

基于多模态特征融合的脑瘤图像分割方法

方新林, 方艳红, 王迪

西南科技大学信息工程学院, 四川 绵阳 621010

【摘要】针对目前大多数医学图像分割方法难以对多模态图像进行特征融合进而完成精确分割任务的问题,提出一种基于编码器-解码器总体架构的多模态脑瘤图像特征融合策略。首先,编码阶段利用孪生网络对不同模态数据进行特征提取,孪生网络结构参数和权值共享的特性可有效减少网络参数量;其次,在进行特征提取的编码阶段加入级间融合,保留不同模态的共性特征的同时强调其互补特征;然后,在解码阶段引入密集跳跃连接思想,最大程度结合不同尺度特征图的低级细节和高级语义信息;最后,设计混合损失函数,在网络生成的预测图受真值图监督的同时让最高级特征融合图也受同倍下采样真值图的监督。所提方法在公开数据集BraTS2019上进行实验,并用图像分割常用的5种指标进行评估。在脑瘤及水肿区域分割任务中得到平均Dice系数为0.884,阳性预测率为0.870,灵敏度为0.898,豪斯多夫距离为3.917,平均交并比达到79.1%,与较先进的算法U-Net和PA-Net相比多项指标均有提升。实验结果说明,级间融合和层间跳跃连接的加入对多模态医学图像的分割效果有所提升,在医学上对脑肿瘤磁共振图像进行病变区域分割具有重要的应用价值和理论意义。

【关键词】多模态图像;脑肿瘤;特征融合;医学图像分割;深度学习

【中图分类号】R318;TP391.4

【文献标志码】A

【文章编号】1005-202X(2022)06-0682-08

Brain tumor image segmentation method based on multi-modal feature fusion

FANG Xinlin, FANG Yanhong, WANG Di

School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

Abstract: Aiming at the problem that most of the current medical image segmentation methods are difficult to perform feature fusion for multi-modal images to achieve accurate segmentation, a multi-modal brain tumor image feature fusion strategy based on encoder and decoder overall architecture is proposed. In the coding phase, twin networks are used to extract features from different modal data. The number of network parameters can be effectively reduced by sharing the structural parameters and weights of twin networks. In addition, the interstage fusion is added in the coding phase of feature extraction to keep the common features of different modes while emphasizing their complementary features. Then, the idea of dense skip connection is introduced in the decoding phase to maximize the combination of low-level details and high-level semantic information of feature maps at different scales. Finally, a mixed loss function is designed, so that the prediction graph generated by the network is supervised by the truth graph, and that the highest-level feature fusion graph is also supervised by the truth graph sampled under the same multiplier. The proposed method is tested on the public data set BraTS2019 and evaluated with 5 commonly used indexes for image segmentation. For the segmentations of brain tumor and edema area, the proposed method are superior to more advanced algorithms U-NET and PA-NET in many indexes, and the average Dice coefficient, positive prediction rate, sensitivity, Hausdorff distance and mean intersection over union of the proposed method are 0.884, 0.870, 0.898, 3.917 and 79.1%, respectively. The experimental results reveal that the addition of interstage fusion and interlayer skip connection can improve the segmentation performance of multimodal medical images, and it has important application value and theoretical significance in the segmentation of brain tumor in magnetic resonance image.

Keywords: multi-modal image; brain tumor; feature fusion; medical image segmentation; deep learning

【收稿日期】2021-12-19

【基金项目】国家重点实验室开放基金(SKLA20200203)

【作者简介】方新林,硕士,研究方向:深度学习、医学图像处理,E-mail: 18140341994@163.com

【通信作者】方艳红,博士,研究方向:智能仪器设备、图像处理与模式识别,E-mail: 19137088@qq.com

前言

神经胶质瘤是一种始于大脑或脊柱的胶质细胞肿瘤,该肿瘤约占所有脑肿瘤和中枢神经系统肿瘤的30%,占有恶性脑肿瘤的80%。作为最常见的原发性脑肿瘤,具有不同层度的侵袭性,会对患者健康

造成巨大威胁。对胶质瘤的精准分割方法的研究在临床诊断上具有重要意义。医学图像分割的目的是半自动或全自动地勾画出感兴趣的解剖结构,如肿瘤、器官、组织等,在临床中有很多应用,如放射学分析、治疗计划、生存分析等^[1-2]。目前,医学图像分割也是一个活跃的研究课题。在过去的数年中,已经提出了数千种医学图像分割方法用于不同医学图像中的各种器官和病变。医学图像分割是医学图像分析中的一个重要领域,是诊断、检测和治疗所必须的。但由于目标组织的大小、形状和位置的变化,医学图像分割是医学图像分析领域最具挑战的任务之一。医学图像分析的主要方法包含计算机断层扫描^[3](Computed Tomography, CT)、核磁共振成像^[4](Magnetic Resonance Imaging, MRI)和阳电子发射断层扫描^[5](Positron Emission Tomography, PET),与单幅图像相比,多模态图像有助于从不同视图中提取特征,并带来互补信息,有助于实现更好的数据表示。多模态图像分割通过融合多个信息来提高分割效果,因其可以提供多个目标肿瘤信息,可以获得更准确的分割结果以便更好地进行诊断,该方法在近年来已成为一种研究趋势。

