

U-Net改进及其在新冠肺炎图像分割的应用

顾国浩,龙英文,吉明明

上海工程技术大学电子电气工程学院,上海 201620

【摘要】CT成像已成为检测新型冠状病毒肺炎(COVID-19)最重要的步骤之一。针对手动分割患者胸部CT图像中毛玻璃混浊区域繁琐的问题提出了一种自注意力循环残差U型网络模型来实现COVID-19患者肺部CT图像的自动分割,辅助医生诊断。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抓取从而提升分割精度。在公开数据集上的分割实验结果显示,该算法的Dice系数、敏感度和特异度分别达到了85.36%、76.64%和76.25%,与其他算法相比具有良好的分割效果。

【关键词】U-Net;新型冠状病毒肺炎;图像分割;循环残差;自注意力机制

【中图分类号】R318;TP391

【文献标志码】A

【文章编号】1005-202X(2022)08-1041-08

Improved U-Net and its application in COVID-19 image segmentation

GU Guohao, LONG Yingwen, JI Mingming

School of Electrical and Electronic Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

Abstract: Computed tomography (CT) has become one of the most important steps in the detection of the corona virus disease 2019 (COVID-19). Aiming at the cumbersome problem of manually segmenting the ground-glass opacity area in the chest CT image, a U-shaped network model with recurrent ResNet and self-attention is proposed to realize the automatic segmentation in the lung CT images of COVID-19 patients and assist doctors in diagnosis. The improved U-Net with recurrent ResNet and self-attention is introduced to enhance the capture of feature information, thereby improving the accuracy of segmentation. The segmentation experiment on the public data set show that the Dice coefficient, sensitivity and specificity of the proposed algorithm reach 85.36%, 76.64% and 76.25%, respectively. Compared with other algorithms, the proposed algorithm has better segmentation performance.

Keywords: U-Net; corona virus disease 2019; image segmentation; recurrent ResNet; self-attention mechanism

前言

新型冠状病毒肺炎(COVID-19)已经是世界上最严重的传染性疾病之一,对全球众多行业都造成了不可估量的损失。由于其传播速度很快,世界卫生组织(WHO)于2020年宣布COVID-19大流行^[1-2]。COVID-19初期检测常用的方法为逆转录聚合酶链反应(RT-PCR)检测^[2]。RT-PCR检测因为检测结果等待时间较长且灵敏度不高等缺点,导致病例不能

及早被发现,甚至由于检测人群的大量聚集和环境的复杂性,容易通过患者和健康个体的相互作用快速传播。而且由于假阴性病例的存在,需要结合胸部影像学检查进一步明确诊断,病毒在肺部的表现是SARS-CoV-2感染的最初迹象之一^[3-5],其CT图像最明显的特征是病灶处具有毛玻璃混浊,将肺部CT征象作为诊断标准可以大幅度降低误诊概率。目前,CT是评估肺形态和检测几种相关病理的金标准成像技术^[3],随着人工智能的高速发展,基于深度学习的方案来进行肺部CT图像的自动分割可以帮助医生更有效率的进行诊断,也为患者争取了治疗的时间。本文针对COVID-19患者肺部病灶处的毛玻璃混浊分割展开研究。

目前COVID-19 CT图像分割面临以下问题:(1)深度学习高度依赖数据的数量,由于时间和隐私问题,COVID-19公开的数据集数量有限,这导致机器

【收稿日期】2021-10-26

【基金项目】国家自然科学基金(61603241)

【作者简介】顾国浩,硕士研究生,研究方向:人工智能,E-mail: gu_guohao97623@163.com

【通信作者】龙英文,博士,副教授,研究方向:人工智能,E-mail: 1825405229@qq.com;吉明明,博士,副教授,研究方向:系统辨识、控制与优化,E-mail: jimingming923@163.com

学习模型的性能不理想,容易出现过拟合。(2) COVID-19 CT图像的病灶特征随时间变化而变化,病灶区域小、边界模糊难以识别,只有少量数据集提供了分割标签,分割时图像的特征信息得不到有效传递。(3)毛玻璃混浊会与原本CT图像中的血管影及气管影等相融合,分割时容易将肺部本身结构与毛玻璃混浊区域一起分割,降低分割精度。

针对上述问题本文的研究贡献如下:(1)通过数据增强技术对数据样本进行扩充以解决数据量少的问题。同时将图像处理为统一尺寸使模型训练更高效。(2)在U-Net的跳跃连接处引入自注意力机制减少对外部信息的依赖,提升内部信息的学习,加强了对特征信息的捕捉。(3)随着网络加深,模型性能也会越来越高,加入循环残差模块使网络加深的同时加快了网络收敛的速度,解决了网络加深带来的梯度爆炸问题,此外这也加强了图像特征信息的传递。后两者的加入使得改进的U-Net网络模型在实验中能更精确地分割出与肺部组织结构相融合的毛玻璃混浊。

