

# 基于 PRAU-Net 的新冠肺炎 CT 图像分割研究

曾庆鹏, 崔 鹏

(南昌大学 数学与计算机学院, 江西 南昌 330031)

**摘要:**针对新冠肺炎 CT 影像病灶区域小、形状结构差异大和噪声等问题,提出一种基于编解码结构的 PRAU-Net 医学图像分割方法。首先,在编码阶段使用一种残差 Inception 注意力卷积模块 (Residual Inception Attention, RIA) 提取特征, RIA 采用残差结构将更深的并行卷积块和通道注意力机制相结合捕获更丰富的特征;其次,将不同尺度的特征通过跳跃连接进行融合,使解码器中特征有更加丰富的全局信息;最后,在解码器中使用全局注意力模块使网络关注相关特征,有效减少了 CT 影像中噪声的影响。为了验证该方法的有效性,分别在三个数据集 (Segmentation dataset nr. 2, CC-CCII 和 COVID19\_1110) 上进行验证,实验结果表明,该方法比经典方法分割结果更加准确,相较于 U-Net 等经典分割方法, Dice 系数提升了 1.12% ~ 14.84%, 敏感度提升了 0.7% ~ 24.63%。为了进一步提高分割性能,使用生成对抗网络对 Segmentation dataset nr. 2 数据集进行了扩充,并利用 PRAU-Net 分割方法和多种经典分割网络进行了验证,结果表明,扩充小样本数据集可以有效地提高分割性能, PRAU-Net 方法的 Dice 系数从 0.836 4 上升到了 0.858 3。

**关键词:**新冠肺炎;医学图像分割;U-Net;残差结构;注意力机制

中图分类号: TP391.41

文献标识码: A

文章编号: 1673-629X(2024)03-0133-07

doi: 10.3969/j.issn.1673-629X.2024.03.020

## Research of COVID-19 CT Image Segmentation Based on PRAU-Net

ZENG Qing-peng, CUI Peng

(School of Mathematics and Computer Sciences, Nanchang University, Nanchang 330031, China)

**Abstract:** Aiming at the problems of small lesion area, large variation in shape structure and noise in CT images of COVID-19, we proposed a parallel residual attention U-Net medical image segmentation method based on encoder-decoder architecture. Firstly, the model extracted features by residual inception attention (RIA) in the encoder. RIA adopted a residual structure to combine deeper parallel convolution block and channel attention mechanisms to capture richer features. Secondly, features of different scales were fused by skip connection to obtain richer global information. Lastly, the global attention module was used in the decoder to enable the network focus on relevant features, which effectively reduced the influence of noise in CT images. To verify the effectiveness of the proposed method, we have conducted experiments on segmentation dataset nr. 2, CC-CCII and COVID19\_1110. Experimental results show that proposed method is more accurate than the classical methods. Compared with classical segmentation methods such as U-Net, the Dice coefficient increases by 1.12% ~ 14.84% and the sensitivity increases by 0.7% ~ 24.63%. To further demonstrate the segmentation performance, the Segmentation dataset nr. 2 is extended by using generative adversarial network, the PRU-NET method and several classical segmentation networks are used to verify the method. It is showed that expanding the small sample dataset can effectively improve the segmentation performance, the Dice coefficient of PRAU-Net method was increased from 0.836 4 to 0.858 3.

**Key words:** COVID-19; medical image segmentation; U-Net; residual structure; attention mechanism

## 0 引言

2019 新型冠状病毒肺炎 (Corona Virus Disease 2019, COVID-19) 被世界卫生组织 (WHO) 于 2020 年 1 月宣布疫情为全球公共卫生事件,据 WHO 统计,截至 2023 年 1 月 29 日,全球累计报告新冠肺炎确诊病例达 7.53 亿,死亡超过 680 万人<sup>[1]</sup>。由于该病毒抗逆

性强且传播方式多,在人群中极容易快速传播,因此,快速、准确地识别新型冠状病毒感染者对于阻断传播、及时对患者进行隔离治疗具有重要意义。

目前,逆转录聚合酶链反应 (RT-PCR) 是 COVID-19 诊断的常规方法,RT-PCR 是从口咽拭子、鼻咽拭子、支气管肺泡灌洗液或气管抽吸物获得样本中的病

收稿日期: 2023-05-21

修回日期: 2023-09-22

基金项目: 国家自然科学基金项目 (62166026)

作者简介: 曾庆鹏 (1974-), 男, 副教授, 通信作者, 研究方向为网络与信息安全、数据挖掘; 崔 鹏 (1996-), 男, 硕士研究生, 研究方向为计算机视觉。

