

# 基于改进DeepLabV3+的COVID-19肺部CT图像语义分割方法

王光宇<sup>1</sup>, 赵曙光<sup>1</sup>, 张笑青<sup>2</sup>, 郭力争<sup>3</sup>

<sup>1</sup>东华大学信息科学与技术学院, 上海

<sup>2</sup>南京理工大学泰州科技学院, 江苏 泰州

<sup>3</sup>河南城建学院计算机与数据科学学院, 河南 平顶山

收稿日期: 2021年11月28日; 录用日期: 2021年12月24日; 发布日期: 2021年12月30日

## 摘要

肺部CT图像是确诊新冠肺炎的必要参考, 但其人工判读的效率较低且专家不足。对CT图像进行高效、准确分割是(半)自动诊断的基础, 本文基于新冠肺炎患者肺部CT图像及其掩膜图构建了数据集, 根据实验选用和改进DeepLabV3+模型, 包括在特征拼接中引入卷积注意力模块, 用深度可分离卷积替换常规卷积, 获得了实验性能突出的语义分割改进模型。

## 关键词

深度学习, 语义分割, 医学影像, 新冠肺炎

# Semantic Segmentation Method of COVID-19 Lung CT Images Based on Improved DeepLabV3+

Guangyu Wang<sup>1</sup>, Shuguang Zhao<sup>1</sup>, Xiaoqing Zhang<sup>2</sup>, Lizheng Guo<sup>3</sup>

<sup>1</sup>College of Information Science and Technology, Donghua University, Shanghai

<sup>2</sup>Taizhou Institute of Sci.&Tech., Nanjing University of Science and Technology, Taizhou Jiangsu

<sup>3</sup>School of Computer and Data Science, Henan University of Urban Construction, Pingdingshan Henan

Received: Nov. 28<sup>th</sup>, 2021; accepted: Dec. 24<sup>th</sup>, 2021; published: Dec. 30<sup>th</sup>, 2021

## Abstract

CT images of the lungs are necessary references for diagnosing COVID-19, but the efficiency of manual judgment is low and there are insufficient experts. In order to use deep learning to improve its diagnostic efficiency and accuracy, this paper constructs a data set consisting of real lung CT images of patients with COVID-19 and corresponding masks. According to the experiment, the DeepLabV3+ model is selected, we introduced convolutional attention module in the feature stitching and used depth separable convolution to replace the original conventional convolution. Finally, we obtained an improved semantic segmentation model with outstanding experimental performance.

