frac{77.9}{23.48}$	79.99	$\frac{87.94}{23.33}$	58.06	$\frac{87.01}{24.16}$	$\frac{74.01}{24.19}$
L-L	71.62	25.29	$\frac{76.96}{25.99}$	$\frac{57.38}{26.67}$	$\frac{72.24}{24.86}$	$\frac{72.66}{26.53}$	$\frac{85.01}{24.11}$	$\frac{55.56}{24.67}$	$\frac{79.04}{25.6}$	$\frac{74.11}{23.88}$
D-S	73.05	26.17	$\frac{78.33}{26.85}$	$\frac{60.03}{26.29}$	$\frac{75.69}{24.08}$	$\frac{73.37}{26.55}$	$\frac{86.91}{26.71}$	$\frac{57.02}{26.44}$	$\frac{79.28}{27.05}$	$\frac{73.77}{25.4}$
S-D	73.17	27.89	$\frac{79.01}{26.42}$	$\frac{60.17}{26.81}$	$\frac{75.56}{29.27}$	$\frac{73.42}{28.1}$	$\frac{87.02}{28.54}$	$\frac{56.91}{27.32}$	$\frac{79.16}{28.09}$	$\frac{74.11}{28.58}$
S-L	73.22	26.45	$\frac{79.42}{27.09}$	$\frac{59.88}{26.31}$	$\frac{76.01}{24.9}$	$\frac{73.78}{27.11}$	$\frac{86.62}{26.12}$	$\frac{56.41}{27.42}$	$\frac{79.79}{26.86}$	$\frac{73.85}{25.8}$
L-S	71.67	31.45	$\frac{76.52}{32.82}$	$\frac{58.21}{30.21}$	$\frac{72.79}{30.92}$	$\frac{72.18}{31.09}$	$\frac{85.23}{32.15}$	$\frac{56.27}{32.86}$	$\frac{77.85}{30.36}$	$\frac{74.31}{31.19}$
S-S	72.97	32.17	$\frac{78.21}{33.5}$	$\frac{58.83}{31.04}$	$\frac{75.17}{32.42}$	$\frac{73.03}{33.42}$	$\frac{85.29}{31.45}$	$\frac{56.27}{30.51}$	$\frac{84.03}{32.99}$	$\frac{72.93}{32.1}$
CDCL-SSLNet(本文)	77.49	22.07	$\frac{83.81}{22.26}$	$\frac{63.47}{21.68}$	$\frac{81.43}{21.16}$	$\frac{79.69}{22.83}$	90.12	$\frac{55.99}{22.6}$	89.45	$\frac{75.96}{21.94}$

从实验结果中可以明显看出, D-L 组合策略, 即 CDCL-SSLNet, 展现出了最佳的分割性能。当类分布公平因子置于 Swin-UNet 上, 而类学习公平因子置于 U-Net 上, 即采用 L-D 策略时, 尽管与本文提出的 CDCL-SSLNet 相比只是因子作用的模型发生了调换, 但最终的分割结果却呈现出较大的差距。具体而言, 其 Dice 系数较 CDCL-SSLNet 低 2.27%, 而 HD95 则较 CDCL-SSLNet 高 9.69%。这一发现揭示了不同类型的因子与模型之间存在着一种最佳的耦合关系, 这种关系能够为模型带来显著的性能提升。

然而, 单纯地引入类分布公平因子并不能有效地改善整体的分割性能。例如, D-D 策略的 Dice 系数较 CDCL-SSLNet 低 4.49%, 而 HD95 则较 CDCL-SSLNet 高 32.21%。这主要是因为某些器官, 尽管分布广泛, 但其学习难度并不仅仅与分布有关, 还与器官自身的特性密切相

关。以胃部为例, 它呈现为一个腔体, 在医学图像上表现为中间镂空状, 这使得模型难以对中间部位的像素进行准确归类。

同样地, 仅仅引入类学习公平因子而忽视类分布因素也难以实现理想的分割效果。甚至在某些情况下, 这种策略的性能可能不如 D-D 策略。SimiS 中提出的伪标签扩充法, 由于其扩充过程缺乏严格的约束, 难以保证生成的伪标签的正确性。因此, 采用 D-S、S-L 和 S-S 策略所获得的性能均未能超越 CDCL-SSLNet。

此外, 通过引入 S 类策略进行对比, 进一步验证了本文提出的类分布公平因子和类学习公平因子的有效性和优越性。为了更直观地展示这两个因子对分割性能提升的作用, 图 5 展示了上述几种策略的分割效果图。