随着深度学习的迅速发展,深度神经网络的自学习性和泛化性不仅对如分类、检测和分割等基础的机器视觉任务提供了更先进的学习手段,还为多模态医学图像领域提供了新的视角。U-Net^[6]提出一种具有跳跃连接的对称结构来提高图像细节特征的保留,成为大部分医学图像分割任务的主流框架,基于U-Net的变体工作^[7-8]进一步提高了分割性能,但其卷积运算的局限性导致其对语义信息的密集预测受到了限制^[9-10];工作建立学习规则对CNN预测肿瘤点的概率进行再判断,虽然分割精度有所提升,但所提算法缺少深度学习的自主性^[11]。上述工作均缺少对多模态特征的结合,会导致脑瘤图像的不完整分割,并且早期的多模态医学图像融合方法^[12],大部分侧重于后续的复杂分割网络体系结构设计,没有考虑不同模态间的关系,同时也没有分析如何融合不同特征信息来提高分割性能,因此在多模态融合中引入深度神经网络模型可使每个模态间的互补信息得以利用。

针对上述问题,根据深度学习和多模态融合的理论提出一种基于编码器-解码器的多模态脑瘤图像分割模型,来实现多模态神经胶质瘤分割任务。

1 多模态医学图像分割相关理论

1.1 医学图像分割

自U-Net图像分割方法于2015年出现以来,针对各种分割任务,有许多基于U-Net改进的医学图像

新分割方法被提出^[13-14]。该类方法对单模态数据来说非常友好,能同时获取低级语义信息和高级语义信息,且分割准确率高、速度快。对于输入有多种模态的情况下,该类方法虽然可以对各个模态特征进行有效提取,但却难以做到去冗余的同时并保留各个模态的突出特征。因此,本文在保留特征提取的主干网络的前提下,提出一种多模态融合特征映射的网络模型,达到有效去除各个模态之间的冗余信息,保留其互补特征的效果。

1.2 孪生网络

目前针对多模态输入图像的主流特征提取过程都是直接将输入数据按照通道叠加的方式送入特征提取网络Backbone(如VGG^[15]、ResNet^[16]),对输出的特征映射赋予相同的权值进行融合或其它操作。本文利用孪生网络对不同模态的图像数据进行联合学习。它的目的是从基于模型的角度发现不同模态之间的共性,它们的信息可以通过反向传播合并到模型参数中,特征提取主干网络共同学习的层次特征被反馈给后续的融合网络部分。由于孪生网络的共享权值特性,分别提取多模态特征的同时,有效减小模型参数量。

1.3 多模态融合

之所以要对不同模态的图像数据进行融合^[17],是因为不同模态的表现方式不一样,看待事物的角度也会不一样,所以存在一些交叉(信息冗余)、互补(比单模态特征更优秀)的现象,甚至模态间可能还存在多种不同的信息交互,如果能合理处理多模态信息,就能得到丰富的特征信息,概括来说多模态的显著特点是冗余性和互补性。

对于经典的基于CNN的多模态视觉融合算法来说,当融合特征来自于某一深度层时,可以根据融合特征所处深度的不同分为4种融合方式,分别为低级融合、中级融合、高级融合和决策级融合^[18-19]。图1为不同级别融合方式的示意图。

对多模态图像来说,不同空间位置和通道的特征对图像中的目标区域或背景区域有着不同的响应映射^[20]。将所有空间位置或者通道都平等考虑,会导致次优结果^[21]。鉴于跳层结构的多层次和多尺度特征的融合能力,可以将其引入到多模态图像检测算法中。利用其多层次的特点,灵活连接图像的高层特征和低层特征,并通过跳层结构的多尺度融合能力,将浅层网络的细节特征和深层网络的语义特征充分结合。因此,不同于上述几种经典融合方式,本文提出一种先层间再跨层融合的机制,对感兴趣目标有较高映射的空间位置或通道所占权重进行灵活分配。本文融合方式如图2所示。

