

基于多尺度特征融合与反向注意力的COVID-19病灶分割

李碧草^{1,2}, 王晶¹, 郭旭伟³, 黄杰¹, 魏苗苗¹, 李盼盼¹

1. 中原工学院电子信息学院, 河南 郑州 450007; 2. 郑州大学信息工程学院, 河南 郑州 450001; 3. 河南科技大学第一附属医院儿科, 河南 洛阳 471000

【摘要】针对新型冠状病毒肺炎(COVID-19)分割问题中感染区域具有高变异性以及病灶与背景对比度低等问题,提出一种基于多尺度特征融合与反向注意力的COVID-19感染分割网络。首先,利用残差网络作为主干网络进行特征提取,并使用全局上下文聚合策略对不同层次特征进行融合得到粗略的分割结果;其次,在网络瓶颈处添加多尺度特征融合模块,利用空洞卷积与多核池化增强网络分割不同尺度病变的能力;最后,设计一种级联结构的反向注意力模块,利用互补区域的细节特征增强背景与目标的对比度。本文方法在COVID-19 CT分割测试集上的准确率、特异性、灵敏度分别达到0.714、0.700和0.958,误检和漏检区域明显减少,细小病灶的分割能力显著提升。

【关键词】COVID-19;肺炎感染分割;全局上下文聚合;多尺度特征融合;反向注意力

【中图分类号】R318;TP391

【文献标志码】A

【文章编号】1005-202X(2023)04-0403-07

COVID-19 lesion segmentation based on multi-scale feature fusion and reverse attention

LI Bicao^{1,2}, WANG Jing¹, GUO Xuwei³, HUANG Jie¹, WEI Miaomiao¹, LI Panpan¹

1. School of Electronic and Information, Zhongyuan University of Technology, Zhengzhou 450007, China; 2. School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China; 3. Department of Pediatrics, the First Affiliated Hospital of Henan University of Science and Technology, Luoyang 471000, China

Abstract: A COVID-19 lesion segmentation network based on multi-scale feature fusion and reverse attention (MFFRA) is proposed to overcome the problems of high variability and low contrast between lesion and background in COVID-19 segmentation. The residual network is used as the backbone network to extract features, and the global context aggregation strategy is adopted to integrate different hierarchical features for obtaining rough segmentation results. In addition, the multi-scale feature fusion module is added at the bottleneck of the network to enable the ability to segment lesions of different scales using atrous convolutions and multi-kernel pooling. Finally, a novel cascaded reverse attention module is designed to improve the low contrast between normal tissue and lesions based on the detailed features of complementary regions. The proposed method has an accuracy, specificity and sensitivity of 0.714, 0.700, 0.958 on the COVID-19 CT test set, reduces the areas of misdetection and missed detection, and enhances the segmentation ability of fine lesions.

Keywords: COVID-19; lung lesion segmentation; global context aggregation; multi-scale feature fusion; reverse attention

前言

新型冠状病毒于2020年初在全球传播,导致全球面临一场新的健康危机。计算机辅助诊断(Computer-Aided Diagnostic, CAD)技术对于抗击新

冠疫情有重要的意义,其中,病灶分割可以观察到患者在不同时期的感染状况,不仅有助于新冠的前期诊断,还有助于后续的治疗。计算机断层扫描(CT)影像技术是一项在医学领域广泛应用的技术。从CT切片上可以观察到典型的感染症状,新型冠状病毒肺炎(COVID-19)早期呈现出毛玻璃形状的混浊,晚期则表现为肺实变。通过对CT切片感染和纵向变化的定性评价,可为病情的诊疗提供重要信息^[1-2]。然而,人工描绘肺部感染是一项繁琐和耗时的工作,此外,放射科医生的感染注释是一项高度主观的任务,经常受到临床经验的影响。因此COVID-19病灶的自动化分割越来越受到人们的重视。

【收稿日期】2022-12-26

【基金项目】国家自然科学基金(61901537);河南省高校科技创新人才支持计划(23HASTIT030);河南省留学人员科研择优项目;中国博士后科学基金(2020M672274);中原工学院学科实力提升青年硕博培育计划(SD202207)

【作者简介】李碧草,博士,副教授,硕士生导师,主要研究方向:人工智能与医学图像处理,E-mail: lbc@zut.edu.cn

近来,针对COVID-19病灶分割的研究取得了显著进展^[3-6]。然而,肺炎感染在纹理、大小和位置上的高度变化给检测带来了极大的挑战。例如,肺实变区域很小,容易导致整个CT切片的假阴性检测。此外,COVID-19感染区域和正常组织间往往对比度较低,边缘模糊,难以准确分割。

