

基于结构重参数化与多尺度深度监督的COVID-19胸部CT图像自动分割

刘金平, 吴娟娟, 张 荣, 徐鹏飞
(湖南师范大学信息科学与工程学院, 湖南长沙 410006)

摘 要: 基于深度学习模型的胸部CT(Computed Tomography)图像自动分割有助于辅助医生诊疗. 但随着网络宽度与深度的加深, 网络训练困难且推理减慢. 为提高隐藏层的学习能力, 深度监督机制被用于网络训练. 但以往的深度监督方法没有考虑模型中多尺度特征图的分层表示以及上采样对参与损失计算的特征图质量的影响. 为加强隐藏层学习过程的直接性同时加快网络推理, 本文提出一种结构重参数化与多尺度深度监督分割网络(Structural Reparameterization and Multi-scale Deep Supervision Network, SR&MDS-Net), 以实现COVID-19(CORONA VIRUS DISEASE 2019)胸部CT图像的高效准确分割. 首先构建一种结构重参数化特征变异(Structure Reparameterized Feature Variation, SRFV)模块将网络的训练与推理进行解耦, 在提高模型表达能力的同时加快推理速度; 然后, 提出一种新颖的多尺度深度监督机制, 以加强网络监督效果, 提高网络性能. 在公开的COVID-19胸部CT图像数据集上进行实验, SR&MDS-Net的灵敏度、特异性、准确率、Dice分别达到了91.5%、99.5%、72.8%、80.1%, 与同类其他方法比较, 具有更优的性能.

关键词: COVID-19; 医学图像分割; 深度学习; U-Net; 结构重参数化; 深度监督学习

基金项目: 国家自然科学基金(No.61971188)

中图分类号: TP391.41

文献标识码: A

文章编号: 0372-2112(2023)05-1163-09

电子学报URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20220368

Toward Automated Segmentation of COVID-19 Chest CT Images Based on Structural Reparameterization and Multi-Scale Deep Supervision

LIU Jin-ping, WU Juan-juan, ZHANG Rong, XU Peng-fei

(College of Information Science and Engineering, Hunan Normal University, Changsha, Hunan 410081, China)

Abstract: Automatic segmentation of chest CT (Computed Tomography) images based on deep learning models is helpful to assist doctors in diagnosis and treatment. However, with the deepening of the network width and depth, the network training is difficult and the inference slows down. In order to improve the learning ability of the hidden layer, the deep supervision mechanism is used in network training. However, previous deep supervision methods did not consider the layered representation of multi-scale feature maps in the segmentation model and the influence of upsampling on the quality of feature maps involved in loss calculation. In order to improve the directness of the hidden layer learning process and speed up the reasoning, this article proposes a structural reparameterization and multi-scale deep supervision network (SR&MDS-Net) for the purpose of achieving the efficient and accurate segmentation results of lesions in COVID-19 (CORONA VIRUS DISEASE 2019) pneumonia chest CT images. Structure reparameterized feature variation (SRFV) module was constructed to decouple the network training and reasoning, which improved the model expression ability and accelerated the reasoning speed. In order to strengthen the effect of network supervision and improve the performance of network, a novel multi-scale deep supervision mechanism is proposed. The experiment was carried out on the public COVID-19 pneumonia chest CT image dataset. The sensitivity, specificity, accuracy, and Dice of SR&MDS-Net reached 91.5%, 99.5%, 72.8%, and 80.1%, respectively, which had better performance compared with other similar methods.

Key words: COVID-19; medical image segmentation; deep learning; U-Net; structure reparameterization; deep supervised learning

Foundation Item(s): National Natural Science Foundation of China (No.61971188)

1 引言

目前,准确的 COVID-19(CORONA VIRUS DISEASE 2019)胸部 CT(Computed Tomography)病灶分割是帮助医生节省读片时间、疾病分级与预后的重要手段^[1]. 纯人工手动胸部 CT 病灶分割依赖医生的临床经验,具有高强度、耗时长特点,无法满足高效、准确对 COVID-19 肺部感染的检出和评估等实际需求,因而 COVID-19 胸部 CT 病灶区域的自动分割在临床实践中变得至关重要^[2].

与自然图像相比,COVID-19 病灶边界模糊、对比度低,给 COVID-19 病灶区域自动分割带来一系列问题. 为实现 COVID-19 胸部 CT 图像病灶的准确自动分割,基于深度学习的分割方法受到国内外学者的广泛重视^[3-8]. Unet 采用端对端的体系结构,通过跳跃连接来融合多尺度信息,是目前医学图像分割领域中最流行的卷积神经网络之一^[3]. 为对 COVID-19 病灶进行有效分割,科研人员研究各种 Unet 变体网络^[4-6]. 此类变体网络通过编码端逐渐进行下采样提取目标特征,解码端不断进行上采样恢复图像尺寸,并采用跳跃连接来弥补下采样期间的信息丢失. 此外,为进一步提高复杂病灶分割精度,残差连接被引入 Unet 变体中^[7,8]. 残差连接的引入能改善权重矩阵退化,提高网络性能,但随着模型层数加深,网络难以训练. 同时网络的训练与推理结构高度内聚,减慢网络推理. 为提高隐藏层学习过程的直接性,深度监督机制被用于训练网络,来加快网络收敛速度. 目前的深度监督机制对每个监督层进

行不同倍率上采样至原始尺寸然后与标签进行损失计算,最后表明在增加深度监督层时可提高分割性能^[9-11]. 但是,图像细节信息的有效获取对于医学图像极为重要,而采用不同倍率进行上采样易造成图像细节信息的侵蚀,最终影响分割准确性^[12]. 目前的深度监督机制既没有考虑上采样过程中噪声引入对参与损失计算的特征图的影响,也没有考虑特征图的分层表示. 为提高网络性能且加快推理速度,本文提出一种新的基于结构重参数化与多尺度深度监督分割方法(Structural Reparameterization and Multi-scale Deep Supervision Network, SR&MDS-Net)来实现 COVID-19 病灶的准确分割. 其中结构重参数化特征变异(Structure Reparameterized Feather Variation, SRFV)模块使模型成为更多较浅模型的隐式集合^[13],提高模型性能的同时加快网络推理速度. 多尺度深度监督机制有效迭代优化模型,加强监督效果并均衡不同样本收敛速度.

