



计算机科学

COMPUTER SCIENCE

基于边缘引导的多尺度医学影像分割方法

姜灏天, 王琦智, 黄扬林, 章雅琴, 胡凯

引用本文

姜灏天, 王琦智, 黄扬林, 章雅琴, 胡凯. [基于边缘引导的多尺度医学影像分割方法](#)[J]. 计算机科学, 2023, 50(11A): 220900059-7.

JIANG Haotian, WANG Qizhi, HUANG Yanglin, ZHANG Yaqin and HU Kai. [Medical Image Segmentation Based on Multi-scale Edge Guidance](#) [J]. Computer Science, 2023, 50(11A): 220900059-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于语义注意力的医学图像超分辨率方法](#)

Medical Image Super-resolution Method Based on Semantic Attention

计算机科学, 2023, 50(11A): 221200107-6. <https://doi.org/10.11896/jsjcx.221200107>

[一种基于因果推理的垃圾分类方法](#)

Novel Method for Trash Classification Based on Causal Inference

计算机科学, 2023, 50(11A): 220800218-6. <https://doi.org/10.11896/jsjcx.220800218>

[接诉即办智能派单业务调度算法研究](#)

Study on Scheduling Algorithm of Intelligent Order Dispatching

计算机科学, 2023, 50(11A): 230300029-7. <https://doi.org/10.11896/jsjcx.230300029>

[基于LSTM神经网络的QPSK智能接收机设计](#)

Design of QPSK Intelligent Receiver Based on LSTM Neural Network

计算机科学, 2023, 50(11A): 230200219-5. <https://doi.org/10.11896/jsjcx.230200219>

[一种噪声容忍的网络流量分类方法](#)

Noise Tolerant Algorithm for Network Traffic Classification Method

计算机科学, 2023, 50(11A): 220800120-7. <https://doi.org/10.11896/jsjcx.220800120>

基于边缘引导的多尺度医学影像分割方法

姜灏天¹ 王琦智¹ 黄扬林¹ 章雅琴² 胡凯¹

1 湘潭大学计算机学院·网络空间安全学院 湖南湘潭 411105

2 中南大学湘雅三医院放射科 长沙 410013

摘要 医学影像的灰阶变化小,分割目标与背景不易区分,因此,进行影像分割是充满挑战性的问题。现有网络模型大多将高频的分割边缘与低频的主体部分统一学习,忽视了高频与低频信息的差异性和两者在图像中占比不同的差别。针对这一问题,提出了基于边缘引导的多尺度卷积神经网络 Edge Guided V-Shape Network (EGV-Net),从低频分割主体和高频分割边缘两个特征角度进行针对性学习。其中,低频特征通过编码-解码方式进行特征传递,学习分割目标的主体部分;高频特征则通过边缘提取方法,首先将高频语义信息从分割图谱中提取出来,再将分割边缘过滤分离。高频边缘通过边缘引导模块指导模型对低频特征做出精准的分割,并恢复边缘细节精度。在肝脏影像与 ISIC2016 数据集上进行的实验结果表明,所提算法对整体分割的把控能力更强,在边缘细节处有更好的分割效果,优于其他模型。

关键词: 深度学习;医学影像分割;多尺度特征;边缘提取;边缘引导

中图分类号 TP391

Medical Image Segmentation Based on Multi-scale Edge Guidance

JIANG Haotian¹, WANG Qizhi¹, HUANG Yanglin¹, ZHANG Yaqin² and HU Kai¹

1 School of Computer Science & School of Cyberspace Science, Xiangtan University, Xiangtan, Hunan 411105, China

2 Department of Radiology, The Third Xiangya Hospital, Central South University, Changsha 410013, China

Abstract Medical images have small gray-scale changes, and segmentation targets and backgrounds are not easy to distinguish, thus image segmentation is full of challenging problems. Most of the existing models unify the segmented high-frequency edges with the low-frequency subjects for learning, ignoring the difference between high-frequency information and low-frequency information and the difference in the proportion of both in the image. To address this problem, edge guided V-shape network (EGV-Net), a multi-scale convolutional neural network based on edge guidance, is proposed to perform targeted learning from two feature perspectives: low-frequency segmented subjects and high-frequency segmented edges. Among them, the low-frequency features are passed through the feature transfer by the encoder-decoder connection method to learn the main part of the segmentation target. The high-frequency features are firstly extracted from the segmentation mapping by edge extraction method, and then the segmentation edges are filtered and separated from it. The segmented edges of high frequency are guided by edge guidance module to make accurate segmentation of low frequency segmented edges and recover edge detail accuracy. Experimental results in liver images and ISIC2016 show that the proposed algorithm has better control over the overall segmentation and better segmentation effect at the edge details than other models.