Fan等^[7]提出的Inf-Net利用并行解码器模块获取全局信息,同时利用注意力模块提高对感染区域的识别能力,从而有效分割肺炎感染区域,但没有关注到病灶分割所需要的多尺度信息。LRSAR-Net采用一种低秩张量自注意力重构模块用于获取长范围的信息,与传统的自注意力模块相比计算复杂度更低。但对比度低的感染区域无法检测的问题依然存在^[8]。多任务学习架构近来被引入到分割任务,利用相关分类等任务提高分割准确性^[9-10]。Wang等^[9]提出了一种有效的多任务网络DeepSC-COVID,分别构建用于特征提取的跨任务特征子网、用于病变分割的病变子网和用于疾病诊断的分类子网,能够同时生成肺炎/非肺炎的二分类结果与感染区域的分割结果。Amyar等^[10]提出了基于编解码器体系结构和多任务学习的深度学习模型,利用不同的数据集联合执行3个学习任务:分割、分类和重建,从而利用包含在多个相关任务中的有用信息来提高分割性能。但该类方法需要大批量训练数据,且没有关注病灶的细节特征。

为改善现有卷积神经网络模型在肺炎图像分割中存在感受野小及细节特征不敏感等问题,本文提出一种基于多尺度特征融合与注意力机制的分割网络,实现COVID-19感染区域的自动化分割。本文贡

献如下:(1)设计一种反向注意力模块,增强背景与目标对比度,使模型更加关注感染区域,减少漏检与误检问题;(2)设计多尺度融合模块,利用空洞卷积与多核池化扩大网络感受野,增强对不同尺度目标的分割能力;(3)使用由粗到细的分割策略,利用全局上下文聚合策略,通过融合所有低层与高层特征生成全局信息与粗略的分割预测;(4)在公开的COVID-19分割数据集上取得了较好的分割结果^[11],漏检与误检现象减少,分割精度明显提升。

1 网络结构

1.1 网络构造

笔者使用了由粗到细的分割策略,即先通过全局上下文的聚合生成粗略的预测分割,再使用多尺度特征融合模块与注意力机制依次进行细化。首先,在特征提取部分,利用ResNet作为主干网络提取不同层次的特征^[12],并使用全局聚合策略对所有特征进行融合得到粗略的分割结果。其次,在网络的瓶颈处添加了多尺度特征融合模块,使网络能够更好地捕捉多尺度特征,增强网络分割不同尺度病变的能力。最后,为了解决正常组织与病灶对比度较低的问题,笔者提出了一种级联结构的反向注意力模块,利用显著性目标检测原理,从输入的高层特征中删除现有的不精确的感染区域,得到互补区域的细节特征,进而增强背景与目标的对比度。其中现有的不完整感染区域首先来自于之前的低高层特征聚合,后续的输入则依赖于上一级级联模块的输出。模型的总体结构如图1所示。

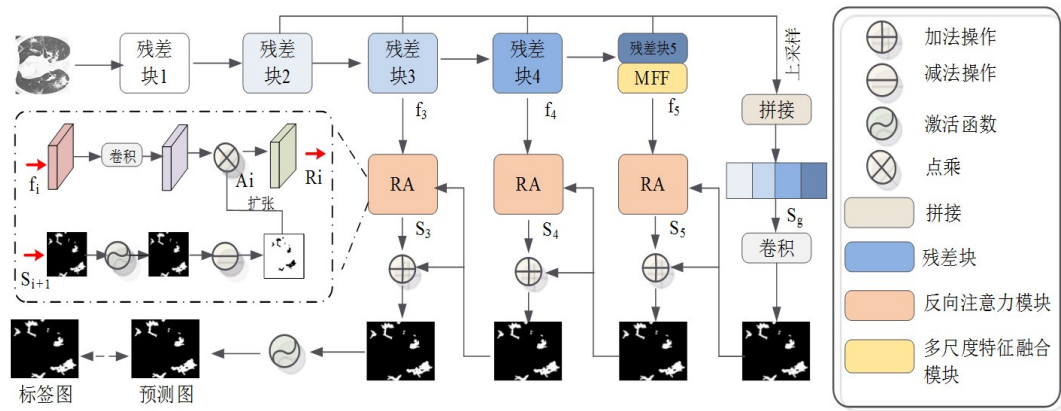


图1 网络整体架构

Figure 1 Overall architecture of the network

1.2 反向注意力模块

ResNet作为主干网络,采用短连接机制来避免网络很深时梯度爆炸和消失的问题,在特征提取上

具有一定的优势。但是普通的残差结构不能在特征提取时很好地关注到感兴趣区域,这会导致模型的误检或漏检。现有的改进大多是增加每个残差块的

卷积层或通道数。网络宽度或深度加倍可能带来一定的精度改善,但这种方式的缺点是复杂度会比较高,而目前计算能力依然是限制神经网络发展的瓶颈。

注意力机制可以抑制非相关区域,以较小的计算量为代价,有效解决病灶与正常组织对比度低的问题。已有研究通常是单独应用空间注意力或通道注意力,或将两者进行并行或串行组合^[13-16]。受Inf-Net启发^[7],笔者设计了一个新的反向注意力模块,利用减法操作从输入特征中删除预测的感染区域,得到互补区域的细节特征,再将其与原始特征图结合,从而增强背景与目标的对比度。