2 本文方法

2.1 SR&MDS-Net

SR&MDS-Net 以编解码结构的 Unet 为骨干,模型结构如图 1 所示. 编码器通过 SRFV 模块、修正线性单元和下采样对图像进行编码,提取表征信息;解码端由上采样、SRFV 模块、修正线性单元组成,提取高层语义信息;引入跳跃连接来实现信息互补,辅助解码端上采样信息尺寸的恢复.

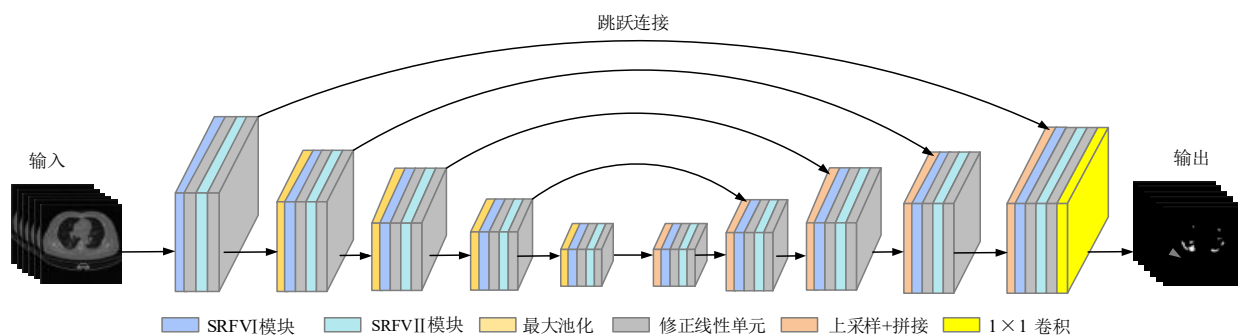


图 1 SR&MDS-Net 网络结构图

SRFV 模块在训练时为多分支拓扑结构,提高模型特征表达能力,推理时转换为单路结构,以加快网络推理速度. 在此基础上,进一步引入协调注意力(Coordinate Attention, CA)机制^[14]对可变异分支提取的特征进行水平和垂直方向分解,以综合通道和位置方向维度最具信息量的特征用于 COVID-19 胸部 CT 图像病灶分割. 这是因为医学图像的通道信息与位置信息对背景复杂、对比度低的医学图像分割具有重要作用. 不同通道映射图贡献度有差异,且每个通道的不同空间位置对目标特征的贡献率也不同.

此外,为加强网络隐藏层的表征能力,采用多尺度深度监督机制用于学习不同尺度层次表示. 对解码端中间层得到的特征信息直接与多尺度标签进行深度监督,在学习多尺度特征图层次表示的过程中,避免解码端中间层特征图不同倍率上采样对参与损失计算的特征图的侵蚀;同时引入动态权重因子,以缓解上采样过程噪声的影响及防止插值后特征图直接融合可能产生的混叠效应,从而起到增强最后层特征图的语义信息,进一步加强监督效果的目的.

2.2 结构重参数化特征变异模块

结构重参数化指训练时将重参数化结构组合到网络中,提高模型表达能力,推理时通过重参数化技巧将复杂结构进行整合,从而提高运算速度.比如,文献[15]在普通卷积层中加入额外的深度卷积,不仅加速网络训练,还在多种计算机视觉任务中取得更好效果.非对称卷积块^[16]采用三个并行的 $d \times d$ 、 $1 \times d$ 、 $d \times 1$ 核代替原来的 $d \times d$ 核,通过一维非对称卷积核来增强方形卷积核,可有效提高准确率.利用增强后的卷积核参数对原始模型进行初始化并部署,在提高模型性能的同时节省算力.随着嵌入式、移动端的发展,模型的效率与内存使用率变的越来越重要.结构重参数化技术与网络性能与模型效率问题提供一种新颖的解决思路.

受文献[13]启发,本文建立的SRFV模块包含SRFV I 二分支结构和SRFV II 三分支结构,如图2所示.SRFV模块训练时采用 1×1 分支、 3×3 分支和恒等分支的多分支拓扑结构,推理时将分支结构全部等价于单路 3×3 卷积拓扑结构,同时对可变异分支提取的特征进一步引入协调注意力机制,进行水平和垂直方向的特征分解,然后进行非线性操作以使通道信息在空间上捕获远程依赖,使网络同时关注位置信息和通道信息.

