

基于深度学习的肺炎影像分割研究

叶端胜, 华云松

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2022年11月1日; 录用日期: 2022年12月23日; 发布日期: 2022年12月31日

摘要

为有效地阻断新型冠状病毒肺炎(COVID-19)疫情传播, 如何快速准确地诊断出COVID-19患者对于疫情防控至关重要。计算机断层扫描(CT)技术是诊断COVID-19患者的最有效手段之一, 为辅助医生进行准确和快速的诊断, 需要将COVID-19患者CT图像中病变区域分割出来。近年来, 深度学习在医学图像分割中的应用越来越广泛, 本文以U-Net语义分割模型为基础框架, 针对U-Net模型在训练过程中出现的神经网络性能退化问题, 在编码器中运用了在ImageNet上预训练好的Resnet网络, 对CT图像进行特征提取。在解码器中引入注意力门机制抑制图像中的无关区域, 突出病变区域以进一步提高分割的精度。通过在相同数据集上的对比验证, 结果表明, 该方法具有良好的分割性能, 能有效分割新冠肺炎病变区域。

关键词

COVID-19, 深度学习, 图像分割, 病变区域

Research on Image Segmentation of Pneumonia Based on Deep Learning

Duansheng Ye, Yunsong Hua

School of Optional Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Nov. 1st, 2022; accepted: Dec. 23rd, 2022; published: Dec. 31st, 2022

Abstract

In order to effectively block the spread of the novel coronavirus disease 2019 (COVID-19), how to quickly and accurately diagnose COVID-19 patients is crucial for epidemic prevention and control. Computed tomography (CT) is one of the most effective methods for the diagnosis of COVID-19 patients. In order to assist doctors in accurate and rapid diagnosis, it is necessary to segment the lesion areas in CT images of COVID-19 patients. In recent years, the application of deep learning in

medical image segmentation has become more and more widespread. In this paper, the U-Net semantic segmentation model is taken as the basic framework, and the pre-trained Resnet network on ImageNet is used in the encoder to extract features from CT images in order to solve the problem of neural network performance degradation in the training process of U-NET model. An attention gate mechanism is introduced into the decoder to suppress irrelevant regions in the image and highlight the diseased regions to further improve the segmentation accuracy. Through comparison and verification on the same data sets, the results show that the proposed method has good segmentation performance and can effectively segment the lesion region of COVID-19.

Keywords

COVID-19, Deep Learning, Image Segmentation, Lesion Area

Copyright © 2022 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

COVID-19 的传染性极强, 尽早检测出被感染患者是阻断疫情传播的重要手段。临床中, CT 的表现是判断患者是否感染和病情发展的一个重要依据, 因此从 CT 图像中分割患者病变区域对提高医生的诊断效率有着重大意义。

由于 COVID-19 患者肺部感染区域在 CT 影像中的形状尺寸、位置等因素多样性, 患者的感染区域分割比较复杂。相比于传统的图像分割算法, 深度学习算法有更好的分割精度。卷积神经网络(Convolutional Neural Network, CNN) [1]的提出, 解决了图像特征提取的难题; Long 等在 2015 年提出将传统 CNN 中的全连接层换为一个个卷积层, 得到了一种全卷积神经网络(Fully Convolutional Network, FCN) [2], 解决了语义级别的图像分割问题; Ronneberger 等提出基于编码器-解码器结构的 U-Net [3]网络模型, 可以在样本量较少的情况下得到较好的分割结果; 在 U-Net 模型提出的基础上, 越来越多的研究者对 U-Net 网络进行改进优化, Zhou 等提出 U-Net++ [4]模型, 通过改进编码器和解码器之间跳跃连接过程中两者之间的细粒度优化网络模型; Oktay 等人在 U-Net 解码器部分使用注意力机制, 提出 Attention-Unet [5], 使分割模型的注意力集中在感兴趣的区域, 并在胰腺的分割中取得了很好的分割效果; Sheng Lian 等将 Residual learning 引入 U-Net, 提出 Res-Unet [6]并在视网膜血管的分割中成功应用; 闫文杰[7]提出基于 U-Net 模型和空洞卷积(Dilated Convolutions, DC)的图像分割算法, 并在网络中添加了卷积层来融合更多层的特征信息, 空洞卷积可以保证参数不增加的同时, 加大了卷积核感受野; 朱海鹏[8]提出一种新的基于 U-Net 的改进型深度学习网络, 称为 Attention Gate-Dense Network-Improved Dilation Convolution-UNET 网络(ADID-UNET), 用于分割 CT 扫描图像的病变区域。Kalene 等[9]人使用 U-Net, 并结合 DCNN, InceptionV3 和 ACNN 等预训练模型对胸部 CT 图像进行训练测试, 以检测感染区域并且取得了不错的分割效果; 宋瑶等[10]人提出改进的 U-Net 网络, 该网络对于新冠肺炎病变区域分割性能良好。