具体来说,首先将输入特征 $\{f_i, i=3, 4, 5\}$ 送入卷积块提取深层特征,再与权重 A_i 进行点乘操作得到输出特征 R_i ,即:

$$R_i = C(F_i) \otimes A_i \quad (1)$$

其中, C 表示卷积操作。受计算机视觉中显著性目标检测的启发,权重 A_i 被定义为:

$$A_i = \varepsilon(S(\sigma(P(S_{i+1})))) \quad (2)$$

其中, P 表示上采样操作, σ 为sigmoid激活函数, S 表示从单位矩阵 E 中减去输入特征的减法操作。符号 ε 表示通道扩展,目的是将反向特征扩展到与原始特征一致。

此外,反向注意力模块是以级联的方式进行的。如图1所示,每级的输入 S_{i+1} 都依赖于上一级模块的输出。最后一级反向注意力模块的输出 S_3 ,经过一个sigmoid激活函数后得到最终预测的肺部感染区域。本文设计的反向注意力模块,能够有效增强背景与病灶的对比度,最终将不精确和粗糙的预测细化为精确和完整的预测图。反向注意力模块的实现如图1左边所示。

1.3 全局上下文聚合策略

由粗到细分割是一种有效的分割方法,可以解决图像扫描范围差异较大、内存占用不足的问题。具体来说,首先使用一个粗糙的全局映射来定位目标,以获得感兴趣区域,然后实现基于感兴趣区域的精细分割。其中,生成

粗略的全局信息尤为重要。

在大多数现有的医学图像分割网络中,通常利用编码器来融合所有的高级和低级特征,迫使网络关注感兴趣的区域^[17-20]。受此启发,笔者利用低级和高级特征的聚合来捕获全局上下文信息。如图1所示,首先将所有特征使用拼接操作进行聚合,再送入一个 1×1 卷积使其与原始特征的维度相匹配,最后将其送入多尺度特征融合模块的第一层级联结构。

1.4 多尺度特征融合模块

在医学图像分析场景中,器官和病变总是表现出多尺度的特点,即使同一类对象的大小往往也有着巨大的差异,比如前期和中晚期的肿瘤图像会呈现尺寸与外观的高度变异性。通过空洞卷积扩大感受野是解决这一问题的常见方法。感受野的变化有利于不同尺度目标的提取。通常,较大感受野的卷积有利于提取较大的目标,而较小感受野的卷积则更有利于提取小目标。通过结合不同速率的空洞卷积,能够提取不同大小目标的特征,解决当前网络感受野固定的问题,进而增强网络对多尺度目标的分割能力。

Deeplabv2受金字塔池化的启发,将4种不同采样率的空洞卷积并行输出^[21]。Deeplabv3在其基础上加入了图像级特征,即先在模型的特征映射上应用全局平均池化,再将其送入并行的空洞卷积块^[22]。CE-Net通过采用不同大小的池化对多尺度上下文特征进行编码,而不需要额外的学习权重^[23]。受以上文献的启发,笔者设计了一个全新的多尺度特征融合模块。与以前的工作不同之处在于该模块主要由两个部分组成:空洞卷积块和多核池化块。如图2所示,在空洞卷积块部分,首先将输入特征送入4个采样率分别为 $\{1, 6, 12, 18\}$ 、卷积核大小为 3×3 的空洞卷积提取多尺度信息,同时添加一个全局平均池化操作获得全局特征,然后经过拼接操作将所得信息融合,并送入接下来的多核池化块。

