

基于双分支编码的闭环分割网络^①

任玉涛, 程远志

(青岛科技大学 信息科学技术学院, 青岛 266061)

通信作者: 程远志, E-mail: yzcheng2007@163.com



摘要: Transformer 模型中, 卷积视觉转换器 (CvT) 具备同时提取图像的局部和全局特征的能力而受到关注. 对于腹部器官分割问题, 为了解决 CNN 模型分割目标边界轮廓模糊不清的问题, 提出了一种新颖的基于 CvT 和 CNN 的双分支闭环分割模型 DBLNet. 模型利用形状先验和预测结果的分割轮廓显式监督并指导网络学习过程. 模型包括: 轮廓提取编码模块 (CEE)、边界形状分割网络 (BSSN) 和闭环结构. CEE 模块首次利用改造的 3D CvT 和 3D 门控卷积层 (GCL) 捕获多层次轮廓特征, 并辅助 BSSN 训练. BSSN 模块设计形状特征融合模块 (SFF) 同时捕获目标区域和轮廓特征, 推动 CEE 训练拟合. 闭环结构使得双分支的分割结果互相反馈并辅助对方的训练. DBLNet 在 BTCV 排行榜上测试, 平均 *Dice* 得分为 0.878, 排名第 13 位; 在临床医院数据进行应用测试, 表现出强大的性能.

关键词: 腹部器官; 边缘轮廓; 双分支编码器; 闭环结构; 卷积视觉转换器; 医学影像处理; 特征融合; 图像分割

引用格式: 任玉涛, 程远志. 基于双分支编码的闭环分割网络. 计算机系统应用, 2024, 33(1): 110-118. <http://www.c-s-a.org.cn/1003-3254/9371.html>

Closed-loop Segmentation Network Based on Dual-branch Encoding

REN Yu-Tao, CHENG Yuan-Zhi

(School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: In the Transformer model, the convolutional vision Transformer (CvT) has caught attention for its ability to extract both local and global features from images simultaneously. However, for abdominal organ segmentation tasks, the blurry object boundaries in CNN models should be addressed. Thus, this study proposes a novel dual-branch closed-loop segmentation model DBLNet based on CvT and CNN. The model employs explicit supervision of segmented contours using shape priors and predicted results to guide the network learning. The DBLNet model includes contour extraction encoding module (CEE), boundary shape segmentation network (BSSN), and closed-loop structure. The CEE module first utilizes modified 3D CvT and 3D gated convolutional layers (GCL) to capture multi-level contour features and assist in BSSN training. The BSSN module contains a shape feature fusion (SFF) module that captures both the object region and contour features to promote CEE training convergence. The closed-loop structure allows mutual feedback of segmentation results between the dual branches, assisting each other's training. Experimental evaluations on the BTCV benchmark show that DBLNet achieves an average *Dice* score of 0.878, ranking 13th. Application tests on clinical hospital data demonstrate the strong performance of the proposed model.

Key words: abdominal organ; edge contours; dual-branch encoder; closed-loop structure; convolutional vision Transformer (CvT); medical image processing; feature fusion; image segmentation

① 基金项目: 国家自然科学基金 (61806107, 61702135)

收稿时间: 2023-07-08; 修改时间: 2023-08-08; 采用时间: 2023-08-31; csa 在线出版时间: 2023-11-28

CNKI 网络首发时间: 2023-11-30

1 引言

1.1 背景

近些年,腹部各类器官病变(肝肿瘤、结肠肿瘤、肾结石、胰腺炎等)一直在发生,每年的病死率也是在增加.病死率居高不下的原因就是患者病变早期医生难以通过外在表现进行检查判断,病变明显时已经达到中晚期,错过最佳治疗时间.腹部器官众多且脆弱,形态结构较为复杂,各类器官发病率较高.如果在发病早期及时发现就能得到有效的治疗,就能极大减少病死率;针对某些轻微病症甚至可以达到根治的效果.腹部器官数量多,疾病往往发生在体内,这就需要进行手术穿刺,对医生的操作技术要求高且需要患者花费更多时间和精力等待.如果检查每个器官是否发生病变,这十分的耗时耗力且有害患者健康.目前,常用的诊治辅助手段就是计算机断层扫描.

在治疗疾病之前,需要专业医师参考医学影像结果作出判断并标注病变部位,然后制定手术方案.腹部器官结构复杂且重叠,标注和识别过程是耗时耗力的.对于解决该问题,可以使用医学图像分割技术.医学图像分割^[1]在临床诊断和疾病治疗中起着至关重要的作用,是计算机视觉研究的重要课题.全自动分割系统在医学图像分析领域得到了广泛的研究.然而,由于复杂的组织关系和结构,目标组织与相邻器官之间的对比度低、边界模糊^[2],现有技术分割精度上仍然存在一些限制.例如,许多方法未能改善器官边界的轮廓分割效果.原因是组织间对比度差,器官间灰度强度不均匀,以及数据集中目标器官和背景不平衡.

目前,有些方法试图解决分割任务中的边界模糊问题,提高分割腹部器官病灶的准确位置.很多专家研发新的模型^[3],设计新的预处理和后处理方法,但是利用好形状先验信息^[4]和网络分割结果也很重要,将网络改进并再次利用分割结果,通过闭环结构促进训练,可以更好地提高分割精度.

1.2 相关工作

已有方法包括基于形状边界和基于 Transformer 的分割方法.