## Keywords

Deep Learning, Semantic Segmentation, Medical Imaging, COVID-19

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

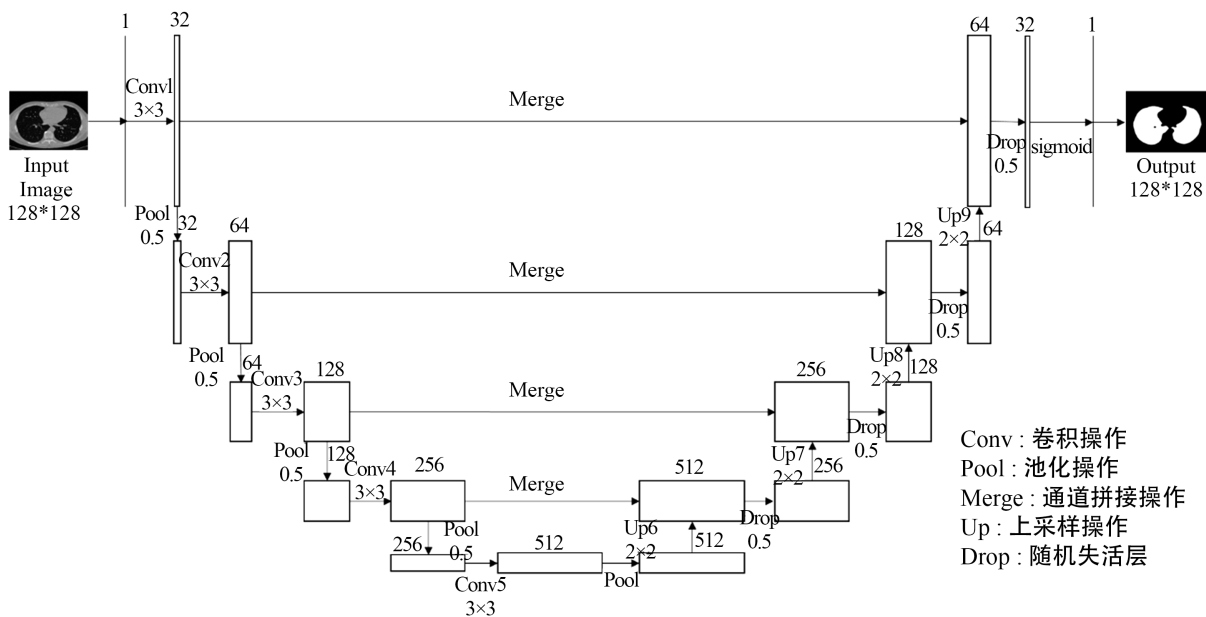
## 1. 引言

核酸检测是大规模筛查新冠病毒感染的主要手段,但新冠肺炎的最终确诊还必须依据对患者肺部 CT 图像的准确判读。大多数新冠肺炎患者的肺部 CT 图像具有共同的特征,包括早期以磨玻璃影为主、晚期以实变影为主[1],故专业医生可通过判读做出诊断。但实际上由于该类图像常存在病灶重叠、背景干扰且分辨率低,医生仅靠肉眼常难以正确判读;时间紧、任务重、专家少的严峻抗疫形势又会进一步增加人工判读因医生疲劳出现诊断错误的可能性。深度学习在医学图像自动分析等的研究、应用上均已展现成功[2],故在实现新冠肺炎高效、准确诊断上也被寄予了厚望。图像语义分割则是其中最基础和重要的步骤。

本文主要研究基于深度学习的新新冠肺炎肺部 CT 图像的图像语义分割,包括:1) 利用新冠肺炎患者肺部 CT 图像及其掩模图(专家标注)构建数据集,并加以预分割。2) 实验比较 U-Net、U-Net++、U-Net+ResNet101、DeepLabV3+的分割性能。3) 选用和改进其中性能较优的 DeepLabV3+,包括引入卷积注意力模块、用深度可分离卷积替换常规卷积,以提高其获取注意力信息的能力和对信息进行去耦。通过实验和主、客观评价、对比,证明了该改进模型的可行性、先进性和有效性。

