

基于上下文特征感知和双频上采样的食管早癌图像分割

孟延宗¹, 李小霞^{1,2}, 周颖玥^{1,2}, 文黎明³, 秦佳敏³, 刘爽利^{1,2}

1. 西南科技大学信息与工程学院, 四川 绵阳 621000; 2. 特殊环境机器人技术四川省重点实验室, 四川 绵阳 621000; 3. 四川绵阳四〇四医院消化内科, 四川 绵阳 621000

【摘要】目的:针对食管早癌图像分割过程中病灶边缘等细节信息丢失的问题,在U-net基础上提出一种基于上下文特征感知和双频上采样的食管早癌图像分割网络。**方法:**利用注意力机制和可分离空洞卷积改进上下文特征感知模块,获取全文上下文信息,提取更多特征细节。提出双频上采样模块,分别从高频和低频进行上采样,有效减少单一上采样因像素插值产生的锯齿效应和转置卷积造成的棋盘效应,减少细节信息的丢失。**结果:**本文方法的平均交并比、敏感度和特异性分别达到80.34%、87.47%和91.53%。**结论:**本文模型优于nnU-Net等主流语义分割模型,保留更多的细节信息,提高食管早癌图像分割精度。

【关键词】食管早癌;上下文特征感知;注意力机制;空洞卷积;双频上采样

【中图分类号】R318;R735.1

【文献标志码】A

【文章编号】1005-202X(2023)08-0957-07

Early esophageal cancer image segmentation based on contextual feature awareness and dual frequency upsampling

MENG Yanzong¹, LI Xiaoxia^{1,2}, ZHOU Yingyue^{1,2}, WEN Liming³, QIN Jiamin³, LIU Shuangli^{1,2}

1. School of Information Engineering, Southwest University of Science and Technology, Mianyang 621000, China; 2. Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, Mianyang 621000, China; 3. Department of Gastroenterology, Sichuan Mianyang 404 Hospital, Mianyang 621000, China

Abstract: Objective To propose a network for early esophageal cancer image segmentation using U-net with contextual feature awareness module and dual frequency upsampling module which solves the problem of loss of detailed information such as lesion edges during image segmentation. **Methods** The contextual feature awareness module improved with the attention mechanism and separable dilated convolution was used to obtain full-text contextual information and extract more feature details. The dual frequency upsampling module was adopted for upsampling from high frequency and low frequency, thereby effectively reducing the aliasing effect caused by pixel interpolation, minimizing the checkerboard effect caused by transposed convolution during single upsampling, and avoiding the loss of detail information. **Results** The mean intersection over union, sensitivity and specificity of the proposed method reached 80.34%, 87.47%, and 91.53%, respectively. **Conclusion** The proposed model is superior to mainstream semantic segmentation models such as nnU-Net for it can retain more detailed information and improve the accuracy of early esophageal cancer image segmentation.

Keywords: early esophageal cancer; contextual feature awareness; attention mechanism; dilated convolution; dual frequency upsampling

前言

【收稿日期】2023-02-11

【基金项目】国家自然科学基金(62071399);四川省科技计划重点研发项目(2021YFG0383, 2023YFG0262)

【作者简介】孟延宗,硕士研究生,研究方向:医学图像语义分割,E-mail: 953590977@qq.com

【通信作者】周颖玥,副研究员,博士,研究方向:图像处理与分析,E-mail: 147256027@qq.com

食管癌属于消化系统的恶性肿瘤,是全球第六大癌症,对人类的健康造成严重的威胁。我国是食管癌高发国家之一,发病率和死亡率都显著高于世界平均水平。对于食管癌,早期手术切除治疗后5年生存率可达90%,而当病情拖延至中晚期时,手术切除治疗后的5年生存率仅有6%~15%^[1]。由于食管癌早期无明显症状,临床上采用内镜检查联合病理组织学活检作为诊断早期食管癌的“金标准”。限于基层医疗机构的设备陈旧、内镜医师缺乏临床诊断经