医学图像分割任务^[5],其中使用形状轮廓完成分割的方法值得研究. Oktay 等人基于 CNN 提出 Attention U-Net^[6],设计包含空间和通道注意力的注意门(AG),学习跳层连接中不同形状的目标区域. Meng 等人基于 GNN 设计用于提取并融合区域和边界特征的注意

力模块(AEM)^[7],在息肉分割任务上效果较好. Lin 等人基于 CNN、ViT 和边界检测算子构建 CTO^[8],它使用边界掩码作为显式监督,指导解码学习过程. Liu 等人提出了一种感知形状和边界的多分支模型^[9],利用一个 V 型模型来进行像素分割图(PSM)预测和有符号距离图(SDM)回归.通过 SDM 和特征融合模块(FFM),进一步提取边界增强特征,并在 BraTS2019 取得较好的结果.

与以往形状特征提取和融合的方法不同: DBLNet 的 BSSN 模块有效利用轮廓先验;设计 SFF 模块处理跳层连接,通过注意力机制逐渐锁定待分割目标区域,使用 GCL 提取边界特征,再通过残差结构融合区域和边界形状等特征.

Vision Transformer: 将 Transformer^[10]用于医学图像分割任务^[11],弥补 CNN 学习全局上下文信息和长程空间依赖性等方面的缺失信息. ViT 模块可用于编码层、瓶颈层或者解码层^[12]等,主要为混合架构或者单一 Transformer 架构. Chen 等人首次将 Transformer 和 CNN 串联构建出强大编码器的 TransUNet^[13],在 BTCV 数据集效果不错. Hatamizadeh 等人为改善卷积的局限性,将 3D 图像分割任务转变为序列的预测问题并设计 UNETR^[14]. 该网络的编码器由多层 Transformer 构成,提取不同尺度的特征图并通过跳层连接输出到解码器,该网络在 BTCV 数据集效果显著. Hatamizadeh 等人^[15]设计 SwinUNETR^[16],它使用了 Swin Transformer 的跨局部和全局特征交互能力以及 UNet 的上采样和下采样结构,语义和空间感知能力更强. 该网络在 BTCV 数据集测试排行榜较优秀.

DBLNet 与以前方法的区别: DBLNet 使用并联形式的双分支闭环结构,其中 CEE 分支使用结合卷积优势的 CvT 和 GCL 构成层级编码器,主要学习轮廓细节和形状等特征;并针对双分支结构的分割结果首次设计闭环结构,实现分支间信息流动.

1.3 本文贡献

我们提出了一种基于 CvT 和 CNN 构建的双分支结构网络,首次将 CvT^[17-19]和 GCL 由 2D 形式改造 3D 形式用于腹部器官分割. 创新点如下: (1) 提出一个形状特征融合模块 SFF 并构建 BSSN 分支, SFF 采用 3D GCL 提取边界特征,并使用注意力提取区域特征,再通过残差结构将两种特征融合. (2) 提出利用 3D CvT 构建多层级的 CEE 分支网络,克服了 ViT 的缺点,

可以充分提取局部和全局特征中的轮廓形状特征。
 (3) 提出基于双分支的 DBLNet 模型, 设计闭环结构充分利用自身分割结果和形状先验信息, 相互促进对方网络的优化收敛, 对最终双分支分割结果采用加权融合方式获得最精确的分割结果。

2 双分支编码结构的闭环分割算法

2.1 DBLNet 算法

本算法网络结构如图 1 所示, 首先需要利用分割掩码提取到形状先验信息, 并将其附加到 ROI 图像中。两者分别输入 CEE 和 BSSN 分支模块, 经过训练并输出分割结果, 分别使用 Canny 和 DPA 算法提取形状轮廓和关键轮廓点。将两者分别输入对方网络, 以此实现间接训练自身网络并帮助直接训练对方网络。双分支模型拟合后或达到设定训练次数, 即可停止训练, 将

两者的分割结果进行加权融合获得最优分割结果。

设计如算法 1 的算法步骤。

算法 1. 双分支编码的闭环分割算法

- 1) 输入原图和掩码, 获取 ROI 图像。
- 2) 根据掩码获取目标器官坐标, 根据位置坐标裁剪提取 ROI 图像 I 。
- 3) 根据掩码图像 I_1 提取形状先验轮廓 I_0 , 将 I_0 和 ROI 融合, 输入分割网络 BSSN。
- 4) while $t < maxiteration$ do:
- 5) DPA: 使用 DPA 关键点检测算法从分割掩码 z 提取关键轮廓点图 Z'' 。
- 6) 提供关键轮廓点图 Z'' 和融合后的 ROI 图像输入 CEE 分支训练并完善分割掩码 Z_1 。
- 7) Canny: CEE 输出分割掩码 Z_1 , 通过 Canny 算法提取轮廓图像 Z' 。
- 8) 提供融合后的 ROI 和分割轮廓图像 Z' 输入 BSSN, 训练并完善初始分割掩码 Z 。
- 9) 双分支模型通过闭环结构促进对方训练, 通过复合损失函数 ($Dice+CE$ Loss) 拟合模型, 直到收敛。
- 10) 通过计算合适的比例系数结合 BSSN 和 CEE 的分割结果, 生成最终预测分割结果。