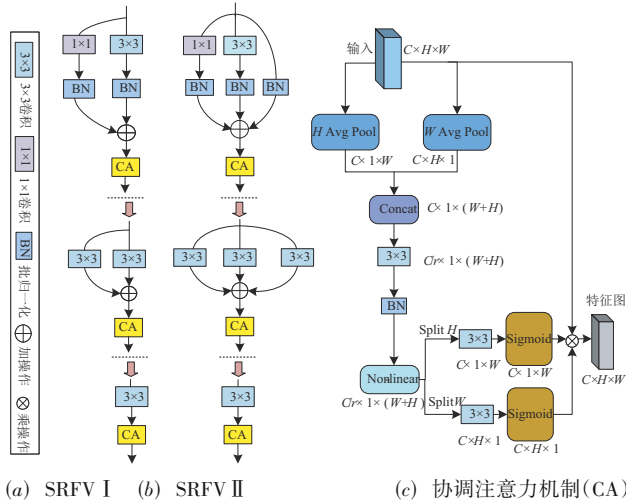


图2 结构重参数化特征变异模块图

将训练好的SRFV I、SRFV II结构重参数化多分支卷积等价转换为单个 3×3 卷积进行推理时,首先将卷积层与批归一化(Batch Normalization, BN)层融合,得到新的卷积表达式. C_1 、 C_2 分别代表输入、输出通道数; μ 、 σ 、 γ 、 β 分别代表BN层的均值、方差、缩放因子和偏置; $\mu^{(0)}$ 、 $\sigma^{(0)}$ 、 $\gamma^{(0)}$ 、 $\beta^{(0)}$ 代表恒等分支BN层的均值、方差、缩放因子和偏置; $\mu^{(1)}$ 、 $\sigma^{(1)}$ 、 $\gamma^{(1)}$ 、 $\beta^{(1)}$ 与 $\mu^{(3)}$ 、 $\sigma^{(3)}$ 、 $\gamma^{(3)}$ 、 $\beta^{(3)}$ 分别代表

因子、偏置. V 代表输入, O 代表输出, $\forall 1 \leq i \leq C_2$.

对于恒等分支BN表达式见式(1)所示:

$$BN_{\gamma, \beta}(V_i, \mu_i, \sigma_i) = (V_i - \mu_i) \frac{\gamma_i}{\sigma_i} + \beta_i \quad (1)$$

对于 1×1 卷积分支与 3×3 卷积分支,BN表达式见式(2)所示:

$$BN_{\gamma, \beta}(V_i * K_i, \mu_i, \sigma_i) = (V_i * K_i') + b_i' \quad (2)$$

其中 $K_i' = \frac{\gamma_i}{\sigma_i} K_i$, K_i 为卷积核参数, $b_i' = \beta_i - \frac{\gamma_i \mu_i}{\sigma_i}$.

对于SRFV I结构重参数化多分支模块,见式(3)所示:

$$O = BN_{\gamma^{(0)}, \beta^{(0)}}(V * K^{(1)}, \mu^{(1)}, \sigma^{(1)}) + BN_{\gamma^{(3)}, \beta^{(3)}}(V * K^{(3)}, \mu^{(3)}, \sigma^{(3)}) \quad (3)$$

则对于SRFV II结构重参数化多分支模块,见式(4)所示:

$$O = BN_{\gamma^{(0)}, \beta^{(0)}}(V * K^{(1)}, \mu^{(1)}, \sigma^{(1)}) + BN_{\gamma^{(3)}, \beta^{(3)}}(V * K^{(3)}, \mu^{(3)}, \sigma^{(3)}) + BN_{\gamma^{(0)}, \beta^{(0)}}(V, \mu^{(0)}, \sigma^{(0)}) \quad (4)$$

卷积与BN融合后,对于恒等分支可看作 1×1 卷积的单位矩阵,这样就得到一个 3×3 的卷积,两个 1×1 的卷积,和三种偏置向量.通过将 1×1 的卷积填充补零到 3×3 卷积,然后与原先 3×3 卷积对应位置相加,并将三个偏置向量进行相加,得到最终的偏置量,这样就得到最终 3×3 卷积核.具体如图3所示.

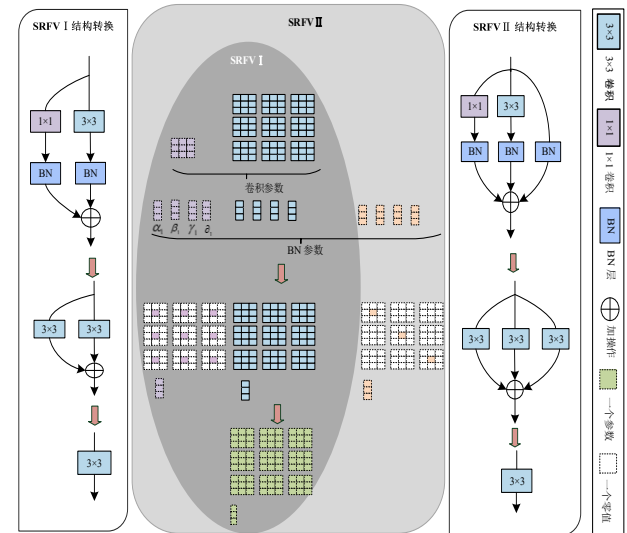


图3 结构重参数化参数转换

2.3 多尺度深度监督机制

深度监督使用单个任务标签对神经网络隐藏层进行监督,其目的是提高隐藏层学习能力,加快网络收敛速度,避免梯度消失.但目前深度监督模型既没有考虑

不同倍率上采样插值对参与损失计算的特征图的影响,也没有考虑隐藏层学习过程的分层表示.同时,如果对全部中间层特征图进行监督,将大大增加运算量.跳跃连接将表征特征与语义特征相结合,对解码端进行深度监督更有意义.为此,本文所提出一种新颖的多尺度深度监督机制,如图4所示.

