

郭逸凡, 林佳成, 潘济, 等. 采用不对称聚焦加权 Dice 损失分割腹部 CT 图像[J]. 智能计算机与应用, 2024, 14(7): 20-28.
DOI: 10.20169/j.issn.2095-2163.240703

采用不对称聚焦加权 Dice 损失分割腹部 CT 图像

郭逸凡¹, 林佳成², 潘济¹, 蒋婷³

(1 西南交通大学希望学院 信息工程系, 成都 610400; 2 厦门理工学院 计算机与信息工程学院, 福建 厦门 361024;
3 吉利学院 智能科技学院, 成都 641423)

摘要: 以 UNet 作为主干网络的少样本学习在医学图像分割领域广泛应用, 但稀少的数据样本带来了数据失衡问题。为了解决该问题, 提出一种采用不对称聚焦加权 Dice 损失用于腹部 CT 图像训练分割。通过在 Dice 损失中引入不对称结构思想和同方差不确定性策略, 用以重新加权损失项, 缓解数据失衡问题, 在 Synapse 和 AbdomenCT-1K 腹部多器官数据集上比联合 Dice 和交叉熵损失函数的基线方法, 平均 Dice 分数分别提升 2.01% 和 2.75%。应用至最先进分割模型和 3D 图像数据集上的实验结果表明, 采用本文所提方法, 不但有更高的分割性能, 而且可视化结果更平滑不嘈杂。

关键词: 医学图像分割; 数据失衡; Dice 损失

中图分类号: TP391.41

文献标志码: A

文章编号: 2095-2163(2024)07-0020-09

Segmentation of Abdominal CT images with Asymmetric Focal Weighted Dice loss

GUO Yifan¹, LIN Jiacheng², PAN Ji¹, JIANG Ting³

(1 Department of Information Engineering, Southwest Jiaotong University Hope College, Chengdu 610400, China;
2 School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361024, Fujian, China;
3 College of Intelligent Science and Technology, Geely University, Chengdu 641423, China)

Abstract: The U-Net, serving as the backbone network, is widely applied in the field of medical image segmentation for few-shot learning. However, lack of adequate data cause data imbalance problem. To address the problem, we propose Asymmetric Focused Weighted Dice Loss function. To solve the problem, we propose a method using Asymmetric Focus Weighted Dice loss for abdominal CT image segmentation. We mitigate the data imbalance by introducing the asymmetric structure idea and homoscedastic uncertainty strategy in Dice loss, which is used to reweight the loss term. our method improves the average Dice scores by 2.01% and 2.75% over the baseline method with joint Dice and cross-entropy loss functions on Synapse and AbdomenCT-1K abdominal multiorgan datasets, respectively. Experimental results on applications to state-of-the-art segmentation models and 3D datasets show that models using our method not only have higher segmentation performance, but also have smoother and less noisy visualization results.

Key words: medical image segmentation; data imbalance; Dice loss;

0 引言

医学图像分割在计算机辅助诊断和智能医疗发挥重要作用。借助于深度学习技术对可疑病灶和肿瘤区域进行精准分割, 能够定性定量地分析患者病情和术后恢复情况, 更好地对疾病诊断与分析, 减轻医生的负担^[1]。常见的医学图像分割任务包括有

脑肿瘤分割、肺分割、心脏分割以及腹部多器官分割等。在这些分割目标中, 针对腹部多器官分割仍然具有一定的挑战性^[2]。首先, 腹部待分割器官通常涉及到胃、肝脏、肾脏、胰腺等多种器官, 数目较多且标签较少; 其次, 待分割器官由于各自解剖学结构和形态学差异, 进一步加剧了分割难度; 最后, 腹部图像涉及到数据失衡问题, 不同器官的相对体积占比

作者简介: 郭逸凡(1997-), 男, 硕士, 助教, 主要研究方向: 医学图像配准, 医学图像分割; 林佳成(1999-), 男, 硕士研究生, 主要研究方向: 医学图像分割, 多模态融合; 潘济(1998-), 女, 硕士, 助教, 主要研究方向: 量子图像安全, 量子图像处理。

通讯作者: 蒋婷(1995-), 女, 硕士, 助教, 主要研究方向: 医学图像分割, 医学图像配准。Email: sillyl@foxmail.com

收稿日期: 2024-01-21

哈尔滨工业大学主办 ◆ 学术研究与应用

不同,小目标区域通常比大目标区域分割性能要差。

随着 UNet 的提出,快速而准确地对医学图像进行语义级别的分割成为可能^[3]。这类 U 型网络架构通常包含对称的编码器—解码器结构和跳跃连接步骤,能够很好地应对医学图像的高分辨率、纹理结构复杂的特点^[4]。凭借这一简单有效的架构,许多 UNet 的变种被提出,包括 Milletari 等^[5]提出的引入残差结构的 VNet;Zhou 等^[6-7]提出的集成不同深度的 UNet++;Huang 等^[8]提出的改进跳跃连接的 UNet3+;Çiçek 等^[9]提出的扩展至 3D 图像的 3DUNet。

在这些研究中,网络结构是影响分割的一个至关重要因素,因此通过改进或者引入更强大的特征提取模块,来捕捉医学图像中感兴趣的区域^[10]。然而,忽视了对图像中数据失衡问题关注,当使用的数据严重失衡时,即使是采用非常强大的网络也可能导致次优的预测结果^[11]。