验等原因,早期食管癌的阳性检出率较低。因此,有必要研究一种基于内镜图像的食管早癌分割方法,通过智能化的手段辅助医生诊断以提高对食管癌的早期检出率和工作效率。

2015年,Ronneberger等基于卷积神经网络(Convolutional Neural Networks, CNN)提出的U-Net拉开了针对医学图像进行语义分割算法研究的序幕,U-Net采用完美的对称结构和跳跃连接,有效地融合低级别和高级别的图像特征,解决普通CNN在医学图像上定位不准确的问题。U-Net不仅赢得了2015年细胞追踪挑战赛的冠军,而且转变了学者们的思维方式,激发了许多有意义的改进,并且在医学图像语义分割领域取得了更加优异的成果^[2-3]。近年来,基于深度学习的人工智能方法在医学各个领域取得显著的进展,尤其是作为一种辅助病灶筛查手段,这些领域包括放射肿瘤学诊断^[4]、皮肤癌分类^[5]、糖尿病视网膜病变分割^[6]、肺结节病变分割^[7]以及使用内窥镜对大肠病变^[8]进行表征。在食管早癌筛查中,基于深度学习的分割方法能有效捕获图像的低层细节特征和高层语义特征,对背景复杂的食管图像分割具有一定的优势。Xue等^[9]通过微血管形态学类型分类进行食管早癌检测,在Caffe中开发一个模型,使用进行特征提取,支持向量机进行分类,开创使用深度学习方法进行食管早癌筛查的先河。Hong等^[10]使用CNN来区分胃生化、肠生化和胃肿瘤,该体系结构由4个卷积层、2个最大池化层和2个完全连接层组成,分类准确率为80.77%。Liu等^[11]通过改进DeepLabV3+网络的特征提取模块,增强深度信息提取能力,在食管癌前病变图像分割上取得了74.43%的Dice系数(Dice Coefficient)。2019年,Luo等^[12]通过样本率可信区间估计法(Clopper-Pearson)鉴别癌性病变方面的诊断准确性,在5个外部验证集

中,诊断准确率为91.5%~97.7%。2021年,Yang等^[13]提出一种基于深度学习模型的食管癌诊断算法,利用6473张经过专业医师标记过的食管窄带内镜图像(Narrow Band Imaging, NBI),通过CNN模型SegNet提取图像特征,构建敏感度和特异性均超过90%的早筛模型。然而,食管早癌类型有着极大的个体差异,病变区域大小和外观也不尽相同,这使得上述技术和方法对食管内镜图像的识别准确率不尽如人意。

在使用U型网络对食管图像进行语义分割过程中,虽然下采样会使语义特征更加明显,但这也导致网络无法完全保留病灶边缘等细节特征,这些细节特征对后续食管早癌病灶的判断具有重要意义。本研究使用文献[14]中的通道和空间双注意力机制,结合可分离空洞卷积(Seperable Dilated Convolution, SDC)改进上下文特征感知模块,防止细节信息的丢失;同时,提出双频上采样模块用于网络解码器,从高频、低频两条支路同时恢复特征图分辨率,维持不同像素间的相对位置,进而保留更多的细节信息,提高食管早癌的分割精度。

1 方法与步骤

本研究提出的网络如图1所示。在编码阶段使用上下文特征感知模块(Contextual Feature Awareness Module, CFAM),利用通道、空间双注意力机制和卷积引导多尺度特征融合,用以感知全局上下文信息,再通过SDC扩大感受野提取更多细节特征^[15]。在解码阶段使用双频上采样模块(Dual Frequency Upsampling Module, DFUM)代替U型网络中的单一上采样,从频率角度减少细节信息和语义信息在上采样过程中的损失,增强网络分割能力。