毒 RNA<sup>[2]</sup>。然而,研究表明 RT-PCR 的灵敏度不高,有时初期的新冠病毒感染者需要多次检测才能被确诊,导致患者不能得到及时有效的治疗,不仅危及患者的生命健康,还加大了新冠病毒的传播可能性,对社会造成更大的生命财产损失。计算机断层扫描(Computed Tomography,CT)作为 COVID-19 筛查的另一方案,对患者肺部进行 CT 扫描可以更加准确地判断受检者是否感染新冠病毒并使放射科医生快速获得患者肺部受损情况和评估疾病程度。COVID-19 的 CT 表现为双肺多发磨玻璃样阴影(Ground Glass Opacity,GGO),伴或不伴实变、血管增粗和小叶间隔增厚<sup>[3]</sup>,但是对 CT 图像的评估和分析需要经验丰富的医生,而且对 CT 图像进行标注更是一项繁重且耗时的工作,导致医疗资源的浪费。利用人工智能技术可以快速且准确地标注出 CT 图像的病灶区域,帮助医生更直接地了解患病程度,不仅减轻了医生的负担,还可以提高医生的诊断效率。

近年来,深度学习在医学影像分割领域受到了广泛的研究与应用,比如肝脏肿瘤分割<sup>[4]</sup>、眼底血管分割<sup>[5]</sup>和脑图像分割<sup>[6]</sup>等。在进行医学影像分割时,相对于传统目标检测方法,深度学习方法不仅注重分割目标的特征提取,还对图像中的上下文信息进行关注研究,从而获得更有效的特征,得到更加准确的分割图像。常见的 U-Net<sup>[7]</sup>, Attention-UNet<sup>[8]</sup>和 UNet++<sup>[9]</sup>等经典的分割方法在提取分割目标的特征时,容易受到图像复杂性和待分割区域不规则性的干扰导致提取到的部分特征缺少关键信息,且网络中下采样池化会导致特征信息丢失,降低了分割的准确率。

针对上述存在的问题,不少学者也提出了改进方案。Kaul 等人<sup>[10]</sup>提出基于注意力的全卷积网络方法,该方法采用了两个并行信息流分支,两个分支都采用编码器-解码器结构,强化了与目标信息有关的特征映射。Gu 等人<sup>[11]</sup>提出了一种上下文编码器网络(Context Encoder Network)用于捕获更多高层信息并保留空间信息,编码阶段的密集空洞卷积(Dense Atrous Convolution,DAC)和解码阶段的残差多核池化(Residual Multi-kernel Pooling,RMP)共同构成了上下文提取模块,该模块有效减少了因连续的池化和跨步卷积运算导致一些空间信息的丢失所造成的影响。Song 等人<sup>[12]</sup>提出一种增加模型复杂性和更多特征融合的方法,在模型中使用许多的跳跃连接和嵌套连接,有效减少了梯度的扩散问题。Feng 等人<sup>[13]</sup>提出了一种新的上下文金字塔融合网络 CPFNet,设计了全局金字塔引导模块(Global Pyramid Guidance,GPG)为解码器提供不同尺度的全局上下文语义信息,尺度感知金字塔融合模块(Scale-Aware Pyramid Fusion,SAPF)则

在深层特征中动态融合多尺度的上下文信息。Liu 等人<sup>[14]</sup>提出一种多感受野神经网络 MRFNet,通过级联的方式将不同层级 U-Net 网络输出的特征映射进行融合,将融合的特征进行卷积操作完成像素级的分类,取得了良好的分割效果。

上述网络分别针对皮肤、眼球和大脑等的医学图像分割已经取得了不错的效果,针对新冠肺炎 CT 影像病变区域的分割,也有学者提出了解决方案。顾国浩等人<sup>[15]</sup>在 U-Net 模型的基础上引入了循环残差模块和自注意力机制来加强对特征信息的提取。任楚岚等人<sup>[16]</sup>在 U-Net 模型的基础上结合残差连接,分层分裂模块(Hierarchical-Split Block),坐标注意力模块和特征内容感知重组上采样来增强模型提取特征能力。宋瑶等人<sup>[17]</sup>对现有的数据集图像和标签同时随机旋转、裁剪和翻转,采用 ImageNet 上预先训练的 EfficientNet-B0 作为特征提取器,并通过 Downsampling 上采样代替传统的上采样结构以改进 U-Net。Fan 等人<sup>[18]</sup>提出一种针对肺部 CT 图像的分割网络 Inf-Net,利用并行部分解码器(Parallel Partial Decoder,PPD)聚合深层特征并生成一个粗略定位肺部感染的全局图,在全局图的指导下,反馈给多个逆向注意模块(Reverse Attention,RA),并结合显示边缘注意力以提高目标区域边界的表征,有效提高了对感染区域边界的识别能力。Rajamani 等人<sup>[19]</sup>提出一种动态可变形网络(Dynamic Deformable Attention Network),将可变形交叉注意力模块引入 U-Net 网络的最深层以连续方式学习注意力系数和注意力偏移,相较于 Fan 等人<sup>[18]</sup>的方法还提高了分割性能。左斌等人<sup>[20]</sup>基于 Fan 等人的方法进行了改进,通过引入通道注意力机制加强网络的特征表示,并运用注意力门模块更好地融合边缘信息。上述方法都是将最深层的特征作为预测结果的关键特征,但由于多次下采样池化和卷积深度增加会丢失部分特征信息,导致最深层的特征信息丢失尤为严重。