医学图像中的数据失衡问题可以划分为两类:第一种是前景和背景之间的比例不平衡,前景是指感兴趣的待分割区域,背景是指除前景以外的无关区域,包括正常组织和其他无关区域;第二种是样本之间难易不平衡。由于不同待分割样本存在解剖学和形态学上的差异,在一定程度上决定了分割的难度。上述问题在 2D 图像分割任务中尤为突出,部分器官不仅占据的像素比例少,有时甚至几乎没有出现,使得在训练期间,总损失被某些类别所占据,并主导了梯度,从而影响了模型的泛化能力^[12]。因此,通过对损失项仔细地选定和改进,可以在不引入新训练数据的前提下,增强待分割区域的关注力度,提高分割精度。

在医学图像分割领域,损失函数的选定通常在 Dice 损失、交叉熵损失或者这二者损失的改进型中选择。在这种情况下,不仅有大量的损失函数可供选择,难以合理地调整不同损失函数的权重^[13]。

受 Focal 损失的启发,特别是其在医学图像处理领域中的应用,包括 Focal Dice 损失和采用不对称结构损失,本文提出一种新颖的应用在 Dice 损失函数,并命名为不对称聚焦加权(Asymmetric Focal Weighted, AFW) Dice 损失函数。为解决在 2D/3D 图像上存在比例不平衡问题,本文构建一种不对称损失结构,以区分感兴趣的前景和背景损失值;对于样本难易不平衡问题,将聚焦因子施加于 Dice 损失上,以达到抑制和强化作用;针对更好地设置不同损失权重问题,采用同方差不确定性策略,在训练过程中能够自学习调整权重值。

1 相关工作

1.1 基于深度学习分割方法

近年来, Dosovitskiy 等^[14]提出的 Vision Transformer (ViT) 在计算机视觉任务取得成功,最近的工作也集中于研究其在医学分割任务中的适用范畴。例如 Chen 等^[15]提出融合 ViT 和卷积网络的 TransUNet 模型;Cao 等^[16]提出使用滑动窗口机制的 Swin-UNet 模型;Hatamizadeh 等^[17]提出应用至 3D 图像上的 UNETR 模型以及 Shaker 等^[18]提出使用高效成对注意模块的 UNETR++ 模型等。这些模型通过对图像特征进行长远距离建模,获取全局上下文信息,在图像分割任务中具有有效性^[19]。然而,ViT 一些固有设置缺陷,特别是补丁块与补丁块之间没有足够的冗余关联,对可变分辨率图像支持力度低,将图像单纯视为线性序列,给处理医学图像带来了挑战^[20-22]。相比之下,全卷积网络所拥有的计算量低、局部建模强、易于训练等特性,仍然受研究者的关注。

通过构建与自注意力机制类似的大核注意力,成为了研究热点。VAN 和 SegNeXt 分别使用可分解的大核卷积、不同尺寸的条状卷积实现的注意力机制,取得了匹敌 Transformer 性能的效果^[23-24]。特别是将大核注意力应用至医学图像分割领域的 D-LKA 和 VTransUNet 证明了由纯卷积实现的网络不仅可以模拟 Transformer 全局感知的特点,同时也能够保留卷积局部感知、计算量低的优势,说明全卷积网络仍然具有不可替代的潜力^[25-26]。

1.2 针对数据失衡问题的损失函数

损失函数定义了分割任务的约束条件,并直接影响训练期间模型的收敛情况。在一定程度上,损失函数的优劣决定了网络模型建模能力的强弱,通过对损失函数进行改进,能够缓解因数据匮乏导致的数据失衡问题,进而让模型稳定发挥性能。交叉熵(Cross Entropy, CE)损失作为使用最广泛的基于分布的损失函数,缺乏应对失衡问题的能力。改进型的加权交叉熵(Weighted Cross Entropy, WCE)通过对每一个子任务的损失项进行重新加权,来抵消类别不平衡^[1]。近年来提出的 Focal 损失作为交叉熵损失的变体,通过引入指数形式的强化因子简单地降低某些类别的贡献,克服了标准交叉熵损失的缺陷^[12]。Li 等^[27]指出 Focal 损失中强化因子不必实施在所有类别中,构建不对称结构的损失函数有能力选择性强化和抑制某些类别。

借助 Focal 损失概念,多位学者将其推广至医学图像分割领域中基于区域的 Dice 损失、Tversky 损失。Sudre 等^[28]提出广义 Dice (Generalised Dice, GD) 损失,按照标签比例占比重新加权损失项权重; Wang 等^[29]提出 Focal Dice 损失,通过简单地使用强化因子,针对性地对脑肿瘤的不同子区域进行关注,提升了分割性能; Wong 等^[30]提出应用在脑部分割任务的指数对数损失,强化了 Dice 损失非线性表达能力; Abraham 等^[31]提出 Focal Tversky 损失,在 Tversky 损失中引入 Focal 损失的思想,帮助集中于难分割类别; Yeung 等^[13]提出 Unified Focal 损失框架,联合多个聚焦形式的损失项,在 2D/3D 图像数据集上进行验证,展现了该策略的优势。然而,如上工作通常在 2 类或者很少的类别(通常小于 4 类)上进行实验,缺乏在更多分割类别的实验结果;其次,引入的损失项,例如 Tversky 函数,包含了多个超参数,这使得调参变得尤为困难。