1 相关工作

1.1 医学图像分割

在2010年之前,用传统方法处理图像一直是主流,如阈值分割、区域分割、聚类等^[6]。但是医学图像因为语义简单、数据量少、多模态等特点,此时用传统方法对其进行分割效果不是很理想。随着深度学习的发展和神经网络模型的出现,各种基于深度学习的图像分割网络模型应运而生,FCN模型最早提出了将CNN用于端到端的图像分割^[7],其定义一个跳跃架构把来自深层、粗层的语义信息与来自浅层、细层的外观信息结合在一起,实现更准确的分割;SegNet提出了解码器结构对其较低分辨率的输入特征图进行上采样,通过最大池化消除了学习增加样本的需要且提高了分辨率^[8];U-Net通过采用编码器+解码器的结构构建了一个U型网络包括获取上下文的收缩路径和支持精确定位的对称扩展路径^[9]。这种网络通过数据增强的技术可以用非常少的图像实现端到端的训练;PSPNet金字塔式的结构,可以对不同层次的特征图进行融合,有效地在场景解析任务中产生高质量的结果^[10];RefineNet明确利用向下采样过程中可用的所有信息,使用远程剩余连接实现高分辨率预测并引入链式剩余池以一种高效的方式捕获丰富的后台上下文^[11]。

1.2 肺部CT图像分割

在肺部CT图像分割中,文献[12]提出了一种高效、自适应的肺CT图像分割方法。该算法采用自适

应均值漂移方法,通过固定带宽估计来估计带宽参数。由于核密度估计方法对带宽参数的依赖性较强,采用粒子群算法对带宽参数进行优化。该方法分割效果较好,能够对肺部小结节进行分割,并能检测出CT图像中的区域。孙华聪等^[13]提出一种基于深度卷积神经网络的胸部CT序列图像3D多尺度肺结节检测方法。组合了压缩激励单元的Res2Net网络为骨干结构,使同一层卷积具有更多的感受野,并引入了上下文增强模块和空间注意力模块。再由Res2Net网络模块和压缩激励单元组成的网络对候选结节进一步分类,以降低假阳性,获得最终结果。刘若愚等^[14]采用DetNet代替原始残差改进了YOLACT模型,解决了检测小型结节的局限性,同时引入了迁移学习解决了数据量太少而引起的过拟合问题。

文献[15]中探索了一种带有ResNet34编码器的U-Net结构来分割多个COVID-19病灶,能够处理问题的高度不平衡性质,以及COVID-19病灶(即大小)的巨大可变性。文献[16]提出了一种基于并行量子自监督网络(PQIS-Net)的集成半监督浅层神经网络框架,用于肺CT图像的全连通层自动分割并证明集成的半监督浅层框架的诊断效率(准确率、精度和AUC)是有前景的。

2 网络模型设计

2.1 网络结构

U-Net模型在2015年ISBI细胞追踪挑战赛获得冠军,如今已经成为处理医学图像的主流网络模型之一。还衍生出了U-Net++^[17]、MultiResUNet^[18]等更多高效的网络模型。原始U-Net的收缩路径重复应用两个3×3卷积,每个卷积后面接一个ReLU和一个2×2最大池操作,用于下采样,每下采样一次特征通道的数量增加一倍;而扩张路径与之相反,所以构成了一个对称结构。COVID-19数据集目前还不够完善,无论是数量还是质量对于神经网络的训练来说都是不够的。所以要利用少量样本获取较好精度是一个难题。从本实验所用数据集样本之一(图1)可以看出肺部当中的部分毛玻璃混浊与血管影、支气管影相融合,要提高分割精度变得尤为困难。

基于U-Net模型改进后的SARes2U-Net结构如图2所示。该模型在保持U-Net原有特点的基础上加入循环残差模块(Recurrent ResNet)^[19]和自注意力机制(Self-Attention, SA)^[20]。模型的总体结构依旧是编码器-解码器结构,将编码器和解码器中每一个模块的最后一个特征通道替换成循环残差模块,并且在原来Concat的过程中引入SA。循环残差模块使