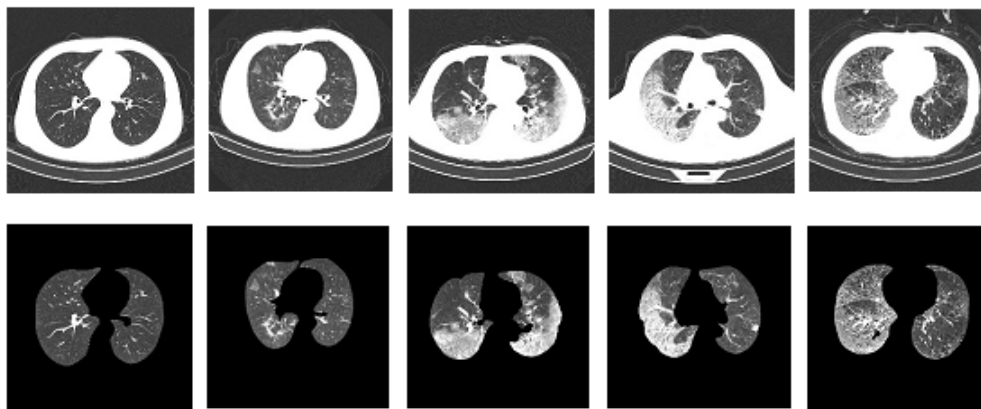
## 2. 数据集的准备

组合使用 Kaggle COVID-19 CT 图像分割竞赛、国家生物信息中心网站 2019 新冠病毒信息库提供的新新冠肺炎患者肺部 CT 图像及其掩模图,本文建立了由 1000 张肺部 CT 图像及其四标签掩模图组成的基本数据集,其图像大小为  $512 \times 512 \times 3$ ,格式为.jpg,掩模图大小为  $512 \times 512$ ,格式为.png。掩模图中标记的区域分别为:0-“肺部磨玻璃影”,1-“肺实变”,2-“肺部其他区域”,3-“背景”。获得基本数据集后,又从国家生物信息中心网站的肺部区域分割数据集中筛选出 2500 张肺部 CT 图像及对应两标签掩模图(肺部和背景)用以训练 LungSeg-Net 模型以实现对基本数据集中 CT 图像的预分割处理,分割预处理能够降低背景对实验的干扰、更好地检验、评价和比较模型、算法的性能。LungSeg-Net 模型的结构见图 1。



**Figure 1.** The structure diagram of LungSeg-Net  
**图 1.** LungSeg-Net 模型结构图

LungSeg-Net 结构类似于 U-Net 模型，模型结构承 U 型，模型左侧部分主要对图像进行下采样操作，使图片尺寸减小，并对特征信息进行提取，右侧部分进行上采样操作将图片恢复至原来的尺寸得到分割结果。模型中引入跳跃连接结构，以将高级特征与低级特征进行融合从而获得更丰富的语义信息。我们按训练集、验证集、测试集 6:2:2 的比例对 LungSeg-Net 进行实验，得出了较好的效果，因此可将其直接用于肺部 CT 图像的预分割中，分割预处理效果参见图 2。



**Figure 2.** Partial results of segmentation preprocessing of CT concentrated images  
**图 2.** 将 CT 集中的图像进行分割预处理后的部分结果

为避免过拟合，又通过移动、旋转、缩放等方式将四标签的基本数据集扩增至 5000 张，并按 6:2:2 的比例分配训练集、验证集和测试集。

### 3. 基于改进 DeepLabV3+网络的肺部 CT 分割算法

本文基于上述数据集，对 U-Net、U-Net++、U-Net+ResNet101、DeepLabV3+等的分割性能进行了实

验比较,从而选用其中较优者 DeepLabV3+作为改进对象。DeepLabV3+的特点是以 Xception 网络[3]为主干网络,并引入了空洞空间金字塔池化结构。空洞卷积能增大感受野,提升该模型捕获上下文信息的能力[4];改变其采样率,可获得多尺度的上下文信息[5]。因该模型涉及通道融合操作,融合特征的卷积深度影响着其抽象程度及后续的特征学习。本文因此将图 3 所示的卷积注意力模块(Convolutional Block Attention Module, CBAM)引入融合操作,以得到更全面可靠的注意力信息[6]。具体地说,见图 4,通道注意力模块(Channel Attention Module, CAM)的输入  $F(H \times W \times C)$  经过全局最大池化和全局平均池化后得到的两个  $1 \times 1 \times C$  的特征图,通过共享神经网络后,再经过加和操作和 sigmoid 激活,得到通道注意力特征;空间注意力模块(Spatial Attention Module, SAM)的输入同样先经过全局平均池化和最大池化,再将所得两个特征图拼接,然后经过  $7 \times 7$  卷积进行降维,最后经 sigmoid 激活生成空间注意力特征;CAM、SAM 的输出相乘,作为 CBAM 的最终输出。