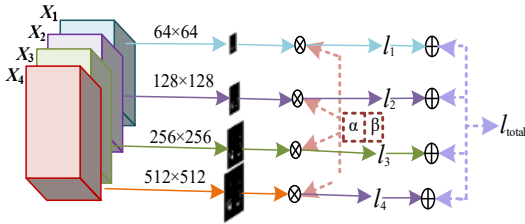


图4 多尺度深度监督机制

与一般分割网络中对最后生成的全分辨率特征图进行监督不同,SR&MDS-Net在每个解码器阶段都产生一个侧面输出,每个侧面输出由多尺度标签数据进行监督.为实现多尺度深度监督,每个解码层通过SRFV I、SRFV II模块,再经过 1×1 卷积,提取多尺度特征并进行降维,输出 X_1, X_2, X_3 .为进一步增强最后层特征的语义信息,加强深度监督效果.将解码器每一层提取的降维后的多尺度特征进行双线性上采样,得到对应全尺度特征图 X'_1, X'_2, X'_3 .全尺度特征图计算式如式(5)所示:

$$X_4 = w_0 X_4 + w_1 X'_1 + w_2 X'_2 + w_3 X'_3 \quad (5)$$

其中, w_0, w_1, w_2, w_3 为动态权重学习因子. X_4 为最后一层解码器对应的特征图.

COVID-19医学图像存在正负样本分布极不平衡,肺部磨玻璃病灶在整个图像上占比极小,使用单一损失函数进行优化时,磨玻璃病灶区域对应的梯度更新容易受到背景区域梯度影响,导致网络训练困难.为此,本文通过组合交叉熵损失函数和Dice相似系数损失,来优化模型.混合损失函数定义式如式(6)所示:

$$l_{\text{seg}} = \alpha l_{\text{ce}} + \beta l_{\text{Dice}} \quad (6)$$

其中 l_{ce} 代表交叉熵损失函数, l_{Dice} 代表Dice相似系数损失函数. α, β 为系数因子.在这篇论文中, $\alpha = 1, \beta = 0.5$ 能获得最好的性能.

对于第一、二、三、四层解码器,损失函数定义见式(7)所示:

$$\begin{cases} l_1 = l_{\text{seg}}(X_1, T_1) \\ l_2 = l_{\text{seg}}(X_2, T_2) \\ l_3 = l_{\text{seg}}(X_3, T_3) \\ l_4 = 2l_{\text{seg}}(X_4, T_4) \end{cases} \quad (7)$$

其中, T_1, T_2, T_3, T_4 依次为多尺度标签数据.

整体网络损失函数计算式为式(8)所示:

$$l_{\text{total}} = l_1 + l_2 + l_3 + l_4 \quad (8)$$

3 实验验证与结果分析

3.1 数据集描述

为验证SR&MDS-Net的分割性能,本文实验综合使用三个公开COVID-19 CT图像数据集,如表1所示.COVID-19 Dataset(<http://medicalsegmentation.com/covid19>)有三种标签:毛玻璃浑浊、实变和胸腔积液,共包含100个CT切片.MosMedData数据集^[17]包含1110卷CT影像,其中50卷共785张切片进行了二值注释.最后,有1844个切片来自Jun数据集^[18].在本实验中,对横断面不是 512×512 分辨率的切片统一调整到 512×512 大小.由于MosMedData数据集和Jun数据集的数据标注并没有区分毛玻璃浑浊、实变和胸腔积液,本文对COVID-19 Dataset的标注统一视为COVID-19病灶.数据集以6:2:2的方式划分为训练集数据、验证集数据和测试集数据.

表1 数据集描述

数据集	CT切片数量/张	病例数/个	图像大小/px
COVID-19 Dataset	100	50	512×512
MosMedData	785	~60	512×512
Jun数据集	1844	20	512×512
			630×401
Our	2792	~130	512×512

3.2 实验设置与评价指标

为了确保模型有效收敛,采用Adam优化器进行网络优化,学习率调整采用warm-up,权重正则化设置为0.0005, batch-size为8, epoch为300.模型是在pytorch框架下用python实现的, GPU为NVIDIA GTX 2080.

本文实验使用了4种用于衡量医学影像分割效果的评估指标对分割效果进行评估,包括灵敏度(sensitivity)、特异性(specificity)、精确率(precision)和Dice系数.定义见式(9)~(12)所示:

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad (9)$$

$$\text{specificity} = \frac{TN}{TN + FP} \quad (10)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (11)$$

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

其中,真阳性(True Positives, TP)指正样本中正确检测的像素数量,假阳性(False Positives, FP)指正样本中错误检测的像素数量,真阴性(True Negatives, TN)指负样本中正确检测的像素数量,假阴性(False Negatives, FN)指负样本中错误检测的像素数量.

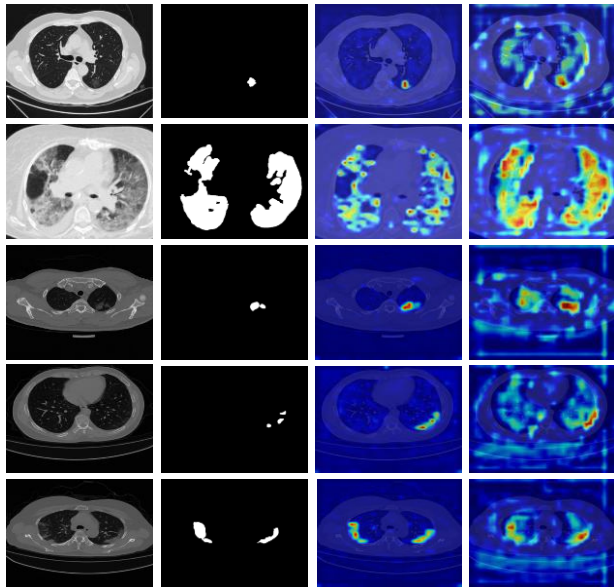
3.3 定性结果

为验证 SR&MDS-Net 训练时多分支结构与推理时单路结构模型效率问题,本文实验对网络训练与网络推理进行量化分析,如表 2 所示. 在实验中,对 512×512 大小的图像迭代 100 次,发现推理网络比训练网络速度快 7.6 fps,且参数量减少 1.98 M.