针对以上问题,该文提出一种并行残差注意力 U-Net(Parallel Residual Attention U-Net,PRAU-Net)分割网络。结合改进的 Inception 模块设计了残差 Inception 注意力卷积模块(Residual Inception Attention Convolution Block,RIA),将改进的 Inception 模块和通道注意力模块融入残差结构中,大大改善模型对特征提取的有效性;同时基于空洞卷积设计了多尺度跳跃连接(Multi-scale Skip Connection,MSC),基于空洞卷积取代了池化下采样,减少了局部特征信息的丢失,多尺度结构聚合了粗粒度语义信息和细粒度语义信息,为解码器提供了更多有助于预测结果的多尺度上下文信息;最后在解码器中利用全局注意力模块(Global

Attention Mechanism, GAM) 提取更关键的特征信息,降低了 CT 影像中噪声等无关特征的影响。面对结构复杂多变新冠肺炎 CT 影像,在减少空间特征信息丢失的同时捕获更加丰富的多尺度上下文信息,从而有效提高了分割的准确度。

## 1 基于 PRAU-Net 的新冠肺炎 CT 影像分割模型

提出的 PRAU-Net 分割模型如图 1 所示。该模型采用 U-Net<sup>[7]</sup> 作为骨干网络,由三个主要部分组成。第一部分是网络的编码部分,编码部分使用 ResNet<sup>[21]</sup> 的网络结构作为特征提取器,将设计的残差 Inception 注意力卷积模块替代了 ResNet 结构中的普通卷积操作来提取特征信息;第二部分是网络的解码阶段,由上采样操作、卷积操作和全局注意力模块组成,将低分辨率的特征信息一步一步恢复到原始图像的大小来预测新冠肺炎 CT 影像中的病灶区域;第三部分是多尺度跳跃连接,将编码器中更浅层和同层的特征信息与来自解码器中更深层的特征信息进行拼接,捕获了多尺度的上下文信息。

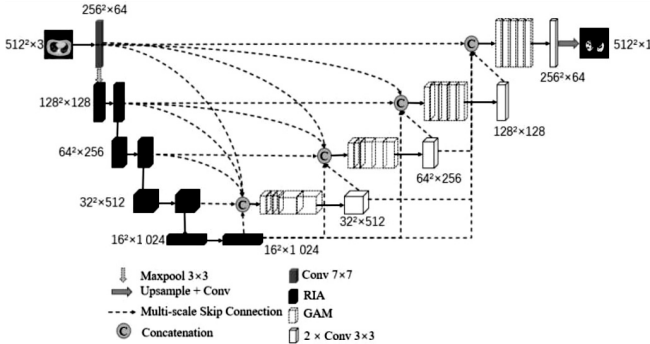


图 1 PRAU-Net 网络模型

### 1.1 残差 Inception 注意力卷积模块

U-Net 网络中的编码器每层使用两个卷积操作提取特征并使用池化下采样降低图像分辨率,两个卷积操作面对存在噪声的大分辨率 CT 影像时难以提取其丰富的特征信息,而池化操作容易丢失特征信息,最终导致分割结果精度差。在 GoogLeNet<sup>[22]</sup> 的启发下,该文使用改进的 Inception 模块来提取特征,改进的 Inception 模块如图 2 所示。该模块包含并行的 4 条特征提取分路,由堆叠的卷积块和池化操作组成。最左边的分路由一个平均池化和 1x1 卷积组成,1x1 卷积可以减少特征的通道数,以此减少网络的参数量,右边 3 条分路首先都通过一个 1x1 的卷积,然后分别通过 1 个、2 个、3 个堆叠的 3x3 卷积,其中 2 个堆叠的 3x3 卷积相当于一个 5x5 的卷积,3 个堆叠的 3x3 卷积相当于一个 7x7 卷积,通过堆叠的不同数量的卷积块提取不同尺度的特征信息,最后将所有分路的特征进行拼