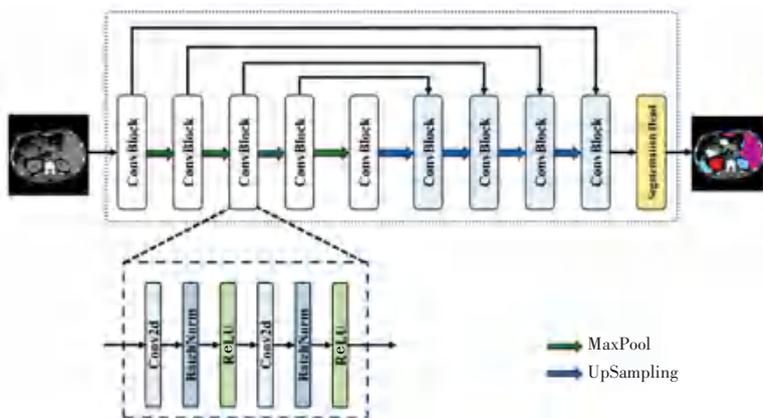


图1 网络结构

Fig. 1 Network structure

2.2 同方差不确定性

贝叶斯概率模型认为,损失项权重的选定与不确定性密切相关,而这种选定是一种动态自学习过程^[32]。不确定性可以由数据引起的(偶然不确定性),例如在数据采集过程中不可避免产生的误差;也可以是模型的自身引起的(认知不确定性),例如模型对新样本的泛化能力不足所产生的误差^[33]。在假定数据的噪声误差分布为常数的条件下(即作同方差),不确定性取决于不同任务的难易程度。基于同方差不确定性实现的损失函数如公式(1)所示:

$$L = \sum_{j=1}^T \left(\frac{1}{\sigma_j^2} L_j + \ln \sigma_j^2 \right) = \sum_{j=1}^T \frac{1}{\sigma_j^2} L_j + \ln \prod_{j=1}^T \sigma_j^2 \quad (1)$$

其中, T 表示任务 (Tasks), 这些任务可以是大大任务, 如配准、重建和模态转换任务, 也可以是小任

2 不对称聚焦加权 Dice 损失分割网络

2.1 网络结构

本文使用原始的 U 型网络结构来简化模型并强调损失函数对性能的重要性, 即仅由卷积层、池化层和上采样层组成。该网络结构如图 1 所示, 其中编码器包含 5 个卷积模块, 解码器包含 4 个卷积模块。每个卷积模块由两个步长为 1, 填充为 1, 卷积核大小为 3×3 的卷积运算、两个 ReLU 激活函数和批量归一化 (BatchNorm) 层级联而成; 编码器和解码器的每个卷积模块后都有一个池化层来减半图像分辨率, 上采样层用于恢复图像分辨率; 通过跳跃连接的方式, 编码器和解码器之间的图像特征信息被拼接融合; 最后, 经过一个分割头 (Segmentation Head) 模块, 得到网络预测分割结果。编码器通道数和解码器通道数分别设置为 32、64、128、256、512 和 256、128、64、32。

务, 如分割任务中不同的分割类别; σ_j^2 本质上为第 j 个任务的方差, 其倒数 $\frac{1}{\sigma_j^2}$ 表示损失项的权重, 同时度量了在第 j 个任务上的不确定性, 这也是一个可训练参数, 在网络训练期间随反向传播而不断更新优化; 最后的对数项可以看成对损失项的修正, 避免权重 $\frac{1}{\sigma_j^2}$ 趋于 0 时, 总损失也趋于 0 的情况。

此外, 对数项可以额外添加 1, 防止损失呈现负值和无法训练的情况发生^[34]。因此, 修改后损失函数如公式(2)所示:

$$L = \sum_{j=1}^T \frac{1}{\sigma_j^2} L_j + \ln \prod_{j=1}^T (\sigma_j^2 + 1) \quad (2)$$

尽管最初的同方差不确定性策略应用在均方误差 (Mean Squared Error, MSE) 损失函数和 CE 损失

函数上。现有在医学图像处理领域工作和实验表明,这种策略仍然可以在 Dice 损失中发挥作用^[35-37]。

值得注意的是,同方差不确定性策略倾向于对低方差的任务分配更多的权重,即这种策略一定程度上优先关注任务中的容易样本而不是困难样本^[38]。

2.3 损失函数

2.3.1 CE 损失函数和 Dice 损失函数

CE 损失函数常用于度量两个不相似的分布,其定义如式(3)所示:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C g_i^c \ln(s_i^c) \quad (3)$$

其中, N 表示像素数目; g_i^c 是一个关于分割类别 c 在像素 i 上的判别符号函数,取值为 1 或者 0; s_i^c 则是对应的预测概率。

Dice 损失函数通过归一化度量后的预测值和真值重叠程度来估计损失值,定义如式(4)所示:

$$L_{Dice} = 1 - 2 \times \frac{\sum_{i=1}^N \sum_{c=1}^C s_i^c g_i^c + \varepsilon}{\sum_{i=1}^N \sum_{c=1}^C (s_i^c + g_i^c) + \varepsilon} \quad (4)$$

其中, ε 表示避免分母为 0 的平滑常数。

2.3.2 不对称聚焦加权 Dice 损失函数

Focal 损失使用同一个强化因子对所有类别进行增强,没有考虑前景和背景之间的差异性,同时也抑制了感兴趣类别的损失^[13,27]。相比之下,本文采用不对称结构移除与背景有关的强化因子,保留感兴趣类别的增强。类似广义 Dice (Generalized Dice, GD) 损失,本文对每个类别提供了对应的权重,定义为式(5):

$$L_{AFWDice} = \sum_{c \neq r}^C w_c L_{Dice_c} + \sum_{c=r}^C w_c (L_{Dice_c})^\gamma \quad (5)$$

其中, w_c 和 L_{Dice_c} 分别表示分割类别 c 的损失权重和 Dice 损失函数; γ 表示调节 Dice 损失函数的强化因子;本文用稀有 (rare) 类别 r 表示聚焦感兴趣的分割器官。

本文提出的损失对感兴趣分割类别设置不同的 γ 值所对应的损失值如图 2 所示。特别地,当 γ 取 1 时,等价于未聚焦,这时退化为经典的 Dice 损失;当 γ 取值介于 0~1 之间时,其对应损失值总比 Dice 损失值要高,表明对感兴趣样本,该策略总会呈现一定程度增强而不是抑制。