表 2 网络图像处理效率对比

网络结构	参数/M	速度/fps
训练网络	19.79	10.69
推理网络	17.81	18.29

第三层下采样块对应热力图可视化结果如图 5 所示,虽胸部 COVID-19 CT 图像背景复杂,但这些背景仍可被有效抑制,使病灶区域得到高亮显示. 也就是说,SR&MDS-Net 能有效突出重点区域,抑制无关区域,从而实现新冠病灶区域更准确的分割.



(a) 原始图像 (b) 标签数据 (c) 本方法热力图 (d) Unet 热力图
图 5 第三层下采样块热力图

为验证协调注意力模块可有效综合通道与位置维度最具信息量的特征进行分割,对具有代表性的通道 SENet (Squeeze-and-Excitation Networks)^[19]、SGE (Spatial Group-wise Enhance) 注意力^[20]、混合通道与空间的 CBAM (Convolutional Block Attention Module)^[21]、PSA (Polarized Self-Attention) 机制^[22]进行对比实验,如表 3 所示. 本文的协调注意力模块比注重在通道维度加权的 SENet 和注重空间维度加权的 SGE 空间注意力机制性能好. 与混合通道与空间的 CBAM 注意力机制、混合通道和空间维度进行双分支自注意力的 PSA 机制对比,本文的协调注意力具有更加优越的表现.

表 3 不同注意力机制对比结果 单位:%

注意力机制	灵敏度	特异性	准确率	DICE
SENet ^[19]	92.2	99.5	71.4	79.9
SGE ^[20]	92.4	99.5	71.2	79.8
CBAM ^[21]	91.4	99.5	71.7	79.5
PSA ^[22]	93.2	99.4	70.5	79.0
SR&MDS-Net	91.5	99.5	72.8	80.1

为验证多尺度深度监督机制的有效性,通过对中间层特征图进行目前常用的上采样操作,以进行深度监督对比,实验结果如表 4 所示,本文所提方法具有明显优势.

表 4 上采样方式对比结果 单位:%

上采样方法	灵敏度	特异性	准确率	DICE
最近邻插值	81.2	99.5	72.4	74.6
双线性插值	92.6	99.4	71.8	79.9
转置卷积	88.1	99.5	71.3	77.2
反池化	91.5	99.4	68.3	76.5
SR&MDS-Net	91.5	99.5	72.8	80.1

对结构重参数化、协调注意力机制、多尺度动态权重深度监督各组件的有效性进行消融实验,结果如表 5 所示. 其中骨干网络是适用于医学图像分割的 Unet 网络,Unet+SRFV-CA 为骨干网络和不含协调注意力机制的 SRFV 模块,Unet+SRFV 为骨干网络和 SRFV 模块.

表 5 消融实验 单位:%

组件	灵敏度	特异性	准确率	DICE
Unet	90.8	99.2	64.5	74.0
Unet+SRFV-CA	89.2	99.4	67.7	75.8
Unet+SRFV	90.6	99.3	67.7	76.0
SR&MDS-Net	91.5	99.5	72.8	80.1

对结构重参数化中每条分支中使用 BN 层与分支之后使用 BN 层的情况进行实验,结果如表 6 所示,每条分支中使用 BN 层可直接将卷积与 BN 进行融合,将数据转换在均值为零,方差为 1 的正态分布状态下,既获得较强非线性表达能力,又避免网络收敛速度慢的问题.

表 6 BN 层位置消融对比 单位:%

BN	灵敏度	特异性	准确率	DICE
Inner BN +CE	89.7	99.4	72.1	77.8
InnerBN+CE+Dice	88.4	99.5	73.3	78.0
InnerBN+CE+0.5Dice	91.5	99.5	72.8	80.1
ExternalBN+CE	90.8	99.4	70.2	77.8
ExternalBN+CE+Dice	89.7	99.4	68.5	76.5
ExternalBN+CE+0.5Dice	88.1	99.4	72.2	76.9

3.4 与现有方法对比

3.4.1 与现有 COVID-19 CT 分割方法对比

为验证 SR&MDS-Net 性能的优越性,与目前先进的 COVID-19 CT 分割模型进行实验对比,如 MiniSeg^[23]、SAU-Net++ (Squeeze excitation residual Atrous spatial pyramid pooling U-Net++)^[24]、D2A U-Net (Dilated Dual Attention U-Net)^[25]、TFCNs (Transformers for Fully Convolutional denseNets)^[26]。对比结果如表 7 所示, SR&MDS-Net 在灵敏度、特异性、准确率、Dice 分别达到了 91.5%, 99.5%, 72.8%, 80.1%。

表 7 SR&MDS-Net 与先进 COVID-19 分割模型对比 单位: %

Method	灵敏度	特异性	准确率	DICE
Miniseg ^[23]	91.0	99.5	72.8	79.6
SAUNet++ ^[24]	80.8	99.3	66.9	71.3
D2A U-Net ^[25]	90.0	99.3	68.9	75.7
TFCNs ^[26]	73.0	99.5	70.1	68.6
SR&MDS-Net	91.5	99.5	72.8	80.1