虽然目前已有一些深度学习方法用于分割病变区域诊断新冠肺炎, 但在新冠肺炎 CT 切片中分割病变区域的相关研究仍然不是很多。且只使用 U-Net 网络对其训练, 存在网络性能退化、图像空间特征利用率低等问题, 使得分割准确率偏低。本文将 Resnet 网络和注意力机制的优点相结合改进 U-Net 网络同时结合迁移学习思想, 提高了 U-Net 网络的分割效果。

2. 深度学习图像分割算法

随着深度学习技术的发展, 基于深度学习的医疗影像分割技术在不断完善。和传统的图像分割算法比较, 深度学习算法对图像的特征提取程度更深, 有更好的分割精度, 且不需要大量人工分析和调参。用深度学习算法替代传统的图像分割算法能得到更好的分割结果。

2.1. 深度学习分割算法流程

本文使用深度学习技术对新冠肺炎患者的 CT 图像中病变区域进行分割, 整体的分割算法流程如图 1 所示。

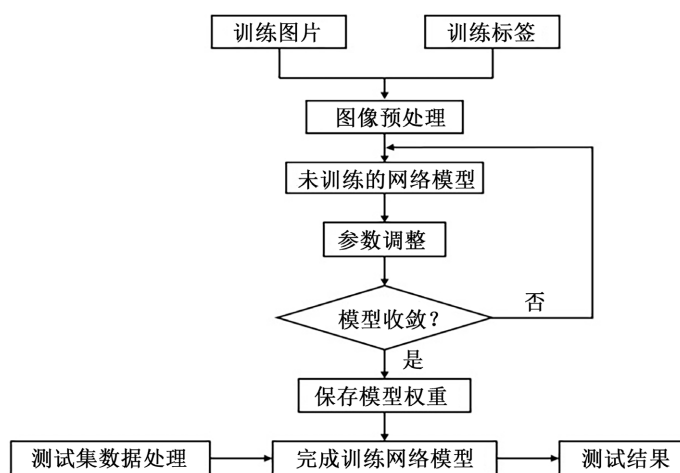


Figure 1. Deep learning image segmentation algorithm process

图 1. 深度学习图像分割算法流程

由图 1 知, 一般情况下利用深度学习技术对图像进行分割要先得到训练图片对应的图像标签, 图像标签是由医生对原始图像进行标注, 即手动分割出病变区域。训练图片及其对应标签构成训练集, 将训练集输入网络中不断迭代运算, 让网络学习标签数据中的特征, 并在每一次迭代后对网络的参数进行更新调整直至模型收敛得到完成训练的模型, 即模型已经无法在从标签中学习到更多特征了, 此时的模型分割能力较好, 损失率最低。测试集用于对模型性能的验证, 也包括原图和对应的标签, 标签用于定量计算模型的分割性能评价指标, 得到的测试结果为评价模型性能的最终结果。

2.2. 肺炎 CT 图像分割算法构建

2.2.1. 编码器 - 解码器结构

在图像语义分割领域, 编码 - 解码器结构是最常用的。本文也是采用这种能够端到端分割整个图片的结构, 以经典的 U-Net 网络模型的 U 型结构为基本框架, 编码器部分实质为下采样, 将原始的输入图像逐层下采样得到分辨率很低的特征图, 解码器实质为上采样, 将编码器中的特征图逐层上采样恢复到较高分辨率并进行逐像素预测最终输出得到预测结果。

2.2.2. 编码模块

在深度学习图像分割任务中, 一定范围内网络深度越深, 网络性能越好。但随着网络深度增大, 网络在训练过程中会出现深度退化问题, 表现为模型训练难以收敛, 模型的梯度消失或爆炸进而模型的分割性能严重下降。随着残差网络(Resnet) [11] [12] [13] 的出现, 残差网络相比于普通网络不容易出现深度