接,使网络模型得到更加全面和多层次的特征表示。

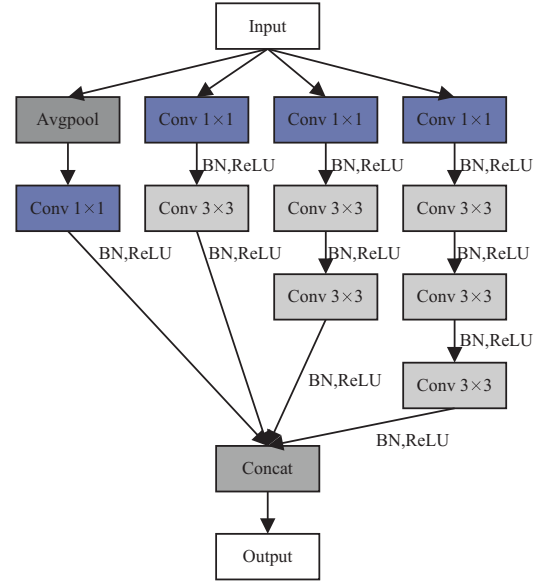


图 2 改进的 Inception 模块

定义该结构输入为  $F_i$ , 通过改进的 Inception 模块,生成的特征图  $F_c$  由公式 1 和公式 2 得到。

$$X_1 = C_{1 \times 1}(\text{AvgPool}(F_i))$$

$$X_2 = C_{3 \times 3}(C_{1 \times 1}(F_i))$$

$$X_3 = C_{3 \times 3}(C_{3 \times 3}(C_{1 \times 1}(F_i)))$$

$$X_4 = C_{3 \times 3}(C_{3 \times 3}(C_{3 \times 3}(C_{1 \times 1}(F_i))))$$

$$F_c = \text{Concat}[X_1, X_2, X_3, X_4]$$

其中,  $C_{1 \times 1}$  表示卷积核大小为 1x1 的卷积操作;  $C_{3 \times 3}$  表示卷积核大小为 3x3 的卷积操作; AvgPool 代表的是平均池化操作,用来减小输入特征的尺寸;  $F_c$  表示特征图在通道方向上拼接。

改进的 Inception 模块可提取到丰富的特征信息,但仍无法解决 CT 影像中噪声带来的干扰和网络池化下采样特征信息丢失的问题。在 Hu 等人<sup>[23]</sup> 提出的 SE-ResNet 模块的启发下,该文设计了残差 Inception 注意力模块,将改进的 Inception 模块和通道注意力模块融入残差结构中,改进的 Inception 模块利用更深的卷积层使网络获得了更大的感受野,通道注意力 SE 模块有效帮助网络关注更相关的特征信息,减少噪声的影响,残差结构则减少了卷积块堆叠变深时网络退化的问题。残差 Inception 注意力模块总体结构如图 3 所示。

残差 Inception 注意力模块的输入特征图为  $F_i$ , 经过改进后的 Inception 模块输出特征图为  $F_c$ ; 接着将  $F_c$  输入到 SE 模块中,SE 模块通过挤压和激励操作实现对特征图通道上的加权,挤压操作将  $H \times W \times C$  的特征图通过全局平均池化得到  $1 \times 1 \times C$  的特征图,再通过一个全连接的神经网络进行激励得到  $1 \times 1 \times C$  的一组权重,将权重乘到每一个输入特征  $F_c$  的通道上实现特



征重新标定,产生特征图  $F_s$ ;最后通过跳跃连接将得到的特征图  $F_s$  和输入特征图  $F_i$  相加得到输出  $F_o$ 。 $F_s$  和  $F_o$  表示为:

$$F_s = \sigma(C_{1 \times 1}(\delta(C_{1 \times 1}(AAP(F_c)))) \otimes F_c \quad (3)$$

$$F_o = F_s + F_i \quad (4)$$

其中,AAP表示自适应平均池化,将  $C \times H \times W$  的全局空间信息压缩到  $C \times 1 \times 1$  的大小,  $\delta$  表示 Relu 激活函数,  $\sigma$  表示 Sigmoid 激活函数,将特征图映射到 0~1 之间,获得一组特征图权重。

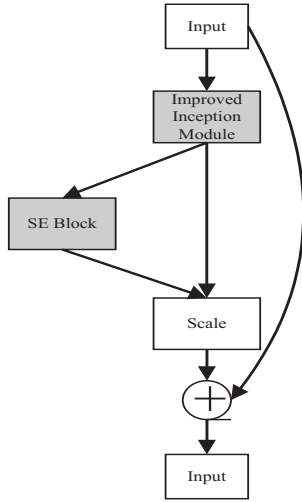


图 3 残差 Inception 注意力模块

## 1.2 全局注意力模块

注意力模块 (Convolution Block Attention Module, CBAM)<sup>[24]</sup>可以同时关注空间域特征和通道域特征,自主学习每个特征的重要程度,通过串联的方式分别对通道和空间上的特征图赋予学习到的权重,可以有效减少噪声的影响,提高网络对感兴趣区域特征的提取。为了提高对新冠肺炎 CT 图像病灶区域的分割性能,该文基于 CBAM 模块设计了全局注意力模块 (Global Attention Module, GAM),不但可以减少无关特征的干扰,还可以更好地聚合空间上下文信息。GAM 模块结构如图 4 所示。

