

引文格式:

毛丽, 李秀丽. 基于深度学习和 CT 影像的新型冠状病毒肺炎病灶分割 [J]. 集成技术, 2020, 9(6): 40-47.

Mao L, Li XL. Segmentation of COVID-19 lesions based on deep learning and CT images [J]. Journal of Integration Technology, 2020, 9(6): 40-47.

基于深度学习和 CT 影像的新型冠状病毒肺炎病灶分割

毛 丽^{1,2} 李秀丽¹

¹(深睿医疗人工智能研究院 北京 100080)

²(中国人民大学统计学院 北京 100086)

摘 要 胸部 CT 图像中新型冠状病毒肺炎 (COVID-19) 病灶的准确分割可以为诊断提供助力。新型冠状病毒肺炎在 CT 影像上的表现包括磨玻璃影、实变、胸腔积液病灶等, 这些征象的纹理、大小和位置变化较大。该研究提出的深度神经网络 RCB-UNet++, 可以用于新型冠状病毒肺炎在 CT 影像上的分割问题。该网络在 UNet++ 网络的基础上, 增加了残差模块和卷积块注意力模块。此架构能有效地提取底层的纹理信息和高层的语义信息, 并基于注意力机制优化模型效果。该研究所提出的 RCB-UNet++ 模型经过在 45 例样本上的训练后, 在 50 例测试集上的 Dice 系数达到了 0.715, 且敏感性和特异性分别达到了 0.754 和 0.952, 超过基于同样数据划分的其他已发表的深度学习模型。这表明所提出的算法有效地提高了分割效果, 具有从 CT 图像中全自动分割新型冠状病毒肺炎病灶的潜力。

关键词 UNet++; 语义分割; 深度学习; 新型冠状病毒肺炎

中图分类号 TP 399 文献标志码 A doi: 10.12146/j.issn.2095-3135.20200921001

Segmentation of COVID-19 Lesions Based on Deep Learning and CT Images

MAO Li^{1,2} LI Xiuli¹

¹(DeepwiseAILab, Beijing 100080, China)

²(School of Statistics, Renmin University of China, Beijing 100086, China)

Abstract Accurate segmentation of COVID-19 pneumonia lesions on chest CT images can facilitate the diagnosis of pneumonia. The CT image finds of which contained the ground-glass opacity, consolidation, pleural effusion, etc. This study proposed a deep neural network RCB-UNet++ for the segmentation of COVID-19 pneumonia lesions in CT images, which exhibit large variations in texture, size and location. The

收稿日期: 2020-09-21 修回日期: 2020-10-20

作者简介: 毛丽, 学士, 研究方向为概率论与数理统计; 李秀丽(通讯作者), 博士, 主任研究员, 研究方向为医学人工智能, E-mail: lixiuli@deepwise.com。

model was built on top of the UNet++ network with an extra residual module and an attention module. This architecture is able to effectively extract low-level texture features and high-level semantic information, thus improving the model performance. The RCB-UNet++ model was trained on 45 samples and tested by another 50 cases. Finally, it achieved a Dice coefficient of 0.715, a sensitivity and specificity of 0.754 and 0.952, outperforming other designed models on the same dataset. The results demonstrate that the proposed algorithm improves the segmentation performance and has potential in fully automatic segmentation of COVID-19 pneumonia lesions on CT images.

Keywords UNet++; semantic segmentation; deep learning; COVID-19

1 引言

2019 年 12 月以来, 新型冠状病毒 (COVID-19) 在全球范围内广泛而快速地传播。根据美国约翰斯·霍普金斯大学系统科学与工程研究中心发布的数据, 截至 2020 年 9 月 18 日, 全球已有 30 065 728 例确诊病例。该疾病传染性强, 严重时可引起严重的急性呼吸困难或多器官功能衰竭。世界卫生组织于 2020 年 1 月 30 日宣布 COVID-19 疫情为“国际关注的突发公共卫生事件”。

研究表明, 计算机断层扫描 (CT) 作为一种非侵入性的影像学检查方法, 对检测 COVID-19 感染患者的肺部病变具有重要价值。CT 在诊断 COVID-19 方面的敏感性远高于反转录聚合酶链反应 (RT-PCR)^[1-2]。因此, CT 可以作为早期发现和诊断 COVID-19 的有效工具。此外, 通过对炎症病灶进行定量分析可以得到 COVID-19 相关的诊断结果^[3-4]。然而, 人工勾画炎症病灶是一项十分繁重的工作, 且高度依赖于专家的临床经验。因此, 构造肺炎病灶的分割模型, 可以提高勾画的效率、准确性和可重复性, 为进一步定量分析奠定良好的基础。