退化问题。为避免 U-Net 网络在训练过程中出现这些问题, 本文用残差网络 Resnet34 替换原始的编码器。Resnet34 由残差单元连接而成, 残差单元如图 2 所示, 公式为:

$$H(x) = F(x) + x \quad (1)$$

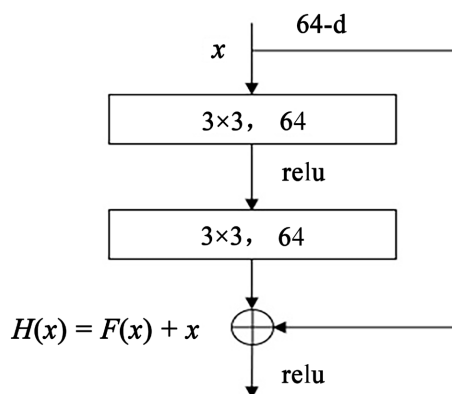


Figure 2. Plot of residual elements
图 2. 残差单元图

Resnet34 网络共分为五个层(Layer 0~4), 每一层的卷积层的卷积核大小和个数具体构成如表 1 所示。

Table 1. Resnet34 is constructed in detail layer by layer
表 1. Resnet34 逐层详细构成

Layer name	Layers	Number
Layer 0	$7 \times 7, 64$	1
Layer 1	$3 \times 3, 64$ $3 \times 3, 64$	3
Layer 2	$3 \times 3, 128$ $3 \times 3, 128$	4
Layer 3	$3 \times 3, 256$ $3 \times 3, 256$	6
Layer 4	$3 \times 3, 512$ $3 \times 3, 512$	3

2.2.3. 解码模块

在解码部分, 图像尺寸在恢复过程中, 保留原始 U-Net 的上采样和跳层连接结构。跳层连接是将下采样过程中的特征图拼接到对称的上采样层后再进行上采样, 目的是用来补充采样过程中丢失的空间特征信息。原始的 U-Net 是直接特征图裁剪后拼接过来, 会带来许多与需要分割的区域无关的背景特征信息, 使最终分割结果准确性较低。注意力机制(Attention Gate, AG)能极大地提高分割精度且已经在实验中得到验证, 在解码过程中通过引入 AG 系统来抑制特征图中的无关紧要的背景信息, 帮助模型更多关注目区域, 减少分割的错误率, 增加模型的灵敏度和预测准确性。AG 的结构如图 3 所示。

其中, g_l 和 x_l 分别为上采样和下采样过程中第 l 层的特征图, g_l 是经过上一层上采样后的, 所以两者尺寸一样, 分别经过一个 1×1 的卷积后点对点叠加, 公式如下:

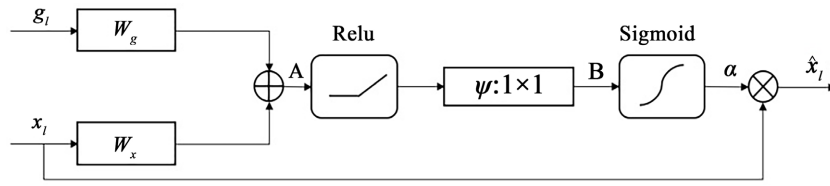


Figure 3. Attention Gate structure diagram
图 3. Attention Gate 结构图

$$A = W_g(g_i) + W_x(x_i) \tag{2}$$

A 经过 Relu 激活函数和另一个 1×1 的卷积 ψ 后得到单通道的 B,

$$B = \psi(\text{Relu}(A)) \tag{3}$$

最后, 经过一个 Sigmoid 函数得到注意力系数 α , 范围为 0 到 1, 即为每个像素点的权重。

$$\alpha = \text{Sigmoid}(B) \tag{4}$$

\hat{x}_i 为经过注意力机制后输出的特征图, 保留了感兴趣区域信息, 由注意力系数与原下采样特征图相乘得到。

$$\hat{x}_i = \alpha \cdot x_i \tag{5}$$

2.3. 网络结构

本文算法是基于 U-Net 改进的, 保留了 U-Net 端对端的原始结构, 在编码器和解码器部分都做了相应的改进, 提高模型的分割性能。网络具体结构如图 4 所示。