Keywords Deep learning, Medical image segmentation, Multi-scale features, Edge extraction, Edge guidance

1 引言

医学影像是针对人体或人体某些部分,采用非侵入方式获取内部组织影像的技术与处理过程。随着科学技术逐渐成熟,计算机断层扫描(Computed Tomography, CT)、磁共振成像(Magnetic Resonance Imaging, MRI)、超声波成像(Ultrasonic Imaging)等多种医学影像方法^[1]应运而生。随着医学影像应用的推广,越来越多的影像需要医生解读与标注,这些均需要专业医生才能完成。目前标注过程以人工标注为主,费时费力且依赖医生的主观认知与判断,不同的医生还

可能会产生不一致的标注结果,造成诊断的不确定性与挑战性。因此,自动化的医学影像分割方法成为目前计算机视觉领域的研究热门。

医学影像的自动化分割,可分为基于边界的分割、基于阈值的分割、基于区域的分割、基于神经网络的分割等^[2]。基于边界的图像分割方法通过差分算子计算图像梯度,区分出图像的前景与背景,目前常用的差分算子包括 Sobel 算子^[3]、Prewitt 算子^[4]、Canny 算子^[5]等;基于阈值的分割方法通过某种方式选取一个或多个不同的阈值,利用阈值把图像二值化,以区分前景与背景,常用的阈值方法^[6]包括 P 参数法、

基金项目:国家自然科学基金(61802328);中国大学生创新创业项目(S202110530024)

This work was supported by the National Natural Science Foundation of China(61802328) and College Students' Innovation and Entrepreneurship Project in China(S202110530024).

通信作者:胡凯(kaihu@xtu.edu.cn)

Ostu 阈值法^[7]等;基于区域的分割方法根据某种相似性准则,对相同或相似性质的像素进行合并,常用的区域分割方法包括区域生长法^[8]与区域分裂合并法^[9]。但是医学影像的灰阶变化往往很小,前景与背景对比度低,且病灶与其他器官的区分并不明显,使分割问题面临挑战。医学影像中含有较多无关的斑点噪声,利用传统机器学习算法进行分割存在较大的局限性,难以对目标区域进行准确分割。

近年来,基于深度学习的分割方法逐渐成为主流。Long 等^[10]首次提出编码-解码结构模型 FCN,奠定了深度学习图像分割的基础。然而,FCN 的分割精度较差,且需要大量的训练数据。Ronneberger 等^[11]提出了跳级连接的 U 型网络模型 Unet,在少量数据集也具有较优秀的表现。在 Unet 的基础上,Zhou 等^[12]提出了密集嵌套连接的 Unet++ 模型,Huang 等^[13]提出了全连接的 Unet3+ 模型。然而,这些模型对于边缘和不连续区域的分割能力有限,在精细化分割中面临挑战。为了引入边缘的指导作用,Zhang 等^[14]提出 Edge-Attention-Network 网络模型,将边缘特征转移到高层特征的表达。为了精确定位边缘信息,Sun 等^[15]提出了 AS-UNet 网络模型,来增强边缘特征的注意力权重。针对分割目标非连续的区域,Valanarasu 等^[16]将 KiNet 与 Unet 结合,提出双分支的网络结构 KiU-Net。上述模型虽然在医学影像分割领域取得了不错的成绩,但也存在一定的不足。首先,现有的医学影像分割模型采用简单的拼接或相加的方式进行特征合并,一部分语义信息被忽视,这些语义信息可能对分割有着指导作用。其次,现有模型对于高低频语义信息的区分不明显,使用相同的方式进行学习。由于忽视了高频与低频信息所占比例不同的问题,高频边缘信息在分割过程中难以有效学习,影响了细节区域的分割效果。

针对以上问题,本文提出了基于边缘引导的多尺度医学影像分割模型 Edge Guided V-shape Network (EGV-Net)。模型从高低频信息两方面展开针对性的学习:首先,针对低频语义,利用编码-解码方式传递信息,学习分割目标的主体内容;其次,针对高频语义,本文提出新型的边缘提取方法,学习分割目标的边缘细节;再次,提出边缘引导模块,功能是将高频的边缘特征与低频的分割主体的特征进行融合,依据边缘空间掩膜指导模型做出精确的分割;最后,在肝脏数据集和 ISIC2016 数据集上与目前已有的医学影像分割方法展开对比,实验结果表明,本文提出的模型在目标不连续区域的分割更加全面,在分割细节处也有更为精确的表现,优于其他模型。

本文的贡献有以下几点:

(1)提出了基于边缘引导的 V 型卷积神经网络 EGV-Net,分离出分割目标的高频边缘特征并引导模型进行分割,用于解决肝脏影像与皮肤病影像的分割问题。

(2)提出多尺度的边缘提取与引导方法。在边缘提取过程中,本文通过边缘强化模块将高频语义信息从整体分割掩膜中分离,通过门控注意力模块对目标边缘进行过滤,得到分割边缘语义信息。在边缘引导过程中,利用高频边缘掩膜与低频分割主体协同优化,引导模型进行精准分割。

2 本文方法

本文方法整体流程如图 1 所示。首先对图像预处理,使用边缘提取算法获取分割目标的边缘,然后通过高效的医学影像分割模型 EGV-Net 训练样本,并使用组合损失函数进行