Miniseg 是一种用于 COVID-19 CT 图像分割的端到端深度学习模型,与 SR&MDS-Net 结果最为接近。该模型编码端由双路径分支通过一系列嵌套的跳跃路径连接而成,其中注意层次空间金字塔模块可有效进行多尺度学习。与本文的模型不同,该模型分支训练与推理解耦高度内聚,无法解耦。此外,Miniseg 虽然也进行了深度监督,但该监督机制没有考虑不同倍率上采样噪声引入对参与损失计算的特征图的影响,也没有考虑隐藏层学习过程的分层表示。

SAUNet++ 采用了一种压缩激励残差模块,该模块将 SENet 通道注意力机制与残差连接相结合,可将更多的权重分配给更重要的通道,并缓解梯度消失的问题。与之不同的是:本文使用协调注意力机制来综合通道和位置最具信息量的特征进行分割,因而能获得病灶区域更关键性特征,获得更好的分割结果。

D2A U-Net 和 TFCNs 都是端到端的 COVID-19 病灶分割模型。D2A U-Net 是一种全卷积神经网络模型,而 TFCNs 是基于卷积和 Transformer 构建起来的分割模型。D2A U-Net 编码端采用 ResNext-50 为特征提取器。

跳跃连接通过注意门控模块增强重点区域。解码端组合膨胀卷积和解码注意力模块,解码注意力模块与跳跃连接处的门控单元类似。与本文的协调注意力不同, D2A U-Net 使用 SENet 来对通道维度进行加权。通过 1×1 卷积进行降维,并使用不同尺寸卷积核进一步提取空间位置特征。D2A U-Net 的注意力模块设计复杂,计算量大。TFCNs 主要是由密集连接块、ResLinear-Transformer、卷积线性注意块组成。其中在跳跃连接处使用卷积线性注意块可增强分割区域特征,编码层末端的 ResLinear-Transformer 有助于增强模型的感受野,但是每一层都使用全连接操作,将大大增加参数量。COVID-19 数据和对应的注释标签获取成本高且困难,易造成模型过拟合。

3.4.2 与经典传统方法对比

同时,为进一步验证 SR&MDS-Net 模型性能,将其与 SegNet^[27]、PSPNet (Pyramid Scene Parsing Network)^[28]、Attention U-Net^[29]、R2U-Net (Recurrent Residual convolutional neural network based on U-Net models)^[30]、不含深度监督的 Unet3+^[31]、含深度监督的 Unet3+^[31] 六种传统方法进行实验对比。所有实验均在相同环境和相同数据划分数据集进行,实验结果如表 8 所示,SR&MDS-Net 在灵敏度、特异性、准确率、Dice 上均表现优异,且 SR&MDS-Net 的计算量 (FLOPs) 与参数量要更少、模型复杂度低。

与 Segnet 网络相比,在灵敏度、特异性、准确率、Dice 上分别提升 4.8 个百分点、0.5 个百分点、15.2 个百分点、12.4 个百分点。这是因为 Segnet 没有充分利用各层网络所提取的特征图,只是单纯进行上采样,忽略了全局上下文信息。

PSPNet 模型主要以金字塔池化为核心,聚合了上下文信息,提高了获取全局信息的能力,但空洞卷积的计算方式,有可能产生棋盘效应,进而影响分割效果。与其相比,本文方法在灵敏度、特异性、准确率、Dice 上分别提升了 3.8 个百分点、0.2 个百分点、6.3 个百分点、6.9 个百分点。

Attention U-Net 通过在 Unet 结构基础上增加注意力门控单元,让模型在上采样过程中通过高维特征信

表 8 SR&MDS-Net 与传统模型对比

Method	灵敏度/%	特异性/%	准确率/%	DICE/%	FLOPs/G	参数量/M
Segnet ^[27]	86.7	99.0	57.6	67.7	160.56	29.44
PSPNet ^[28]	87.7	99.3	66.5	73.2	159.48	27.50
Attention U-Net ^[29]	88.5	99.3	65.1	74.2	333.53	40.11
R2U-Net ^[30]	83.7	98.4	48.6	59.5	611.69	39.09
Unet3+ ^[31]	90.6	99.2	63.7	73.2	799.72	26.97
Unet3+_DeepSupvise ^[31]	92.7	99.2	64.1	74.6	800.26	27.01
SR&MDS-Net	91.5	99.5	72.8	80.1	157.98	17.81

息对低维特征进行指导,但高维特征与低维特征存在语义鸿沟,有可能导致信息丢失.与之相比,本文方法在灵敏度、特异性、准确率、Dice上分别提升了3个百分点、0.2个百分点、7.7个百分点、5.9个百分点.

R2U-Net将残差连接和循环卷积结合起来,替换了原来Unet模型中普通的卷积.残差单元有助于提高网络性能,循环残差卷积层进行特征积累,可以更好地表示分割任务的特征.但高内聚的多分支结构使网络训练难度加大,且不利于网络推理.与本文方法相比,SR&MDS-Net在灵敏度、特异性、准确率、Dice上分别提升了7.8个百分点、1.1个百分点、24.2个百分点、20.6个百分点.

Unet3+充分利用全尺度跳跃连接来结合不同尺度特征图的底层细节和高层语义来提高分割精度;同时该论文提出了一种全尺度特征图深度监督机制,解码端每一侧都产生一个全尺度特征预测,然后由标签数据进行监督,以此实现对隐藏层的监督.本文对该模型进行两方面对比:不含深度监督的Unet3+和含深度监督的Unet3+.实验结果表明深度监督机制可以提高隐藏层的学习能力.本文方法在特异性、准确率、Dice指标上均有提升.