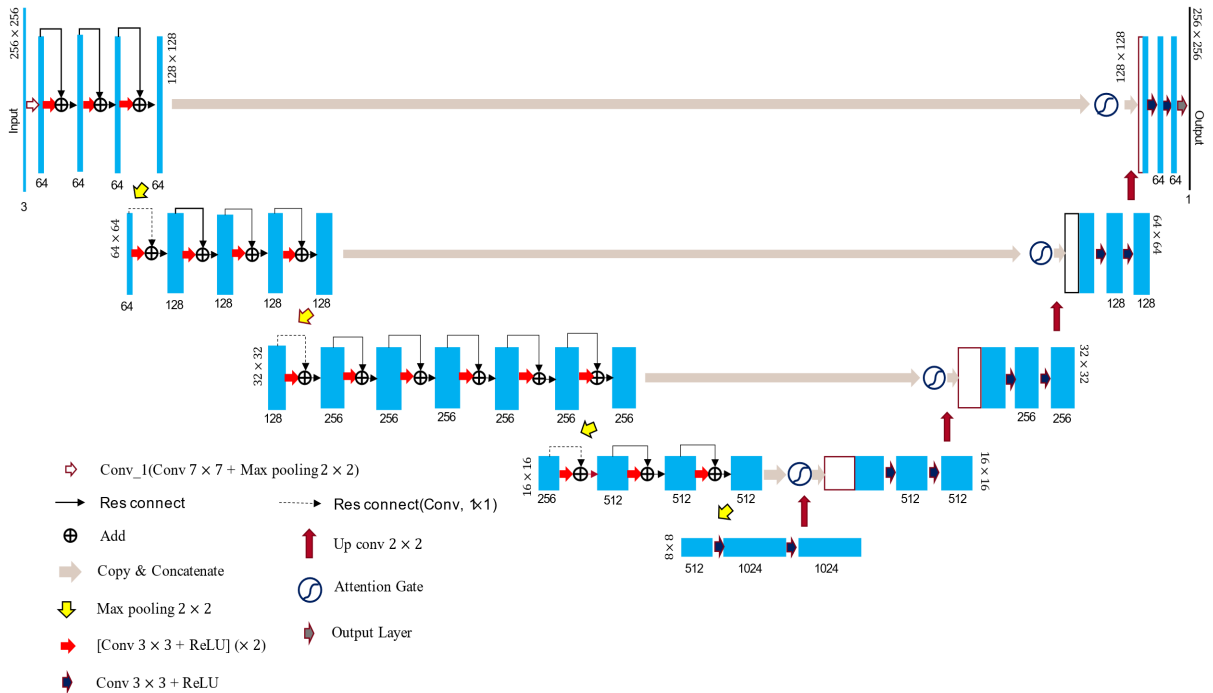


Figure 4. Diagram of the network structure
图 4. 网络结构图

该网络的编码模块由 Resnet34 构成, 利用迁移学习, Resnet34 的初始权重为在 ImageNet [14] 上预训练过的权重。初始输入图片尺寸为 256×256 , 通道数为 3, 经过卷积层 Conv_1, 即为 Resnet 的 Layer 0 层, 输出特征图尺寸为 128×128 , 通道数为 64, 此后经过 Resnet34 的 Layer 1~4 层逐层下采样, 每下采样一层, 特征图的尺寸减半, 通道数变为原来两倍。图中残差连接有两种, 实现箭头为常规残差连接, 每残差单元的输出和输入的通道数一样, 可直接叠加, 虚线箭头为升维残差连接, 由于两层之间卷积核数量不同, 每层的初始残差单元输入和输出通道数不同, 输出是输入的两倍, 进行残差连接时需要输入用 1×1 的卷积进行升维, 使两者通道数一致方可进行叠加; 解码模块由 4 组上采样层和卷积层组成, 与编码模块相对应, 在每层编码模块和解码模块之间是带有 AG 模块的跳层连接, 图像随着逐层上采样, 尺寸逐层恢复, 通道数也逐层减小, 每层尺寸大小和通道数均与对应的编码模块层相同, 方便跳层连接时两者的叠加。为了得到最后的分割结果与输入图片尺寸一致, 最后的输出层再进行一次上采样和一次 1×1 的卷积, 使图片尺寸恢复为与原图一致, 通道数降为 1, 最终得到尺寸为 256×256 的单通道分割结果。