目前, 深度学习已被初步应用于 COVID-19 的分割和诊断任务^[5-6]。然而, 由于 COVID-19 的病灶纹理、大小和位置变化较大, 且与正常组

织间差异较小, 因此为分割模型的构造带来了较大的挑战。此外, 数据的收集和标注也十分困难。尽管, Inf-Net^[7]提出了新的网络结构和 Loss 形式, 并通过挖掘更多病灶边界的信息, 提高了模型的分割能力。但是, 该方法较为复杂, 需要提前计算分割区域的边界, 无法进行端到端的分析。Zhang 等^[8]构造的 COVID-19 病灶分割模型, 可以用于分割 6 种类型的肺部结构, 包括肺部区域、磨玻璃病灶、肺实变病灶、肺纤维化、间质增厚和胸腔积液。但是, 该方法需要大量的训练数据。

本研究在 UNet++ 网络模型^[9]的基础上, 融合了残差模块和卷积块注意力模块, 提出 RCB-UNet++ 模型。RCB-UNet++ 可以利用较少的数据进行训练, 以实现端到端的 COVID-19 炎症病灶区域分割。

2 材料和方法

2.1 数据

本文基于公开的 COVID-19 CT 分割数据集^[10]进行训练和测试。该数据集由意大利医疗和介入放射协会收集, 并由一位放射科医生逐像素勾画, 其中勾画病灶包括磨玻璃、肺实变和胸腔积液。因为图像是由不同的设备和采集协议所得, 所以数据集在分辨率和图像质量上呈现多

样化。该数据集共 100 例数据，来自超过 40 个确诊为 COVID-19 的病例。其中 45 例数据用于模型训练，5 例用于模型验证，50 例用于测试，划分方式和 Inf-Net 模型^[7]的构造过程一致。此外，将所有磨玻璃、肺实变和胸腔积液病灶合并为炎症病变区域，作为模型的分割目标。

2.2 方法

2.2.1 RCB-UNet++网络构造

UNet++^[9]优化了 U-Net 的网络拓扑结构，提高了对不同层次特征的捕获能力。为进一步提高模型的参数利用效率和表达能力，本文使用 Residual Block^[11]替代原始的 VGG 结构。对于医学图像而言，不同器官本身的相对位置可以提供更多的先验信息。例如，肺炎病灶一定是在肺部区域，而不会在隔膜上发生。因此，引入注意力机制，可以有效地提高模型的表现。在模型的下采样过程中，每一个残差模块都连接卷积块注意力模块 (CBAM)^[12]，确保提取有效的空间信息和

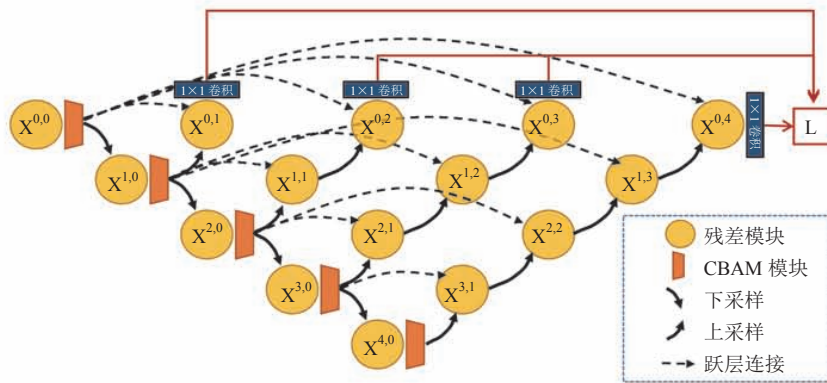
特征通道。此外，基于深度监督的方式，能进一步提高梯度的传播能力。RCB-UNet++模型的总体网络结构如图 1(a) 所示。

2.2.2 残差模块

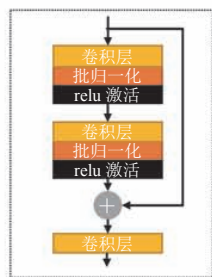
残差模块被设计用于处理卷积神经网络较深时出现的梯度消失问题，具体如图 1(b) 所示。残差模块通过跃层链接 (Shortcut Connection)，把模块的输入连接到后面的层，使得后面的层可以直接学习残差。这种结构优化了模型的学习目标，降低了学习难度，使得在模型较深时也可以保持较好的拟合效果。

2.2.3 卷积块注意力模块

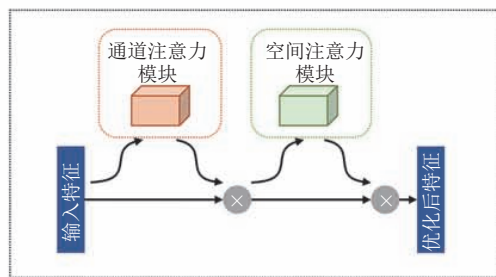
对于网络结构中间的特征图，注意力机制可以应用在空间和通道两个维度，分别推断出注意力权重，然后将注意力权重和原特征图相乘，得到新的特征图。卷积块注意力模块可以对模型特征图的空间和通道的重要性进行自适应调整。由于卷积块注意力模块是一个轻量级的模块，它可