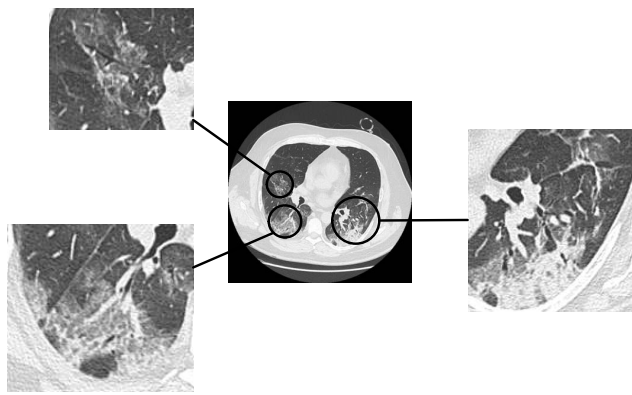


图1 样本示意图
Figure 1 One of the samples

特征信息前后传递更加流畅;自注意力机制有着对局部和全局信息的掌握性,减少特征信息的丢失,还有着并行性、参数少的特点,加快模型的训练。二者的引入都是为了加强特征信息的传递,从而缓解现有 COVID-19 数据集样本少、质量差且部分样本病灶与肺部气管血管相融的问题,提高了分割的精度。

2.1.1 循环残差模块 在残差网络出现之前,高速公路网络被用以解决训练深度较深的网络而被提出^[21],其定义为:

$$y = F(x, w_f) \cdot T(x, w_t) + x \cdot C(x, w_c) \tag{1}$$

其中, $y = F(x, w_f)$ 为常规卷积神经网络, $T(x, w_t)$ 控

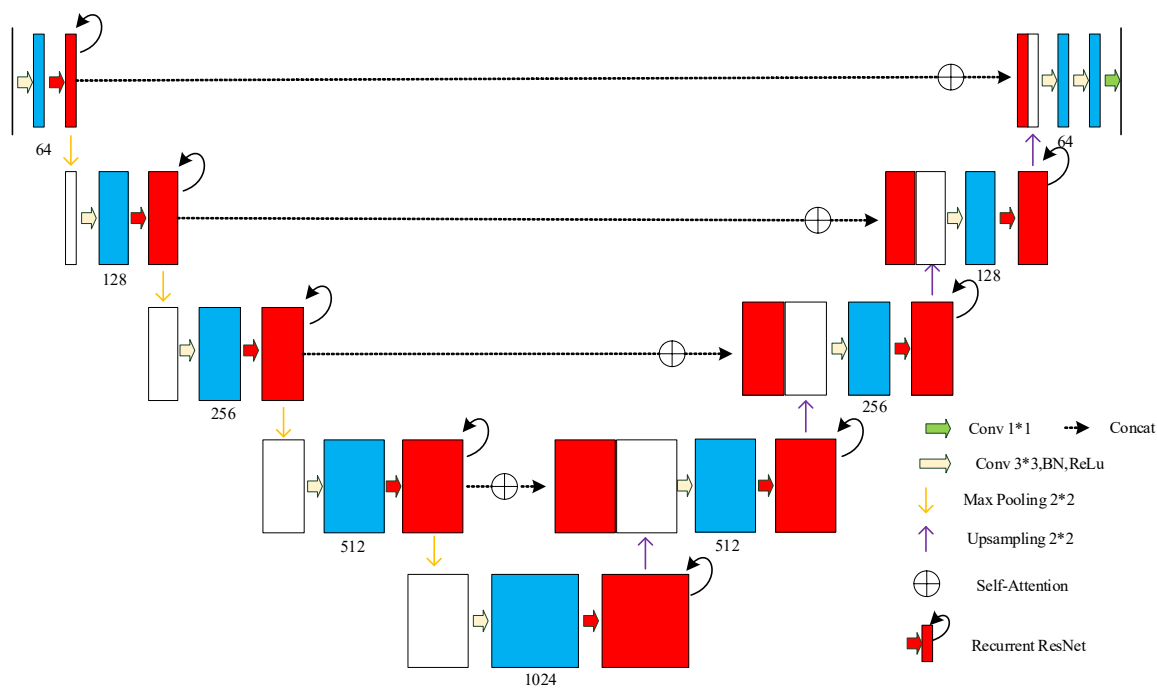


图2 SARes2U-Net结构图
Figure 2 SARes2U-Net structure

制变换强度, $C(x, w_c)$ 控制输入信号的保留强度,二者都为非线性变换。He 等^[22]提出了残差操作的思想,将高速公路网络中的两个非线性变换变为恒等映射,即令 $T(x, w_t) = 1$ 、 $C(x, w_c) = 1$, 此时就得到了残差网络,其定义为:

$$y = F(x, w_f) + x \tag{2}$$

网络的加深和加宽都会提高其性能,但是随着深度变深,问题也随之而来:梯度消失或梯度爆炸。残差网络的提出有效解决了神经网络因为深度加深而导致的梯度爆炸以及网络退化等问题,它用恒等映射代替了复杂的非线性变换,减少了参数使用从而减轻了计算负担,加快了网络收敛速度的同时也降低了过拟合风险。此外它还加强了图像特征信息