3. 实验及结果分析

3.1. 实验数据集

本文采用的 COVID-19CT 图像分割数据集[15], 其中包含 20 例患者的 CT 数据, 共有 1000 多个切片和对应的分割标签, 分割标签中感染区域的标注由专业的放射科医生完成和验证。由于同一患者的不同图层的 CT 切片之间具有相似性, 最终选出 224 张切片用于训练, 10 张切片用于测试。

3.2. 实验内容

本实验是基于 Pytorch 框架实现的, 数据集拆分为 224 张作为训练集和验证集, 比例为 9:1, 另外 10 张设置为测试集。按照 2.1 中深度学习分割算法流程完成对模型的训练和测试。训练过程中, 选用二元交叉熵(BCE With Logits Loss)作为网络的损失函数, 公式如下:

$$Loss = -\frac{1}{n} \sum_{k=1}^n q_k \log(\sigma(p_k)) + (1 - q_k) \log(1 - \sigma(p_k)) \quad (6)$$

其中, n 表示每张图片的像素个数, p_k 和 q_k 表示预测的像素值和真实值, σ 表示 Sigmoid 函数, 用来将 p_k 值映射到(0, 1)之间。

本文采用 Adam 优化器优化网络模型, batch size 值设置为 2, 学习率设置为 0.0001, 模型训练迭代次数设置为 200, 每轮训练会计算网络的损失函数值, 通过反向传播调整参数, 损失函数值不断减小最终保持不变, 此时模型收敛。保存训练过程中在验证集上分割效果最好的模型参数。

为验证改进的 ResAttUnet 模型分割性能, 本文采用对比实验, 将本文算法与原始 U-Net、只引入残差卷积的 Res-Unet 和只有注意力机制的 Attention-Unet 模型分割性能进行对比。具体实验为设置相同的网络参数, 用相同的数据集分别对以上几种网络进行训练和测试, 得到实验结果。

3.3. 评估指标

为从多个角度验证本文算法的性能, 计算测试集上各个网络分割结果的评估指标并作对比本文采用了准确率(Acc)、精密度(Pc)、敏感性(Sen)、特异性(Sp)和 F1 评分(F1), 计算公式如下。计算这些指标, 需要生成混淆矩阵, 并定义真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN), 如下表 2 所示。

准确率 Accuracy (Acc): 表示正确预测的病变像素个数与图片像素总数的比值。计算公式如公式(7); 精密度 Precision (Pc): 正确预测病变占预测病变像素总数的比例计算公式如公式(8); 敏感性 Sensitivity (Sen): 正确预测的病变像素个数与实际病变像素总数的比值, 计算公式如公式(9); 特异性 Specificity (Sp): 正

确预测的非病变像素个数占实际非病变像素总数比例, 计算公式如公式(10); F1 评分(F1): 综合了精确度和敏感性结果得出的综合评价指标, F1 值越大, 网络分割的效果越好, 计算公式如公式(11)。

Table 2. Definitions of TP, TN, FP, FN

表 2. TP、TN、FP、FN 的定义

区域类别	实际病变	实际非病变
预测病变	True Positives (TP)	Flase Positives (FP)
预测非病变	Flase Negatives (FN)	True Negatives (TN)

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$Pc = \frac{TP}{TP + FP} \quad (8)$$

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

$$Sp = \frac{TN}{TP + FN} \quad (10)$$

$$F1 = 2 \times \frac{Pc \times Sen}{Pc + TN} \quad (11)$$

3.4. 实验结果分析

本文针对 U-Net 网络在 COVID-19 患者 CT 图像病变区域分割过程中分割效果不佳的问题, 对网络模型进行了相应的改进并成功完成了模型的训练。图 5 为模型在训练集上的损失曲线。

由图 5 可以看出, 模型在迭代次数达到 100 左右时, 损失函数值基本趋于不变, 稳定在 0.02 左右, 代表模型基本收敛。图 6 为模型训练过程中在验证集上的损失函数曲线, 验证集可以真实反应训练后的模型实际。当模型迭代 90 次后, 损失值始终保持在较低水平, 基本稳定在 0.03 左右。说明模型训练完成。