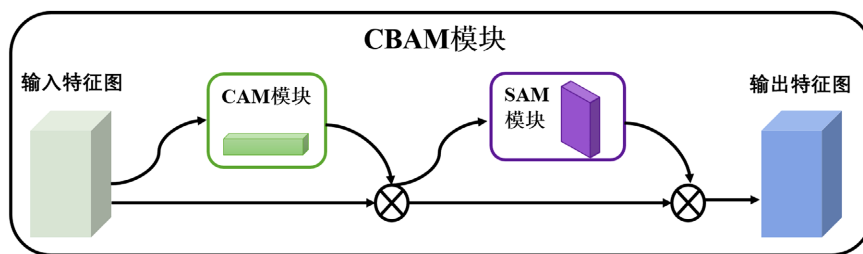


Figure 3. The structure diagram of CBAM  
图 3. CBAM 结构图

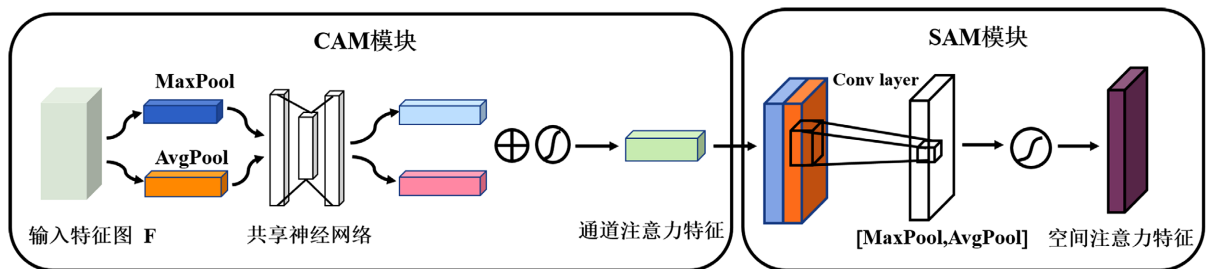


Figure 4. The structure diagrams of CAM and SAM  
图 4. CAM 与 SAM 结构

此外,为进行信息去耦,减少信息丢失[7],本文将 DeepLabV3+解码器中的常规卷积代之以深度可分离卷积,即在进行深度卷积时,一个卷积核对一个通道操作,之后通过逐点卷积将深度卷积的结果进行加权组合[8]。改进后的 DeepLabV3+模型结构参见图 5。模型编码器部分通过 ASPP 块以对多尺度图像信息进行编码,而解码器模块通过一系列反卷积以及特征图拼接操作得出细化的分割结果。由于在特征图的通道拼接操作中引入了 CBAM 模块并在输出部分引入深度可分离卷积块,从而进一步提高了模型捕获特征信息的能力,有助于得到更加优秀的分割结果。

#### 4. 实验及结果分析

实验基于 Google Colaboratory 进行,采用: TensorFlow: 2.6.0; 操作系统: Ubuntu 18.04.5; GPU: Tesla K80; 显存 12 GB; 硬盘 68 GB; Python: 3.7; CUDA: 11.2。首先对改进 DeepLabV3+模型等进行了训练和验证。训练集和验证集的 Loss、f-score 曲线参见图 6,从中可见 DeepLabV3+训练到约 40 代已达收敛。