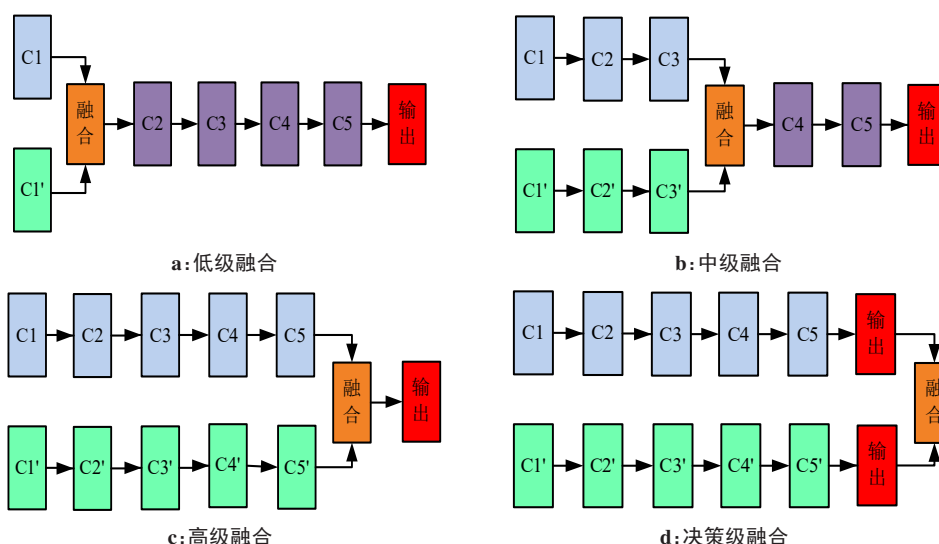


图1 不同融合方式

Figure 1 Different fusion methods

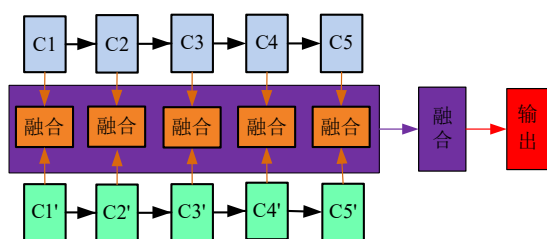


图2 本文融合策略

Figure 2 The proposed fusion strategy

2 多模态融合的脑瘤图像分割方法

2.1 问题描述

本文的目标是建立一个分割多模态神经胶质瘤的模型,该模型在分别提取各模态特征的同时,还能实现跨模态特征融合。本文基于BraTS2019数据集^[22]构建了训练集和测试集,由于存在多模态数据,因此每个样本由4个单通道模态数据加上对应的真值图像组成。两个集合中的样本组不存在重复,本文分割类别有3类,3类训练集和测试集中模态图片不变的情况下改变每类的真值图像,训练目标为最小化且收敛的像素级交叉熵损失。通过训练获得的网络分割模型,在测试集上进行多模态肿瘤分割性能,评估标准包括Dice相似系数(Dice Similarity Coefficient, DSC)、阳性预测率(Positive Prediction Value, PPV)、灵敏度(Sensitivity, SE)、平均交并比(mean Intersection Over Union, mIOU)和豪斯多夫距离(Hausdorff Distance, HD)。

2.2 网络框架

本文提出的网络模型总体框架如图3所示。它采用经典的编解码器网络架构^[23-24],利用孪生网络对多模态图像进行特征提取的特点是由于孪生网络的

网络结构和网络权值是共享的,因此针对多模态输入数据可以实现跨模态知识共享的同时能极大地降低网络模型的参数量。

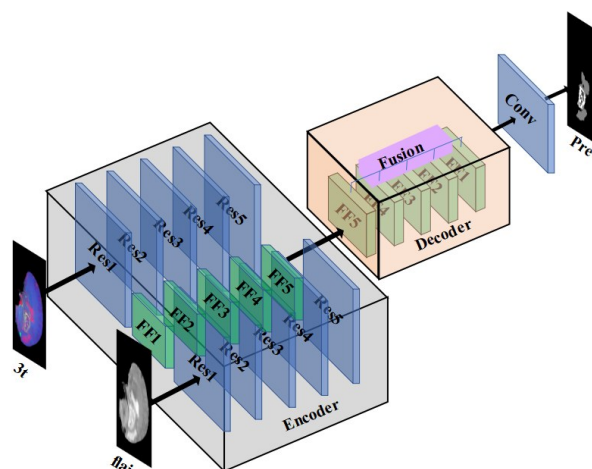


图3 网络总体架构

Figure 3 Overall framework of the network

其中编码器部分将输入的两个模态 3t、flair 图像进行逐级特征提取,提取出来的特征图通过 FF 融合模块进行级间融合。解码器部分通过密集连接策略进行二次融合和上采样。最后通过 1*1 的卷积生成预测分割图,该预测图由真值图(GT)进行监督。

2.3 融合策略

鉴于多模态图像融合的精髓在于保留共性特征的同时要强调互补特征^[21]。因此,本文在融合策略上分两步走。第一步在提取的同级特征基础上进行完全融合,以达到留互补、去冗余特征的效果;第二步在解码器部分进行上下级跳跃连接,连接方式采