此外,图 2 还表明,当 Dice 值趋于 1 时, γ 值为

0.25 的曲线出现剧烈陡峭现象;当 Dice 值趋于 0 时,该曲线较为平滑;说明训练期间容易样本在 Dice 指标上些微变化将产生较大的波动,而困难样本的 Dice 值变化较平稳。从同方差不确定性策略的角度看,这种大波动意味着具有较大的方差。因此,不确定策略将帮助抑制这些容易样本的权重,变相强化困难样本的权重。比较设置不同 γ 值产生的曲线,将强化因子 γ 值设置为 0.25 可能具有一定的优势。

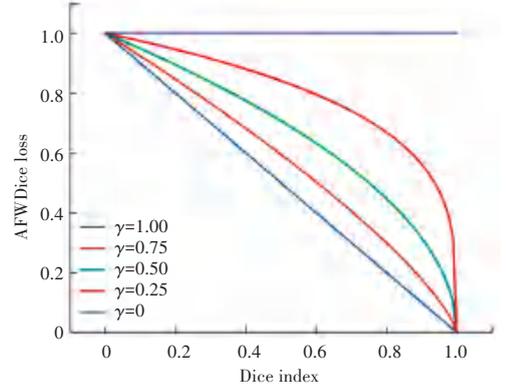


图 2 改变 γ 值对 AFWDice 损失的影响

Fig. 2 Effect of changing γ with the AFWDice loss

2.3.3 同方差不确定性联合损失函数

本文将 CE 损失函数与所提出的 AFWDice 损失函数联合起来,并使用修正后的同方差不确定性策略来学习各自类别的权重。具体步骤是将 w_c 替换为 $\frac{1}{\sigma_c^2}$, 并追加修正项 $\ln(\prod_{c=1}^C \sigma_c^2 + 1)$, 其定义如式(6)所示:

$$L_{AFWDice} = \sum_{c \neq r}^C \frac{1}{\sigma_c^2} L_{Dice_c} + \sum_{c=r}^C \frac{1}{\sigma_c^2} (L_{Dice_c})^\gamma + \ln(\prod_{c=1}^C \sigma_c^2 + 1) \quad (6)$$

采用这种自学习策略,一方面避免了在高维空间中对最优损失超参数的搜索,摊销了时间和算力开销;另一方面还可以达到抑制容易类别权重的作用,变相增大困难类别权重。

最后,联合交叉熵损失和所提出不对称聚焦加权 Dice 损失函数,在模型训练时对分割网络进行约束,如公式(7)所示:

$$L = L_{CE} + L_{AFWDice} \quad (7)$$

3 实验与结果分析

3.1 数据预处理与实验设置

本文使用 3 个腹部图像数据集来评估所提出的方法。第一个是 CHAOS (Combined CT-MR Healthy

Abdominal Organ Segmentation)数据集,包含40张病人扫描图像及其对应的4个器官(liver:肝脏, kidney:肾脏, pancreas:胰腺, spleen:脾脏)的注释,其横截面像素间距在 $0.56 \times 0.56 \text{ mm}^2$ 至 $0.79 \times 0.79 \text{ mm}^2$ 之间^[39];第二个是从TransUNet获取的Synapse数据集,包含30组病人扫描图像以及9种器官的注释(aorta:主动脉, gallbladder:胆囊, spleen:脾脏, left kidney:左肾, right kidney:右肾, liver:肝脏, pancreas:胰腺, spleen:脾脏, stomach:胃),其横截面像素间距在 $0.54 \times 0.54 \text{ mm}^2$ 至 $0.94 \times 0.94 \text{ mm}^2$ 之间^[15];最后一个AbdomenCT-1K数据集,拥有超过1000张病人扫描图像,这些图像的标签是由多种深度学习方法预测生成,并被医学专家检查和校正,其横截面像素间距在 $0.44 \times 0.44 \text{ mm}^2$ 至 $1.04 \times 1.04 \text{ mm}^2$ 之间^[40]。通过在上述3个数据集上进行评估,验证本文所提方法的分割性能和泛化能力。

本文在CHAOS数据集上搜索在损失函数中的最优强化因子 γ 和最佳稀有类别 r ,而在Synapse和AbdomenCT-1K数据集上训练并测试性能。本文按照3:1的比例将CHAOS数据集划分训练和验证集,用于模型的调参;对于Synapse数据集,将18组病人案例用于训练,其余12组用于测试;从AbdomenCT-1K数据集随机选择100组,并按照6:1:3的比例用于训练、验证和测试。本文将所有的3D原始图像亨氏单元(Hounsfield Unit, HU)值限制在 $[50, 350]$ 范围内,以关注腹部器官,然后将图像归一化到 $[0, 1]$ 内,保证像素各向同性($1.0 \times 1.0 \text{ mm}^2$)并沿着水平面分割成若干数量的分辨率为 224×224 的2D切片。为了增加数据集的多样性,本文使用一些简单的数据扩增算法对这些图像进行扩充,包括镜像(沿横轴方向或者沿纵轴方向)、随机旋转($-30^\circ \sim +30^\circ$)。

本文在训练过程中使用一台12 G的NVIDIA GeForce RTX 2080Ti的GPU,每轮的批量设置大小为24,总共训练150轮;为了快速收敛,本文动量设置为0.9的SGD优化器;初始学习率 lr 设置为0.01,并且每轮学习率按照指数衰减关系 $lr \times (1 - \frac{epoch}{total_epoch})^{0.9}$ 进行衰减,实验环境为Python 3.6、PyTorch 1.7框架和CUDA 10.0。

对于同方差不确定性策略,在实际网络训练过程中,也将可训练参数和网络模型参数一并放入SGD优化器中,并通过用 $\sigma_c^2 = \exp(s)$ 进行简单地变形,避免分式中分母为0的情况。