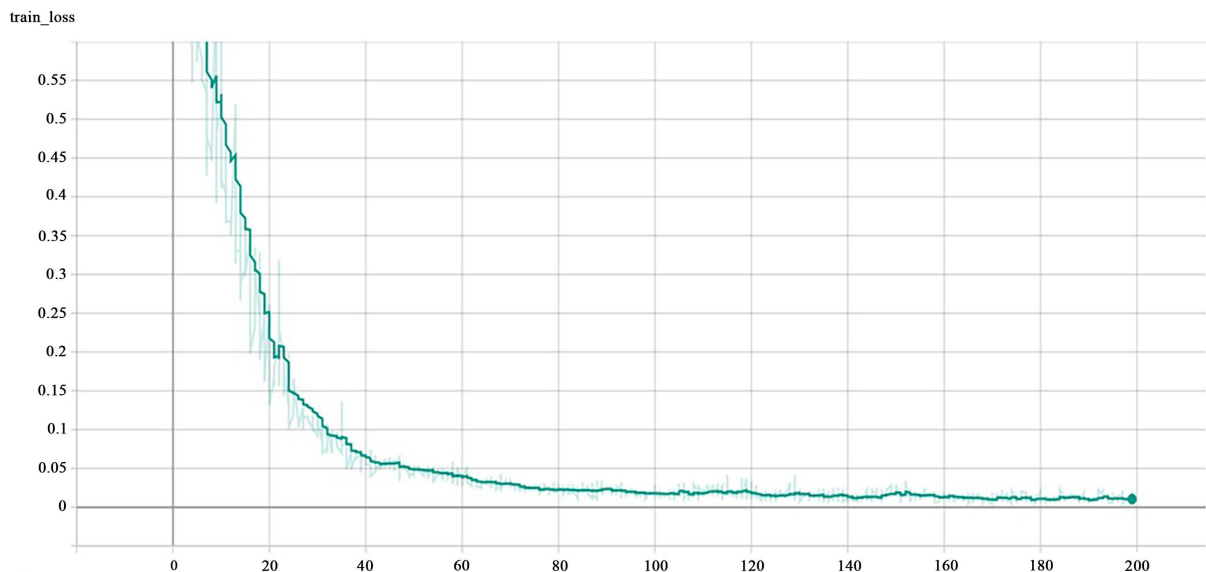


Figure 5. Training set loss curve

图 5. 训练集损失曲线

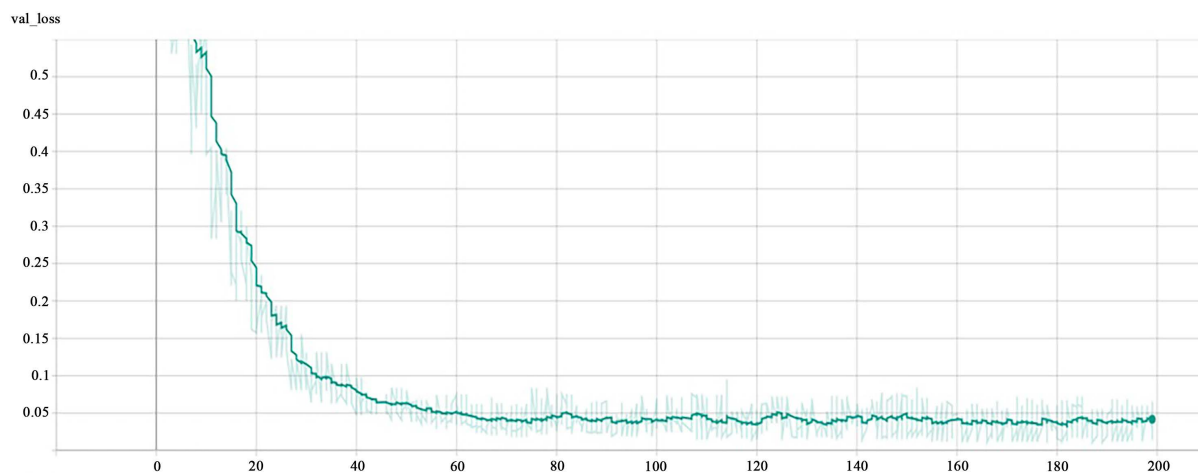


Figure 6. Verify the set loss curve
图 6. 验证集损失曲线

训练好的模型在测试集分割后得到结果, 部分视觉效果图如下图 7 所示。由左向右, 各列依次为(a) 新冠肺炎患者 CT 原图、(b) 标签、(c) Unet 结果、(d) Res-Unet 结果、(e) Attention-Unet 结果和(f) ResAttUnet (本文)结果。