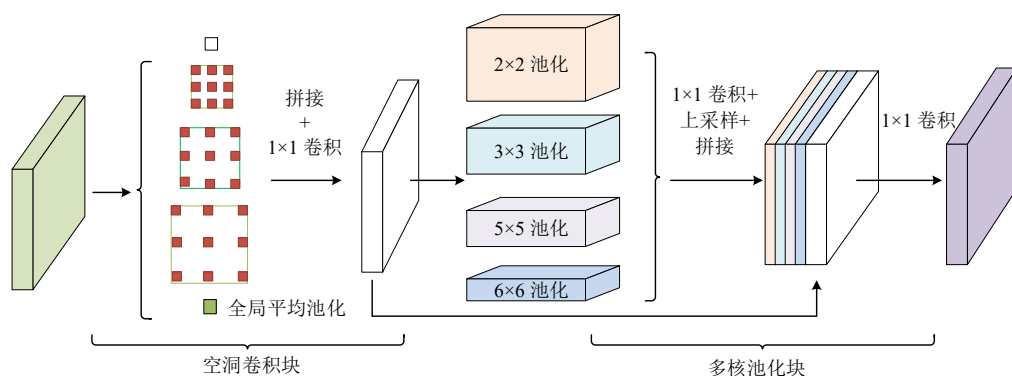


图2 多尺度特征融合模块

Figure 2 Multi-scale feature fusion module

在多核池化块部分,首先将其送入4个并行的核尺寸不同的池化层。使用不同视野大小的核同样有助于检测不同尺寸的目标。将池化后的输出送入 1×1 卷积后得到通道数相同大小不同的4张特征图,然后用双线性插值将4张特征图扩展到相同分辨率。最后使用残差结构,将其与原始特征图拼接在一起,送到 1×1 卷积后得到最终输出。

1.5 损失函数

损失函数由3部分组成。首先,用标准的二值交叉熵损失测量由低层特征生成的边缘映射 S_e 与由真实标签图得到的边缘映射 G_e 之间的差异,将其定义为边缘损失,公式如下:

$$L_{\text{edge}} = -\sum_x \sum_y^w \left[G_e \log(S_e) + (1 - G_e) \log(1 - S_e) \right] \quad (3)$$

其中, (x, y) 为每个像素在预测的边缘映射和真实的边缘映射中的坐标。 G_e 由真实标签图的梯度计算得到, S_e 由低层特征经过卷积层后生成, w 和 h 分别表示相应的宽度和高度。

其次,将损失函数 L_{seg} 定义为交并比损失 L_{IoU} 和加权二值交叉熵损失 L_{BCE} 的组合,公式如下:

$$L_{\text{seg}} = L_{\text{IoU}}^w + \lambda L_{\text{BCE}}^w \quad (4)$$

其中, λ 是权重,在实验中设为1。与在分割任务中广泛采用的标准二值交叉熵损失不同,加权二值交叉熵损失增加了像素的权重,以突出不同像素的重要性。

最后,将3个高层特征(S_3 、 S_4 和 S_5)和全局特征 S_g 向上采样到与真值图相同的大小,并计算其损失。因此,总损失可推广到:

$$L_{\text{total}} = L_{\text{seg}}(G_s, S_g^{\text{up}}) + L_{\text{edge}} + \sum_{i=3}^5 L_{\text{seg}}(G_s, S_i^{\text{up}}) \quad (5)$$

2 结果

2.1 数据集与实验设置

本文基于公开的COVID-19分割数据集进行实验^[11]。所有CT图像均由意大利医学和介入放射学学会收集。放射科医师使用不同的标签对CT图像进行分段,以识别肺部感染。该数据集由来自不同COVID-19患者的100张轴向CT图像组成,50例数据用于模型训练与验证,50例用于测试,数据划分方式和Inf-Net一致^[7]。在训练过程中,所有图片统一裁剪为 256×256 。此外,将原数据集中两种不同类型的肺部感染合并为炎症病变区域,作为分割目标,如图3所示。