根据图 5 的实验结果可以看出, 本模型的分割性能优异。具体来说, L-L、L-D、D-S 和 L-S 模型在肝脏分割中表

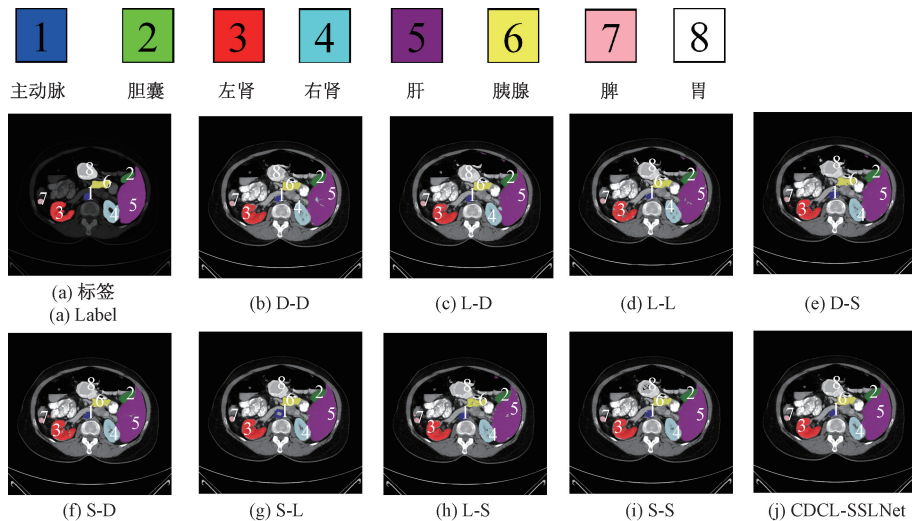


图 5 各种组合策略在 Synapse 多器官分割数据集上的分割结果图

Fig. 5 Segmentation results of various combination strategies on the Synapse multi-organ segmentation dataset

现出分割不精确的现象,而 S-D 模型则在左肾的分割上出现了过度分割。相比之下,本模型分割精确,无明显的过度分割问题。尤其值得注意的是,本模型在小目标和复杂环境下的主动脉分割效果显著优于其他模型。与 S-L 模型相比,S-L 模型在主动脉分割上表现不佳,分割结果呈现中空环状,反映出其在边缘像素判断方面的不足,导致边缘模糊不清。而本模型能够准确、完整地分割边缘。此外,除本模型外,其他模型均存在分割结果不完整的情况。综合对比来看,本文提出的基于类分布公平因子和类学习公平因子的 CDCL-SSLNet 模型在分割性能上优于其他模型。

3 结 论

本文构建的 CDCL-SSLNet 融合了 U-Net 与 Swin-UNet 子模型,实现了高效的交叉学习。针对医学图像资源短缺以及其中普遍存在的类分布不均衡问题,提出了类分布公平因子对损失函数进行加权处理。同时,考虑到胃部、胰腺等特殊类别器官的学习难度,又引入了类学习公平因子,进一步对损失进行加权调整。通过实验探索,找到了两种公平因子与子模型之间的最佳耦合关系,从而显著提升了分割性能。CDCL-SSLNet 可根据数据中语义类别的分布及学习特点动态地调整损失比重。这种策略有效地减小了模型的学习偏差,显著提升了模型的整体分割性能,特别是在处理小目标和复杂场景时展现出较好的性能。

参考文献

[1] 曹玉红,徐海,刘荪傲,等. 基于深度学习的医学影像分割研究综述[J]. 计算机应用, 2021, 41(8): 2273-2287.
CAO Y H, XU H, LIU S AO, et al. A review of

medical image segmentation based on deep learning[J]. Journal of Computer Applications, 2021, 41(8): 2273-2287.