非均衡边缘与分割目标的学习。

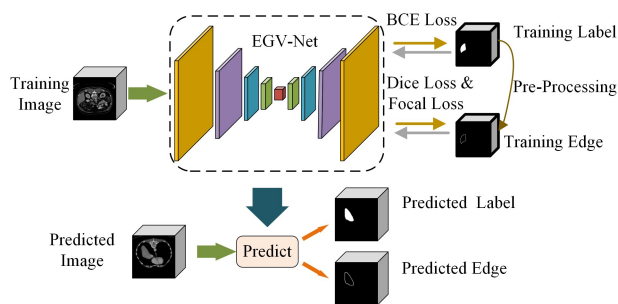


图 1 本文方法整体流程图

Fig. 1 Overall flow chart of our method

2.1 预处理

为了获得像素级别的边缘空间信息,本文通过边缘检测算法,提取分割标签的边缘轮廓。遍历分割标签的所有像素点,如果一个像素点的灰度值为 0,并且其 8 邻域内的值不全为 0,则将该点标记为边界点。对以此得到的所有边界点进行连通处理,保证每一个连通区域内的边界点都是连续的,从而得到光滑连续的边界。

2.2 网络模型

本文模型结构如图 2 所示,模型采用编码-解码结构,使用跳级连接传递编码层与解码层间的语义信息。其中,编码层由两个 3×3 的卷积层构成,后与一个残差结构相连接。另一方面,解码器使用连续的上采样操作复原分辨率,学习不同尺度的边缘特征语义。

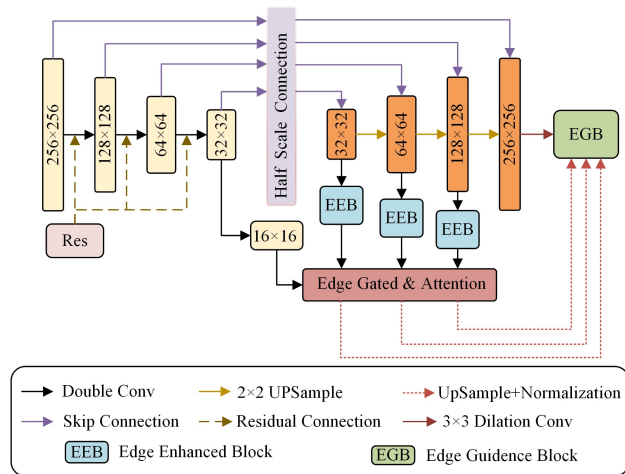


图 2 EGV-Net 模型的整体结构图

Fig. 2 Architecture of EGV-Net

本模型分为 3 个阶段:低频主体学习、高频边缘提取与边缘引导。低频主体通过编码-解码结构学习,高频边缘由边缘提取模块进行提取,最后利用边缘特征指导模型做出细致的分割。

下面对低频主体学习、高频边缘提取、边缘引导模块进行详细介绍。

2.2.1 低频主体学习

为了学习分割任务的低频主体区域,本文提出了多尺度的编码-解码结构,解码单元与所有浅层编码特征相融合,学习低频分割语义。

以第 3 层解码层为例,结构如图 3 所示。首先第 1 层和第 2 层编码元素经过 3×3 卷积层、最大池化层、 3×3 卷积层,得到与第 3 层编码层相同尺度的特征图谱。第 3 层编码

元素通过 3×3 的卷积层,第4层解码元素经过 3×3 的卷积层、尺度为2的双线性插值上采样操作、 3×3 的卷积层。对统一尺度的特征图谱进行 Concat 操作,再通过 3×3 的卷积层,得到特征结果。

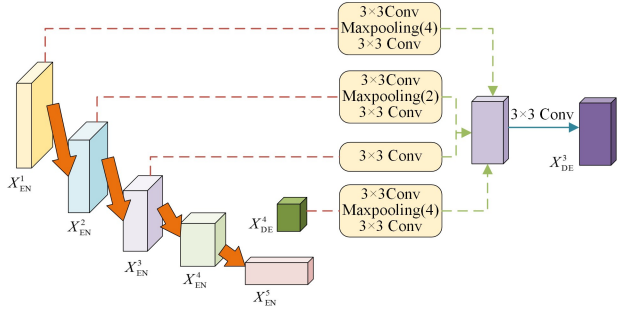


图3 第3层多尺度跳级连接的示例图

Fig. 3 Example of multi-scale skip connection of layer 3

2.2.2 高频边缘提取

医学影像的对比度与灰度范围较小,分割目标与背景的差异不明显且含有大量无关噪声,给边缘检测造成了困难。针对以上问题,本文提出一种边缘提取方法,该方法包含边缘强化模块与门控注意力模块。边缘强化模块用来划分边缘区域,将高频语义从整体特征中分离。门控注意力模块则进一步进行提取,强制滤除无关边缘与其他高频噪声,仅保留分割边缘的空间特征。下面进行详细介绍。

(1) 边缘强化模块

边缘强化模块如图4所示,对于第 k 层解码层元素 X_{De}^k ,首先进行最大池化下采样操作,与深层解码元素采样至相同尺度。对于所有的深层解码元素 $X_{De}^g (g=k+1, \dots, 5)$,以腐蚀方式内缩,得到低频特征向量。将原始特征与低频特征向量相消融,得到高频边缘特征。最后使用双线性插值上采样操作还原到原解码层 X_{De}^k 的尺度,得到高频处理结果 T_k^g :