本文使用两个常用分割指标Dice相似系数(Dice Similarity Coefficient, DSC)和Hausdorff距离(Hausdorff Distance, HD)来评估分割后的性能。

3.2 超参数 γ 调控

本文引入了一个超参数作为损失函数的强化因子。为了选择合适的值,本文在CHAOS数据集上进行搜索,如图3所示。以采样间隔为0.25去验证在CHAOS数据集上的DSC性能,当取值为0时,总损失近似退化为仅使用CE损失,在此条件下,模型表现最差;当取值为1.0时,本文所提改进Dice函数退化为对称的结构,次于不对称结构的性能;当取值0.25时,性能达到最佳。特别地,对比 γ 取值为1.0、0.75和0.50、0.25的性能,前两者要弱于后两者,初步验证同方差不确定性策略能够在一定程度补偿AFWDice所带来的缺陷。本文假设在所提方法中设置较低的 γ 值,有助于改善数据的难易失衡问题。因此在后续实验中,本文设置0.25作为超参数的默认值。

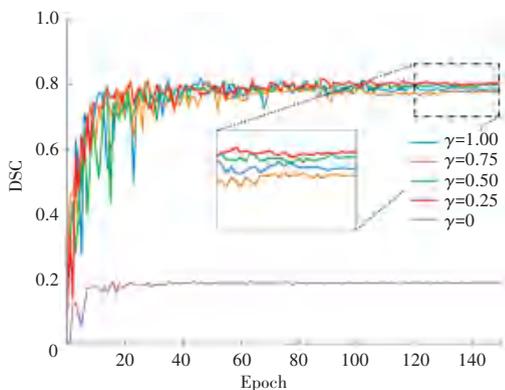


图3 使用不同 γ 在CHAOS数据集的验证结果

Fig. 3 Validation results on the CHAOS dataset by using different γ

3.3 超参数 r 调控

本文引入另一个超参数稀有类别 r 。对于CHAOS数据集,稀有类别 r 可从{liver, kidney, pancreas, spleen}中选择。为了选择合适数量的稀有类别 r ,本文在数据集上依次强化聚焦4种器官、3种器官、2种器官、一种器官和无任何器官,其在验证集性能如图4所示,当对所有4个器官均强化聚焦时,性能达到最优,并且二者基本满足线性关系,即聚焦器官数量越多,模型从中获得的增益越大。特别地,当不选择任何器官时,退化为对称的Dice结构,这时取得最差性能。因此,本文假设对所有待分割器官进行聚焦,能够强化Dice损失的非线性表达能力,有助于防止模型过早收敛到次优解,并达到稳定提升模型性能的目的。

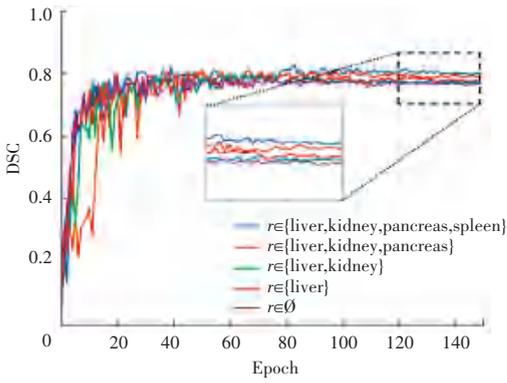


图 4 使用不同 r 在 CHAOS 数据集的验证结果

Fig. 4 Validation results on the CHAOS dataset by using different r

表 1 在 Synapse 数据集上的分割结果

Table 1 Segmentation results on the synapse dataset

损失函数	DSC/%	HD/mm	aorta	kidney/l	gallbladder	kidney/r	liver	pancreas	spleen	stomach
$L_{Focal} \& L_{Dice}$	77.92	28.97	87.46	83.01	63.01	79.58	93.45	56.42	86.15	74.92
$L_{CE} \& L_{GD}$	75.60	36.66	86.94	79.17	58.07	76.37	91.63	57.81	84.19	70.61
$L_{CE} \& L_{Dice}$	77.46	33.31	88.54	81.41	67.93	72.28	93.26	58.71	82.74	74.85
Ours	79.47	28.41	89.00	84.47	64.73	81.24	94.41	58.07	86.72	77.13

不同器官的分割可视化结果如图 5 所示。第一列是真值标签 (Ground Truth, GT); 第二列是联合 Focal 损失和 Dice 损失的分割图; 第三列是联合 CE 损失和 GD 损失的分割图; 第四列是联合 CE 损失和 Dice 损失的分割图; 第五列是本文方法的分割图。从图 5 可以看出, 联合 CE 损失和 GD 损失最容易产生错分割, 如第三行第三列对脾脏欠分割; 而联合 CE 损失和 Dice 损失也存在错误分割现象, 如第一行第四列错误地分割出胆囊; 联合 Focal 损失和 Dice 损失相对能够较好把握大目标器官, 但对小目标的器官, 如第三行第二列中的左肾, 其分割结果不够准确; 本文方法能够很好地缓解欠分割和错分割问题, 提升了分割结果的准确性。此外, 本文还观察到模型在分割胃时性能不稳定, 这是因为胃内部的液体、空气与外界环境具有相似的灰度值, 导致胃表面与其内部空间之间的差异较大, 影响网络分割效果。

3.4.2 在 AbdomenCT-1K 数据集的实验结果

本文在 AbdomenCT-1K 数据集上进行实验, 计

表 2 在 AbdomenCT-1K 数据集的实验结果

Table 2 Segmentation results on the AbdomenCT-1K dataset

损失函数	DSC/%	HD/mm	liver	kidney	pancreas	spleen
$L_{Focal} \& L_{Dice}$	84.84	28.63	94.27	88.28	65.49	91.30
$L_{CE} \& L_{GD}$	84.86	35.19	94.02	89.51	65.96	89.92
$L_{CE} \& L_{Dice}$	83.18	32.86	94.10	88.81	61.46	88.34
Ours	85.93	24.88	95.14	91.34	65.34	91.90