为清晰的展示所提方法与以上方法的实验效果,对测试集中的六张图像进行了可视化显示,结果如图6所示.由图6可知,本文方法具有更好的分割效果.

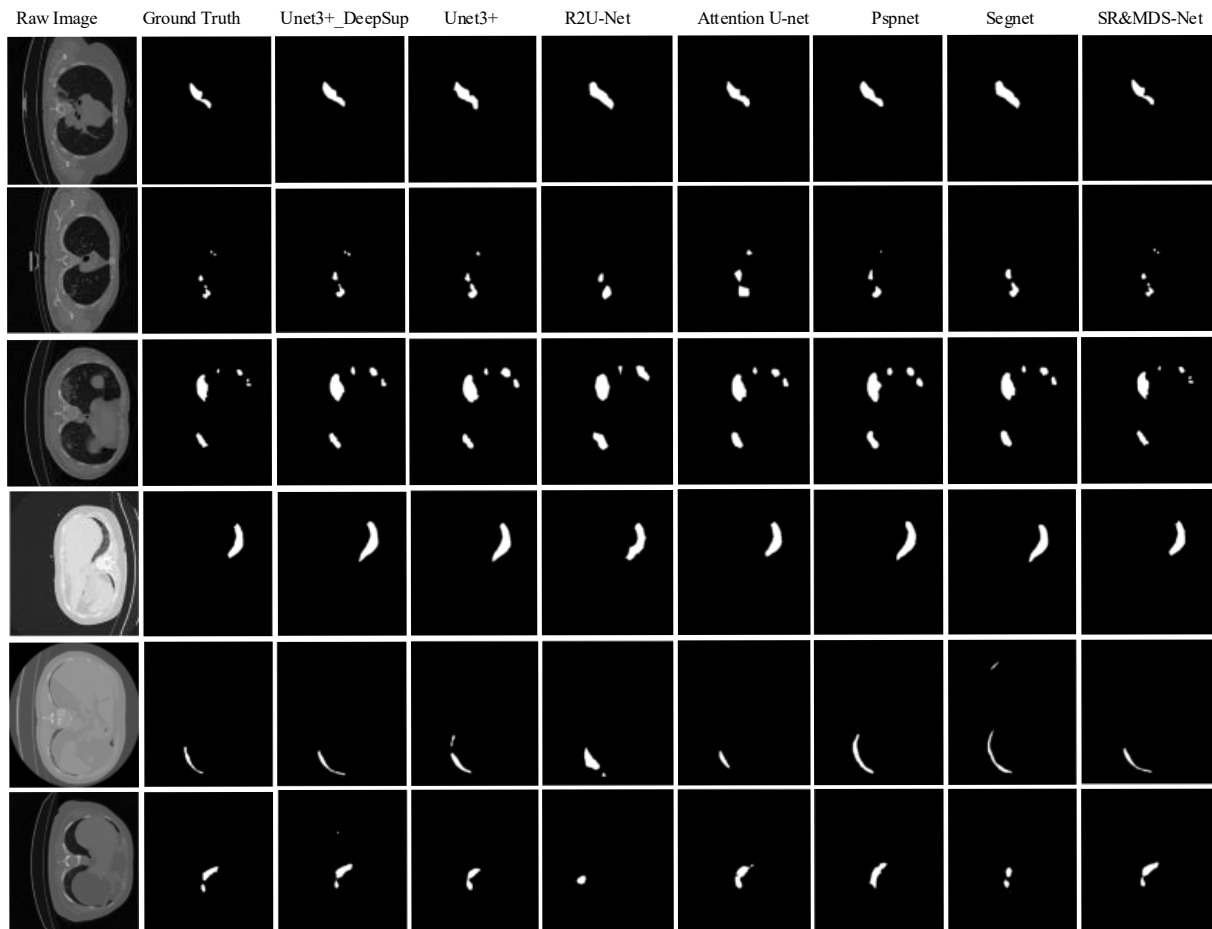


图6 不同模型在测试集上的分割结果

4 总结

针对多分支结构难以训练且不利于推理的问题,本文提出了一种新的基于结构重参数化与多尺度深度监督的 COVID-19 分割网络,结果表明该网络具有潜在临床应用价值.然而,目前大型开源的 COVID-19 分割数据集仍然缺乏,SR&MDS-Net的鲁棒性需要更多数据

集进行实验验证.同时,SR&MDS-Net并没有考虑 CT 图像的三维信息,CT 图像空间信息对 COVID-19 病灶分割很重要.而数据集有限,三维分割易造成过拟合,后续将进一步考虑 2.5 维分割,以改进模型性能.