用 Dense-Net^[23]中提到的密集连接方式,用以将不同尺度特征图的低级细节与高级语义结合起来,下面将分别介绍两种融合策略。

2.3.1 级间融合 该阶段融合操作在图3中的FF模块中进行,主干网络提取特征的5个层级中每个层级都进行融合。融合的操作如图4所示,C1表示输入的3t模态经 Res1 生成的特征映射,C1'表示 flair 模态经 Res1 生成的特征映射。该融合过程的操作如下所示:

$$FF(\{C1, C1'\}) = C1 \oplus C1' \oplus (C1 \otimes C1') \quad (1)$$

其中,“ \oplus ”表示元素相加,强调特征互补,“ \otimes ”表示元素相乘,强调特征共性。这两种运算性质在跨膜态融合中起到不可替代的重要作用。同理,两种模态经主干网络后续生成的特征映射图分别为 C2、C2'、C3、C3'、C4、C4'、C5、C5',并和上述融合过程一样最终生成 FM2、FM3、FM4、FM5。

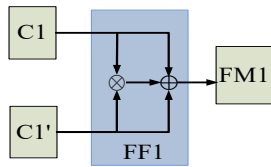


图4 级间融合

Figure 4 Interstage fusion

2.3.2 跳跃连接 卷积网络如果在靠近输入的层和靠近输出的层之间包含较短的连接,就可以更深入、更准确、更有效地进行训练。Huang 等^[23]在此理论上提出密集卷积网络(DenseNet),该网络以前馈的方式将每一层连接到其他每一层。对于每一个层,它之前层的所有 feature-map 被用作为输入,它本身的 feature-map 被用作它所有后续层的输入。由于图像数据通常具有尺度自相似性,该网络在不同深度的连接层实际上同时包含和使用了不同尺度的图像信息。因此,在低级特征和高级特征融合过程中起到非常好的效果。DenseNet 具有几个令人着迷的优势:减轻梯度消失问题,增强特征传播,强调特征重用,同时由于瓶颈层的 1*1 卷积,有效减小了参数数量。

本文参考 UNet++^[25]在 UNet 的基础上改进的策略,在网络解码器部分引入 DenseNet 思想,通过将来自不同尺度特征图的低级细节与高级语义结合起来,完成跳级融合图像特征。解码器部分网络结构如图5所示。

2.4 损失函数

由于本文的融合过程涉及到级间融合以及跳跃连接两部分,因此,本文训练的损失函数为混合损失

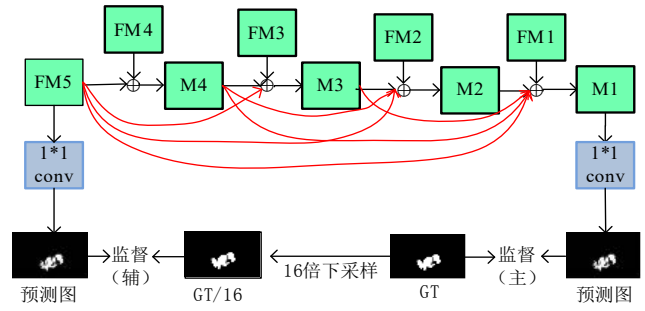


图5 解码器部分结构图

Figure 5 Framework of decoder

函数。如图5所示,以解码器最终预测分割图和GT的监督为主,记为 L_{final} ,最高级融合结果图 FM5 与下采样的 GT 监督为辅,记为 L_{fm5} 。共同组成的混合损失函数为:

$$L_{all} = L_{final} + \lambda L_{fm5} \quad (2)$$

其中, λ 用于平衡两个损失之间的权重关系,同时两个损失函数均使用语义分割领域广泛使用的交叉熵损失函数:

$$L = -\sum [G_i \log(S_i) + (1 - G_i) \log(1 - S_i)] \quad (3)$$

其中, i 表示像素索引, G 表示来自 GT 图像的监督, S 表示预测图。

3 实验与结果分析

3.1 数据集

本文采用 BraTS2019 数据集,包含高级神经胶质瘤(High-Grade Glioma, HGG)和低级神经胶质瘤(Low-Grade Glioma, LGG)两类。每例病人包含 flair、f1、f1ce、f2 共 4 个模态的三维 MR 图像及其对应的 GT 掩膜图像,每张图像的分辨率为 155×240。上述 4 个模态,也叫 4 种不同 MR 序列。其中 flair 序列是 MR 的一种常用的序列,全称是液体衰减反转回复序列,也称为水抑制成像技术。在该序列上,脑脊液呈现低信号(暗一些),实质性病灶和含有结合水的病灶显示为明显的高信号(亮一些)。t1、t2 是由于测量电磁波的物理量不同而产生的两种不同的序列,通过 t1 图像可以看出各种断层解剖图;t2 信号伴随着水含量的变化,反映出病灶周围的水肿区域,所以通过 t2 图像可以看出病变区域;通过 t1ce 序列可以看出血液饱和程度。