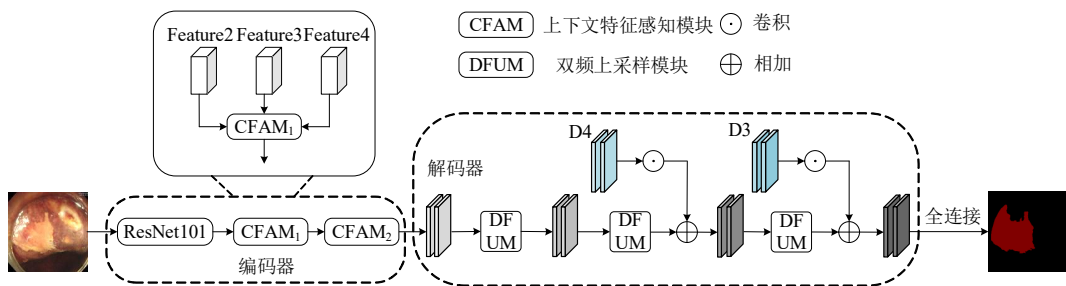


图1 本文提出的网络结构

Figure 1 Proposed network structure

1.1 CFAM

由于食管病灶目标较小,在内镜图像中呈现边界等细节模糊,影响对病灶的精准分割。因此本研

究利用双注意力机制和卷积引导多尺度特征融合,并结合SDC构造CFAM,其结构如图2所示。

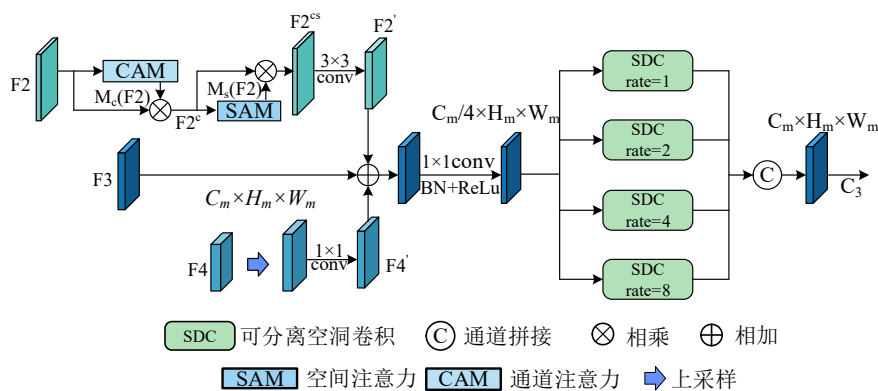


图2 上下文特征感知模块

Figure 2 Contextual feature awareness module

首先在细节信息最丰富的初始层引入通道注意力机制(Channel Attention Mechanism, CAM)和空间注意力机制(Spatial Attention Mechanism, SAM),便于感知病灶边缘特征表达;然后通过卷积和上采样的方式将多阶段特征图引导至中间层特征图,再通过相加增强特征信息交融能力,获取图像语义上下文;最后并联SDC,维持计算量的同时扩大感受野,获取图像细节信息。

CFAM以2、3、4阶段为例,编码器Stage2产生的特征映射 $F2 \in R^{C^2 \times H_2 \times W_2}$ 首先经过CAM,获得通道注意力权重 $M_c(F2)$,将 $F2$ 与 $M_c(F2)$ 相乘得到 $F2^c \in R^{C^2 \times H_2 \times W_2}$,完成通道注意力对原始特征 $F2$ 的重标定,生成的 $F2^c$ 再经过SAM得到空间注意力 $M_s(F2)$,最后执行 $F2^c$ 与 $M_s(F2)$ 的相乘操作得到 $F2^{cs}$,在空间和通道两个维度上完成病灶边缘等细节特征在特征图 $F2$ 上的重标定。

为了同时维持编码器提取到的语义信息和细节信息,选用中间层作为信息引入过渡层,将得到的 $F2^{cs}$ 和 $F4$ 特征映射到与中间层 $F3$ 相同维度大小 $R^{C_m \times H_m \times W_m}$,通过相加进行跨通道信息交融,获取全局上下文语义信息。但编码器获得的所有特征并不能全部有效用于分割,使用 1×1 卷积和归一化操作将特征图降维至 $\frac{C_m}{4} \times H_m \times W_m$,这样的操作既能在后续

的每条支路获取分割所需的精确特征,捕获局部信息,又可以维持总通道数 C_m 不变。对降维后的特征图并联卷积核为 3×3 空洞率为 1、2、4、8 的 SDC,维持计算量同时扩大感受野,使网络能在大感受野中准确定位并提取细节特征,最后通过通道拼接升维至 C_m 输出,弥补空洞卷积采样稀疏造成的远距离相关性的损失。CFAM₁ 计算过程如下:

$$C_3 = \text{Cat}(\underset{\text{rate} = 1, 2, 4, 8}{\text{Conv}}_{\text{DS}}(\text{Conv}_{1 \times 1}[F'_2 \oplus F_3 \oplus F'_4])) \quad (1)$$