3.4 实验结果

3.4.1 在 Synapse 数据集上的实验结果

本文在 Synapse 数据集上对比了由 Focal 损失和 Dice 损失、CE 损失和 GD 损失、CE 损失和 Dice 损失以及本文所提方法训练的 UNet 模型的性能, 结果见表 1。可见本文所提出方法在多个器官的分割性能上均有所提升, 相较于联合 CE 损失和 Dice 损失, 本文方法在平均 DSC 值提升了 2.01%, HD 降低了 4.90 mm; 相较于次优的 Focal 损失和 Dice 损失, DSC 值提升了 1.55%, HD 降低了 0.56 mm, 说明对 Dice 损失实施聚焦, 要比 CE 损失实施聚焦 (即 Focal 损失) 更有优势。

算了平均 DSC, HD 以及 4 种器官的 DSC, 其结果见表 2。相比于联合 CE 损失和 Dice 损失, 本文方法的平均 DSC 较于联合 CE 损失和 Dice 损失, 提升了 2.75%, HD 降低了 7.98 mm; 相较于次优的 Focal 损失和 Dice 损失, 平均 DSC 提升了 1.09%, HD 也减少了 3.75 mm。



图 5 在 Synapse 数据集的可视化结果

Fig. 5 Visualization results on the Synapse dataset

可视化了不同器官的分割结果如图6所示,呈现两个比较明显的现象。与图5类似,所有方法对大目标器官如肝脏的分割表现良好;对于小目标的器官,其余方法容易产生错误分割,如第一行第三列和第四列所示,联合CE损失、Dice损失以及联合CE损失、GD损失均将部分肝脏错误识别为脾脏。

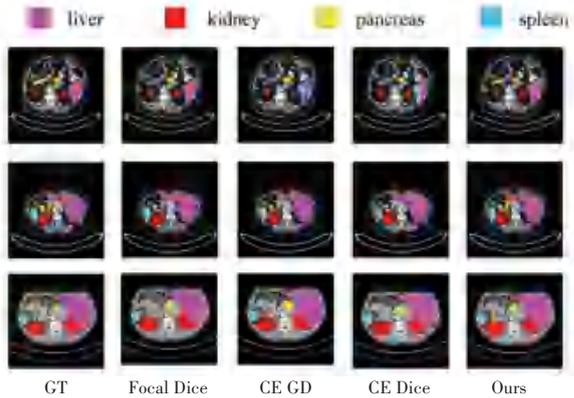


图6 在 AbdomenCT-1K 数据集的可视化结果

Fig. 6 Visualization results on the AbdomenCT-1K dataset

3.5 消融实验

为了验证所提方法分割性能的有效性和泛化能力,本文进行了消融实验,并对实验结果进行分析。消融实验包括:同方差不确定性策略对性能的影响、不对称的Dice损失函数结构对性能的影响。

3.5.1 同方差不确定性策略对性能的影响

本文在两个数据集进行消融实验,其结果见表3。将权重自学习策略修改为等值权重策略,即Dice损失项中各权重均设置相同,修改后DSC性能下降0.7%左右,表明同方差不确定性有能力自动调控不同分割类中权重,克服难易类别失衡问题;此外,在实验中使用同方差不确定性策略,引入的额外参数量微乎其微,还能使分割更加精准。

表3 同方差不确定性策略对性能的影响

Table 3 Influence of homoscedastic uncertainty strategy on performance

同方差不确定性策略	Synapse		AbdomenCT-1K	
	DSC/%	HD/mm	DSC(%)	HD(mm)
不使用	78.72	32.91	85.28	24.84
使用	79.47	28.41	85.93	24.88

3.5.2 不对称Dice损失结构对性能的影响

本文采用对称(Symmetric)的Dice损失函数结构测试在两个数据集上的性能,其性能见表4。与不对称的Dice损失函数相比,二者均在DSC性能上下降了1.4%左右,说明不对称的结构能够克服对感兴趣类别的有害抑制,同时保留对背景元素的抑制。

表4 对称结构对性能的影响

Table 4 Influence of symmetric structure on performance

结构	Synapse		AbdomenCT-1K	
	DSC/%	HD/mm	DSC/%	HD/mm
对称	78.07	33.15	85.48	27.74
不对称	79.47	28.41	85.93	24.88

3.6 补充实验

3.6.1 扩展至其他2D分割模型

为了展示所提方法的优点,本文在UNet++^[7]、TransUNet^[15]以及VATransUNet^[26]3个开源模型上进行实验,实验的默认损失函数及权重遵循原论文设置(即CE损失和Dice损失),其实验结果见表5。这些模型相较于使用默认的联合CE和Dice损失,使用本文所提损失函数的UNet++平均DSC性能在两个数据集上分别提升1.30%、2.05%;使用本文损失函数的TransUNet分别提升2.45%、1.38%;使用本文损失函数VATransUNet分别提升3.65%、0.88%;并且在HD指标上,各个模型均有下降,这表明本文工作具有一定泛化能力,能够推广应用到其他2D分割模型中。

表5 在不同模型上的性能

Table 5 Performance on different models

方法	Synapse		AbdomenCT-1K	
	DSC/%	HD/mm	DSC(%)	HD(mm)
UNet++	77.84	30.76	84.05	38.83
TransUNet	77.79	27.78	84.12	24.52
VATransUNet	79.11	31.54	87.51	21.48
UNet++ & Ours	79.14	26.59	86.10	29.16
TransUNet & Ours	80.24	21.81	85.50	17.16
VATransUNet & Ours	82.76	18.66	88.39	14.45