BraTS2019 数据集包括来自 BraTS2013 的 20 个病例、来自 CBICA 的 120 个病例、来自 TCIA 的 102 个病例以及来自 TMC 的 8 个病例。其中 BraTS2013 的图像和标签由专业临床医生进行手动标注;TCIA 数据集由专业医生进行放射学评估。这些图像和对应的标签已被各自领域的专家所认可。

本文将上述 BraTS2019 数据集分为训练集和测试集。其中,训练集由 BraTS2013 和 CBICA 的 140 个病例数据组成,预处理且经挑选后的二维序列切片总计共 15 325 张,整合模态后共 3 065 组;测试集由 TCIA 和 TMC 的 110 个病例数据组成,预处理且经挑选后的二维序列切片总计共 7 165 张,整合模态后共 1 433 组。

3.2 预处理

(1)二维切片: BraTS2019 数据集官方以 .nii 格式的三维数据给出,首先对三维 MR 数据进行二位切片,每个病例以冠状方向切割可以获得 155 张分辨率为 155×240 的单通道切片。原始 4 种模态序列图片和 GT 掩膜切片如图 6a 所示。

(2)整合模态:通过第 3.1 节对 4 种不同序列图像的功能描述,本文将 t1、t2、t1ce 3 种序列通过通道叠加变为三通道图像,作为孪生网络主路特征提取的对象,叫作 3t 模态; flair 序列单独作为孪生网络支路特征提取的对象,以此减少网络模型参数和降低网络模型复杂度。整合之后的双模态如图 6b 所示。

(3)标签分类:图 6d 给出了每组切片对应病变区域的 3 类标签,分别是为浮肿区域 (Peritumoral Edema, ED)、增强肿瘤区域 (Enhancing Tumor, ET)、坏疽 (Tumor Center, TC)。本文参考 MICCAI 的分割任务,将上述标签合并为 3 个嵌套的子区域分别为 WT、ET、TC,其中 WT=ED+ET+TC,处理前后标签对应关系如表 1 所示。标签分离后的 GT 图像如图 6c 所示。

(4)裁剪:由于二维切片中含有大量对分割结果不产生影响的黑色背景区域,因此对每个病例的序列切片进行裁剪是有必要的,这样可以扩大肿瘤区域占比,减少不必要的卷积操作。各模态数据通过遍历每个病例的上下左右边界最大值进行裁剪,最后统一 Resize 成 320×320 的大小,Resize 过程采用双线性插值。测试过程的预测结果会映射回原始输入图像的大小,便于肉眼观察其与 GT 之间的差异。

(5)Z-score 标准化:由于神经网络学习的本质是为了学习数据分布,所以最后对上述裁剪好的图像进行 Z-score 标准化处理,主要目的是将不同量级的数据统一转化为同一个量级,保证数据之间的可比性。同时可以提高网络的训练速度,加快网络收敛。

3.3 评价指标

为了定量评估 WDNET 模型的肿瘤分割性能,本文采用医学图像分割常用的 5 个指标进行度量。涉及相关指标计算的有以下几个参数: TP (True Positive): 真阳性,指预测结果是 TRUE,且 GT 也是

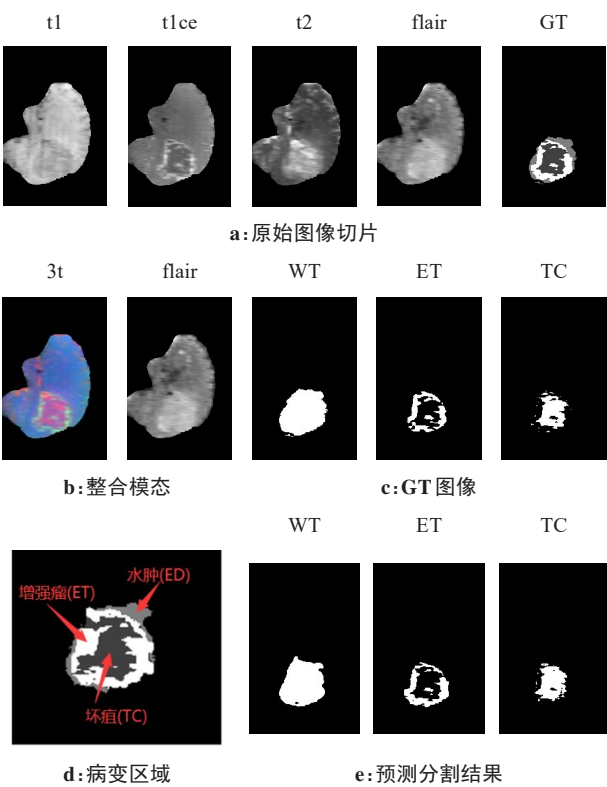


图 6 图像数据处理与分割结果

Figure 6 Image data processing and segmentation results

表 1 3 类标签包含的病变区域

Table 1 Diseased area contained in 3 types of labels

分类	TC	ED	ET
WT	√	√	√
TC	√		
ET			√

TRUE 的像素数量; TN (True Negative): 真阴性, 表示预测结果是 FALSE, 且 GT 也是 FALSE 的像素数量; FN (False Negative): 表示预测结果是 FALSE, 但 GT 是 TRUE 的像素数量; FP (False Positive): 表示预测结果是 TRUE, 但 GT 是 FALSE 的像素数量。