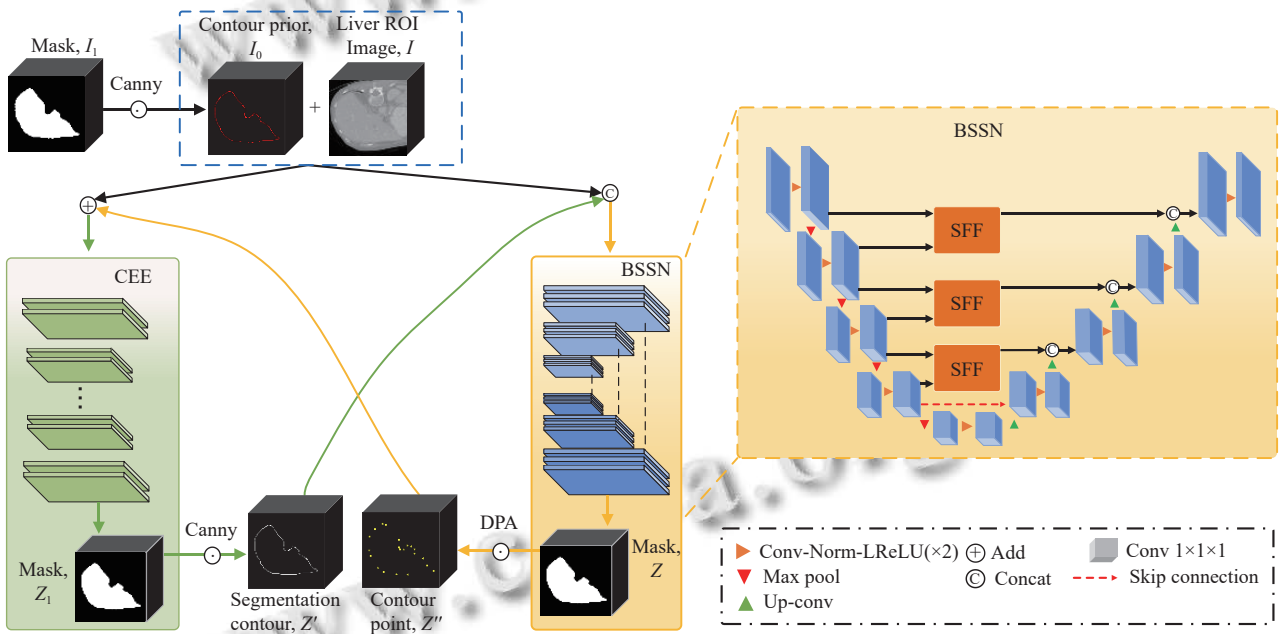


图 1 DBLNet 网络结构图

2.2 提取形状先验和 ROI 图像

根据掩码坐标位置裁剪输入图像, 裁剪掉部分背景区域, 保留掩码中目标区域。对应到输入图像坐标, 裁剪后就能获取到 $224 \times 224 \times 32$ 的 ROI 图像 I 。

获取轮廓先验 I_0 : 使用边缘检测算法提取轮廓。我们扩展三维形式的 Canny 算法^[20]提取轮廓先验 I_0 , 使接近决策边界的像素取值 1; 而非边界像素取值 0。将输入的三维图像进行高斯平滑, 以减小噪声对边缘检测的影响。使用一阶导数对平滑后的图像计算梯度幅

值和方向, 梯度幅值和方向通过式 (1) 计算:

$$G_x = I_2 \times H_x, G_y = I_2 \times H_y, G_z = I_2 \times H_z$$

$$G = \sqrt{G_x^2 + G_y^2 + G_z^2}, \theta = \arctan\left(\frac{G_z}{\sqrt{G_x^2 + G_y^2}}\right) \quad (1)$$

其中, I_2 为分割掩码 I_1 平滑后的图像, H_x 、 H_y 和 H_z 分别为水平、竖直和深度方向上的 Sobel 算子。然后使用非极大值抑制, 对梯度方向上的局部极大值进行保留, 减小非轮廓点的响应。通过双阈值方法确定边缘轮廓,

进行边缘连接,形成连续的边缘。

2.3 CEE 分支

轮廓提取编码模块 (CEE module): 由 3D CvT、3D GCL 和反卷积构成, 结构如图 2 所示。CvT Encoder 有 4 层, 注重提取图像轮廓特征信息。算法描述: ROI 图像 I 、轮廓先验图 I_0 和 BSSN 的轮廓点图 Z'' 融合后输入 CvT Encoder, 通过反卷积输出分割掩码。再使用 Canny 算法提取分割轮廓 Z' , 然后将其输入 BSSN 分支。CEE 输出特征存在很多非目标轮廓和无关特征, 可以借助环路中 BSSN 的轮廓点图 Z'' 过滤掉部分异常特征, 最后将分割轮廓 Z' 反馈给 BSSN 分支。

2.4 BSSN 分支

边界形状分割网络 (BSSN): 网络各层结构如图 1 右侧所示, 在 UNet 基础上设计 SFF 模块提取特征图中边界形状和区域特征, 使分割掩码更精确。算法描述:

轮廓先验 I_0 、分割轮廓 Z' 和 ROI 图像 I 融合后输入 BSSN, BSSN 学习如何使用 CEE 模块的分割轮廓 Z' , 最终使用轮廓逼近算法中的 Douglas-Peucker Algorithm^[21] 从分割掩码 Z 中提取轮廓点图 Z'' , 然后将其输入 CEE 分支。DPA 算法减少轮廓点数目并保持轮廓形状, 通过环路辅助 CEE 去除一些异常轮廓特征。BSSN 的分割掩码受 CEE 的分割轮廓影响, 使用轮廓信息限定目标区域并改善分割掩码边缘。

形状特征融合模块 (SFF): SFF 如图 3 所示, 主要由 AGs^[7] 和 3D GCL 组成, 用于提取内部区域和边界形状等特征并融合输出。具体实现: 首先将相邻编码层尺度为 (N, C, H, W, D) 的特征图 S_3 和 $(N, 2 \times C, H/2, W/2, D/2)$ 的 S_4 , 通过线性插值恢复到同一尺度并使用 $1 \times 1 \times 1$ 的卷积进行通道降维得到 (N, C, H, W, D) 的 S'_4 , 然后将两者输入注意门 (AGs) 计算。