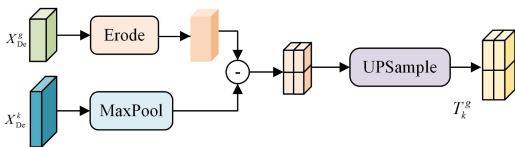
$$T_k^g = U_k^g(D_k^g(X_{De}^k) - \Theta(X_{De}^g)) \quad (1)$$

其中, $\Theta(\cdot)$ 表示特征内缩, $U_k^g(\cdot)$ 表示对第 k 层解码元素进行尺度为 g 的上采样操作, $D_k^g(\cdot)$ 表示对第 k 层解码元素进行尺度为 g 的下采样。

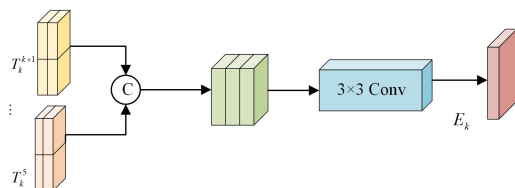
对以此得到的 $T_k^g (g=k+1, \dots, 5)$ 进行拼接,融合不同感受野的特征,再通过 3×3 卷积层得到边缘强化的结果 $E_k (k=2, 3, 4)$:

$$E_k = \tau^{3 \times 3}([\cdot T_k^g]), g=k+1, \dots, 5 \quad (2)$$

其中, $\tau^{3 \times 3}(\cdot)$ 表示 3×3 卷积、标准化、激活函数为 ReLU 的卷积操作, $[\cdot]$ 表示对所有特征进行 Concat 操作。



(a) 边缘降采样机制



(b) 多尺度边缘融合机制

图4 边缘强化模块结构图

Fig. 4 Edge enhancement module structure

(2) 门控注意力模块

模块结构如图5所示,通过门控模板 Gate feature 的门控注意力掩膜过滤输入特征 Input feature。本文将第5层编码层 X_5 作为 Input feature,其含有最大的感受野,对分割区域的整体把控能力最强。将含有丰富细节的边缘强化图谱作为 Gate feature,可对于特征过滤起到指导模板作用。首先对 Input feature 进行全局平均池化操作,再通过 3×3 卷积操作,最后使用 Sigmoid 函数得到该特征的注意力分布 α_k :

$$\alpha_k = \sigma(f^{3 \times 3}(gavpool(X_5))) \quad (3)$$

其中, $\sigma(\cdot)$ 表示 Sigmoid 函数, $f^{3 \times 3}(\cdot)$ 表示 3×3 的卷积操作, $gavpool(\cdot)$ 表示全局平均池化, X_5 为解码层第5层的特征向量。再对 Input feature 进行双线性插值上采样操作,得到与解码器第2,3,4层相同尺度的 G_5^k :

$$G_5^k = BN(f^{3 \times 3}(UP_k(X_5))) \quad (4)$$

其中, $BN(\cdot)$ 表示标准化, $UP_k(\cdot)$ 表示尺度为 k 的上采样操作, $f^{3 \times 3}(\cdot)$ 表示 3×3 卷积。对两特征向量相减,再经过 3×3 的卷积操作,通过 Sigmoid 函数得到边界的注意力分布 β_k :

$$\beta_k = \sigma(f^{3 \times 3}(Re(G_5^k - E_k))) \quad (5)$$

其中, $Re(\cdot)$ 表示 ReLU 激活函数。然后通过 GL 层进行特征流的输入限制,得到过滤去除噪声与无关特征信息的边缘特征 Ga_k :

$$Ga_k = \tau^{3 \times 3}((X_{De}^k \odot \alpha_k) \oplus (X_{De}^k \odot \beta_k) \oplus X_{De}^k) \quad (6)$$

其中, $\tau^{3 \times 3}(\cdot)$ 表示使用 3×3 卷积核,标准化,激活函数为 ReLU 的卷积层; \odot 表示特征向量间的点乘, \oplus 表示特征向量间的加法。

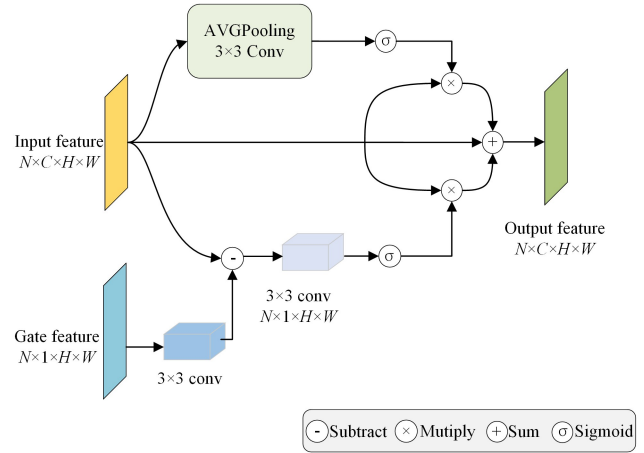


图5 门控注意力模块结构图

Fig. 5 Gated and attention module structure

由于边缘区域灰度范围小,因此本文采取深监督的方式进行训练,针对第2,3,4层的边缘特征,使用双线性插值方法还原分辨率,再通过两个 3×3 的卷积操作得到边界分割结果 $Edge_k$:

$$Edge_k = f^{1 \times 1}(\tau^{3 \times 3}(UP(Ga_k))) \quad (7)$$

其中, $f^{1 \times 1}(\cdot)$ 表示 1×1 卷积核, $\tau^{3 \times 3}(\cdot)$ 表示使用 3×3 卷积核、标准化、ReLU 激活函数的卷积层; $UP(\cdot)$ 表示双线性插值上采样方法。对于边缘提取结果,使用 $Dice\ loss^{[17]}$ 与 $Focal\ loss^{[18]}$ 组合损失函数进行训练。