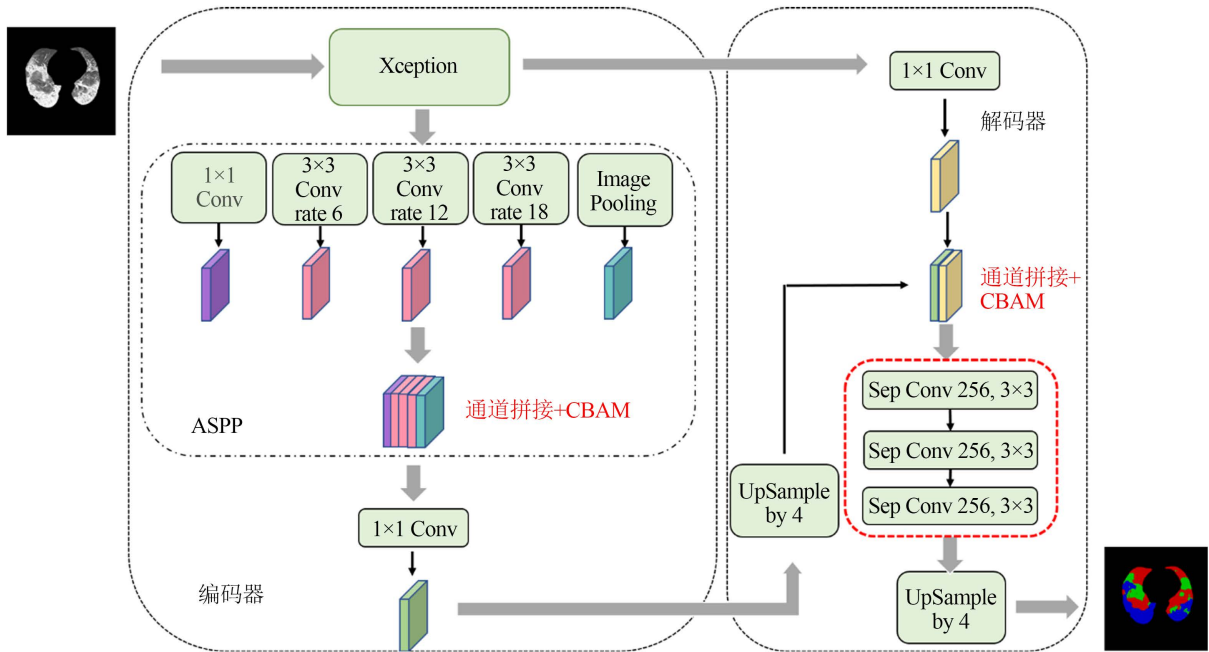


Figure 5. The structure diagram of improved DeepLabV3+  
图 5. 改进的 DeepLabV3+模型

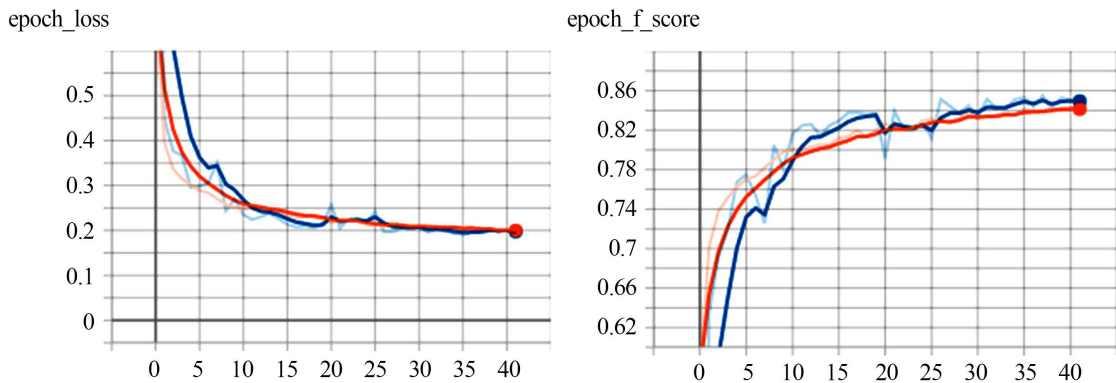
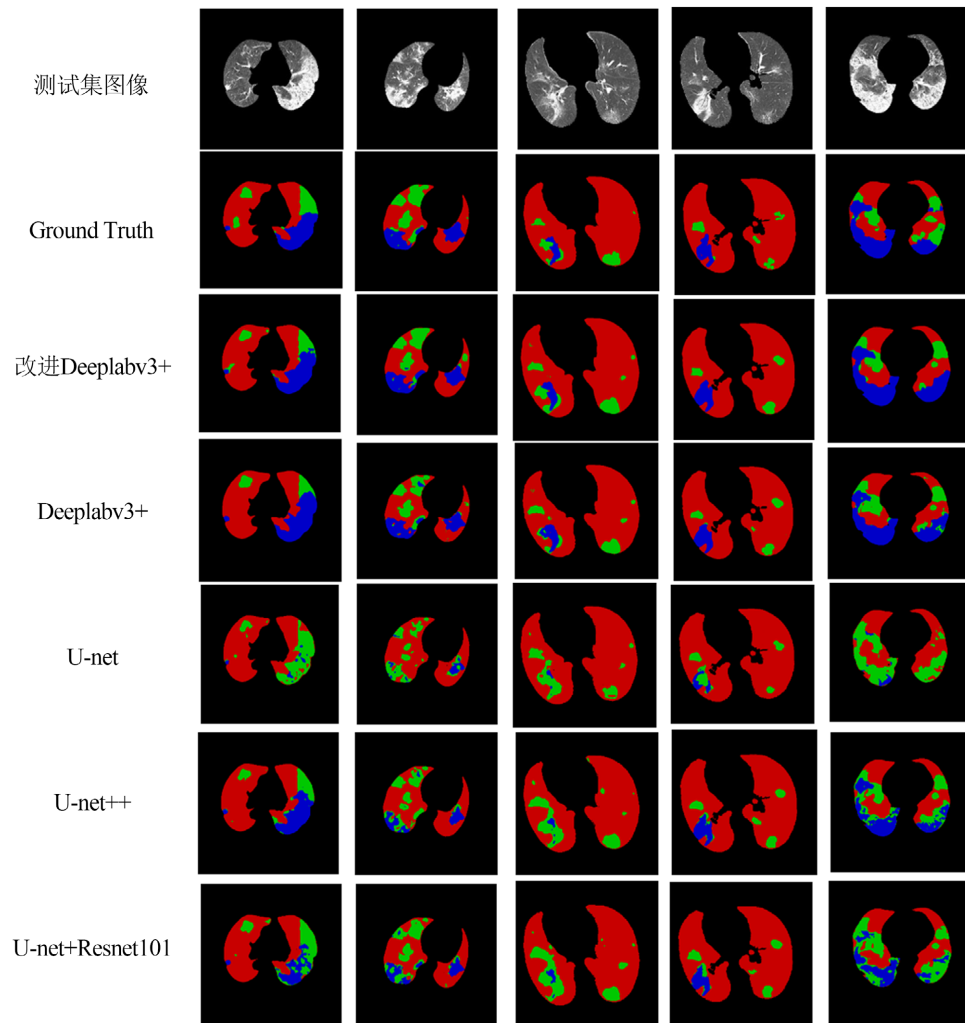