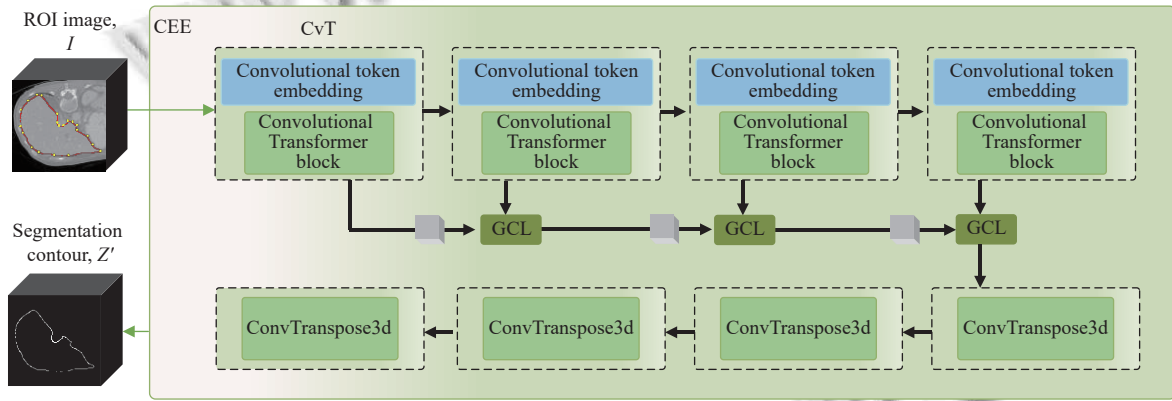


图 2 CEE 分支网络结构

$$\begin{cases} S_3 = \text{Attention}(S_3, S_4) \\ S'_4 = \text{Upsample}(S_4, \text{scale} = 2) \\ S'_3 = \text{CovBlock}(\text{concatenate}(S_3, S'_4)) \end{cases} \quad (2)$$

$$S''_3 = (S''_4 \times w_t)_{(i,j,k)} = ((S''_{4(i,j,k)} \times \alpha_t(i,j,k)) + S''_{4(i,j,k)})^T w_t \quad (4)$$

提取大致区域特征, 再将提取的区域特征图 S'_3 和通道降维后的形状特征图 S'_4 输入 3D GCL 提取轮廓边界信息得到尺度为 (N, C, H, W, D) 的 S''_3 。使用 GCL, 我们首先将 S'_3 和 S'_4 拼接起来, 然后经过归一化的 $1 \times 1 \times 1$ 卷积层 $C_{1 \times 1 \times 1}$, 再通过 Sigmoid 函数 σ , 得到一个注意力图 α_t :

$$\alpha_t = \sigma(C_{1 \times 1 \times 1}(S'_3 \| S'_4)) \quad (3)$$

其中, $\|$ 表示特征图的连接。给定注意力图 α_t , GCL 被应用于形状特征图 S'_4 , 逐元素乘积与注意力图 α_t , 然后进行残差连接和通道权重化, 使用内核 w_t 在每个像素点 (i, j, k) 上计算。GCL 公式如下:

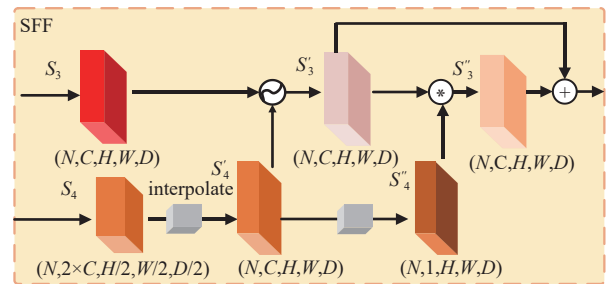


图 3 SFF 模块的结构流程图

最终通过残差结构将区域特征融合边界轮廓信息后输出到对应解码层。

SFF 各阶段可视化热力图, 如图 4 所示: 对于某种医学图像数据 (以 BTCV 中的 Liver 为例), 图 4(a) AGs 热力图: 特征图输入 AGs 中, 重点关注 Liver 的区域。

图 4(b) GCL 热力图: 从 AGs 获取得到区域经过 3D GCL 处理后重点关注边界特征. 图 4(c) SFF 热力图: AGs 和 GCL 提取的特征通过残差结构实现特征融合.

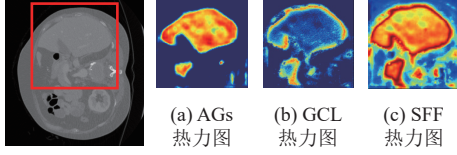


图 4 各阶段热力图

2.5 输出结果

DBLNet 考虑到各自模块具有的分割特点和优势, 最终输出分割掩码采用两种模型输出结果的加权融合. 计算式 (5) 如下:

$$Z_{DBLNet} = \lambda Z + (1-\lambda)Z_1 \quad (5)$$

其中, Z 和 Z_1 表示分支模块的分割双掩码, λ 表示分割掩码的权重, 取值为 0–1 之间的值. 根据加权分割结果 Z_{DBLNet} 和真实标签 GT 进行研究计算, 记录实验超参数 λ 对 $Dice$ 和其他指标的影响. 根据实验结果计算得 λ 的最佳取值为 0.86, 此时的分割结果 Z_{DBLNet} 最好, $Dice$ 精确度较为稳定.

3 实验分析

为了验证 DBLNet 的有效性, 我们使用 BTCV (CT) 和医院临床数据集进行实验.