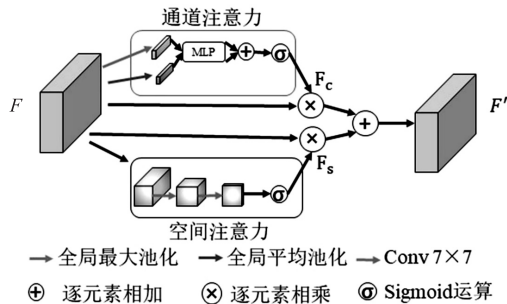


图 4 全局注意力模块

GAM 将通道注意力模块和空间注意力模块并联,在通道注意力模块中,输入特征图首先分别经过自适应平均池化和自适应最大池化,接着通过一个共享多

层感知机获得两个通道方向的输出特征图,再将两个输出特征图逐元素相加得到融合特征图,最后通过 Sigmoid 激活函数输出关于通道方向的权重,将权重沿通道方向与输入特征图逐元素相乘得到最终输出特征图。通道注意力计算公式如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AAP}(F)) + \text{MLP}(\text{AMP}(F))) \quad (5)$$

$$F_c = M_c(F) \otimes F \quad (6)$$

其中,AAP表示自适应平均池化,AMP表示自适应最大池化,AAP和AMP分别抽取不同通道特征图的平均值和最大值,将不同通道特征图的大小由  $C \times H \times W$  压缩到  $C \times 1 \times 1$ ;MLP表示多层感知机,用于实现通道的压缩和扩张; $\sigma$ 表示 Sigmoid 激活函数; $M_c$ 表示通道方向的权重; $F_c$ 是经过通道注意力模块的输出特征图。

GAM的空间注意力模块使用卷积操作替换了原始的池化操作,缓解了因池化操作导致特征信息弥散的问题。具体来讲,就是使用两个  $7 \times 7$  普通卷积替换了平均池化和最大池化操作,两层  $7 \times 7$  的卷积操作通过其更大的感受野扩大了空间信息的交互,可以帮助网络更精确地提取特征图的空间位置信息。空间注意力计算公式如下:

$$M_s(F) = \sigma(C_{7 \times 7}(C_{7 \times 7}(F))) \quad (7)$$

$$F_s = M_s(F) \otimes F \quad (8)$$

其中, $C_{7 \times 7}$ 表示卷积核大小为  $7 \times 7$  的卷积操作, $M_s$ 表示空间方向的权重, $F_s$ 是经过空间注意力模块的输出特征图。

输入特征图  $F$  经过 GAM 的通道子模块和空间子模块后得到两个不同的输出特征图  $F_c$  和  $F_s$ ,最后将两者相加得到最终输出特征图  $F'$ ,如公式 9 所示。

$$F' = F_c + F_s \quad (9)$$

## 1.3 多尺度跳跃连接

为了进一步缓解由于网络过深导致的特征信息丢失的问题,将 U-Net 中的同层跳跃连接改为多尺度的跳跃连接,将包含更多空间信息的浅层特征和有更丰富细节信息的深层特征融合。当编码器浅层特征通过跳跃连接传递到解码器更深层时,由于编码器中浅层的特征信息分辨率更大,在跳跃连接中使用空洞卷积代替了下采样池化操作,减少了池化操作导致空间信息的丢失,使更多的空间信息被传递到更深层中。相比于传统卷积,空洞卷积可以在相同的参数量时获得更大的感受野,在解码器中将含有更多细节信息的更深层特征进行上采样传递到浅层的网络中,丰富的空间信息和细节信息共同组成全局特征信息使网络对病灶区域有更精确的预测结果,解决了新冠肺炎 CT 影像病灶区域小与形状不规则造成分割结果差的问题。多尺度跳跃连接过程如公式 10:

$$X_{De}^j = [D_c^{3 \times 3}(X_E^{i < j}), X_E^{i=j}, C_{3 \times 3}(U(X_D^{i > j}))] \quad (10)$$

其中,  $i \in \{1, 2, 3, 4, 5\}$ ,  $j \in \{1, 2, 3, 4\}$ ,  $i$  表示待融合特征的层数,  $j$  表示融合后特征的层数;  $X_{De}^j$  表示第  $j$  层融合后的特征;  $D_c^{3 \times 3}$  表示卷积核大小为  $3 \times 3$  的空洞卷积;  $X_E$  表示编码器中的特征;  $C_{3 \times 3}$  表示卷积核大小为  $3 \times 3$  的普通卷积;  $U$  表示上采样操作;  $X_D$  表示解码器的特征;  $[]$  表示沿通道维度的拼接。

## 2 实验与分析

### 2.1 数据集和评价指标

该文选取 3 个数据集验证模型的有效性: Segmentation dataset nr. 2 (数据集 1) 源于向用户免费开放的放射科数据库 Radiopaedia, 其中包括 9 例新冠肺炎患者的 829 张 CT 切片, 其中 373 张切片被标记为感染新冠肺炎并进行了标注。CC-CCII (数据集 2) 来自中国胸部 CT 图像调查协会, 其中包括 150 例新冠肺炎患者的 750 张 CT 切片, 对其中 455 张切片进行了标注。COVID19\_1110 (数据集 3) 由莫斯科医学提供, 包含 1 110 例新冠肺炎患者三维肺部 CT 图像, 其中 50 例带有分割标签, 处理成 2 维切片后共 600 张  $512 \times 512$  的 CT 图像。实验中仅选择已标注数据, 具体数据划分如表 1 所示。