2.2.3 边缘引导模块

为了在原有尺度下增大模型的感受野,并且恢复由于编码-解码结构缺失的边缘细节,本文提出边缘引导模块,模块

结构如图 6 所示。使用膨胀卷积替代池化操作,在不影响特征尺度的情况下增大感受野。

本文在 ASPP^[19]的基础上进行改进,在第一层解码器中使用膨胀率分别为 4,8,12 的膨胀卷积,且与高频边缘 $Edge_k$ 相结合,组成特征金字塔。利用多尺度的空间边缘掩膜与膨胀特征残差连接,将目标区域的关注度重新分布,恢复边缘细节并指导低频区域的分割。特征金字塔的公式表示如下:

$$M_k = \begin{cases} \tau^{3 \times 3}(X_{Dc}^k), & k=1 \\ \tau^{3 \times 3}(Re(P_k(X_{Dc}^k) \oplus Edge_k)), & k=2,3,4 \end{cases} \quad (8)$$

其中, $\tau^{3 \times 3}(\cdot)$ 表示含有 3×3 卷积、标准化、ReLU 激活函数的卷积层, $P(\cdot)$ 表示膨胀卷积, \oplus 表示特征向量加法, $Re(\cdot)$ 表示 ReLU 激活函数,得到边缘引导结果 M_k 。对所有的 M_k 进行拼接,经过两个 3×3 的卷积层获得最终的引导结果。

$$P = \mathcal{F}^{3 \times 3}([\mathcal{M}_k]), k=1, \dots, 4 \quad (9)$$

其中, $\mathcal{F}^{3 \times 3}(\cdot)$ 表示由两个 3×3 卷积层组成的双卷积操作, $[\cdot]$ 表示对所有特征进行 Concat 操作。

通过上述操作,边缘特征的空间位置能够指导模型在小范围内对目标区域进行分割学习,获取更多的分割细节。

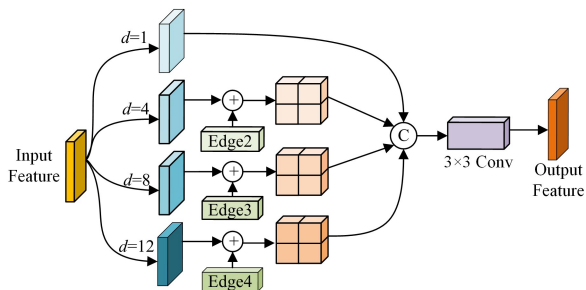


图 6 边缘引导模块结构图

Fig. 6 Edge guidance module

2.3 损失函数

本文模型采用深监督策略的学习方式,针对不同的训练目标使用不同的损失函数。具体来说,对于最终分割结果,使用 3×3 的卷积操作保持原通道数,再通过 1×1 的卷积进行降维,得到最终结果。在训练中,使用 $BCELoss$ 作为训练的损失函数。对于多尺度边缘的训练,首先通过双线性插值进行上采样操作,将特征的尺度恢复到原图像,再通过 3×3 的卷积操作, 1×1 的卷积操作,输出相应的边缘分割结果。针对样本特征分布非均衡的问题,本文使用 $Dice Loss$ ^[17] 和 $Focal Loss$ ^[18] 组合损失作为边缘训练的损失函数,公式表示如下:

$$L_{edge}^k = \omega_k * l_{focal}(p, y) + h_k * l_{dice}(p, y) \quad (10)$$

其中, ω_k 和 h_k 分别为两损失函数的权重,本文中均设置为 1。总体的损失函数如下:

$$L_{total} = \alpha L_{bce} + \beta \sum_{k=2}^4 L_{edge}^k \quad (11)$$

其中, α 和 β 为各损失的权重, L_{bce} 为 BCE 损失,在本实验中, α 设为 1, β 设为 0.2。

3 实验结果与分析

3.1 实验数据集与评估指标

3.1.1 数据集

本文使用的数据集为专家进行标注的肝脏 CT 数据集与 ISIC2016 皮肤数据集^[20]。肝脏数据集包含 420 张 CT 影像

和标注结果,本实验将其中 340 幅影像用于模型的训练,60 幅用于模型的验证,20 幅用于模型的测试。皮肤数据集由 1279 幅影像及标注结果组成,本实验将其中 800 幅影像用作训练集,100 幅用作验证集,379 幅用作测试集。

3.1.2 评估指标

本文使用语义分割中常用的评价指标进行分割质量的评估,分别为 Dice 系数、平均准确度 (ACC)、灵敏度 (SEN)、IOU 系数。评估指标计算式如下:

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (12)$$

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \quad (13)$$

$$SEN = \frac{TP}{TP + FN} \quad (14)$$

$$IOU = \frac{TP}{TP + FN + FP} \quad (15)$$

其中, TP 代表真实值为正,预测值为正的个数; FP 表示真实值为负,预测值为正的个数; FN 表示真实值为正,预测值为负的个数; TN 表示真实值为负,预测值为负的个数。

3.2 实验环境

本文利用深度学习开源框架 Pytorch 进行网络模型的搭建与训练。首先将肝脏影像经过窗宽与窗位的调整,再插值调整为 256×256 的分辨率;对于分割标签,首先通过预处理获取分割边缘,再将边缘图像与标签图像经过归一化操作,作为模型的训练真实值。本文在 NVIDIA TITAN Xp 上使用自适应梯度优化 (Adam) 算法进行训练,训练迭代的次数设置为 50 次,批处理数设置为 4,学习率为 0.001。

3.3 与其他算法的对比

3.3.1 在肝脏数据集上的对比结果

将 EGV-Net 模型与 Unet^[11], Unet++^[12], Unet3+^[13], KiU-Net^[16], ResUnet^[21] 等先进图像分割模型在肝脏数据集上进行实验比较,实验结果如表 1 所列。

表 1 不同网络模型在肝脏数据集的分割性能

Table 1 Segmentation performance of different models on liver dataset

Model	Params/($\times 10^6$)	Acc/%	Dice/%	Sen/%	Iou/%
Unet ^[11]	118.42	98.95	93.63	91.62	88.10
Unet++ ^[12]	34.51	99.02	94.12	90.70	89.02
Unet3+ ^[13]	21.11	99.08	94.53	90.72	89.67
KiU-Net ^[16]	139.25	99.01	94.04	90.71	88.76
ResUnet ^[21]	124.41	98.98	94.44	90.80	89.53
Ours	35.05	99.27	96.10	95.02	92.53

本文模型在以上 4 种评估指标上均优于现有模型。具体来说,网络模型的平均像素精度、Dice 系数、Sensitivity 系数、IOU 系数分别为 99.27%, 96.10%, 95.02%, 92.53%。其相较于 Unet 模型,分别提升了 0.32%, 2.47%, 3.40%, 4.43%;相较于 Unet++ 模型,分别提升了 0.25%, 1.98%, 4.32%, 3.51%;相较于 Unet3+ 模型,分别提升了 0.19%, 1.57%, 4.30%, 2.86%;相较于 KiU-Net 模型,分别提升了 0.26%, 2.06%, 4.31%, 3.77%;相较于 Res-Unet 模型,分别提升了 0.29%, 1.66%, 4.22%, 3.00%。对比结果如图 7 所示,可以看出,本文模型在分割细节处表现较好,由于边界的引导作用,模型在分割过程中较好地表达出分割目标的整体特征。

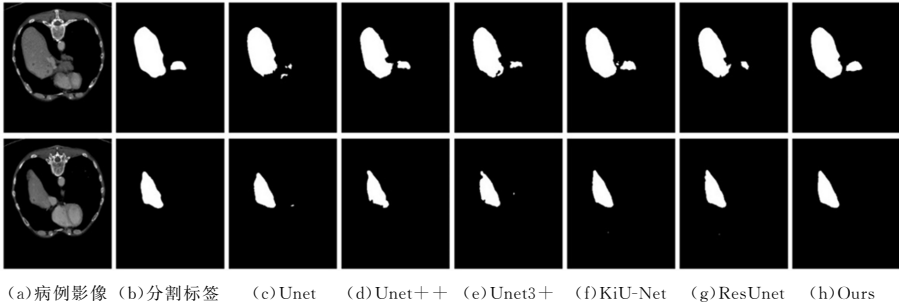


图 7 不同模型在肝脏数据集的对比结果

Fig. 7 Comparison of segmentation performance of different models on liver dataset