(1) Dice 系数, 它是一种集合相似度度量函数, 通常用于计算两个样本的相似度值, 其值范围为 [0, 1], 计算公式如下所示:

$$\text{Dice} = \frac{2TP}{2TP + FN + FP}$$

(4)

(2) PPV, 即准确率, 指在所有样本的预测结果中, 所有阳性样本中真阳性所占比例, 计算公式如下所示:

$$\text{PPV} = \frac{TP}{TP + FP}$$

(5)

(3) SE, 即灵敏度, 指在所有阳性样例中被预测出为真阳性的样例所占比例, 计算公式如下所示:

$$\text{SE} = \frac{TP}{TP + FN}$$

(6)

(4)mIOU,即平均交并比,它是语义分割的标准度量,用于计算所有样本预测值和真实值GT两个集合的交集与并集之比的平均值:

$$\text{mIOU} = \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP}}$$

(7)

(5)HD,即豪斯多夫距离,它是描述两组点集之间相似程度的一种量度,它是两个点集之间距离的一种定义形式:假设有两组集合 $A=\{a_1, \cdots, a_p\}, B=\{b_1, \cdots, b_q\}$,则这两个点集之间的豪斯多夫距离定义为:

$$H(A, B) = \max(h(A, B), h(B, A))$$

(8)

其中,

$$h(A, B) = \max_{a \in A} \{ \min_{b \in B} \|a - b\| \}$$

(9)

$$h(B, A) = \max_{b \in B} \{ \min_{a \in A} \|b - a\| \}$$

(10)

其中, $\|\cdot\|$ 为点集 A 和 B 点集间的距离范式。式(8)称为双向豪斯多夫距离,是最基本形式,式(9)和式(10)分别指从 A 到 B 和从 B 到 A 的单向豪斯多夫距离。可以看出,双向豪斯多夫距离 $H(A, B)$ 是单向距离 $h(A, B)$ 和 $h(B, A)$ 两者中的较大者,它度量了两个点集间的最大不匹配程度。本文统一采用双向豪斯多夫距离作为评价指标。

3.4 实验结果

3.4.1 实验配置及参数设置 本实验运行的硬件平台配置为处理器 Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz,内存RAM为16 GB,GPU为NVIDIA GeForce RTX 2080 Ti,显存为11 GB,系统环境为Windows 10。集成开发环境 (IDE, Integrated Development Environment) 为 Spyder4.2.3, 搭载 Python3.8.0+ Pytorch1.8.0+torchvision0.9.0+cuda10.2。训练过程采用Adam优化器作参数优化,初始学习率设置为0.000 05,权重衰减设置为0.000 5,动量 momentum 设置为0.99,

batch_size 设置为1,训练次数 epoch 为100。3种标签的分割任务训练 loss 曲线如图7所示,可以看出,均在 epoch=30 左右开始收敛。

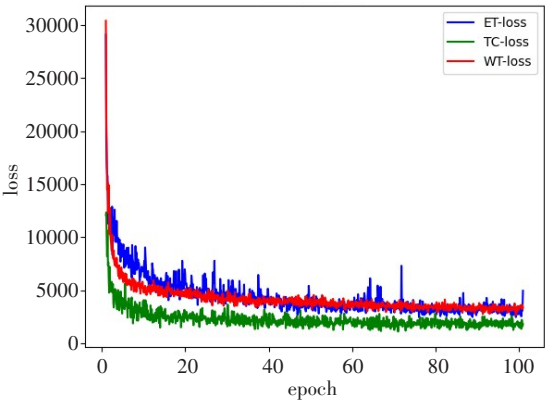


图7 3类标签训练损失曲线
Figure 7 Training loss curves of 3 types of labels

3.4.2 本文分割效果 为了验证本文提出的基于多模态特征融合的脑瘤图像分割网络模型的性能,利用TCIA和TMC两个数据集组成的1433组测试集对训练好的网络模型进行测试。预测结果如图6e所示,可以看出,该模型预测的各类分割结果(图6e)与GT(图6c)图像比较接近。

为了验证本文提出的网络模型在多模态脑瘤图像分割任务上具有先进性,对比试验采用当前的两个SOTA算法U-Net和PA-Net。在使用相同训练集进行模型训练的基础上,用相同的测试集对3种模型进行平行测试,与本文方法进行对比实验。以WT标签分割任务为例,3种网络的WT分割结果和GT图像的对比如图8所示。