参考文献

[1] 孙捷,张永梅.基于动静态特征双输入神经网络的咳嗽

- 声诊断 COVID-19 算法[J]. 电子学报, 2023, 51(1): 202-212.
- SUN Jie, ZHANG Yong-mei. A dynamic-static dual input deep neural network algorithm for diagnosing COVID-19 by cough[J]. Acta Electronica Sinica, 2023, 51(1): 202-212. (in Chinese)
- [2] SHI F, WANG J, SHI J, et al. Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19[J]. IEEE Reviews in Biomedical Engineering, 2021, 14: 4-15.
- [3] 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.
- [4] 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.
- [5] PALURU N, DAYAL A, JENSSEN H B, et al. Anam-Net: Anamorphic depth embedding-based lightweight CNN for segmentation of anomalies in COVID-19 chest CT images [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(3): 932-946.
- [6] GAO K, SU J, JIANG Z, et al. Dual-branch combination network (DCN): Towards accurate diagnosis and lesion segmentation of COVID-19 using CT images[J]. Medical Image Analysis, 2021, 67: 101836.
- [7] WANG X, YUAN Y, GUO D, et al. SSA-Net: Spatial self-attention network for COVID-19 pneumonia infection segmentation with semi-supervised few-shot learning[J]. Medical Image Analysis, 2022, 79: 102459.
- [8] LIU J, DONG B, WANG S, et al. COVID-19 lung infection segmentation with a novel two-stage cross-domain transfer learning framework[J]. Medical Image Analysis, 2021, 74: 102205.
- [9] ZHU Q, DU B, TURKBHEY B, et al. Deeply-supervised CNN for prostate segmentation[C]//2017 International Joint Conference on Neural Networks (IJCNN). Anchorage: IEEE, 2017: 178-184.
- [10] 陆倩杰, 柏正尧, 樊圣澜, 等. COVID-19肺部CT图像多尺度编解码分割[J]. 中国图象图形学报, 2022, 27(3): 827-837.
- LU Qian-jie, BAI Zheng-yao, FANG Sheng-lan, et al. Multiscale codec network based CT image segmentation for human lung disease derived of COVID-19[J]. Journal of Image and Graphics, 2022, 27(3): 827-837. (in Chinese)
- [11] MAJI D, SIGEDAR P, SINGH M. Attention Res-UNet with guided decoder for semantic segmentation of brain tumors[J]. Biomedical Signal Processing and Control, 2022, 71: 103077.
- [12] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [13] DING X, ZHANG X, MA N, et al. RepVGG: Making VGG-style convnets great again[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13733-13742.
- [14] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13713-13722.
- [15] CAO J, LIY, SUN M, et al. DO-Conv: Depthwise over-parameterized convolutional layer[J]. IEEE Transactions on Image Processing, 2022, 31: 3726-3736.
- [16] DING X, GUO Y, DING G, et al. ACNET: Strengthening the kernel skeletons for powerful cnn via asymmetric convolution blocks[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 1911-1920.
- [17] MOROZOV S, ANDREYCHENKO A, PAVLOV N, et al. Mosmeddata: Chest CT scans with COVID-19 related findings dataset[EB/OL]. (2020-05-13) [2022-04-07]. <https://arxiv.org/abs/2005.06465>.
- [18] JUN M, CHENG G, YIXIN W, et al. COVID-19 CT lung and infection segmentation dataset (Verson 1.0)[DS/OL]. (2020-04-20) [2022-04-07]. <https://zenodo.org/record/3757476#.Y3wvnnZBxPY>.
- [19] HUJ, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [20] LI X, HU X, YANG J. Spatial group-wise enhance: Improving semantic feature learning in convolutional networks[EB/OL]. (2019-05-25) [2022-04-07]. <http://arxiv.org/abs/1905.09646>.
- [21] WOO S, PARK J, LEE J, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 3-19.
- [22] LIU H, LIU F, FAN X, et al. Polarized self-attention: To-

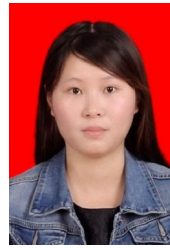
wards high-quality pixel-wise regression[EB/OL]. (2021-07-08)[2022-04-07]. <https://arxiv.org/abs/2107.00782>.

- [23] QIU Y, LIU Y, LI S, et al. MiniSeg: An extremely minimum network for efficient COVID-19 segmentation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(6): 4846-4854.
- [24] XIAO H, RAN Z, MABU S, et al. SAUNet++: An automatic segmentation model of COVID-19 lesion from CT slices[J/OL]. The Visual Computer, 2022: 1-14. DOI: <https://doi.org/10.1007/s00371-022-02414-4>.
- [25] ZHAO X, ZHANG P, SONG F, et al. D2A U-Net: Automatic segmentation of COVID-19 CT slices based on dual attention and hybrid dilated convolution[J]. Computers in Biology and Medicine, 2021, 135: 104526.
- [26] LI Z, LI D, XU C, et al. TFCNs: A CNN-Transformer hybrid network for medical image segmentation[C]//Artificial Neural Networks and Machine Learning-ICANN 2022. Cham: Springer, 2022: 781-792.
- [27] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [28] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 6230-6239.
- [29] OKTAY O, SCHLEMPER J, FOLGOC L, et al. Attention U-Net: Learning where to look for the pancreas[EB/OL]. (2018-05-20)[2022-04-07]. <http://arxiv.org/abs/1804.03999>.
- [30] ALOMMZ, HASAN M, YAKOPCIC C, et al. Recurrent-residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation[EB/OL]. (2018-05-29)[2022-04-07]. <https://arxiv.org/abs/1802.06955>.
- [31] 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). Barcelona: IEEE, 2020: 1055-1059.

作者简介



刘金平 男,1983年9月出生于湖南省邵阳市.现为湖南师范大学信息科学与工程学院教授、博士生导师,主要从事智能信息处理、过程检测、故障诊断等相关研究.
E-mail: ljp202518@163.com



吴娟娟 女,1996年5月出生于湖南省株洲市.目前在湖南师范大学信息科学与工程学院攻读硕士学位.研究方向为生物医学图像处理.
E-mail: Ajuan0527@hunnu.edu.cn



张荣 女,1999年12月出生于湖南省长沙市.目前在湖南师范大学信息科学与工程学院攻读硕士学位.研究方向为生物医学图像处理.
E-mail: cpstzr@163.com



徐鹏飞(通讯作者) 男,1977年9月出生于湖南省岳阳市.现为湖南师范大学信息科学与工程学院副教授,主要从事智能信息处理方面的研究.
E-mail: xupf@hunnu.edu.cn