的传递。

常用的残差模块分为两种:常规残差模块和瓶颈残差模块。常规残差模块由两个 3×3 卷积层构成,在常规残差模块的基础上增加一层卷积层,并且使残差模块由两个 1×1 和一个 3×3 的卷积层构成,这就是残差模块的改进模块——瓶颈残差模块,其参数量仅为常规残差模块的 6%,上下两个 1×1 卷积层为升维或降维操作从而方便中间 3×3 的卷积层进行计算,进一步提高了计算效率。

通过在瓶颈残差模块中增加一个恒等映射,将第一次的卷积结果重新代入进行二次卷积运算从而构成了一个循环残差模块,其定义为:

$$g = F(y, w_f) + x \tag{3}$$

其中, g 为循环之后的残差预测。新构成的循环残差模块由于递归操作, 其参数量没有增加, 在计算效率没有很大影响的同时加强了特征信息的积累, 使网

络对特征信息有了更好的提取。这 3 种残差模块如图 3 所示。

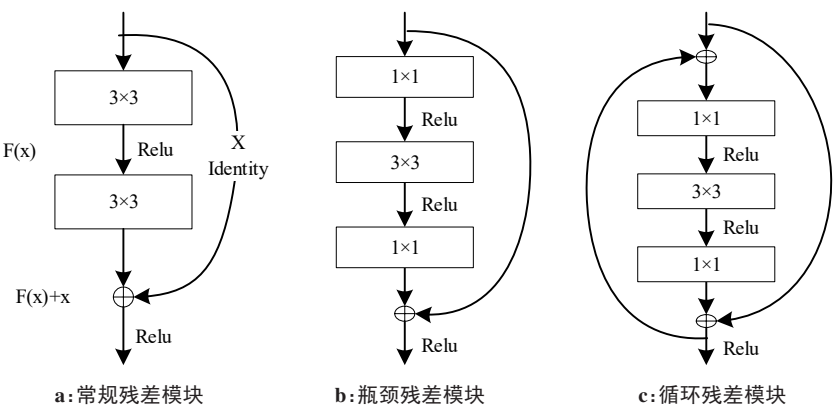


图3 3种残差模块结构图

Figure 3 Three kinds of recurrent ResNet structures

2.1.2 自注意力机制 神经网络中注意力机制的灵感来源于人类视觉中的注意力机制。当人类在观察一个场景时, 不会从场景的一边仔细地向另一边遍历似地探索, 而是依据自身的需求去关注特定的一部分, 并且如果在一个场景里频繁地出现那个想关注的事物时, 那么在下次类似场景中也会不自觉地关注那个事物。而神经网络中的注意力机制是利用权重来实现其特定的关注。

注意力机制其本质(图4)是一个查询(Q)到键值对(K - V)的映射, 也可以看作是一个软寻址操作: 把输入看作储存器中的内容, 由地址 K (键) 和值 V 构成, 通过 $K=Q$ 的查询条件取出对应的值 V 。这里不一定非要满足 $K=Q$, 只需计算 K 与 Q 的相似度即可, 常用的方法为点积:

$$F(q, k_i) = q \cdot k_i \tag{4}$$

将得到的相似度进行 SoftMax 归一化处理得到对应的注意力得分 a_i :

$$a_i = \text{softmax}(F(q, k_i)) = \frac{e^{F(q, k_i)}}{\sum_{i=1}^N e^{F(q, k_i)}} \tag{5}$$

最后对注意力得分 a_i 加权求和即可得到 Attention:

$$\text{Attention}(Q, K, V) = \sum_{i=1}^N a_i v_i \tag{6}$$

SA 特殊之处在于 K, Q, V 三者相等且在计算相似度时引入了尺度 $\sqrt{d_k}$ (式7), $\sqrt{d_k}$ 的引入使内积不会太大, 从而确保了梯度的稳定性。其结构如图5所示, 这种方法也叫放缩点积 Attention (Scaled Dot-Product Attention)^[20]:

$$\text{Self Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

其中, K^T 为 K 的转置, $Q = K = V \in R^{l \times n}$ (l 为单个向量维度, n 为向量数量)。自注意力机制减少了外部信息的依赖, 提升内部信息的学习, 加强了对特征信息的捕捉。