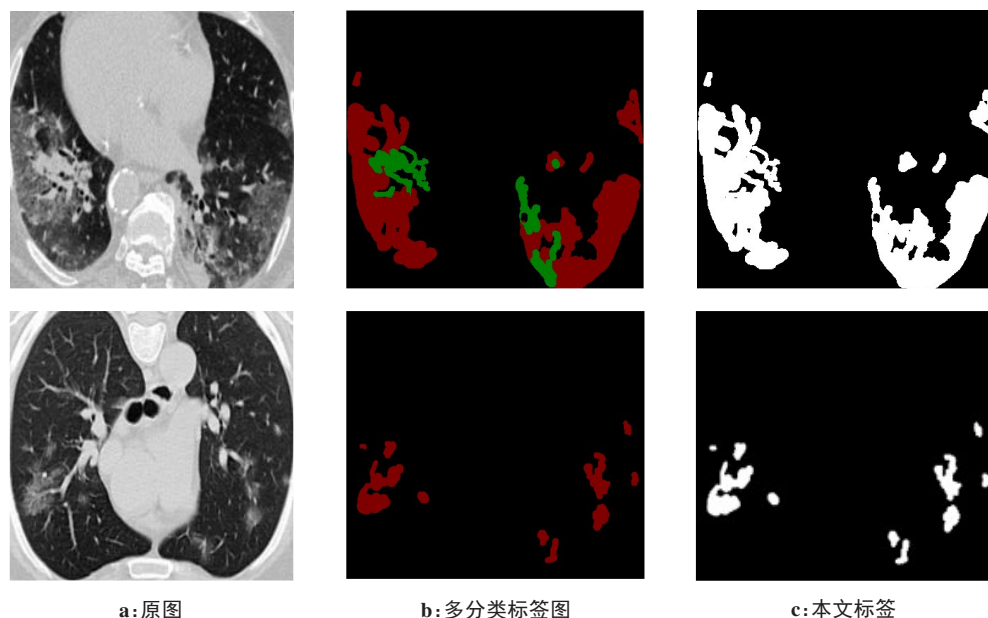


图3 新冠肺炎分割数据集中不同感染程度肺炎患者的CT轴向切片,红色和绿色分别代表毛玻璃影和肺结节区域
Figure 3 Varying degrees of CT axial sections from people infected with COVID-19 on COVID-19 segmentation dataset, where the red and green masks denote the ground-glass opacity and consolidation, respectively

本文主干网络使用基于ImageNet预训练的ResNet50,利用公共的PyTorch平台实现。训练和测试使用12 G内存的NVIDIA-GTX1080 GPU显卡。在训练过程中,采用Adam算法优化模型^[24],初始学

习率设为 $1e^{-4}$,权重衰减系数为0.001,训练轮次为100。

2.2 评价指标

本实验使用了6个被广泛采用的指标,除了分割领域中常见的Dice系数、敏感度、特异性,还增加了

结构度量 (Structure Measure, SM)、配准增强度量 (Enhanced-Alignment Measure, EM) 和平均绝对误差 (Mean Absolute Error, MAE) 3 个指标作为最终的评估结果。

SM 用于评估模型预测结果和真实标签之间的相似性, 公式如下:

$$SM = (1 - \alpha) \times S_o(S_p, G) + S_r(S_p, G)$$

(6)

EM 可以同时度量两个二值矩阵局部特征和整体特征的相似性, 公式如下:

$$EM = \frac{1}{w \times h} \sum_x \sum_y [S_p(x, y), G(x, y)]$$

(7)

MAE 测量了模型的预测结果和真实分割之间的误差, 公式如下:

$$MAE = \frac{1}{w \times h} \sum_x \sum_y |S_p(x, y) - G(x, y)|$$

(8)

2.3 对比实验

为了比较本文分割网络的性能, 笔者将其与 5 种经典的模型 UNet^[25]、UNet++^[26]、Att-UNet^[27]、Gate-UNet^[28]、Inf-Net^[7] 进行对比实验。定量结果见表 1。可以看出, 基于同样的数据集和同样的测量方法, 本文所提方法在 Dice 系数、敏感度和特异性等方面都优于 UNet 和 Inf-Net 等经典模型。例如, 本文方法的 Dice 系数与特异性分别达到了 0.714 与 0.958, 明显优于 Inf-Net 的 0.682 与 0.943。笔者将性能的提升归因于反向注意力模块、多尺度特征融合模块与全局聚合策略。本文所设计模块不仅增强了感染区域与正常组织间的对比度, 有利于捕获与任务相关的特征, 并且分别对多尺度信息与全局信息进行聚合, 增强了网络对不同尺度病灶的分割能力。

表 1 不同模型在 COVID-19 CT 测试集上的评价指标对比
Table 1 Diagnostic performances of different models on COVID-19 CT segmentation test set

模型	Dice	敏感度	特异性	SM	EM	MAE
UNet ^[25]	0.439	0.534	0.858	0.622	0.625	0.186
UNet++ ^[26]	0.581	0.672	0.902	0.722	0.720	0.120
Att-UNet ^[27]	0.583	0.637	0.921	0.744	0.739	0.112
Gate-UNet ^[28]	0.623	0.623	0.623	0.623	0.623	0.623
Inf-Net ^[7]	0.682	0.692	0.943	0.781	0.838	0.082
本文方法	0.714	0.700	0.958	0.795	0.869	0.069

为了更直观地进行对比, 笔者将不同网络的分割图进行可视化, 结果如图 4 所示。可以看到, 本文方法明显优于基线方法。从整体效果来看本文方法产生的分割结果更加接近真实标签, 具有更准确的

边界, 而其他方法给出了相对模糊的边界, 特别是在细小的感染区域; 从结果的第 2、3 行可以看出, 肺部感染漏检区域和误检区域相对减少, 相比之下, 其他经典模型和基线方法存在大量的错误分割区域。这证明了本文模型的有效性。