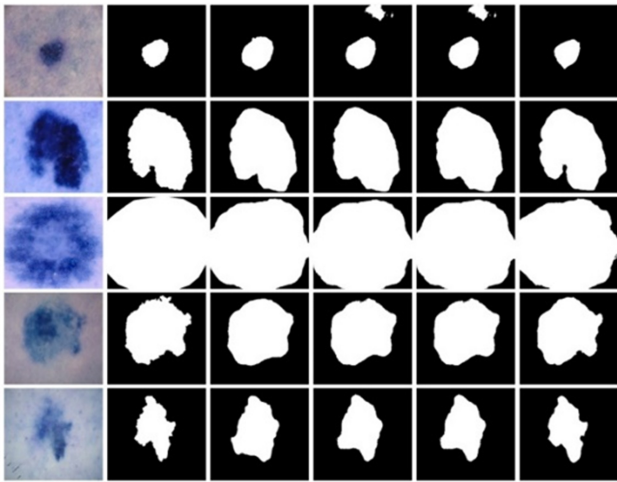
3.3.2 在 ISIC2016 数据集的对比结果

为了验证 EGV-Net 在不同数据集的分割情况,我们在 ISIC2016 数据集进行进一步的测试。ISIC2016 是由 ISIC 组织的皮肤病黑色素分割数据集,由患者的病例图像与医生的实际标注组成。该数据集由 900 张训练图像与 379 张测试图像组成,每张图像分辨率为 718×542 到 4288×2848 不等。在进行训练与测试时,本文将图像分辨率调整到 256×256 ,进行归一化操作,将训练集与分割标签的像素值调整为 0 到 1 之间,并对部分训练图像进行水平翻转。使用不同的模型进行测试,训练次数为 200 次,学习率为 0.0001,结果如表 2 所列,分割结果如图 8 所示。在不同模型的分割对比中,本模型在 ISIC2016 数据集中也取得了较为先进的结果。相较于其他网络模型,EGV-Net 可以较好地提取出边界处细节,即使在边缘稍模糊的医学影像中也有较为精确的分割效果,在分割细节处也有较好的分割表现。

表 2 不同网络模型在 ISIC2016 的分割性能

Table 2 segmentation performance of different models on ISIC2016

Model	Params/($\times 10^6$)	Acc/%	Dice/%	Sen/%	Iou/%
Unet ^[11]	118.42	93.56	89.29	89.45	81.42
Unet++ ^[12]	34.51	94.64	89.41	92.06	82.32
Unet3+ ^[13]	21.11	94.97	90.08	91.35	83.30
Ours	35.05	95.08	90.64	90.90	83.99



(a) 病例影像 (b) 分割标签 (c) Unet (d) Unet++ (e) Unet3+ (f) Ours

图 8 模型在 ISIC2016 数据集上的结果

Fig. 8 Performance of different models on ISIC2016

3.3.3 多尺度编码-解码跳级连接性能分析

为了验证本文方法的性能,我们仅修改编码器与解码间的信息传递结构,分别使用 Skip connection, Nested connec-

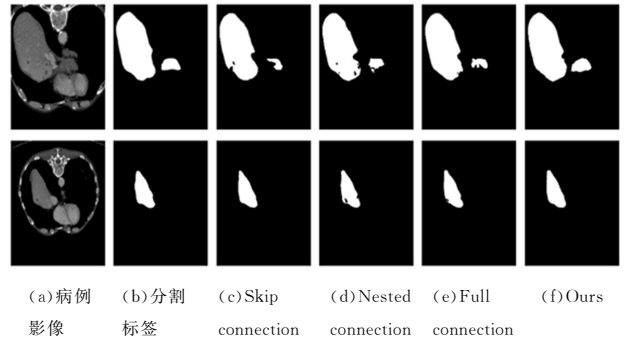
tion, Full connection 和本文方法进行对比,结果如表 3 所列。

表 3 特征连接方式的性能对比

Table 3 Performance comparison of connection methods

Model	Params($\times 10^6$)	Acc/%	Dice/%	Sen/%	Iou/%
Skip -connection	226.44	99.15	95.26	91.62	90.99
Nested -connection	34.99	99.05	94.44	90.82	89.45
Full -connection	77.69	99.20	95.64	92.66	91.66
Ours	35.05	99.27	96.10	95.02	92.53

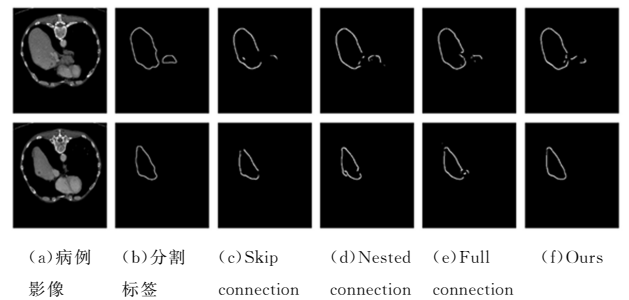
通过改变编码器与解码器间的信息传递的结构,来验证本文方法的有效性。本文提出的编码-解码连接方式在以上指标中均处于先进水平,优于其他算法。分割对比结果如图 9 所示,边缘对比结果如图 10 所示。从对比结果中可以看出,本文方法可以更准确地提取出边缘空间位置信息,更好地指导网络进行精准细致的分割。



(a) 病例影像 (b) 分割标签 (c) Skip connection (d) Nested connection (e) Full connection (f) Ours

图 9 连接方式的分割对比结果

Fig. 9 Segmentation comparison of connection methods



(a) 病例影像 (b) 分割标签 (c) Skip connection (d) Nested connection (e) Full connection (f) Ours

图 10 连接方式的边缘对比结果

Fig. 10 Edge comparison of connection methods

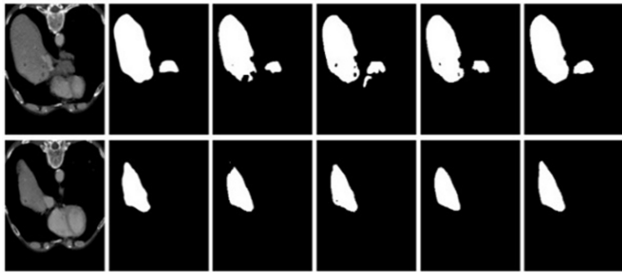
3.3.4 门控注意力模块性能分析

为了探究门控注意力模块对分割性能的影响,本文将门控注意力模块替换为卷积模块、注意力机制 SE 模块^[22]、门控 GSCNN 模块^[23]进行实验,实验结果如表 4、图 11 所示。