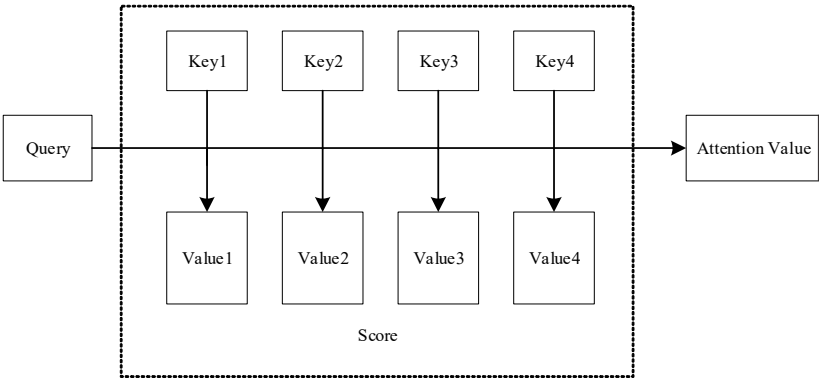


图4 注意力机制本质结构

Figure 4 Essential structure of attention mechanism

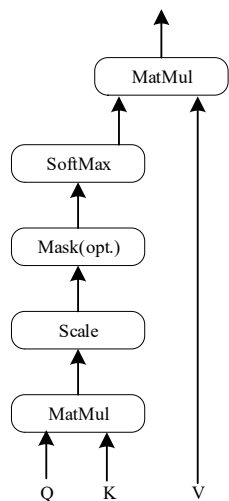


图5 自注意力机制计算结构
Figure 5 Calculation structure of self-attention mechanism

2.2 评价指标

为了客观的说明模型的综合性能,采用4种常用的评价指标来衡量其性能:准确率(Accuracy)、骰子系数(Dice)、灵敏度(Sensitivity)和特异度(Specificity)。其定义如下:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{9}$$

$$\text{Specificity} = \frac{TN}{FP + TN} \tag{10}$$

$$\text{Dice} = \frac{2|A \cap B|}{|A| + |B|} \tag{11}$$

灵敏度和特异度是在医学角度常关注的两个指标,其中TP为真阳性(预测为阳性且实际也为阳性的样本数);TN为真阴性(预测为阴性实际也为阴性的样本数);FP为假阳性(预测为阳性实际为阴性的样本数);FN为假阴性(预测为阴性实际为阳性的样本数)。Dice系数用来衡量两个集合的相似度,数值为0~1,数值越大表示集合相似度越高,这里的集合A和集合B分别表示分割结果像素构成和实际标签像素构成。

3 实验结果

3.1 数据集及实验环境

使用 COVID-19 CT segmentation dataset 作为实验训练用的数据集,此数据集分为两部分:第一部分全是 COVID-19 患者的肺部 CT 图像,来自 40 名患者总计 100 张图像,这些图像由公开的 JPG 图像转化而来;第二部分将阳性案例和阴性案例放一起,共计 829 张图像,其中 373 张被标记为阳性。初始数据集

共计 929 张图像,且所有图像都经过专业的放射科医生进行了标签分割,其中 mask 为 1 的是磨玻璃结节,mask 为 2 的是实变,mask 为 3 的是胸腔积液。

将数据集的所有图像统一处理为 256×256 的大小,并且使用数据增强(随机旋转、随机裁剪、随机翻转)来扩充训练集中的 image 和 mask,数据增强解决了训练样本不够的问题,最终数据集扩充至 2 000 张。

硬件环境:处理器为 Intel Core i7-8750H,显卡为 GeForce GTX 1070 with Max-Q Design(8 GB),操作系统为 Windows 10;软件环境:PyCharm,python3.7,实验框架为 tensorflow 和 keras。

3.2 实验对比

3.2.1 不同损失函数下的性能对比 损失函数可以用来评估模型预测结果和真实结果不一样的程度,不同的损失函数针对不同特点的数据集有着不同效果,甚至效果差距很大。本文采用了 Dice Loss、二元交叉熵(BCE)、焦点损失(Focal Loss)以及对 Dice Loss 和改进后的 BCE 进行加权求和得到的组合损失函数(Combo Loss)这 4 种损失函数进行了对比实验,实验结果如表 1 所示。

表 1 损失函数对比				
Table 1 Comparison of loss functions				
函数	准确率	Dice 系数	灵敏度	特异度
Dice Loss	0.753 4	0.826 1	0.758 8	0.748 6
BCE	0.758 1	0.831 1	0.761 5	0.755 1
Focal Loss	0.764 2	0.835 3	0.763 3	0.761 6
Combo Loss	0.771 6	0.843 6	0.772 4	0.769 8