3.1 实验环境

硬件配置: Intel(R) Core(TM) i9-10900K CPU @ 3.70 GHz; 128 GB 内存; GeForce RTX 3090 24 GB 显卡. 软件环境: Ubuntu 18.04 LTS 操作系统; Python 3.8 编程语言; PyTorch 1.7 深度学习框架. 算法训练 epoch 数量为 2000, batch_size 为 1, 初始学习率为 0.01, 动量为 0.99 的 SGD 优化器. BSSN 的损失函数^[22]采用 $Dice$ loss 和 Cross entropy loss 的复合形式, CEE 使用 Boundary loss 训练.

3.2 数据集

BTCV (CT)^[16]: 对 50 名患者进行腹部 CT 扫描获得, 腹部 13 个器官在范德比尔特大学医学中心放射科医生的监督下进行注释. 将 30 个 CT 图像划分为训练集 24 个和验证集 6 个, 测试集 20 个病例. 每张 CT 面内分辨率为 0.54 mm×0.54 mm 至 0.98 mm×0.98 mm, 而切片厚度范围为 2.5–5.0 mm. 主要分割 Spleen、Right Kidney、Esophagus 和 Liver 等 13 个器官, 器官种类较多且形态各异.

临床医院的腹部数据集: 来源于大型合作医疗机构, 通过专业设备扫描得到并由多位专家手工标注并矫正. 每张 CT 由 120–210 个切片组成, 面内分辨率为 0.63 mm×0.63 mm 至 0.80 mm×0.80 mm, 切片厚度范围为 1.0–1.5 mm, 尺寸大小为 512×512 像素. 其中数据集共 50 个病例, 划分为训练集 30 个和验证集 10 个, 测试集 10 个病例. 本次主要分割食管、肾上腺、静脉、胰腺和肝脏等一些体积小、低对比度等特点的器官, 评估本方法的适应性.

3.3 评价指标

本次实验使用常见的医学图像分割指标来评估分割结果的精度, 主要有 $Dice$ score 和 95% Hausdorff distance (HD). $Dice$ 系数是一种相似性度量指标, 范围在 [0, 1], 分割最好结果为 1, 最差为 0. HD_{95} 也叫豪斯多夫距离, 相比于 $Dice$ 来说, 它对分割出的边界更敏感, HD_{95} 使用真实标签和预测标签表面点集之间的距离的第 95 百分位. 因此, 在计算 HD_{95} 时, 可使得离群值中非常小的子集的影响最小化. $Dice$, HD 以及 HD_{95} 公式如下:

$$\left\{ \begin{array}{l} Dice(G, P) = \frac{2 \sum_{i=1}^I G_i P_i}{\sum_{i=1}^I G_i + \sum_{i=1}^I P_i} \\ HD(G', P') = \max\{\max_{g' \in G'} \min_{p' \in P'} \|g' - p'\|, \\ \max_{p' \in P'} \min_{g' \in G'} \|p' - g'\|\} \\ HD_{95}(G', P') = \max\{HD(G', P'), HD_{95}(G', P'_{0.95})\} \end{array} \right. \quad (6)$$

其中, G 是真实标签, P 是预测标签, G_i 和 P_i 是位置体系 i 的数值. G' 是真实标签的表面点集, P' 是预测标签的表面点集.

ASD 通常用于衡量分割结果与真实分割之间的距离. 它衡量的是两个表面之间的平均距离, ASD 数值越小表示两者越相似. Average surface distance (ASD) 公式如下:

$$ASD = \frac{1}{N_{G'} + N_{P'}} \left(\sum_{i=1}^{N_{G'}} \min_{j=1}^{N_{P'}} \|G'_i - P'_j\| + \sum_{i=1}^{N_{P'}} \min_{j=1}^{N_{G'}} \|P'_i - G'_j\| \right) \quad (7)$$

其中, $N_{G'}$ 和 $N_{P'}$ 分别为真实标签和预测标签的表面点集中的点数, $\sum_{i=1}^{N_{G'}} \min_{j=1}^{N_{P'}} \|G'_i - P'_j\|$ 表示真实标签表面上的所有点与预测标签表面上最近的点之间的欧氏距离之和, $\sum_{i=1}^{N_{P'}} \min_{j=1}^{N_{G'}} \|P'_i - G'_j\|$ 表示预测标签表面上的所有点与真实标签表面上最近的点之间的欧氏距离之和.

3.4 各项数据集对比实验

BTCV 的实验结果: DBLNet 对比其他模型, 可视化结果如图 5 所示, 结果表明模型具有强大的分割能力. 对大器官较好, 对小目标器官的边缘和形状分割比其他模型更加细致精确, 使得各个器官分割准确率有

所提高. 各器官 *Dice* 得分如表 1 所示, DBLNet 的平均 *Dice* 为 0.878, 比 PaNN^[23]、CoTr 和 TransUNet 的方法高出 2.4%、3.4% 和 4.0%. 在小器官(胆囊、胰腺)的分割中, 我们的方法显著优于 PaNN, *Dice* 要高出 4.2% 和 4.6%.