表4 门控注意力模块的性能对比

Table 4 Performance comparison of gate-attention modules

Model	Acc/%	Dice/%	Sen/%	Iou/%
Conv	99.11	95.13	91.99	90.79
SE Block ^[22]	99.18	95.49	94.08	91.39
GSCNN ^[23]	99.21	95.60	93.06	91.61
Ours	99.27	96.10	95.02	92.53



(a)病例影像 (b)分割标签 (c)Conv (d)SE Block (e)GSCNN (f)Ours

图11 门控注意力模块的对比结果

Fig. 11 Comparison of gate-attention module

从各指标可以看出,更换为不同的模块后分割精度都有降低。本文方法对边界更加敏感,能提取出更为准确的整体轮廓,分割效果更优。

结束语 本文提出了一种高效的医学影像分割模型,依据高低频语义在图像的不同作用,在分割的不同阶段进行针对性的学习。其中,低频语义通过编码-解码方式传递信息,分割出目标的大致内容;高频语义通过边缘提取方法进行提取,首先利用边缘强化模块分离出高频语义并增强其空间表现力,然后通过门控注意力模块过滤出分割目标的边缘特征。高频语义与低频语义通过边缘引导模块进行残差连接,模型受到高频边缘的指导作用,能获得更精准的分割结果。

本文方法仍具有较大的进步空间,模型在处理分割区域不连续的目标时,会出现断点或分割边缘缺失的情况。考虑不连续目标区域的分割,提升高频边缘信息的分割精度,使用多任务学习策略,都是我们未来可能的研究方向。

参考文献

[1] ZHENG G Y, LIU X B, HAN G H. A review of computer-aided detection and diagnosis systems for medical imaging[J]. Journal of Software, 2018, 29(5): 1471-1514.

[2] LIN Y, TIAN J. A review of medical image segmentation methods[J]. Pattern Recognition and Artificial Intelligence, 2002, 15(2): 192-204.

[3] CHIEN Y. Pattern classification and scene analysis [J]. IEEE Transactions on Automatic Control, 1974, 19(4): 462-463.

[4] PREWITT J M S. Object enhancement and extraction[J]. Picture Processing and Psychopictorics, 1970, 10(1): 15-19.

[5] CANNY J. A computational approach to edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986(6): 679-698.

[6] SIVAKUMAR V, MURUGESH V. A brief study of image segmentation using thresholding technique on a noisy image[C]//

International Conference on Information Communication and Embedded Systems(ICICES2014). IEEE, 2014: 1-6.

[7] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66.

[8] LEVINE B F, BETHEA C G, THURMOND C D, et al. An organic crystal with an exceptionally large optical second-harmonic coefficient: 2-methyl-4-nitroaniline[J]. Journal of Applied Physics, 1979, 50(4): 2523-2527.

[9] PITMAN J. Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition[J]. Combinatorics, Probability and Computing, 2002, 11(5): 501-514.

[10] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440.

[11] 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.

[12] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKSHI N, et al. Unet++: A nested u-net architecture for medical image segmentation[M]// Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3-11.

[13] HUANG H, LIN L, TONG R, et al. Unet 3+: A full-scale connected unet for medical image segmentation[C]// ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 1055-1059.

[14] ZHANG Z, FU H, DAI H, et al. Et-net: A generic edge-attention guidance network for medical image segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2019: 442-450.

[15] SUN J M, GE Q Q, LI X M, et al. A medical image segmentation network with edge enhancement feature[J]. Journal of Electronics & Information Technology, 2022, 44: 1-10.

[16] VALANARASU J M J, SINDAGI V A, HACIHALILOGLU I, et al. Kiu-net: Overcomplete convolutional architectures for biomedical image and volumetric segmentation[J]. IEEE Transactions on Medical Imaging, 2021, 41(4): 965-976.

[17] MILLETARI F, NAVAB N, AHMADI S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation[C]// 2016 fourth international conference on 3D vision (3DV). IEEE, 2016: 565-571.

[18] LIN T Y, GOYAL P, GIRHICK R, et al. Focal loss for dense object detection[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.

[19] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.

[20] GUTMAN D, CODELLA N, CELEBI E, et al. Skin lesion analysis toward melanoma detection: A challenge at the international

symposium on biomedical imaging (ISBI) 2016[C]// Hosted by the International Skin Imaging Collaboration (ISIC). 2016.

- [21] XIAO X, LIAN S, LUO Z, et al. Weighted res-unet for high-quality retina vessel segmentation[C]// 2018 9th International Conference on Information Technology in Medicine and Education(ITME). IEEE, 2018:327-331.
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:7132-7141.
- [23] TAKIKAWA T, ACUNA D, JAMPANI V, et al. Gated-scnn: Gated shape cnns for semantic segmentation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019:5229-5238.



JIANG Haotian, born in 2001, undergraduate. His main research interests include deep learning and medical image processing.



HU Kai, born in 1984, Ph.D, associate professor, is a member of China Computer Federation. His main research interests include machine learning, pattern recognition, bioinformatics, and medical image processing.