2.4 消融实验

本节探究了网络中不同模块对于分割精度的影响, 以证明模块的有效性。用模块一代表反向注意力模块, 模块二代表多尺度特征融合模块, 模块三代表全局上下文融合策略。如表 2 所示, 与基线方法相比, 添加不同模块后各个指标均有不同程度的提升。其中, 添加反向注意力模块和多尺度特征融合模块后 Dice 系数分别提升 0.5% 和 1.5%。而模块的组合也对性能的提升有着极大的影响, 例如, 在同时添加模块一和模块二之后, Dice 系数提升到了 0.708, 与基线方法相比提升了 1.2%。在同时添加模块一、模块二和模块三之后, Dice 系数提升到 0.714, 与基线相比提升了 1.8%。表 2 的结果验证了本文创新的有效性。

表 3 展示了模块三在选择不同的融合策略的实验结果。基线方法是在只添加模块一+模块二的情况下进行的。在融合策略的实验中, 笔者分别进行 3 组不同实验。融合策略一如图 5a 所示, 首先将所有低层与高层特征使用拼接操作进行聚合, 其次将聚合后特征送入 1×1 卷积, 目的是使其与原始特征的维度相匹配。最后将其送入多尺度特征融合模块的第一层级联结构。融合策略二、三代表分别在融合策略一中添加因式分解通道注意力模块 (Factorized Channel Attention, FCA)^[29] 与尺度注意力模块 (Scale Attention, SA)^[15] 后的结果, 具体如图 5 所示。显而易见, 与基线方法相比融合策略二、三在 Dice 系数上分别降低了 0.7% 和 0.9%, 更是与最终的融合策略一相差甚远。这表明本文网络中最简单的融合策略一反而起到了最佳的效果。笔者最终选定融合策略一作为最终的融合策略。

3 结束语

COVID-19 感染区域的自动化分割对于临床诊断和治疗具有非常重要的应用价值。本文提出了一种基于反向注意力机制和多尺度特征聚合的肺炎分割框架。反向注意力模块可以通过提取高层特征中的互补信息来改善物体与组织之间的对比度。多尺度特征聚合模块可以扩展接受域, 进一步提取更丰富的语义特征。全局上下文聚合策略可以通过聚合低级和高级特性来捕获全局信息。与 U-Net++ 模型和其他先进的分割模型相比, 本文模型在识别低