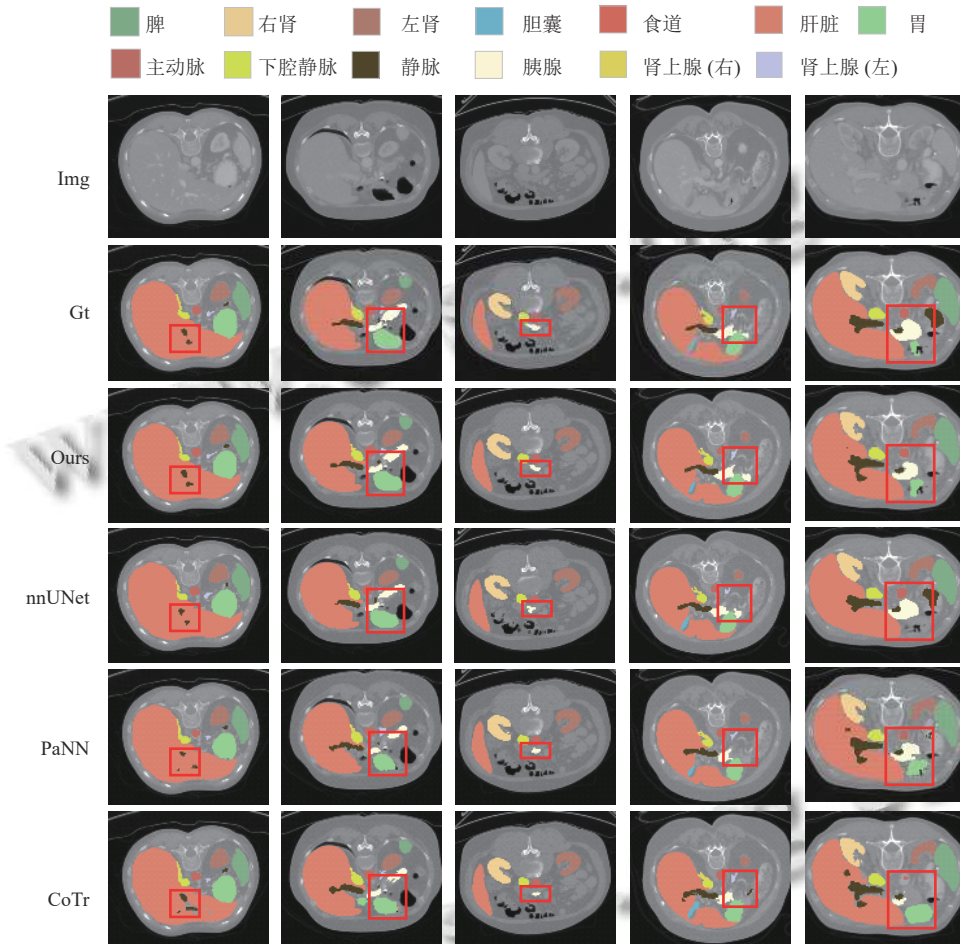


图 5 在 BTCV 数据集上可视化对比分割结果

表 1 BTCV 标准挑战赛的测试集的 *Dice* 得分

方法	脾	右肾	左肾	胆囊	食道	肝脏	胃	主动脉	下腔静脉	静脉	胰腺	肾上腺	平均
SETR MLP	0.930	0.889	0.894	0.650	0.762	0.953	0.819	0.872	0.739	0.720	0.716	0.614	0.796
ASPP	0.935	0.892	0.914	0.689	0.760	0.953	0.812	0.918	0.807	0.695	0.720	0.629	0.811
TransUNet	0.952	0.927	0.929	0.662	0.757	0.969	0.889	0.920	0.833	0.791	0.775	0.637	0.838
CoTr	0.958	0.921	0.936	0.700	0.764	0.963	0.854	0.920	0.838	0.787	0.775	0.694	0.844
RandomPatch	0.963	0.912	0.921	0.749	0.760	0.962	0.870	0.889	0.846	0.786	0.762	0.712	0.844
PaNN ^[23]	0.966	0.927	0.952	0.732	0.791	0.973	0.891	0.914	0.850	0.805	0.802	0.652	0.854
DBLNet	0.965	0.933	0.957	0.774	0.836	0.977	0.920	0.922	0.873	0.831	0.848	0.781	0.878

根据实验结果测试可知, 本文提出的方法总体平均 *Dice* 为 0.878, 排名第 13 位, *ASD* 得分为 1.329, 排名第 11 位, *HD* 得分为 18.276, 排名第 10 位. 总体来看, 超越大多数方法, 这些数据证实了本文提出的

SFF 和双分支闭环结构是有效的. 图 5 为腹部多器官的分割结果, 可以说明本文的方法占有优势. 与表 1 结果一致, 本文分割方法在提取各器官形状边界方面效果更好. 与现存方法相比, 本文方法甚至可以更完整的

提取到小器官和其他器官的形状. 根据测试结果, 仔细观察图 5 可知, 除了分割结果外, 我们还提供了真实标签方便更好的比较. 随机选 5 个样本数据, 前 4 列表明: 对小目标区域的分割完整性和精度较高且形状更规则, 同时改善了 CoTr 具有残缺孔洞的缺点, 并且边缘形状比其他模型更好. 第 5 列中 PaNN 预测的胰腺和胃出现边缘像素残缺, CoTr 和 nnUNet 中甚至丢失预测目标, 而 DBLNet 分割的目标区域相对准确且完整.

临床医院数据的实验结果: *Dice* 和 *HD₉₅* 计算结果

如表 2 所示. 其中肾上腺的 *Dice* 为 0.847, 胰腺的 *Dice* 为 0.886, 肝脏的 *Dice* 为 0.986. 除肝脏外, 其他器官 *Dice* 比 Swin UNETR 均提高了约 12% 以上, 并且 *HD₉₅* 大幅降低, DBLNet 的平均 *HD* 减小到 3.208, 超过其他众多方法. 观察个体病例可视化结果如图 6, 可知 DBLNet 具备分割细小或狭长物体的能力; 其他方法分割的食道、静脉和胰腺存在边缘轮廓多余像素粘连的问题, 肾上腺出现边角像素残缺问题. 而 DBLNet 分割的目标区域和边缘轮廓相对完整, 克服了上述问题.