Figure 6. Loss and F-score curves, where the red line represents the training set and the blue line represents the verification set

图 6. Loss 和 F-score 变化曲线，其中红线代表训练集，蓝线代表验证集

之后对改进模型进行测试，并与 DeepLabV3+、U-net、U-net++、U-net+ResNet101 模型的预测结果进行了对比。U-net 模型是医学分割领域具有代表性的分割模型，其基本思想是先通过卷积、池化操作实现特征提取，再通过反卷积操作将图像尺寸逐步还原。由于 U-net 模型在医学图像分割任务中表现出色，后来出现了许多基于 U-net 的改进模型。U-net++对不同深度的 U-net 模型进行集成并且改进了跳跃连接结构以解决 U-net 存在的模型最佳深度先验未知及拼接过程具有局限性的问题。U-net+ResNet101 将 U-net 中主干网络替换成 ResNet101 以提高模型提取特征的能力。各模型预测的部分典型实例见图 7。其中黑色区域为背景，红色区域为正常的肺部，蓝色区域为肺实变病灶，绿色部分为肺部磨玻璃影病灶。

由图 7 可知，与其它模型相比，DeepLabV3+识别肺部和病灶区域的能力更强，但在两个病灶边缘处的预测结果不够理想，即过于平滑，细节有所丢失；本文提出的改进 DeepLabV3+，则较好地预测了病灶的细节特征，反映其在处理复杂图像时鲁棒性更好。



**Figure 7.** Typical examples of prediction results of each model. The image in the first line is the original image of the test set, the second line is the correct label, and the images in the third to seventh lines are the prediction results of each model