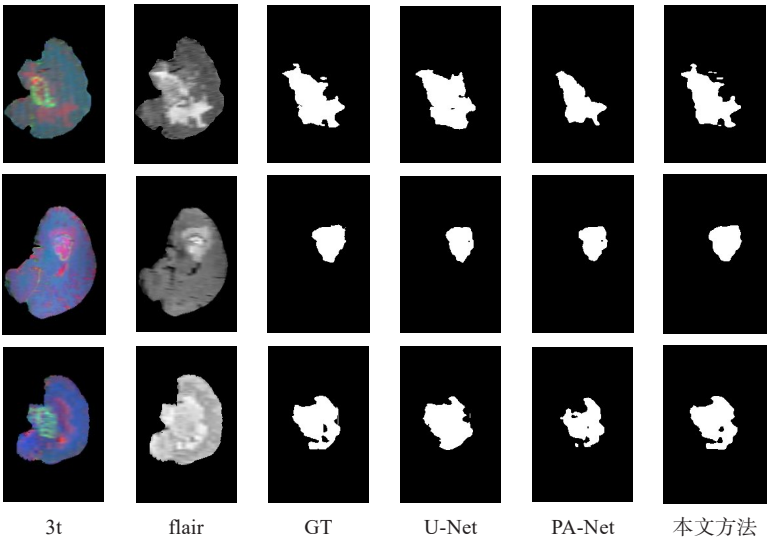


图8 不同算法分割结果对比
Figure 8 Comparison of segmentation results of different algorithms

第1、2列分别是输入模型的两模态3t和flair图像,第3列是真值GT图像,第4列是U-Net分割结果,第5列是PA-Net分割结果,第6列是本文方法的分割结果。由第3行案例可以看出,PA-Net模型容易出现漏检的情况,而U-Net难以将目标区域内部的背景与目标区分开来,而本文方法预测的分割结果与金标准比较相似。相比本文提出的方法,U-Net没有利用残差模块,卷积特征采用层层递进的方式进行传递,因此特征利用率较低。而PA-Net虽然引入FPN并且添加了路径增强模块,能有效融合低级特征与高级特征,但是对于多模态通道特征仍然没有做到去冗余、留互补。相对而言,本文方法在多模态脑瘤MR图像分割任务上解决了其他算法的过分割和欠分割问题,对目标肿瘤区域可以进行相对准确的分割。通过定量计算预测样本与对应GT之间的差异,得到3种模型在各类分割任务的平均分割精度指标,对比结果如表2所示。

表2 模型分割结果对比
Table 2 Comparison of model segmentation performances

标签	模型	Dice	PPV	SE	HD	mIOU/%
WT	U-Net	0.736	0.621	0.904	4.927	58.3
	PA-Net	0.801	0.792	0.809	4.611	66.8
	本文方法	0.884	0.870	0.898	3.917	79.1
ET	U-Net	0.691	0.574	0.806	5.072	52.8
	PA-Net	0.718	0.681	0.759	4.923	56.1
	本文方法	0.792	0.738	0.856	4.136	64.7
TC	U-Net	0.624	0.589	0.717	4.391	45.3
	PA-Net	0.681	0.695	0.675	4.672	51.6
	本文方法	0.714	0.629	0.825	3.868	55.5

由表2可以看出,3种分割任务中WT的分割精度相对ET和TC来说更高,原因可能是增强瘤区域的结构分布较为杂乱,导致肿瘤核心及其外围增强肿瘤相对水肿区域更加难以区分。另外,除了U-Net在WT分割任务中的SE指标值最高外,本文提出的网络模型指标明显高于其它两种,这说明本文方法在多模态脑瘤图像分割任务中具有更好的效果。但是,由于本文方法的网络模型采用了分路特征提取以及繁复的跨级特征融合等措施,导致网络模型体量相对其它两种模型来说较大,本文提出的网络模型参数高达1.2亿,大小为490 M,在第3.4.1节的实验配置中推理速度为17 FPS,说明该方法在占用更多硬件算力和内存的情况下,提高了医学图像分割任务的精度。因此,该方法的另一个弊端是很难部署到低算力AI嵌入式设备上。

4 结 语

本文针对现有图像分割算法对于多模态医学图像分割效果较差的问题,提出一种基于孪生网络的多模态特征融合分割方法。该方法的特点是在编码器阶段加入级间融合来对各模态的特征进行加强,以及在解码器阶段引入跨层密集跳跃连接来融合不同尺度的多模态特征,充分融合各级特征映射图的同时也对不同模态图像进行了去冗余信息、保留互补信息的操作。实验结果表明,该模型能够准确分割3类脑部神经胶质瘤病变区域。和其他算法相比,在Dice、mIOU、HD等评价指标上都有较大提升。但是,该模型体量过大,推理速度较慢,因此,在后续工作中,针对其进行轻量级优化研究是有必要的。

【参考文献】

[1] 苏会芳,周国锋.放射组学的兴起和研究进展[J].中华医学杂志,2015,95(7):553-556.
Su HF, Zhou GF. Research progress of radiomics[J]. Chinese Medical Journal, 2015, 95(7): 553-556.