- [2] 曾黎,汤红忠,王蔚,等. 结合多任务学习的半监督病理图像分割方法[J]. 协和医学杂志, 2023, 14(2): 416-425.
ZENG L, TANG H ZH, WANG W, et al. Semi-supervised pathological image segmentation method combined with multi-task learning[J]. Medical Journal of Peking Union Medical College Hospital, 2023, 14(2):416-425.
- [3] EGGER J, GSAXNER C, PEPE A, et al. Medical deep learning a systematic meta-review[J]. Computer Methods and Programs in Biomedicine, 2022, 221:106874.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, 2015: 234-241.
- [5] YIN X X, SUN L, FU Y H, et al. U-Net-based medical image segmentation[J]. Journal of Healthcare Engineering, 2022,2022: 4189781.
- [6] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, DOI: 10.48550/arXiv.1706.03762.
- [7] FANG J, YANG CH, SHI Y T, et al. External attention based TransUNet and label expansion strategy for crack detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10):

- 19054-19063.
- [8] CAO H, WANG Y Y, CHEN J, et al. Swin-Unet: Unet-like pure transformer for medical image segmentation[C]. European Conference on Computer Vision, 2022: 205-218.
- [9] 刘坤, 王典, 荣梦学. 基于半监督生成对抗网络 X 光图像分类算法[J]. 光学学报, 2019, 39(8): 0810003. LIU K, WANG D, RONG M X, et al. X-ray image classification algorithm based on semi-supervised generative adversarial network [J]. Acta Optica Sinica, 2019, 39(8): 0810003.
- [10] JIAO R SH, ZHANG Y CH, DING L, et al. Learning with limited annotations: A survey on deep semi-supervised learning for medical image segmentation[J]. Computers in Biology and Medicine, 2023, 169: 107840.
- [11] 王鹏举, 张晓, 冀振燕, 等. 面向医学图像分割的傅里叶半监督学习方法[J]. 北京邮电大学学报, 2024, 47(3): 69. WANG P J, ZHANG X, JI ZH Y, et al. Fourier semi-supervised learning method for medical image segmentation[J]. Journal of Beijing University of Posts and Telecommunications, 2024, 47(3): 69.
- [12] CHEN X K, YUAN Y H, ZENG G, et al. Semi-supervised semantic segmentation with cross pseudo supervision[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 2613-2622.
- [13] LUO X D, HU M H, SONG T, et al. Semi-supervised medical image segmentation via cross teaching between CNN and transformer [C]. International Conference on Medical Imaging with Deep Learning, PMLR, 2022: 820-833.
- [14] YU L Q, WANG SH J, LI X M, et al. Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation [C]. Medical Image Computing and Computer Assisted Intervention-MICCAI 2019: 22nd International Conference, 2019: 605-613.
- [15] GUO L ZH, LI Y F. Class-imbalanced semi-supervised learning with adaptive thresholding [C]. International Conference on Machine Learning, PMLR, 2022: 8082-8094.
- [16] WANG Y D, YU ZH H, WANG J D, et al. Exploring vision-language models for imbalanced learning [J]. International Journal of Computer Vision, 2024, 132(1): 224-237.
- [17] LIN Y Q, YAO H F, LI Z ZH, et al. Calibrating label distribution for class-imbalanced barely-supervised knee segmentation [C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, Cham: Springer Nature Switzerland, 2022: 109-118.
- [18] TAHA A A, HANBURY A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool[J]. BMC Medical Imaging, 2015, 15(1): 29.
- [19] QIN F CH, GUO J M, LANG F K. Superpixel segmentation for polarimetric SAR imagery using local iterative clustering[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(1): 13-17.
- [20] QURESHI I, YAN J H, ABBAS Q, et al. Medical image segmentation using deep semantic-based methods: A review of techniques, applications and emerging trends[J]. Information Fusion, 2023, 90: 316-352.
- [21] CHEN ZH CH, YANG J, CHEN L F, et al. Efficient railway track region segmentation algorithm based on lightweight neural network and cross-fusion decoder [J]. Automation in Construction, 2023, 155: 105069.
- [22] JIAO R SH, ZHANG Y CH, DING L, et al. Learning with limited annotations: A survey on deep semi-supervised learning for medical image segmentation[J]. Computers in Biology and Medicine, 2023, 169: 107840.
- [23] DALEN J, HOFFMANN A, DICKEN V, et al. A novel iterative method for lesion delineation and volumetric quantification with FDG PET[J]. Nuclear Medicine Communications, 2007, 28(6): 485-493.
- [24] SORET M, BACHARACH S L, BUVAT I. Partial-volume effect in PET tumor imaging[J]. Journal of Nuclear Medicine, 2007, 48(6): 932-945.

作者简介

武丽, 硕士, 副教授, 主要研究方向为人工智能、模式识别。

E-mail: 79901056@qq.com

丁琴, 硕士研究生, 主要研究方向为人工智能、图像处理、目标检测。

E-mail: 1640543204@qq.com

葛彩成(通信作者), 硕士研究生, 主要研究方向为机器学习、图像处理。

E-mail: 853136455@qq.com