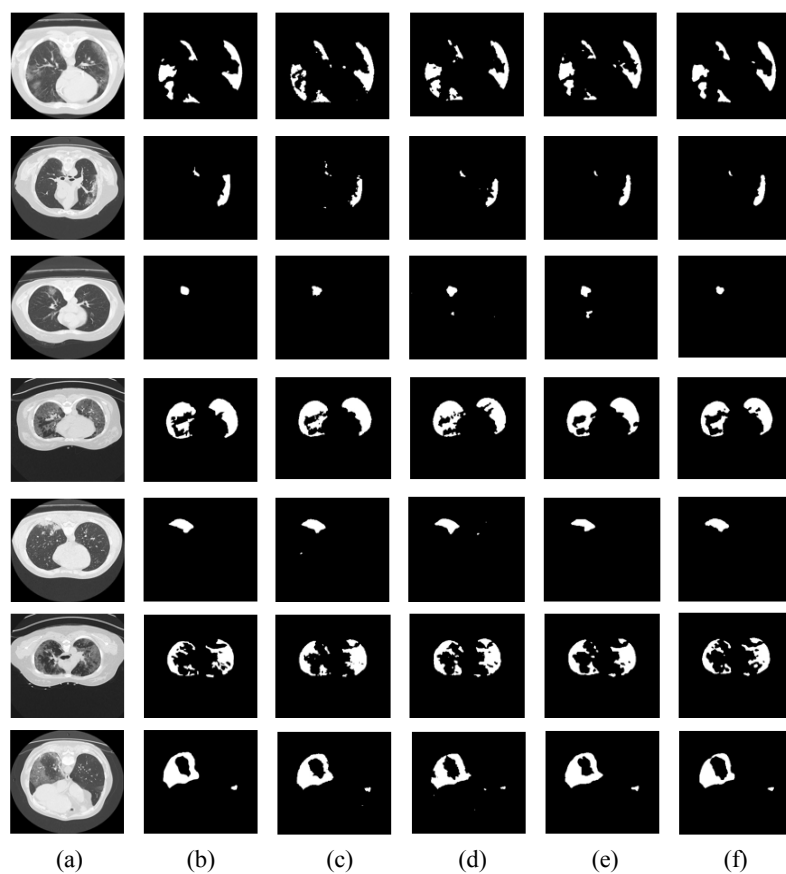


Figure 7. Segmentation results for different models
图 7. 不同模型的分割结果

从图 7 中几种模型的分割结果视觉对比图可以看出, 几种方法在同样的数据集上均存在一定漏分割和错误分割的现象。图(f)为本文提出的算法在测试集上的分割结果。从图中的视觉对比来看, 未改进的网络容易受背景区域影响, 错误将背景区域预测为病变区域, 病变区域的细节分割也不理想。对比之下, 本文改进的算法改善了这些问题, 分割更加精确, 与专家分割的标签(b)的事实基本相当。

为进一步验证改进的算法分割性能是否提升, 定量计算这几种模型在测试集上的评价指标, 如表 3 所示, 由表中的数据可以看出, 相比于其他模型, 本文模型在相同测试集上各项指标均有一定的优势, 说明本文算法在新冠肺炎 CT 图像病变区域的分割中达到较强的分割水平。

Table 3. Quantitative evaluation of segmentation of different models
表 3. 不同模型的分割的定量评估指标

Method	Acc	Pc	Sen	Sp	F1
Unet	0.982	0.695	0.769	0.965	0.731
Res-Unet	0.983	0.766	0.738	0.986	0.752
Attention-Unet	0.986	0.770	0.897	0.971	0.829
ResAttUnet (本文)	0.994	0.811	0.929	0.995	0.867