其中, F'_2 、 F_3 、 F'_4 分别表示不同阶段的特征映射,特

征图大小均为 $C_m \times H_m \times W_m$, $\text{Conv}_{1 \times 1}$ 表示 1×1 卷积, Conv_{DS} 表示空洞率分别为 1、2、4、8 的 SDC, $\text{Cat}_{\text{rate}=1,2,4,8}$ 表示拼接操作, C_3 为 CFAM₁ 输出。

为了利用所有阶段产生的多尺度特征,使模型能感知语义上下文和细节上下文,网络分别在编码器的2、3、4阶段和3、4、5阶段使用CFAM,即CFAM₁和CFAM₂。

1.2 DFUM

语义分割网络在上采样阶段常使用插值法、转置卷积^[16]、亚像素卷积(Sub-Pixel)^[17]获得高分辨率特征图。但对于病灶图像语义分割任务,需要在上采样阶段同时维持图像的高频信息(病灶前景和背景灰度差大的边缘信息)和低频信息(图像灰度缓慢变化的病灶整体信息)。单独使用转置卷积或亚像素卷积会因采样率设置过大引入棋盘效应,造成低频信息丢失;而单独使用插值法则会因随机改变像素大小引入锯齿效应,影响病灶图像边缘的识别,造成图像高频信息丢失。

针对上述问题,本研究提出 DFUM,该模块将在高频和低频维度获取网络所需特征。如图3所示,低频支路首先通过 1×1 卷积扩大输入特征图通道数,使特征维度不会因亚像素卷积改变,之后使用亚像素卷积增大分辨率。该方法先通过卷积的方式得到 r^2 个通道的特征图(r 为图像扩大倍率);再通过周期筛选的方式将低分辨率图像中的每一个像素的 r^2 个通道展开成 $r \times r$ 的大像素重组在一起,所以通道数 $r^2 C$ 缩减成 C ,图像尺寸(H, W)扩展成($r \times H, r \times W$),得到低频输出 $F1$ 。

Odena 等^[18]证实即便叠加使用转置卷积也会造成棋盘效应,不利于特征融合和图像重建,因此高频支路采用双线性插值迭加卷积组的方式获取图像高频信息。如图 3 所示,高频支路首先进行双线性插值增大图像分辨率,此时输出特征图为 F_d ,接着将 F_d 送

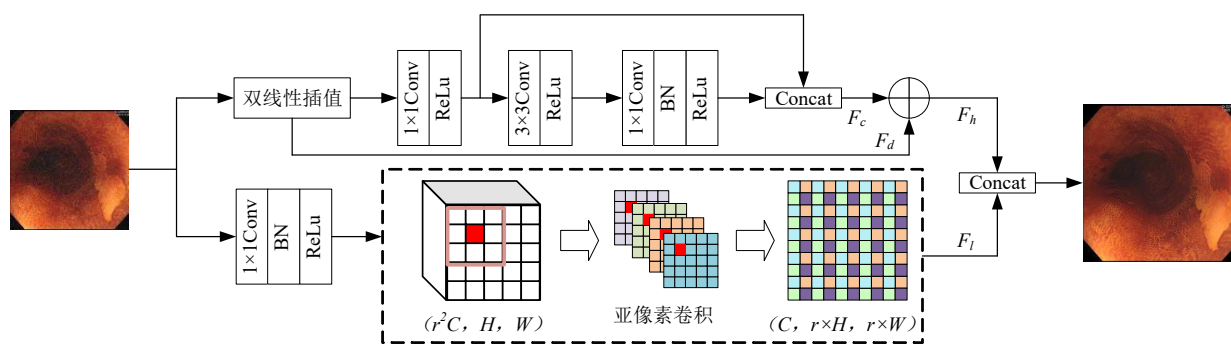


图3 双频上采样模块

Figure 3 Dual frequency upsampling module

入后续的卷积层进行细节特征提取,对卷积层得到的二维特征图进行通道拼接,强化细节特征的表达,得到与特征图 F_d 大小相同的二维特征向量 F_c 。最后将 F_c 和 F_d 相加得到输出 F_h ,即高频信息输出,该模块最后的输出如下:

$$F_{out} = F_1 + F_h \quad (2)$$

DFUM从高频、低频两条支路上获取病灶边缘等细节特征和病灶概貌等语义特征,减少上采样过程中空间上下文信息的损失,使网络生成更清晰的高分辨率特征图,提升分割精度。

2 实验内容

2.1 数据集与数据处理

国内专家推荐使用白光内镜检查联合卢戈氏液染色内镜(Lugol Chromoendoscopy, LCE)及NBI图像作为食管癌内镜筛查的首选方法^[19]。基于此,本研究使用白光内镜图像、LCE图像和NBI内镜图像联合组成食管早癌数据集(图4)。食管早癌数据集共3 503张(白光图像783张,NBI图像791张,碘染色图像1 929张),由四川绵阳四0四医院提供,所有病例均经组织学证实。数据均为医生在做胃镜检查时采集,具有随机性、一般性和真实性。数据经过医院的消化科专家按照内镜报告确定病灶区域进行标记,再由本团队通过Lableme软件对数据进行精细标注工作,将图像裁剪为256×256后进行图像预处理。

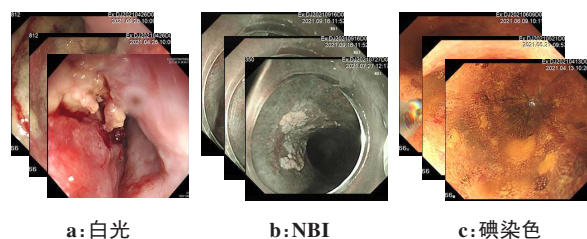


图4 自建食管早癌数据集

Figure 4 Self-built early esophageal cancer data set

为减少反光和低对比度对模型的影响,加快模型训练效率,增强模型的泛化能力,针对食管内镜图像本身的特点进行如下预处理:(1)使用随机水平翻转和随机剪切,使食管早癌出现在不同位置,减轻模型对病灶出现位置的依赖性;(2)将RGB图像转化为HSV图像;(3)为减少反光和强光对内镜图像的影响,对步骤(2)处理后的图像进行亮度和对比度的随机调整(0.8~1.2),降低模型对高亮和低对比度的敏感度;(4)使用标准正态分布的方法进行数据标准化,加快模型收敛。

2.2 评价指标

在医学图像辅助诊断中,常用敏感度(Sensitivity, Se)和特异性(Specificity, Sp)分别表示漏诊率和误诊率,指标越高,漏诊率和误诊率越低。在图像分割领域,常用平均交并比(Mean Intersection over Union, MIoU)表示模型的像素级分割精度,F1-Score综合表示精确度和召回率。本研究使用敏感度、特异性、MIoU、F1-Score这4类指标评价分割算法的性能,具体计算公式如下:

$$Se = \frac{TP}{TP + FN} \quad (3)$$

$$Sp = \frac{TN}{TN + FP} \quad (4)$$

$$MIoU = \frac{TP}{TP + FP + FN} \quad (5)$$

$$F1-Score = \frac{2 \times Pr \times Se}{Pr + Se} \quad (6)$$

其中,TP为真阳性样本个数,TN为真阴性样本个数,FP为假阳性样本个数,FN为假阴性样本个数。Pr为精确率指标,计算公式如下:

$$Pr = \frac{TP}{TP + FP} \quad (7)$$

2.3 实验参数设置

实验依托Pytorch平台搭建神经网络,版本为1.8.0,Python版本为3.6.5。实验平台配置如下:GPU型号为NVIDIA GTX3080,显存大小为8 GB。

训练策略如下:将数据集按照7:2:1划分训练集、验证集和测试集,训练过程中使用SGD作为优化器,初始学习率设置为 0.5×10^{-3} ,并在前10个训练轮次中使用预热学习^[20]来加快模型收敛。每迭代1轮保存1次网络模型,共迭代300轮,保存测试结果最好的模型。因食管早癌数据正负样本分布不平衡,选用Focal loss^[21]作为损失函数,增加难分正样本在损失函数中的权重,提高总体正样本的分割准确度。 p_i 反映了与真实值接近程度, p_i 越大说明越接近真实值,即分类越准确, γ 为可调节因子,取值在0到1之间, L_{fl} 表示Focal loss,计算方式如下:

$$L_{fl} = -(1 - p_i)^{\gamma} \log(p_i)$$

(8)

3 实验结果与分析

3.1 消融实验结果与分析

为验证本研究提出的CFAM和DFUM对食管早癌分割任务的有效性,使用MIoU、敏感度和特异性评价在自建数据集上进行消融实验,对比基础模型添加两个模块前后的分割性能。消融实验结果如表1所示,从实验结果可以看出,CFAM对模型分割效果提升明显,DFUM也在一定程度上提升了分割性能。本文算法较基础网络的MIoU、敏感度和特异性分别提升4.8%、7.24%和2.82%。