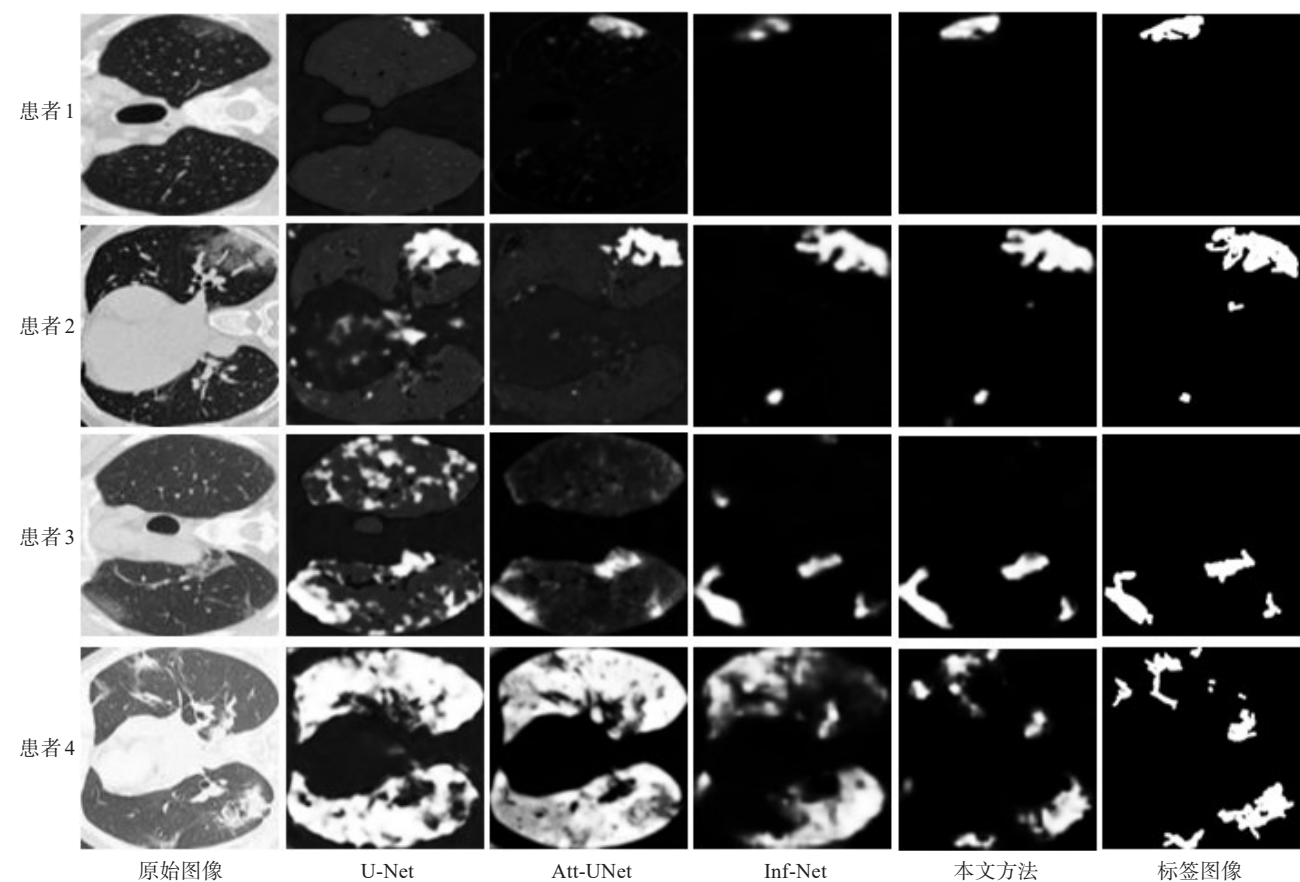


图4 不同模型在 COVID-19 CT分割测试集上的可视化效果对比

Figure 4 Visual comparison among different models on COVID-19 CT segmentation test set

表2 消融实验						
Table 2 Ablation study						
模型	Dice	敏感度	特异性	SM	EM	MAE
基线	0.696	0.699	0.949	0.786	0.850	0.075
+模块一	0.701	0.698	0.952	0.785	0.857	0.074
+模块二	0.711	0.705	0.955	0.797	0.861	0.072
+模块一、二	0.708	0.705	0.953	0.792	0.864	0.073
+模块一、二、三	0.714	0.700	0.958	0.795	0.869	0.069

表3 全局上下文聚合策略消融实验						
Table 3 Ablation study of global context aggregation strategy						
模型	Dice	敏感度	特异性	SM	EM	MAE
基线	0.708	0.705	0.953	0.792	0.864	0.073
融合策略一	0.714	0.700	0.958	0.795	0.869	0.069
融合策略二	0.701	0.721	0.943	0.766	0.870	0.078
融合策略三	0.699	0.709	0.940	0.784	0.849	0.082

对比度目标和检测不同尺度的病变方面具有更好的性能。在未来,一个可能的想法是将网络扩展到其他相关的医疗任务,例如息肉和皮肤病分割。

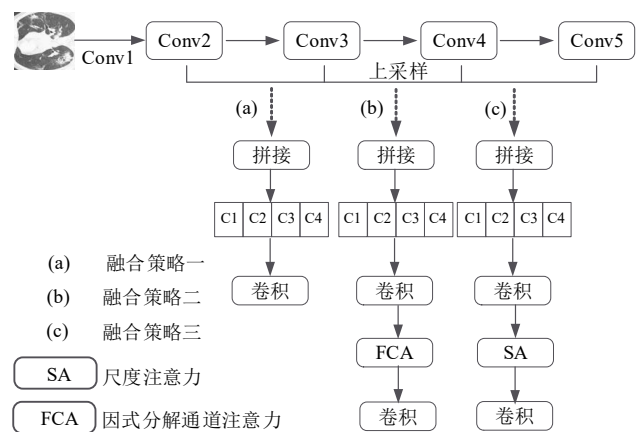


图5 全局上下文聚合策略

Figure 5 Global context aggregation strategy

【参考文献】

[1] Wu Q, Wang S, Li L, et al. Radiomics analysis of computed tomography helps predict poor prognostic outcome in COVID-19 [J]. Theranostics, 2020, 10(16): 7231-7244.

[2] Wei W, Hu XW, Cheng Q, et al. Identification of common and severe COVID-19: the value of CT texture analysis and correlation with clinical characteristics[J]. Eur Radiol, 2020, 30(12): 6788-6796.

[3] Chaganti S, Balachandran A, Chabin G, et al. Quantification of tomographic patterns associated with COVID-19 from Chest CT[J]. arXiv, arXiv: 2004.01279v5, 2020.