3.6.2 扩展至3D图像数据

为了展示对3D图像数据的分割性能,本文使用开源的模型UNETR^[17]和UNETR++^[18]作为基线模型,从AbdomenCT-1K数据集中随机选择200组图像按照7.5:1:1.5的比例,进行训练集、验证集和测试集划分。本文将这些图像分辨率重采样为 $256 \times 256 \times 128$,并设置图像体素大小为 $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ 。

本实验在另一台配备了80G显存NVIDIA A100显卡的服务器上进行3D图像分割实验,所有实验均训练1500轮,批量大小为6。实验环境为Python 3.6,PyTorch 1.7框架和CUDA 11.7。每个实验训练平均花费12h、75G显存资源。

在AbdomenCT-1K的3D图像数据集上的性能

见表6,UNETR在肝脏、肾脏以及脾脏性能均表现良好,但在胰腺上性能欠佳。使用本文提出损失函数的UNETR比联合CE和Dice损失函数的UNETR在平均DSC值提升1.89%;UNETR++模型则表现更优,在胰腺器官精度达到69.00%以上;结合本文所提损失函数,在平均DSC上比联合CE和Dice损失函数的UNETR++提升0.31%,表明本文提出的损失函数对3D分割模型亦有效果,并能促进网络对医学图像的特征提取能力。

表6 在AbdomenCT-1K 3D数据集上性能

Table 6 Segmentation results on the AbdomenCT-1K 3D dataset

方法	DSC/%	liver	kidney	pancrea	spleen
UNETR	76.41	92.00	83.73	42.88	87.03
UNETR++	87.70	94.60	93.82	69.00	93.36
UNETR&Ours	78.30	93.70	88.21	46.07	85.21
UNETR++&Ours	88.01	96.08	94.87	69.17	91.90

本文在该数据集进行3D可视化如图7所示。图7有两个较为明显的现象:1)采用默认联合CE和Dice损失的UNETR有更为嘈杂的分割结果;2)采用默认联合CE和Dice损失的UNETR++在器官边缘处不平滑。作为对比,采用本文损失函数的UNETR模型和UNETR++模型由于使用指数形式的Dice损失,强化了损失函数的非线性表达能力并提高分割的准确性和精度。

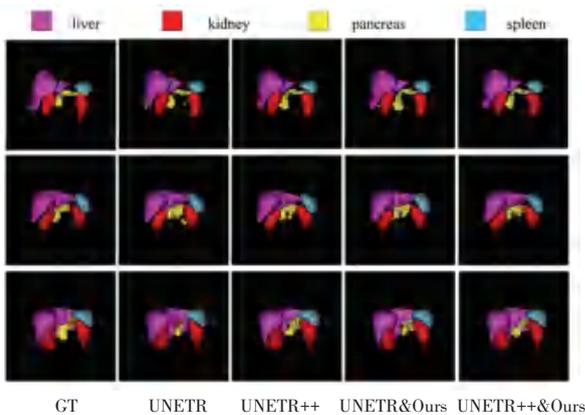


图7 在AbdomenCT-1K数据集的3D可视化结果

Fig. 7 3D visualization results on the AbdomenCT-1K dataset

4 结束语

本文通过将强化因子和不对称结构引入Dice损失中,并以自学习方式更新调整各自类别的权重,提出一种新的基于不对称聚焦加权的Dice损失。进一步的图像分割实验中表明,本文提出损失函数能够在其他的模型中稳定提升性能;实验性能和可视化结果表明,本文提出损失函数的性能优于经典

的CE损失和Dice损失的组合函数以及其他损失函数,表明损失函数中引入类似Focal损失的强化因子和同方差不确定性策略,可以在不扩充数据量的前提下,稳定提升性能。在未来的工作中,本文将进一步扩展该损失函数到其他器官数据集中。项目代码将在<https://github.com/MischiefGhostOgre/HyperSegUNet>发布。

参考文献

- [1] WANG R, LEI T, CUI R, et al. Medical image segmentation using deep learning: A survey[J]. IET Image Processing, 2022, 16(5): 1243-1267.
- [2] WANG W, CHEN C, DING M, et al. TransBTS: Multimodal brain tumor segmentation using transformer[C]//Proceedings of Medical Image Computing and Computer Assisted Intervention. IEEE, 2021: 109-119.
- [3] RONNBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Proceedings of Medical Image Computing and Computer Assisted Intervention. IEEE, 2015: 234-241.
- [4] YIN X, SUN L, FU Y, et al. U-Net-based medical image segmentation[J]. Journal of Healthcare Engineering, 2022: 4189781.
- [5] MILLETARI F, NAVAB N, AHMADI S-A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation[C]//Proceedings of 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016: 565-571.
- [6] ZHOU Z, SIDDIQUEE M M R, TAJBAKSH N, et al. UNet++: A nested u-net architecture for medical image segmentation[C]//Proceedings of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3-11.
- [7] ZHOU Z, SIDDIQUEE M M R, TAJBAKSH N, et al. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 39: 1856-1867.
- [8] HUANG H, LIN L, TONG R, et al. UNet 3+: A Full-scale connected unet for medical image segmentation[C]//Proceedings of ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 1055-1059.
- [9] ÇIÇEK Ö, ABDULKADIR A, LIENKAMP S S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation[C]//Proceedings of Medical Image Computing and Computer-Assisted Intervention. IEEE, 2016: 424-432.
- [10] KUNHIMON S, SHAKER A, NASEER M, et al. Learnable weight initialization for volumetric medical image segmentation[J]. Artificial Intelligence in Medicine, 2024, 151: 102863.
- [11] ZHAO R, QIAN B, ZHANG X, et al. Rethinking dice loss for medical image segmentation[C]//Proceedings of 2020 IEEE International Conference on Data Mining (ICDM). IEEE, 2020: 851-860.
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. IEEE, 2017: 2980-2988.