3.2 实验结果对比与分析

为验证本文方法的有效性,将所提出模型与其他分割模型进行对比实验^[22-24]。对比实验采用相同的优化算法、损失函数和超参数设置。实验结果见表2。由表2可以看出,本文算法在MIoU、敏感度、特异性和F1-Score中均表现最佳。除了量化的实验数据,本研究还使用Grad-CAM^[25]可视化特征热力图对结果进行定性分析,图5的Grad-CAM显示本文算法与其他方法相比能更好地覆盖目标病灶,表明本文算法能更好地完成病灶区域分割任务。

本研究选择表2中除本文网络外最好的网络模型nnU-Net(图6第3行)以及华西医院近期研究成果“深度学习辅助早期食管癌诊断模型共享平台”(http://huaxi-inno.innovsight.com)(图6第4行)与本文模型(图6第5

表1 自建数据集上的消融实验结果(%)

Table 1 Ablation experiment results on self-built dataset (%)

实验	CFAM	DFUM	MIoU	敏感度	特异性
1	×	×	75.54	80.23	88.71
2	√	×	78.57	85.63	90.27
3	×	√	77.10	83.92	89.33
4	√	√	80.34	87.47	91.53

表2 主流算法在自建数据集上的实验结果(%)

Table 2 Experimental results of mainstream algorithms on self-built dataset (%)

方法	MIoU	敏感度	特异性	F1-Score
TransU-Net ^[22]	78.21	86.36	90.87	87.40
DeepLabV3+ ^[23]	77.55	82.60	89.52	85.91
SegNet* ^[10]	76.82	84.45	88.95	85.62
nnU-Net ^[24]	80.08	87.05	91.23	88.82
本文算法	80.34	87.47	91.53	90.60