[2] 曲滨鹏,魏晓洁,缪佳,等.图像分割技术在医学图像处理中的应用实践[J].科技创新与应用,2021,11(16):178-180.
Qu BP, Wei XJ, Miao J, et al. Application of image segmentation technology in medical image processing[J]. Science & Technology Innovation and Application, 2021, 11(16): 178-180.

[3] Lei T, Sewchand W. Statistical approach to X-ray CT imaging and its applications in image analysis. II. A new stochastic model-based image segmentation technique for X-ray CT image[J]. IEEE Trans Med Imaging, 1992, 11(1): 62-69.

[4] Bauer S, Wiest R, Nolte LP, et al. A survey of MRI-based medical image analysis for brain tumor studies[J]. Phys Med Biol, 2013, 58(13): R97.

[5] Boellaard R. Standards for PET image acquisition and quantitative data analysis[J]. J Nucl Med, 2009, 50(Suppl 1): 11S-20S.

[6] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer, 2015: 234-241.

[7] Li X, Chen H, Qi X, et al. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes[J]. IEEE Trans Med Imaging, 2018, 37(12): 2663-2674.

[8] Guan S, Khan AA, Sikdar S, et al. Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal[J]. IEEE J Biomed Health, 2019, 24(2): 568-576.

[9] 王翔宇,李海生,吕丽君,等.基于U-net和可见光谱图像的黄瓜褐斑病分割[J].光谱学与光谱分析,2021,41(5):1499-1504.
Wang XY, Li HS, Lü LJ, et al. Segmentation of cucumber brown spot based on U-NET and visible spectrum images[J]. Spectroscopy and Spectral Analysis, 2021, 41(5): 1499-1504.

[10] 殷晓航,王永才,李德英.基于U-Net结构改进的医学影像分割技术综述[J].软件学报,2021,32(2):519-550.
Yin XH, Wang YC, Li DY. Review of medical image segmentation techniques based on improved U-Net structure[J]. Journal of Software, 2021, 32(2): 519-550.

[11] 师冬丽,李镔,关欣.结合卷积神经网络和模糊系统的脑肿瘤分割[J].计算机科学与探索,2018,12(4):608-617.
Shi DL, Li Q, Guan X. Brain tumor segmentation based on convolutional neural network and fuzzy system[J]. Computer Science and Exploration, 2018, 12(4): 608-617.

[12] Zhou T, Ruan S, Canu S. A review: deep learning for medical image segmentation using multi-modality fusion[J]. Array, 2019, 3: 100004.

[13] 殷宁波,黄冕,刘利军,等. MS-UNet++: 基于改进UNet++的视网膜血管分割[J].光电子·激光,2021,32(1):35-41.

- Yin NB, Huang M, Liu L J, et al. Ms-unet++: retinal vascular segmentation based on improved UNet++ [J]. Journal of Optoelectronics·Laser, 2021, 32(1): 35-41.
- [14] 蒋宏达, 叶西宁. 一种改进的I-Unet网络的皮肤病图像分割算法[J]. 现代电子技术, 2019, 42(12): 52-56.
- Jiang HD, Ye XN. An improved skin disease image segmentation algorithm based on I-UNET network [J]. Modern Electronics Technology, 2019, 42(12): 52-56.
- [15] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409. 1556, 2014.
- [16] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [17] Zhu Z, Yin H, Chai Y, et al. A novel multi-modality image fusion method based on image decomposition and sparse representation[J]. Inform Sci, 2018, 432: 516-529.
- [18] He A, Luo C, Tian X, et al. A twofold siamese network for real-time object tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 4834-4843.
- [19] Li C, Wu X, Zhao N, et al. Fusing two-stream convolutional neural networks for RGB-T object tracking[J]. Neurocomputing, 2018, 281 (15): 78-85.
- [20] Liu HP, Sun FC. Fusion tracking in color and infrared images using joint sparse representation[J]. Science China Information Sciences, 2012, 55(3): 590-599.
- [21] Zhou T, Fan DP, Cheng MM, et al. RGB-D salient object detection: a survey[J]. Computational Visual Media, 2021, 7(1): 37-69.
- [22] Li X, Luo G, Wang K. Multi-step cascaded networks for brain tumor segmentation [C]//International MICCAI Brainlesion Workshop. Cham: Springer, 2019: 163-173.
- [23] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference On Computer Vision and Pattern Recognition. 2017: 4700-4708.
- [24] Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Trans Pattern Anal, 2017, 39(12): 2481-2495.
- [25] Zhou Z, Siddiquee MM, Tajbakhsh N, et al. Unet++: redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Trans Med Imaging, 2019, 39(6): 1856-1867.

(编辑:陈丽霞)