表 2 腹部器官平均 *Dice* 和平均 *HD₉₅* 得分

方法	食道 (Eso)		静脉 (Vein)		胰腺 (Pancreas)		肾上腺 (AG)		肝脏 (Liver)		平均 (Average)	
	<i>Dice</i> ↑	<i>HD₉₅</i> ↓	<i>Dice</i> ↑	<i>HD₉₅</i> ↓	<i>Dice</i> ↑	<i>HD₉₅</i> ↓	<i>Dice</i> ↑	<i>HD₉₅</i> ↓	<i>Dice</i> ↑	<i>HD₉₅</i> ↓	<i>Dice</i> ↑	<i>HD₉₅</i> ↓
3DUNet	0.672	13.251	0.683	11.725	0.720	9.532	0.462	8.792	0.948	14.852	0.697	11.630
TransUNet	0.641	8.526	0.763	9.617	0.692	11.528	0.564	10.826	0.951	9.634	0.722	10.026
Swin-Unet	0.526	17.820	0.652	13.274	0.715	10.105	0.635	13.395	0.942	15.146	0.694	13.948
UNETR	0.611	8.394	0.604	15.103	0.665	12.496	0.465	9.749	0.934	14.833	0.656	12.115
Swin UNETR	0.754	20.172	0.723	10.530	0.751	11.781	0.631	11.906	0.943	13.591	0.760	13.596
DBLNet	0.877	3.149	0.886	3.586	0.886	4.967	0.847	2.702	0.986	1.638	0.896	3.208

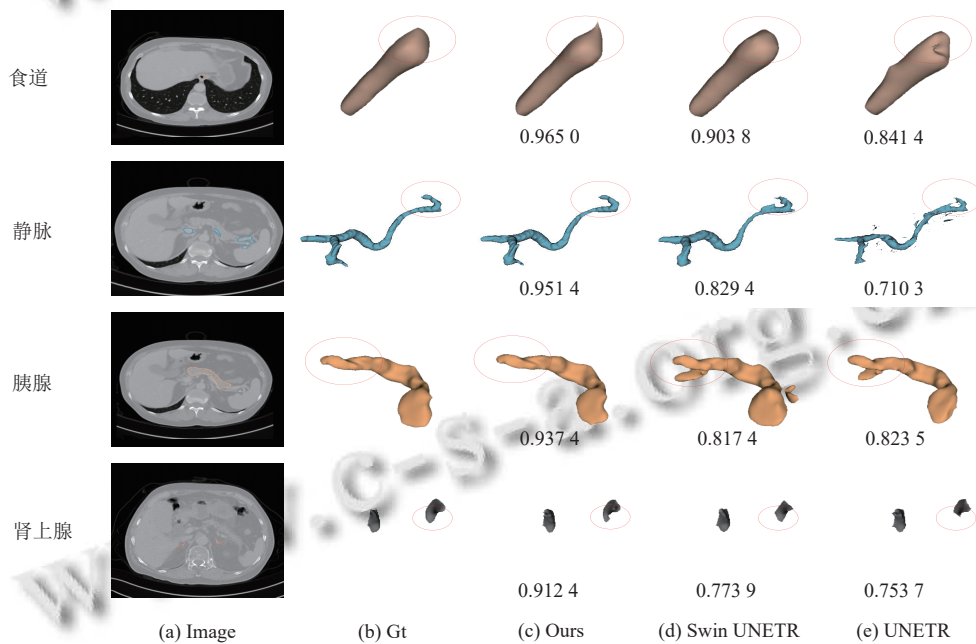


图 6 个体病例可视化各器官三维分割结果和 *Dice*

3.5 消融实验

为研究不同因素对模型的影响, 我们使用 BTCV 数据集进行消融研究. 具体研究分支 CEE module 和 BSSN 中 SFF module 对最终分割结果的影响. 我们考虑以下 3 种情况, 研发并试验以下模型: BL: 基线 UNet+批规范化. DBLNet*: UNet+CEE. DBLNet: UNet+CEE+SFF.

CEE module: UNet 增加 CEE 分支改造为并联的

双分支结构, 并将 CEE 的分割轮廓输入 BSSN 模块. 与 BL 比较, 结果如表 3 所示, 多器官分割准确率提高, 证明提取的轮廓特征发挥了作用. 多器官平均 *Dice* 比 BL 提高 0.403, 平均 *HD* 减少了 4.5556.

BSSN module: BSSN 中增加 SFF 模块处理相邻码层的跳层连接, 提取掩码的轮廓点输入 CEE 模块. 分割结果如表 3 所示, SFF 模块也提高了多器官的

Dice, 使得模型分割精度平均 *Dice* 得分比 BL 提高 0.753, 平均 *HD* 减少 4.983 4, 故 BSSN 中的 SF 模块是有效的, 具有重要的作用。

表3 对 CEE 和 BSSN 模块进行的消融实验

方法	<i>Dice</i>	<i>ASD</i>	<i>HD</i>
BL	87.086	1.589	23.2508
DBLNet*	87.489	1.4387	18.6952
DBLNet	87.839	1.3294	18.2674

实验结果表明, DBLNet 在分割精度方面优于单分支. CEE 和 SF 进一步提高了模型分割的精度. DBLNet 比 DBLNet* 的 *Dice* 提高了 0.35, *ASD* 和 *HD* 分别降低 0.1093 和 0.4278.

4 结论与展望

本文介绍了一种基于 CvT 和 CNN 的双分支编码结构的闭环网络, 命名为 DBLNet. 我们在 BTCV 数据集进行实验, 结果表明该方法对多器分割官的有效性; 并在临床医院数据进行应用, 发挥了模型的应用价值. 总之, DBLNet 为提取目标轮廓特征以及区域特征而设计并联的双分支结构, 并为实现信息的流动而首次设计闭环结构. 虽然本算法精度不是最高, 但有效解决了边界轮廓模糊不清问题, 下一步需要进行的研究工作是: (1) 精简优化模型, 减小计算量; (2) 进一步提高具有复杂结构的器官的预测精度.