结果显示,不同损失函数对模型的性能确实有着不一样的效果,单个损失函数不如组合损失函数的性能好。原因为 Dice Loss、BCE 和 Focal Loss 这 3 种单一损失函数在应对这种不平衡样本(不同阳性样本中的毛玻璃混浊在整张肺部 CT 图像中占比差距悬殊)表现不是很好,Dice Loss 由于本身的反向传播性差于另外两种且在训练时容易造成参数的震荡,所以表现最差;Focal Loss 作为 BCE 的变化,通过减少权重使其更容易应对困难样本的学习,使其在 3 种单一损失函数中表现最好。Combo Loss 很好的将 Dice Loss 和 BCE 结合起来,加强了反向传播的同时对困难样本有着不错的训练效果,因此其分割性能最好。

3.2.2 不同方案下的性能对比 本文对模型进行了诸多改进并选取了其中 3 种来与最终方案对比(都是基于 U-Net),分别为:只引入残差模块(ResNet);引入

残差模块和SE注意力模块(SE+ResNet)^[23];引入SA模块和残差模块(SA+ResNet)。分割性能如表2所示。实验发现,注意力机制的引入对分割性能有比较大的提升,且SA模块比SE模块的性能稍好一点,循环残差模块(Recurrent ResNet)比普通残差模块的性能要好一点。

表2 不同方案对比
Table 2 Comparison of different schemes

方案	准确率	Dice系数	灵敏度	特异度
ResNet	0.746 3	0.801 2	0.748 6	0.744 3
SE+ResNet	0.751 8	0.826 9	0.752 3	0.751 4
SA+ResNet	0.755 1	0.835 4	0.758 3	0.754 6
SA+ Recurrent ResNet	0.764 8	0.854 8	0.767 8	0.762 5

分析原因:长距离的信息传递时信息会逐渐弱化,注意力机制的出现很好的优化了这一点,其加强了对远程依赖信息的捕捉且能高效的分配信息处理资源。通过比较权重的大小来实现重要信息的聚焦,并且可以随时调整权重,有着很高的可伸缩性和稳健性。而SA是注意力机制运用的一个新阶段,所谓SA,可以理解为自己学习自己。它减少了对外部信息的依赖,更加关注特征信息的内部关系,对特征信息来说有着非常好的传递性。残差网络有效解决了因为网络加深而出现的梯度爆炸或退化问题,而循环残差通过递归让网络可以加深的同时,特征信息也有了积累,进而使分割性能更好。

3.2.3 不同模型下的性能对比 将本文的最终模型SARes2U-Net与FCN、U-Net、U-Net++和ResUNet进行了性能对比,batch_size、epoch、学习率、损失函数和激活函数等基本参数保持一致,实验结果如表3所示。结果显示,本文的方法在COVID-19分割上有着不错的效果,Dice系数、灵敏度和特异度分别比FCN模型提高了11.54%、8.75%和8.72%。

FCN模型虽然可以将深浅信息结合得到更准确的分割,但是对图像中的细节不够敏感,缺乏空间一

致性,忽略了像素之间的关系,所以分割结果不够精细。与U-Net相比,SARes2U-Net模型Dice系数、灵敏度和特异度分别提高了5.18%、3.76%和4.13%。U-Net与FCN相比,同时获取上下文的收缩路径和支持精确定位的对称扩展路径,使上下文信息可以向更高分辨率传播,此外其引入的跳跃连接也加强了信息的传递。原始的U-Net在数据处理过程中加强了镜像边缘和弹性形变操作,使其在应对细胞分割时有着非常好的效果,但由于本数据集有有的阳性案例的毛玻璃混浊非常小且数据质量也不高,训练起来比较吃力。

U-Net++利用长短连接相结合将U-Net结构填满,同时利用浅层特征和深层特征,无论是大的视野还是小的视野都能够感受到,在提升精度的同时利用深监督大幅度缩减了参数量。ResUNet解决了网络加深带来的网络退化问题,可以训练更深的网络,带来了更好的效果。本文方法与二者相比,在Dice系数上也有4.61%和3.73%的提升,对病灶分割更精确。

图6展示了4个不同CT图像上这5种算法的分割效果(每张结果图右下角的局部放大图用来提升可视化效果),可以看出每种算法分割效果都有不尽如人意的地方。这4个图像中的毛玻璃混浊依次由小变大,当混浊较小时,易受背景干扰,背景的一部分会被错当成混浊区域一起分割,FCN分割的效果最差,其余4种效果一般;随着混浊区域变多(病情较重),5种模型的分割效果逐渐趋近,但本文提出的模型在一些肺部血管影、气管等与毛玻璃混浊相接甚至叠加的区域以及边缘细节(如第4张结果图)分割效果更好。