行)进行分割结果可视化,在随机挑选的5张图片中可以看出本文算法的分割结果更接近真实值(图6第2行)。

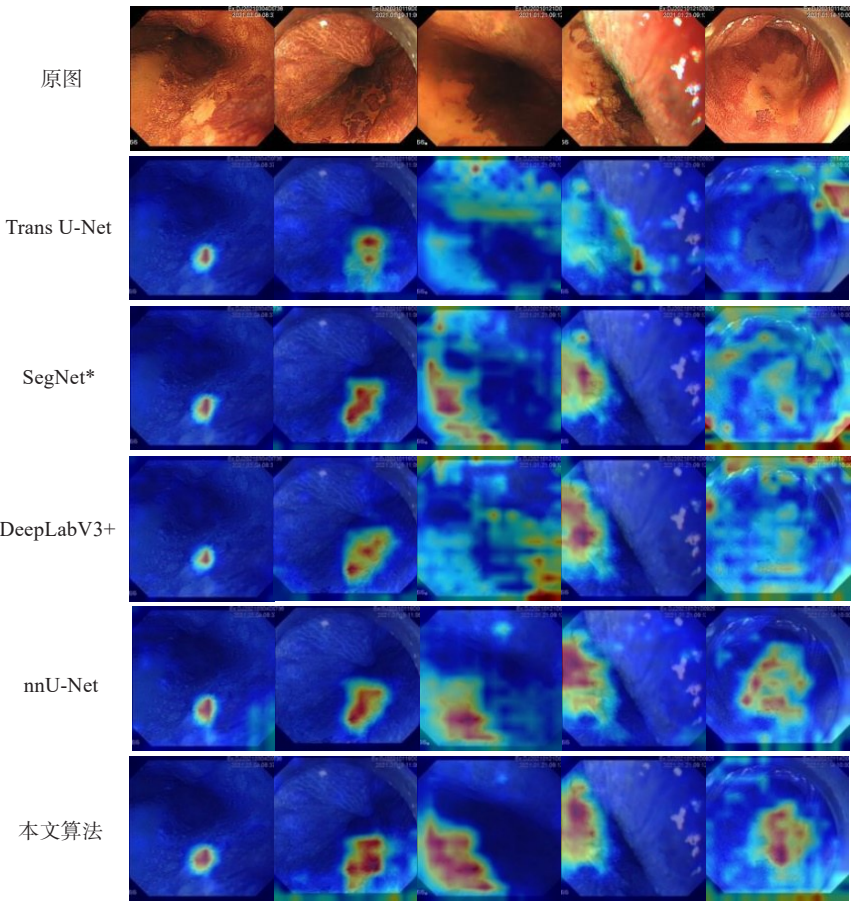


图5 热力图

Figure 5 Heat maps

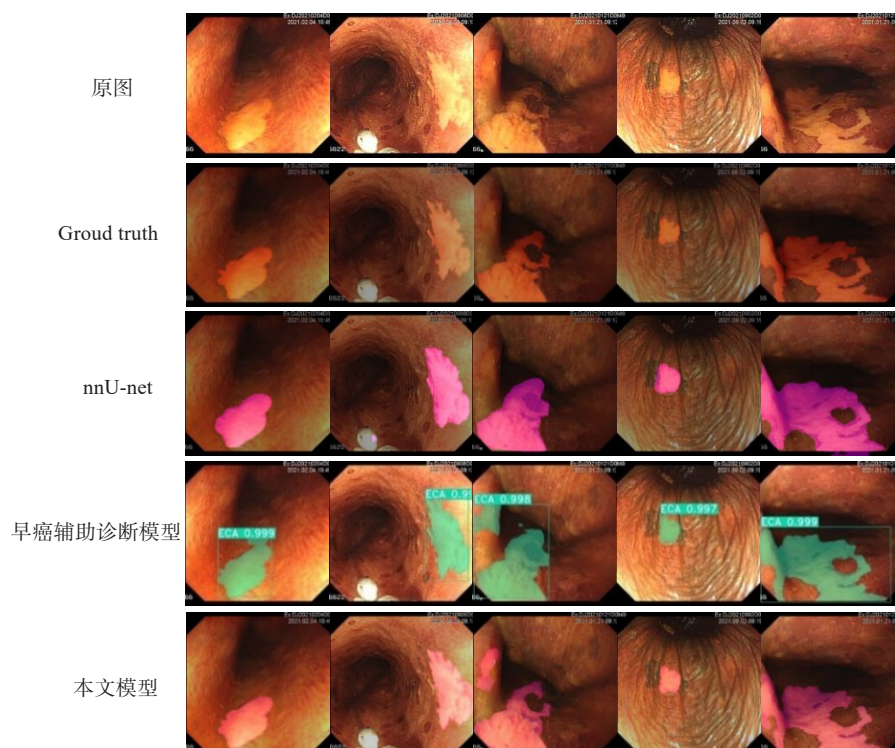


图6 分割结果

Figure 6 Segmentation results

4 结论

针对食管早癌图像分割过程中病灶边缘等细节特征丢失的问题,本研究提出一种基于CFAM和DFUM的食管图像语义分割方法。利用双注意力机制和SDC搭建CFAM,获取细节上下文,并采用DFUM有效结合特征图中低频信息和高频信息来恢复图像分辨率,减少细节信息的丢失。实验结果表明本研究所提出网络模型在MIoU、敏感度和特异性上较其他网络都有不同程度的提高,满足目前内镜检查对食管早癌的分割要求。

【参考文献】

- [1] 李国仁,戴建华.我国早期食管癌筛查的研究进展[J].中华胸心血管外科杂志,2021,37(1):52-58.
Li GR, Dai JH. Research progress of early esophageal cancer screening in China [J]. Chinese Journal of Thoracic and Cardiovascular Surgery, 2021, 37(1): 52-58.
- [2] Ibtihaz N, Rahman MS. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Networks, 2020, 121: 74-87.
- [3] Guo S, Sun N, Pei Y, et al. 3D-UNet-LSTM: a deep learning-based radar echo extrapolation model for convective nowcasting [J]. Remote Sensing, 2023, 15(6): 1529.
- [4] Huang D, Bai H, Wang L, et al. The application and development of deep learning in radiotherapy: a systematic review [J]. Technol Cancer Res Treat, 2021, 20: 15330338211016386.
- [5] Brinker TJ, Hekler A, Enk AH, et al. Deep neural networks are superior to dermatologists in melanoma image classification [J]. Eur J Cancer B Oral Oncol, 2019, 119: 11-17.
- [6] 吕念祖,李小霞,肖娟,等.结合注意力和条件生成对抗网络的视网膜血管分割[J].小型微型计算机系统,2021,42(6):1288-1292.