表 1 数据集划分信息

数据集	总样本数	训练集 样本数	测试集 样本数	样本 分辨率
数据集 1	373	299	74	$512 \times 512$
数据集 2	455	364	91	$512 \times 512$
数据集 3	600	480	120	$512 \times 512$

为了从多个角度说明所提方法对新冠肺炎病灶区域的分割性能, 使用 4 种评估指标, 包括 Dice 系数、敏感性 (Sensitivity)、特异性 (Specificity) 和准确率 (Accuracy)。其中, Dice 系数是一种基于区域的度量, 用于衡量两个样本中正样本的重叠率, 如公式 11 所示; 敏感性也称召回率, 用于度量所有正样本中被正确识别的比例, 如公式 12 所示; 特异性是指负样本中被正确识别的部分所占的比例, 如公式 13 所示; 准确率指的是在所有样本中正样本和负样本被正确识别的占比, 如公式 14 所示。

$$\text{Dice} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (11)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (12)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (13)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (14)$$

其中, TP (True Positives) 表示被正确检测为正样本的像素数量; FP (False Positives) 表示被错误检测为正样本的像素数量; FN (False Negatives) 表示被错误检测为负样本的像素数量; TN (True Negatives) 表示被正确检测为负样本的像素数量。

### 2.2 实验设置

实验采用 5 倍交叉验证的方法, 将数据集随机分为 5 个子集, 训练集和测试集按 4 : 1 比例分配, 重复进行 5 次该过程的实验, 取 5 次实验的平均值作为实验结果。所有输入模型的 CT 影像图统一大小为  $512 \times 512$ , 批处理大小设置为 4, 并随机对图像进行旋转提高数据的多样性。其中 Epoch 设置为 70, 采用 RMSProp 优化器训练网络模型, 初始学习率为 0.000 01, 具体参数如表 2 所示。

表 2 PRAU-Net 网络参数

Parameter	Value
训练集随机采样数	299
训练集重采样	299
输入层	$512 \times 512 \times 3$
隐藏层	$16 \times 16 \times 1\ 024$
输出层	$512 \times 512 \times 1$
激活函数	ReLU
损失函数	BCEWithLogitLoss
动量 (momentum)	0.9
梯度优化器	RMSprop
Batch size	4
学习率	0.000 01
训练次数	70

### 2.3 分割性能实验

为了验证文中方法的分割性能, 对比实验以 Dice 系数、灵敏度 (Sensitivity)、特异性 (Specificity) 和准确率 (Accuracy) 作为评价指标, 以五种网络分割方法作为比较基线: U-Net<sup>[7]</sup>, Attention-UNet<sup>[8]</sup>, UNet++<sup>[9]</sup>, Res-UNet<sup>[25]</sup>, DeepLabV3+<sup>[26]</sup>。在 Segmentation dataset nr. 2 数据集上的实验可视化结果如图 5 所示, 可以看出文中方法在边缘区域和细小的区域分割表现优于其它方法。

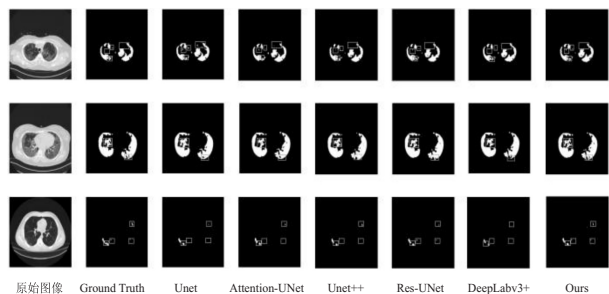


图 5 分割结果可视化

基于评价指标的实验结果如表 3 所示。由表 3 可

以发现, PRAU-Net 网络模型在数据集 1 中的 Dice 系数、Sensitivity 以及 Accuracy 都比其它网络的优, 只有 Specificity 指标略低于 DeepLabV3+, 与经典的 U-Net 网络模型相比较, Dice 系数、Sensitivity、Specificity 和 Accuracy 分别提升了 5.65%, 8.10%, 0.02% 和 0.04%, 其中 Dice 系数和 Sensitivity 提升较为明显, 相较于其它对比网络, Dice 系数有 2.35% ~ 6.02% 的提升, Sensitivity 有 1.89% ~ 10.02% 的提升, 但 Specificity 略低于 DeepLabV3+ 方法。在数据集 2 的实验结果中, 文中方法的 Dice 系数和 Sensitivity 指标都优于所有对比方法的, 相较于 DeepLabV3+, Dice 系数和 Sensitivity 分别提升了 8.96% 和 14.68%, 相较于其