4. 总结

通过 COVID-19 患者 CT 图像病变区域的分割对患者诊断及治疗和阻断疫情的传播有着重要意义。传统的 U-Net 虽然能较准确地分割出感染区域, 但网络退化、空间特征利用不充分等问题使得 U-Net 网络的分割性能难以提升。为改善这些问题, 本文以 U-Net 网络的编码-解码结构为框架, 用残差网络 Resnet34 作为编码器, 结合注意力机制优点, 提出了一种改进的深度卷积网络模型。该模型通过残差网络代替传统卷积网络提高编码器部分的特征提取能力, 同时利用注意力机制抑制无关的背景部分来提高模型的分割性能。通过对比实验, 本文模型在新冠肺炎病变区域的分割性能有了一定的提升。下一步可以考虑用不同深度的残差网络和性能更好的注意力模块对网络进一步优化。

参考文献

- [1] Sharif Razavian, A., Azizpour, H., Sullivan, J. and Carlsson, S. (2014) CNN Features Off-the-Shelf: An Astounding Baseline for Recognition. 2014 *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Columbus, 23-28 June 2014, 806-813. <https://doi.org/10.1109/CVPRW.2014.131>
- [2] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- [3] Ronneberger, O., Fischer, P., & Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. Lecture Notes in Computer Science*, Vol. 9351, Springer, Cham, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [4] Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N. and Liang, J. (2018) Unet++: A Nested U-Net Architecture for Medical Image Segmentation. Architecture for Medical Image Segmentation. In: Stoyanov, D., Taylor, Z., Carneiro, G., Syeda-Mahmood, T., et al., Eds., *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Lecture Notes in Computer Science*, Vol. 11045, Springer, Cham, 3-11. https://doi.org/10.1007/978-3-030-00889-5_1
- [5] Oktay, O., Schlemper, J., Folgoc, L.L., et al. (2018) Attention U-Net: Learning Where to Look for the Pancreas. ArXiv: 1804.03999.
- [6] Lian, S., Li, L., Lian, G., Xiao, X., Luo, Z. and Li, S. (2019). A Global and Local Enhanced Residual U-Net for Accurate Retinal Vessel Segmentation. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **18**, 852-862.

-
- <https://doi.org/10.1109/TCBB.2019.2917188>
- [7] 闫文杰. 基于深度学习的肺部 CT 影像分割算法的研究[D]: [硕士学位论文]. 大连: 大连理工大学, 2019. <https://doi.org/10.26991/d.cnki.gdllu.2019.002214>
- [8] 朱海鹏. 一种基于 ADID-UNET 模型的 COVID-19 肺部 CT 图像感染区域分割方法研究[D]: [硕士学位论文]. 汕头: 汕头大学, 2021. <https://doi.org/10.27295/d.cnki.gstou.2021.000304>
- [9] Kalane, P., Patil, S., Patil, B.P. and Sharma, D.P. (2021) Automatic Detection of COVID-19 Disease Using U-Net Architecture Based Fully Convolutional Network. *Biomedical Signal Processing and Control*, **67**, Article ID: 102518 <https://doi.org/10.1016/j.bspc.2021.102518>
- [10] 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.
- [11] Targ, S., Almeida, D. and Lyman, K. (2016) ResNet in ResNet: Generalizing Residual Architectures. ArXiv: 1603.08029.
- [12] Zhang, Q., Cui, Z., Niu, X., Geng, S. and Qiao, Y. (2017) Image Segmentation with Pyramid Dilated Convolution Based on ResNet and U-Net. In: Liu, D., Xie, S., Li, Y., Zhao, D. and El-Alfy, E.S., Eds., *Neural Information Processing. Lecture Notes in Computer Science*, Vol. 10635, Springer, Cham, 364-372. https://doi.org/10.1007/978-3-319-70096-0_38
- [13] Bekele, E., Narber, C. and Lawson, W. (2017) Multi-Attribute Residual Network (MA ResNet) for Soft-Biometrics Recognition in Surveillance Scenarios. 2017 12th *IEEE International Conference on Automatic Face & Gesture Recognition*, Washington DC, 30 May-3 June 2017. <https://doi.org/10.1109/FG.2017.55>
- [14] Deng, J., Dong, W., Socher, R., *et al.* (2009) ImageNet: A Large-Scale Hierarchical Image Database. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 20-25 June 2009, 248-255. <https://doi.org/10.1109/CVPR.2009.5206848>
- [15] Ma, J., Wang, Y., An, X., *et al.* (2020) Towards Efficient COVID-19 CT Annotation: A Benchmark for Lung and Infection Segmentation. ArXiv: 2004.12537.