- Lü NZ, Li XX, Xiao J, et al. Segmentation of retinal blood vessels combining attention and conditional generative adversarial network [J]. Journal of Chinese Mini-Micro Computer Systems, 2021, 42(6): 1288-1292.
- [7] 江悦莹,施一萍,翁晓俊,等.融合Vnet和边缘特征的肺结节分割算法[J].中国医学物理学杂志,2022,39(6):705-712.
Jiang YY, Shi YP, Weng XJ, et al. Lung nodule segmentation algorithm fused with Vnet and edge features [J]. Chinese Journal of Medical Physics, 2022, 39(6): 705-712.
- [8] Rao H, Trieu JA, Nair P, et al. Artificial intelligence in endoscopy: more than what meets the eye in screening colonoscopy and endosonographic evaluation of pancreatic lesions [J]. Gastrointest Endosc, 2022, 3(3): 16-30.
- [9] Xue DX, Zhang R, Feng H, et al. CNN-SVM for microvascular morphological type recognition with data augmentation [J]. J Med Biol Eng, 2016, 36(6): 755-764.
- [10] Hong J, Park BY, Park H. Convolutional neural network classifier for distinguishing Barrett's esophagus and neoplasia endomicroscopy images [C]//2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017: 2892-2895.
- [11] Liu D, Jiang H, Rao N, et al. Depth information-based automatic annotation of early esophageal cancers in gastroscopic images using deep learning techniques [J]. IEEE Access, 2020, 8: 97907-97919.
- [12] Luo H, Xu G, Li C, et al. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study [J]. Lancet Oncol, 2019, 20(12): 1645-1654.
- [13] Yang H, Hu B. Recent advances in early esophageal cancer: diagnosis and treatment based on endoscopy [J]. Postgrad Med, 2021, 133(6): 665-673.
- [14] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module [C]//Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.
- [15] Park H, Yoo Y, Seo G, et al. Concentrated-comprehensive convolutions for lightweight semantic segmentation [J]. ArXiv: 1812.04920, 2018.
- [16] 卢梁,曾晓天,董国昭,等.血流动力学响应函数在脑疾病中的研究进展[J].中国医学物理学杂志,2022,39(4):479-483.
Lu L, Zeng XT, Dong GZ, et al. Advances in hemodynamic response function in brain disease research [J]. Chinese Journal of Medical

- Physics, 2022, 39(4): 479-483.
- [17] Shi W, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1874-1883.
- [18] Odena A, Dumoulin V, Olah C. Deconvolution and checkerboard artifacts[J]. Distill, 2016, 1(10): e3.
- [19] 国家消化系统疾病临床医学研究中心(上海), 中华医学会消化内镜学分会, 中国医师协会内镜医师分会消化内镜专业委员会, 等. 中国食管鳞癌前状态及癌前病变诊治策略专家共识[J]. 中华消化内镜杂志, 2020, 37(12): 853-867.
- National Clinical Research Center for Digestive System Diseases (Shanghai), Digestive Endoscopy Branch of Chinese Medical Association, Digestive Endoscopy Professional Committee of Endoscopy Branch of Chinese Medical Doctor Association, et al. Chinese expert consensus on diagnosis and treatment of precancerous conditions and lesions of esophageal squamous cell carcinoma[J]. Chinese Journal of Digestive Endoscopy, 2020, 37(12): 853-867.
- [20] Xiong R, Yang Y, He D, et al. On layer normalization in the transformer architecture[C]//International Conference on Machine Learning. PMLR, 2020: 10524-10533.
- [21] Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection [C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2980-2988.
- [22] Chen J, Lu Y, Yu Q, et al. Transunet: transformers make strong encoders for medical image segmentation[J]. Arxiv. 2102.04306, 2021.
- [23] Chen LC, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]// Proceedings of the European Conference on Computer Vision. 2018: 801-818.
- [24] Isensee F, Jaeger PF, Kohl SA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature Methods, 2021, 18(2): 203-211.
- [25] Selvaraju RR, Cogswell M, Das A, et al. Grad-cam: visual explanations from deep networks *via* gradient-based localization[C]//Proceedings of the IEEE international Conference on Computer Vision. 2017: 618-626.

(编辑:谭斯允)