它几种对比网络, Dice 系数仍有 1.12% ~ 3.15% 的提升, Sensitivity 则有 0.71% ~ 5.63% 的提升, 而 UNet++ 方法的 Accuracy 指标优于所有方法的, 相较于文中方法 Accuracy 提升了 0.02%, 但其网络参数量远比文中方法的高, DeepLabV3+ 在 Specificity 指标表现同样优于所有方法, 相较于文中方法有 0.06% 的提升。在数据集 3 的实验结果中, 文中方法的 Dice 系数、Sensitivity 和 Accuracy 都比其它几种经典方法的优, 尤其是 Dice 系数和 Sensitivity 表现较好, 相较于 U-Net 分别提升了 9.52% 和 14.22%。实验结果表明, 文中方法有效地提升了在新冠肺炎 CT 图像数据集上的分割性能。

表 3 不同模型的结果指标对比

方法	数据集 1				数据集 2				数据集 3			
	Dice	Sensitivity	Specificity	Accuracy	Dice	Sensitivity	Specificity	Accuracy	Dice	Sensitivity	Specificity	Accuracy
U-Net	0.791 7	0.749 7	0.998 5	0.996 1	0.675 3	0.635 5	0.997 3	0.991 5	0.642 9	0.584 2	0.999 4	0.996 9
Attention-UNet	0.793 2	0.768 0	0.998 2	0.995 9	0.673 4	0.621 4	0.997 6	0.991 6	0.670 8	0.637 5	0.999 1	0.997 0
UNet++	0.794 4	0.757 5	0.998 4	0.996 0	0.679 8	0.621 7	0.997 7	0.991 7	0.688 3	0.645 5	0.999 3	0.997 1
Res-UNet	0.817 2	0.795 4	0.998 4	0.996 2	0.686 9	0.651 8	0.997 1	0.991 5	0.673 9	0.629 1	0.999 2	0.997 1
DeepLabV3+	0.788 9	0.736 6	0.999 0	0.995 3	0.637 5	0.572 4	0.997 9	0.990 6	0.613 1	0.535 4	0.999 5	0.996 8
Ours	0.836 4	0.810 4	0.998 7	0.996 5	0.694 6	0.656 4	0.997 3	0.991 5	0.704 1	0.667 3	0.999 2	0.997 1

## 2.4 数据扩充实验

由于医学图像数据集具有一定的隐私性且对医学图像进行标注需要耗费大量人力, 拥有大量样本的数据集通常难以获取。文中的数据集 1 仅有 373 张被标记的新冠肺炎 CT 图像, 小样本的数据图像可能无法使网络学习到足够的特征, 因此, 该文利用循环生成对抗网络<sup>[27]</sup>对数据集 1 进行了扩充, 生成了 373 张伪新冠肺炎 CT 图像, 采用半监督学习的方式对生成的新冠肺炎 CT 图像进行标注, 最后将其加入到数据集 1 的训练集中辅助模型训练。为了验证扩充后的数据集对网络模型的分割性能是否有影响, 使用多种经典网络及提出的方法进行了实验, 实验结果如表 4 所示。

表 4 扩充数据集后实验结果对比

方法	数据集 1		数据集 1+生成数据	
	Dice	Sensitivity	Dice	Sensitivity
U-Net	0.791 7	0.749 7	0.807 1	0.769 6
Attention-UNet	0.793 2	0.768 0	0.814 1	0.781 5
UNet++	0.794 4	0.757 5	0.818 6	0.784 3
Res-UNet	0.817 2	0.795 4	0.825 9	0.800 1
DeepLabV3+	0.788 9	0.736 6	0.826 8	0.797 4
Ours	0.836 4	0.810 4	0.858 3	0.827 5

由表 4 可以发现, 使用扩充数据集的分割网络模型在 Dice 系数和 Sensitivity 评价指标上均有一定程度的提升, 其中 U-Net 网络的 Dice 系数和 Sensitivity 分别为 0.807 1 和 0.769 6, 各自提升了 1.95% 和

2.65%, Attention-UNet 的 Dice 得分达到了 0.814 1, 提升了 2.63%, UNet++ 和 Res-UNet 的 Dice 得分分别提升了 3.05% 和 1.06%。其中 DeepLabV3+ 表现出最明显的性能提升, Dice 系数提升了 4.8%, 而 Sensitivity 有 8.25% 的提升, 虽然提出的 PRAU-Net 在 Dice 系数和 Sensitivity 评价指标中分别只提升了 2.62% 和 2.11%, 但所提方法相较于对比网络仍有最好的分割性能。总体实验结果表明: 扩充小样本数据集可以帮助分割网络获得更好的分割性能。

## 3 结束语

基于 U-Net 网络模型, 该文提出了一种并行残差注意力网络模型 PRAU-Net。该模型在编码阶段采用嵌入通道注意力的残差 Inception 卷积模块捕获输入特征的不同尺度语义信息并对其进行特征重标定, 并采用了残差结构, 能够在提取更丰富的特征信息的同时解决网络退化问题; 在解码阶段将多尺度跳跃连接中捕获的多尺度上下文信息输入全局注意力模块中对其特征进行重新加权, 加强了网络对病灶区域特征的关注。分别在不同的三个数据集进行实验, 实验结果证明, 该方法有效降低了 CT 影像中噪声在分割任务中的影响, 相较于经典分割方法, 提高了对病灶区域的分割准确度。针对数据集样本较少的问题, 验证了对 Segmentation dataset nr. 2 数据集使用生成对抗网络进行扩充后, 使用多种网络模型验证了扩充样本数据集