**图 7.** 各模型预测结果的典型实例。其中，第 1 行为测试集原图，第 2 行为正确的标签，第 3 行至第 7 行的图像为各模型的预测结果

在客观评价方面，本文采用 MIoU (Mean Intersection over Union, 均交并比)和 MPA (Mean Pixel Accuracy, 均像素准确率)作为评价指标。MIoU 是图像分割的标准度量指标，评价的是真实值和预测值的交集和并集之比，MIoU 越大说明分割效果越好。MPA 评价的则是图像中每个像素被分类正确的概率，MPA 越大则表示模型越精确。计算公式如式(1)、(2)所示。

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (1)$$

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{\sum_{i=0}^k P_{ii}}{\sum_{i=0}^k \sum_{j=0}^k P_{ij}} \quad (2)$$

式中， $k$  表示类别数、 $P_{ii}$ 、 $P_{ij}$  依次表示预测正确、错误的像素数。相应的各模型预测性能对比参见表 1、表 2。

**Table 1.** IoU performance of each model  
**表 1.** 各模型的 IoU 性能

网络结构	IoU/%				
	背景	肺部	肺实变	磨玻璃影	MIoU
U-net	99.93	95.64	43.93	40.42	69.98
U-net++	<b>99.96</b>	<b>96.73</b>	55.45	65.34	79.37
Unet+ResNet101	99.93	96.64	55.63	66.57	79.69
DeepLabV3+	99.80	95.90	56.52	70.12	80.58
改进 DeepLabV3+	99.86	96.49	<b>59.76</b>	<b>74.13</b>	<b>82.56</b>

**Table 2.** PA performance of each model  
**表 2.** 各模型的 PA 性能

网络结构	PA/%				
	背景	肺部	肺实变	磨玻璃影	MPA
U-net	<b>100.0</b>	<b>98.89</b>	45.65	53.08	74.41
U-net++	100.0	98.63	70.49	71.23	85.09
Unet+ResNet101	99.99	98.55	70.08	73.65	85.57
DeepLabV3+	99.95	97.55	72.46	80.40	87.59
改进 DeepLabV3+	99.96	98.22	<b>73.98</b>	<b>86.25</b>	<b>89.60</b>

由表 1、表 2 可见, 各模型对背景和肺部区域的分割效果均较理想, 但 DeepLabV3+对病灶的分割性能明显好于 U-net 类模型, 而改进 DeepLabV3+模型的分割性能进一步提升, 证明 CBAM 模块和深度可分离卷积确实能提升模型的性能, 得到更精确的分割结果。

## 5. 结论

本文针对 COVID-19 肺部 CT 图像的语义分割, 提出了改进的 DeepLabV3+模型, 并通过实验对比验证了其先进性。今后将进一步研究新冠肺炎 CT 图像的自动分析。

## 参考文献

- [1] 管汉雄, 熊颖, 申楠茜, 等. 新型冠状病毒肺炎(COVID-19)临床影像学特征[J]. 放射学实践, 2020, 35(2): 125-130.
- [2] 蒋西然, 蒋韬, 孙嘉瑶, 等. 深度学习人工智能技术在医学影像辅助分析中的应用[J]. 中国医疗设备, 2021, 36(6): 164-171.
- [3] 张潜, 桑军, 吴伟群, 等. 基于 Xception 的细粒度图像分类[J]. 重庆大学学报, 2018, 41(5): 85-91.
- [4] Chen, L.C., Papandreou, G., Kokkinos, I., et al. (2014) Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *Computer Science*, **4**, 357-361.
- [5] Chen, L.C., Papandreou, G., Kokkinos, I., et al. (2018) DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 834-848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- [6] Yu, Y., Liu, M., Feng, H., et al. (2020) Split-Attention Multiframe Alignment Network for Image Restoration. *IEEE Access*, **8**, 39254-39272. <https://doi.org/10.1109/ACCESS.2020.2967028>
- [7] Chollet, F. (2017) Xception: Deep Learning with Depthwise Separable Convolutions. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 1800-1807. <https://doi.org/10.1109/CVPR.2017.195>
- [8] 王亚刚, 郝怡媛, 潘晓英. 改进 DeepLabv3+网络的肠道息肉分割方法[J]. 计算机科学与探索, 2019, 14(7): 8-16.