4 结束语

CT成像是COVID-19检测中的重要一环,有效的分割出病灶对于COVID-19的治疗有着重要帮助。针对部分病患肺部CT图像上毛玻璃混浊占比小且易受背景干扰的特点,本文结合循环残差和SA,提出了一种改进的U-Net网络。循环残差在加深网络提高性能的同时有效地抑制了过拟合;SA减少对外部信息的依赖,加强对局部信息的捕捉。二者的引入都是为了保证特征信息在传递时得到有效保留,这样才能得到更好的训练效果。在对比实验中,本文算法在应对毛玻璃混浊较小或与气管/血管影相融时具有更好的分割效果。下一步将继续研究算法的改进和数据的预处理,在应对更加复杂多类的数据情况时加强分割效果,减少资源占比。

表3 不同模型对比
Table 3 Comparison of different models

模型	准确率	Dice系数	灵敏度	特异度
FCN	0.676 8	0.738 2	0.678 9	0.675 3
U-Net	0.724 8	0.801 8	0.728 8	0.721 2
U-Net++	0.723 6	0.807 5	0.722 5	0.725 2
ResUNet	0.736 4	0.816 3	0.737 8	0.735 1
SARes2U-Net	0.763 5	0.853 6	0.766 4	0.762 5

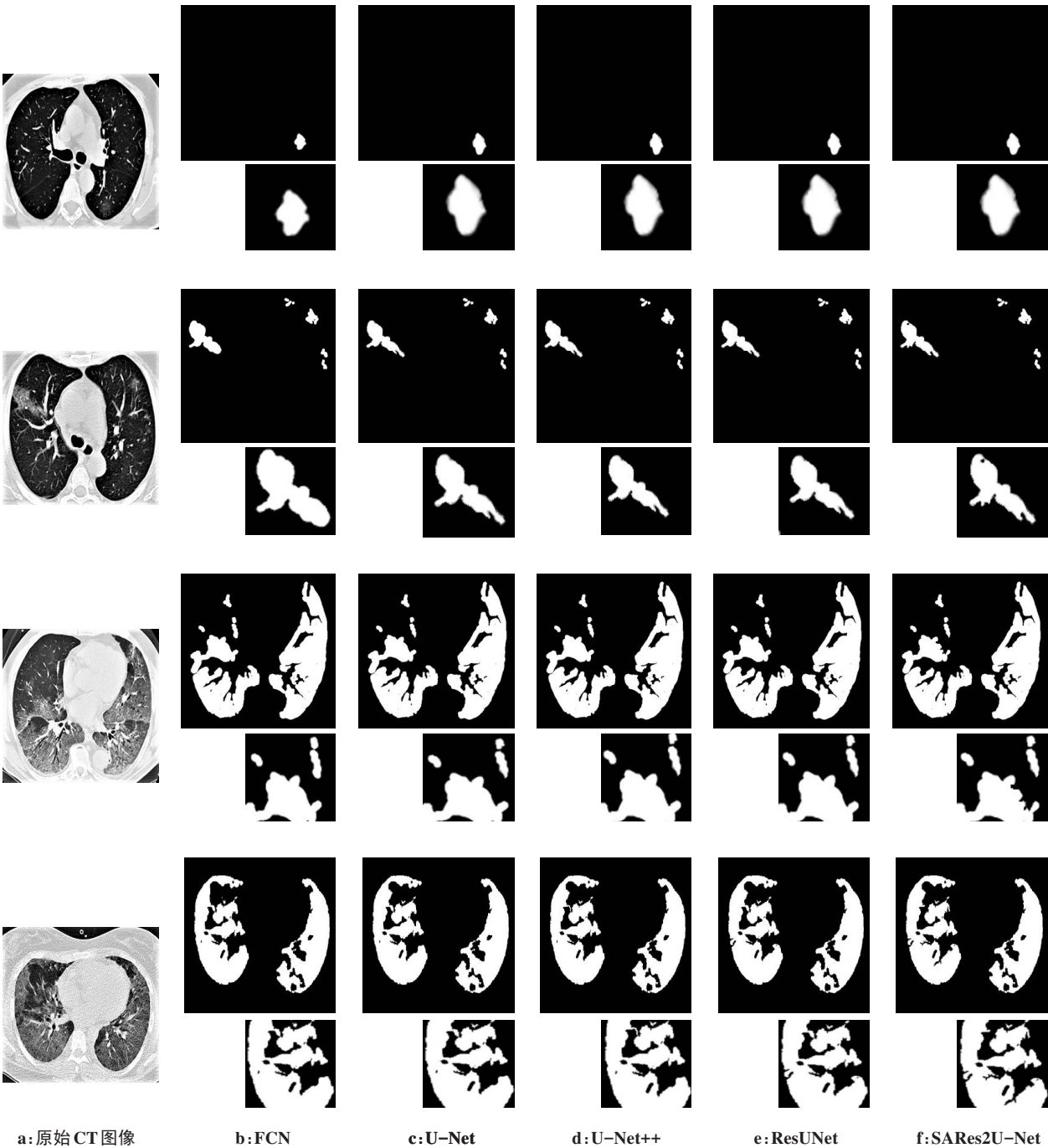


图6 不同模型的分割效果图

Figure 6 Segmentation results obtained by different models

【参考文献】

[1] Dey N, Rajinikanth V, Fong SJ, et al. Social group optimization-assisted kapur's entropy and morphological segmentation for automated detection of COVID-19 infection from computed tomography images[J]. Cognit Comput, 2020, 12(5): 1011-1023.

[2] Wang WL, Xu YL, Gao RQ, et al. Detection of SARS-CoV-2 in different types of clinical specimens[J]. JAMA, 2020, 323(18): 1843-1844.

[3] Bellani G, Mauri T, Pesenti A. Imaging in acute lung injury and acute respiratory distress syndrome[J]. Curr Opin Crit Care, 2012, 18(1): 29-34.

[4] Penarrubia L, Ruiz M, Porco R, et al. Multiple assays in a real-time RT-PCR SARS-CoV-2 panel can mitigate the risk of loss of sensitivity by new genomic variants during the COVID-19 outbreak[J]. Int J Infect Dis, 2020, 97: 225-229.

[5] Bhandary A, Prabhu GA, Rajinikanth V, et al. Deep-learning framework to detect lung abnormality-a study with chest X-ray and lung CT scan images[J]. Pattern Recogn Lett, 2020, 129: 271-278.

[6] 周莉莉, 姜枫. 图像分割方法综述研究[J]. 计算机应用研究, 2017, 34(7): 1921-1928.

[7] Zhou LL, Jiang F. Survey on image segmentation methods [J]. Application Research of Computers, 2017, 34(7): 1921-1928.

[8] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2015, 39(4): 640-651.

[9] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(12): 2481-2495.

[10] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer

- International Publishing, 2015: 234-241.