[4] 顾国浩, 龙英文, 吉明明. U-Net改进及其在新冠肺炎图像分割的应

- 用[J]. 中国医学物理学杂志, 2022, 39(8): 1041-1048.
- Guo GH, Long YW, Ji MM. Improved U-Net and its application in COVID-19 image segmentation[J]. Chinese Journal of Medical Physics, 2022, 39(8): 1041-1048.
- [5] Han Z, Wei B, Hong Y, et al. Accurate screening of COVID-19 using attention based deep 3D multiple instance learning[J]. IEEE Trans Med Imaging, 2020, 39(8): 2584-2594.
- [6] Wang XG, Deng XB, Fu Q, et al. A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT[J]. IEEE Trans Med Imaging, 2020, 39(8): 2615-2625.
- [7] Fan DP, Zhou T, Ji GP, et al. Inf-Net: automatic COVID-19 lung infection segmentation from CT images[J]. IEEE Trans Med Imaging, 2020, 39(8): 2622-2637.
- [8] 张桃红, 郭徐徐, 张颖. LRSAR-Net语义分割模型用于新冠肺炎CT图片辅助诊断[J]. 电子与信息学报, 2022, 44(1): 48-58.
- Zhang TH, Guo XX, Zhang Y. LRSAR-Net semantic segmentation model for computer aided diagnosis for COVID-19 CT image[J]. Journal of Electronics & Information Technology, 2022, 44(1): 48-58.
- [9] Wang X, Jiang L, Li L, et al. Joint learning of 3D lesion segmentation and classification for explainable COVID-19 diagnosis[J]. IEEE Trans Med Imaging, 2021, 40(9): 2463-2476.
- [10] Amyar A, Modzelewski R, Li H, et al. Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: classification and segmentation[J]. Comput Biol Med, 2020, 126: 104037.
- [11] COVID-19 CT segmentation dataset[EB/OL]. [2020-10-27]. <https://medicalsegmentation.com/covid19>.
- [12] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.
- [13] Park J, Woo S, Lee JY, et al. BAM: bottleneck attention module[J]. arXiv, arXiv: 1807.06514v1, 2018.
- [14] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module[C]//European Conference on Computer Vision. Springer, Cham, 2018.
- [15] Gu R, Wang GT, Song T, et al. CA-Net: comprehensive attention convolutional neural networks for explainable medical image segmentation[J]. IEEE Trans Med Imaging, 2020, 40(2): 699-711.
- [16] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Trans Pattern Anal Mach Intell, 2019, 42(8): 2011-2023.
- [17] 张永梅, 彭炯, 马健喆, 等. 基于改进U-Net的肺癌识别方法[J]. 计算机工程与设计, 2021, 42(1): 256-262.
- Zhang YM, Peng J, Ma JZ, et al. Lung cancer recognition method based on improved U-Net[J]. Computer Engineering and Design, 2021, 42(1): 256-262.
- [18] 王原, 马瑜, 江妍, 等. U-Net改进的视网膜血管图像分割算法[J]. 计算机工程与设计, 2021, 42(10): 2884-2893.
- Wang Y, Ma Y, Jiang Y, et al. Improved retinal vascular image segmentation algorithm based on U-Net[J]. Computer Engineering and Design, 2021, 42(10): 2884-2893.
- [19] Singh VK, Abdel-Nasser M, Rashwan HA, et al. FCA-Net: adversarial learning for skin lesion segmentation based on multi-scale features and factorized channel attention[J]. IEEE Access, 2019, 7: 130552-130565.
- [20] Khan A, Kim H, Chua L. PMED-Net: pyramid based multi-scale encoder-decoder network for medical image segmentation[J]. IEEE Access, 2021, 9: 55988-55998.
- [21] Chen LC, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Trans Pattern Anal Mach Intell, 2018, 40(4): 834-848.
- [22] Chen LC, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[J]. arXiv, arXiv: 1706.05587, 2017.
- [23] Gu ZQ, Cheng J, Fu HZ, et al. CE-Net: context encoder network for 2D medical image segmentation[J]. arXiv, arXiv: 1903.02740, 2014.
- [24] Kingma D, Ba J. Adam: a method for stochastic optimization[J]. arXiv, arXiv: 1412.6980, 2014.
- [25] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.
- [26] Zhou ZW, Siddiquee MM, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation[C]//DLMI 2018, ML-CDS 2018: Deep Learning in Medical Image Analysis and Multi-modal Learning for Clinical Decision Support, 2018: 3-11.
- [27] Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: learning where to look for the pancreas[J]. arXiv, arXiv: 1804.03999, 2018.
- [28] Schlemper J, Oktay O, Schaap M, et al. Attention gated networks: learning to leverage salient regions in medical images[J]. Med Image Anal, 2019, 53: 197-207.
- [29] Singh VK, Abdel-Nasser M, Rashwan HA, et al. FCA-Net: adversarial learning for skin lesion segmentation based on multi-scale features and factorized channel attention[J]. IEEE Access, 2019, 130552-130565.

(编辑: 薛泽玲)