- [13] YEUNG M, SALA E, SCHÖNLIEB C B, et al. Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation [J]. *Computerized Medical Imaging and Graphics*, 2022, 95: 102026.
- [14] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words; Transformers for image recognition at scale [J]. *arXiv preprint arXiv:2010.11929*, 2020.
- [15] CHEN J, LU Y, YU Q, et al. Transunet: Transformers make strong encoders for medical image segmentation [J]. *arXiv preprint arXiv:2102.04306*, 2021.
- [16] CAO H, WANG Y, CHEN J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation [C] // *Proceedings of European Conference on Computer Vision*. Cham: Springer, 2022: 205–218.
- [17] HATAMIZADEH A, YANG D, ROTH H R, et al. UNETR: Transformers for 3D medical image segmentation [C] // *Proceedings of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2021: 1748–1758.
- [18] SHAKER A M, MAAZ M, RASHEED H, et al. UNETR++: Delving into efficient and accurate 3D medical image segmentation [J]. *arXiv preprint arXiv: 2012.04497*, 2022.
- [19] SHAMSHAD F, KHAN S, ZAMIR S W, et al. Transformers in medical imaging: A survey [J]. *arXiv preprint arXiv: 2201.09873*, 2022.
- [20] XIAO T, SINGH M, MINTUN E, et al. Early convolutions help transformers see better [J]. *arXiv preprint arXiv: 2106.14881*, 2021.
- [21] CHU X, ZHANG B, TIAN Z, et al. Do we really need explicit position encodings for vision transformers [J]. *arXiv preprint arXiv:2102.10882*, 2021.
- [22] WU H, XIAO B, CODELLA N C F, et al. CvT: Introducing convolutions to vision transformers [C] // *Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, 2021: 22–31.
- [23] GUO M H, LU C Z, LIU Z N, et al. Visual attention network [J]. *Computational Visual Media*, 2023, 9(4): 733–752.
- [24] GUO M H, LU C Z, HOU Q, et al. Segnext: Rethinking convolutional attention design for semantic segmentation [J]. *arXiv preprint arXiv: 2209.08575*, 2022.
- [25] AZAD R, NIGGEMEIER L, HÜTTEMANN M, et al. Beyond self-attention: Deformable large kernel attention for medical image segmentation [C] // *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. IEEE, 2024: 1287–1297.
- [26] JIANG T, XU T, LI X. VA-Trans UNet: A U-shaped medical image segmentation network with visual attention [C] // *Proceedings of the 2022 11th International Conference on Computing and Pattern Recognition*. IEEE, 2022: 128–135.
- [27] LI Z, KAMNITSAS K, GLOCKER B. Overfitting of neural nets under class imbalance: Analysis and improvements for segmentation [C] // *Proceedings of the International Conference on Medical Image Computing and Computer - Assisted Intervention*. IEEE, 2019: 402–410.
- [28] SUDRE C H, LI W, VERCAUTEREN T K M, et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations [C] // *Proceedings of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, and 7th International Workshop, Held in Conjunction with MICCAI*. IEEE, 2017: 240–248.
- [29] WANG P, CHUNG A C S. Focal dice loss and image dilation for brain tumor segmentation [C] // *Proceedings of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. IEEE, 2018: 119–127.
- [30] WONG K C L, MORADI M, TANG H, et al. 3D segmentation with exponential logarithmic loss for highly unbalanced object sizes [C] // *Proceedings of Medical Image Computing and Computer Assisted Intervention-MICCAI 2018: 21st International Conference*. Granada, Spain: IEEE, 2018: 612–619.
- [31] ABRAHAM N, KHAN N M. A novel focal tversky loss function with improved attention U-Net for lesion segmentation [C] // *Proceedings of 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019: 683–687.
- [32] KENDALL A, GAL Y. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? [C] // *Proceedings of the 31st International Conference on Neural Information*. IEEE, 2017: 5580–5590.
- [33] KENDALL A, GAL Y, CIPOLLA R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics [C] // *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2017: 7482–7491.
- [34] LIEBEL L, KÖRNER M. Auxiliary tasks in multi-task learning [J]. *arXiv preprint arXiv:1805.06334*, 2018.
- [35] ELMAHDY M S, BELJAARDS L, YOUSEFI S, et al. Joint registration and segmentation via multi-task learning for adaptive radiotherapy of prostate cancer [J]. *IEEE Access*, 2021, 9: 95551–95568.
- [36] WANG L, JU L, ZHANG D, et al. Medical matting: A new perspective on medical segmentation with uncertainty [C] // *Proceedings of the International Conference on Medical Image Computing and Computer - Assisted Intervention*. IEEE, 2021: 573–583.
- [37] DANGI S, LINTE C A, YANIV Z. A distance map regularized CNN for cardiac cine MR image segmentation [J]. *Medical Physics*, 2019, 46(12): 5637–5651.
- [38] VANDENHENDE S, GEORGIOULIS S, VAN GANSBEKE W, et al. Multi-task learning for dense prediction tasks: A survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(7): 3614–3633.
- [39] KAVUR A E, GEZER N S, BARIŞ M, et al. Chaos challenge-combined (ct-mr) healthy abdominal organ segmentation [J]. *Medical Image Analysis*, 2021, 69: 101950.
- [40] MA J, ZHANG Y, GU S, et al. Abdomenct-1k: Is abdominal organ segmentation a solved problem? [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(10): 6695–6714.