- [10] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 6230-6239.
- [11] Lin G, Milan A, Shen C, et al. RefineNet: multi-path refinement networks for high-resolution semantic segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 5168-5177.
- [12] Fazli S, Jafari M, Safaei A. Automated lung CT image segmentation using kernel mean shift analysis[C]//2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP). IEEE, 2014: 392-396.
- [13] 孙华聪, 彭延军, 郭燕飞, 等. 3D 多尺度深度卷积神经网络肺结节检测[J]. 中国图象图形学报, 2021, 26(7): 1716-1725.
- Sun HC, Peng YJ, Guo YF, et al. 3D multi-scale deep convolutional neural networks in pulmonary nodule detection[J]. Journal of Image and Graphics, 2021, 26(7): 1716-1725.
- [14] 刘若愚, 刘立波. 基于改进全卷积网络模型的肺结节检测[J]. 激光与光电子学进展, 2020, 57(16): 174-183.
- Liu RY, Liu LB. Detection of pulmonary nodules based on improved full convolution network model[J]. Laser & Optoelectronics Progress, 2020, 57(16): 174-183.
- [15] Rocha J, Pereira S, Campilho A, et al. Segmentation of COVID-19 Lesions in CT Images [C]//2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI). IEEE, 2021: 1-4.
- [16] Konar D, Panigrahi BK, Bhattacharyya S, et al. Auto-diagnosis of COVID-19 using lung CT images with semi-supervised shallow learning network[J]. IEEE Access, 2021, 9: 28716-28728.
- [17] Zhou ZW, Siddiquee MM, Tajbakhsh N, et al. UNet++: redesigning skip connections to exploit multiscale features in image segmentation [J]. IEEE Trans Med Imaging, 2020, 39(6): 1856-1867.
- [18] Ibtehaz N, Rahman MS. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Netw, 2020, 121: 74-87.
- [19] Alom MZ, Hasan M, Yakopcic C, et al. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation[J]. ArXiv, 2018, arXiv: 1802.06955.
- [20] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. ArXiv, 2017, arXiv: 1706.03762.
- [21] Srivastava RK, Greff K, Schmidhuber J. Highway networks[J]. ArXiv, 2015, arXiv: 1505.00387.
- [22] 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: 770-778.
- [23] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42(8): 2011-2023.

(编辑: 薛泽玲)