对网络分割性能的有效性。需要说明的是,该文采用的数据集是二维的肺部 CT 图像切片,进一步的研究方向可尝试对 3D 图像数据进行分割研究。

#### 参考文献:

- [1] WHO coronavirus disease (COVID-19) dashboard[EB/OL]. (2023)[2023-01-29]. <https://covid19.who.int>.
- [2] FANG Y,ZHANG H,XIE J,et al. Sensitivity of chest CT for COVID-19;comparison to RT-PCR[J]. Radiology,2020,296(2):115-117.
- [3] 邓燕芳,胡磊,余芳,等.新冠肺炎(COVID-19)的临床特征及 CT 表现[J]. 现代医用影像学,2021,30(6):1012-1016.
- [4] TIAN H,JI B,QUAN W,et al. MPA-Net:multi-scale pyramid attention network for liver tumor segmentation[C]//2021 international conference on electronic information engineering and computer science (EIECS). Changchun: IEEE, 2021: 658-661.
- [5] GUO C,SZEMENYEI M,YI Y,et al. Dense residual network for retinal vessel segmentation [C]//ICASSP 2020 - 2020 IEEE international conference on acoustics,speech and signal processing (ICASSP). Barcelona:IEEE,2020:1374-1378.
- [6] MOESKOPS P,VIERGEVER M A,MENDRIK A M,et al. Automatic segmentation of MR brain images with a convolutional neural network[J]. IEEE Transactions on Medical Imaging, 2016,35(5):1252-1261.
- [7] RONNEBERGER O,FISCHER P,BROX T. U-net: convolutional networks for biomedical image segmentation[C]//Medical image computing and computer-assisted intervention - MICCAI 2015;18th international conference. Munich: Springer,2015:234-241.
- [8] OKTAY O,SCHLEMPER J,FOLGOC L L,et al. Attention U-Net:learning where to look for the pancreas[J]. Medical Image Analysis,2019,53:197-207.
- [9] ZHOU Z,SIDDIQUEE M M R,TAJBAKHS N,et al. Unet+:redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Transactions on Medical Imaging,2019,39(6):1856-1867.
- [10] KAUL C,MANANDHAR S,PEARS N. Focusnet:an attention-based fully convolutional network for medical image segmentation[C]//2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019). Venice: IEEE, 2019: 455-458.
- [11] GU Z,CHENG J,FU H,et al. Ce-net:context encoder network for 2d medical image segmentation[J]. IEEE Transactions on Medical Imaging,2019,38(10):2281-2292.
- [12] SONG T,MENG F,RODRIGUEZ-PATON A,et al. U-next;a novel convolution neural network with an aggregation u-net architecture for gallstone segmentation in ct images[J]. IEEE Access,2019,7:166823-166832.
- [13] FENG S,ZHAO H,SHI F,et al. CPFNet:context pyramid fusion network for medical image segmentation[J]. IEEE Transactions on Medical Imaging,2020,39(10):3008-3018.
- [14] LIU L,WU F X,WANG Y P,et al. Multi-receptive-field CNN for semantic segmentation of medical images[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(11): 3215-3225.
- [15] 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.
- [16] 任楚岚,张亨. 基于改进 U-Net 模型的新冠肺炎图像分割[J]. 网络安全技术与应用,2022(10):38-41.
- [17] 宋瑶,刘俊. 改进 U-Net 的新冠肺炎图像分割方法[J]. 计算机工程与应用,2021,57(19):243-251.
- [18] FAN D P,ZHOU T,JI G P,et al. Inf-net:automatic covid-19 lung infection segmentation from ct images[J]. IEEE Transactions on Medical Imaging,2020,39(8):2626-2637.
- [19] RAJAMANI K T,SIEBERT H,HEINRICH M P. Dynamic deformable attention network (DDANet) for COVID-19 lesions semantic segmentation[J]. Journal of Biomedical Informatics, 2021,119:103816.
- [20] 左斌,李菲菲. 基于注意力机制和 Inf-Net 的新冠肺炎图像分割方法[J]. 电子科技,2023,36(2):22-28.
- [21] 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. Las Vegas: IEEE, 2016:770-778.
- [22] SZEGEDY C,VANHOUCKE V,IOFFE S,et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas:IEEE,2016:2818-2826.
- [23] HU J,SHEN L,SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City: IEEE,2018:7132-7141.
- [24] WOO S,PARK J,LEE J Y,et al. Cbam:convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). Munich:Springer,2018:3-19.
- [25] 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). Hangzhou:IEEE,2018:327-331.
- [26] CHEN L C,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 (ECCV). Munich:Springer,2018:801-818.
- [27] ZHU J Y,PARK T,ISOLA P,et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//Proceedings of the IEEE international conference on computer vision. Venice:IEEE,2017:2223-2242.