参考文献

- 黄晓鸣, 何富运, 唐晓虎, 等. U-Net 及其变体在医学图像分割中的应用研究综述. 中国生物医学工程学报, 2022, 41(5): 567–576. [doi: 10.3969/j.issn.0258-8021.2022.05.006]
- He AL, Wang K, Li T, *et al.* H2Former: An efficient hierarchical hybrid Transformer for medical image segmentation. IEEE Transactions on Medical Imaging, 2023, 42(9): 2763–2775. [doi: 10.1109/TMI.2023.3264513]
- Wu HP, Xiao B, Codella N, *et al.* CvT: Introducing convolutions to vision Transformers. Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 22–31. [doi: 10.1109/ICCV48922.2021.00009]
- Liu TR, Huang JJ, Dai TH, *et al.* Gated multi-layer convolutional feature extraction network for robust pedestrian detection. Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona: IEEE, 2020. 3867–3871. [doi: 10.1109/ICASSP40776.2020.9054437]
- 徐光宪, 冯春, 马飞. 基于 UNet 的医学图像分割综述. 计算机科学与探索, 2023, 17(8): 1776–1792. [doi: 10.3778/j.issn.1673-9418.2301044]
- Oktay O, Schlemper J, Folgoc LL, *et al.* Attention U-Net: Learning where to look for the pancreas. arXiv:1804.03999, 2018.
- Meng YD, Zhang HR, Zhao YT, *et al.* Graph-based region and boundary aggregation for biomedical image segmentation. IEEE Transactions on Medical Imaging, 2022, 41(3): 690–701. [doi: 10.1109/TMI.2021.3123567]
- Lin Y, Zhang D, Fang X, *et al.* Rethinking boundary detection in deep learning models for medical image segmentation. Proceedings of the 28th International Conference on Information Processing in Medical Imaging. San Carlos de Bariloche: Springer, 2023. 730–742. [doi: 10.1007/978-3-031-34048-2_56]
- Liu XW, Hu YK, Chen JG, *et al.* Shape and boundary-aware multi-branch model for semi-supervised medical image segmentation. Computers in Biology and Medicine, 2022, 143: 105252. [doi: 10.1016/j.combiomed.2022.105252]
- Raffel C, Shazeer N, Roberts A, *et al.* Exploring the limits of transfer learning with a unified text-to-text Transformer. The Journal of Machine Learning Research, 2020, 21(1): 140.
- Xu GP, Wu XR, Zhang X, *et al.* LeViT-UNet: Make faster encoders with transformer for medical image segmentation. arXiv:2107.08623, 2021.
- Zhang JP, Xie YT, Wang Y, *et al.* Inter-slice context residual learning for 3D medical image segmentation. IEEE Transactions on Medical Imaging, 2021, 40(2): 661–672. [doi: 10.1109/TMI.2020.3034995]
- Chen JN, Lu YY, Yu QH, *et al.* TransUNet: Transformers make strong encoders for medical image segmentation. arXiv:2102.04306, 2021.
- Hatamizadeh A, Tang YC, Nath V, *et al.* UNETR: Transformers for 3D medical image segmentation. Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa: IEEE, 2022. 1748–1758. [doi: 10.1109/WACV51458.2022.00181]
- Hatamizadeh A, Nath V, Tang YC, *et al.* Swin UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images. Proceedings of the 7th International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Philadelphia: Springer, 2021. 272–284. [doi: 10.1007/978-3-031-08999-2_22]
- Larsson M, Zhang YH, Kahl F, *et al.* Robust abdominal

- organ segmentation using regional convolutional neural networks. *Applied Soft Computing*, 2018, 70: 465–471. [doi: [10.1016/j.asoc.2018.05.038](https://doi.org/10.1016/j.asoc.2018.05.038)]
- 17 Liu Z, Lin YT, Cao Y, *et al.* Swin Transformer: Hierarchical vision transformer using shifted windows. *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021. 9992–10002. [doi: [10.1109/ICCV48922.2021.00986](https://doi.org/10.1109/ICCV48922.2021.00986)]
- 18 Yuan L, Chen YP, Wang T, *et al.* Tokens-to-token ViT: Training vision transformers from scratch on ImageNet. *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021. 538–547. [doi: [10.1109/ICCV48922.2021.00060](https://doi.org/10.1109/ICCV48922.2021.00060)]
- 19 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- 20 李祥健, 朱家明, 徐婷宜. 基于改进 Canny 算子的双水平集医学图像分割. *无线电通信技术*, 2021, 47(2): 226–231. [doi: [10.3969/j.issn.1003-3114.2021.02.015](https://doi.org/10.3969/j.issn.1003-3114.2021.02.015)]
- 21 Xie Z, Wang HM, Wu L. The improved Douglas-Peucker algorithm based on the contour character. *Proceedings of the 19th International Conference on Geoinformatics*. Shanghai: IEEE, 2011. 1–5. [doi: [10.1109/GeoInformatics.2011.5981173](https://doi.org/10.1109/GeoInformatics.2011.5981173)]
- 22 陈英, 张伟, 林洪平, 等. 医学图像分割算法的损失函数综述. *生物医学工程学杂志*, 2023, 40(2): 392–400. [doi: [10.7507/1001-5515.202206038](https://doi.org/10.7507/1001-5515.202206038)]
- 23 Zhou YY, Li Z, Bai S, *et al.* Prior-aware neural network for partially-supervised multi-organ segmentation. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE, 2019. 10671–10680. [doi: [10.1109/ICCV.2019.0107](https://doi.org/10.1109/ICCV.2019.0107)]

(校对责编: 牛欣悦)