(a) 改进的 Unet++ 网络结构图



(b) 残差模块内部结构图



(c) 卷积块注意力模块示意图

图 1 模型流程图

Fig. 1 The flowchart of the proposed model

以嵌入到任何卷积神经网络架构中。卷积块注意力模块的示意图如图 1(c) 所示。该模块由通道注意力模块和空间注意力模块组成。通道注意力模块为特征图的每一个通道分配一个权重。其中, 特征图的每一个通道可以看作一个特征提取器, 通过对通道增加注意力机制, 可以使模型更注重有效的特征, 大大提高了网络的表达能力。空间注意力模块考虑特征图的哪一个区域应该有更高的响应, 将空间注意力模块的输出和特征图进行像素级别的点乘, 即可得到加权后的特征图。具体如公式(1)~(2)所示:

$$\mathbf{F}' = M_C(\mathbf{F}) \otimes \mathbf{F} \quad (1)$$

$$\mathbf{F}'' = M_S(\mathbf{F}') \otimes \mathbf{F}' \quad (2)$$

其中, M_C 为通道维度上的注意力图像提取操作; M_S 为空间维度上的注意力图像提取操作; \otimes 为像素级别的点乘; \mathbf{F} 为输入的特征; \mathbf{F}' 为经过通道维度的权重调整后得到的特征; \mathbf{F}'' 为经过空间维度的权重调整后得到的特征。

2.2.4 损失函数

二分类交叉熵损失函数是分割模型常用的损失函数, 等效于最小化相对熵的差异, 具体如公式(3)所示:

$$L_{\text{bce}} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1-y_i) \log(1-\hat{y}_i)] \quad (3)$$

其中, L_{bce} 为二分类交叉熵损失函数; N 为像素点的个数; y_i 为某个像素的真实类别; \hat{y}_i 为像素被预测为阳性(属于肺炎病灶)的概率。

Dice 损失函数直接优化模型的目标, 使预测和真实分割之间的 Dice 系数最大化, 具体如公式(4)所示:

$$L_{\text{dice}} = 1 - \frac{2|\mathbf{X} \cap \mathbf{Y}| + 1}{|\mathbf{X}| + |\mathbf{Y}| + 1} \quad (4)$$

其中, L_{dice} 为 Dice 损失函数; \mathbf{X} 为模型的预测结果张量; \mathbf{Y} 为真实的分割结果张量; $|\mathbf{X} \cap \mathbf{Y}|$ 由 \mathbf{X} 和 \mathbf{Y} 逐像素相乘计算获得; $|\mathbf{X}|$ 和 $|\mathbf{Y}|$ 分别为预测结果和真实分割的张量的模。

相比于交叉熵损失, Dice 损失更加直接, 往往可以取得更好的结果。而且, Dice 损失可以更好地处理类别不均衡的问题, 即预测前景像素个数和背景像素个数差异较大的问题。但是, 其梯度的形式不如二分类交叉熵损失函数良好。对于 Dice 损失而言, 在极端场景下, 梯度会变得非常大, 导致模型的训练更加不稳定。因此, 本实验综合二分类交叉熵损失和 Dice 损失, 具体如公式(5)所示:

$$L = 0.5 \times L_{\text{bce}} + L_{\text{dice}} \quad (5)$$

为了提高模型隐藏层学习的直接性和透明度, 本模型采用深度监督的方式训练网络。首先, 对每个隐藏层引入伴随目标函数, 即对节点 $X^{0,1}$ 、 $X^{0,2}$ 、 $X^{0,3}$ 、 $X^{0,4}$ 均添加输出层(1×1 卷积层)。然后对于每一个输出层, 基于如公式(5)所述的损失函数进行优化。

2.2.5 模型训练

由于训练数据较少, 因此, 进行数据增强可以提高网络的鲁棒性, 降低过拟合。本实验所用数据增强方法包括随机扩大和缩小、随机裁剪、随机亮度和对比度调整。本实验首先将图像统一为 128×128 的像素大小, 然后按 0.8~1.2 的随机比例缩放, 再随机裁剪为 128×128 的像素大小, 并以 0.3 的概率进行随机亮度和对比度调整。

本模型基于 Adam 算法^[13]来优化模型参数。实验共训练 500 个轮次, 设置 Adam 优化算法的初始学习率为 0.001。最后, 在整个训练过程中, 选择验证集上交并比 (IOU) 最大的模型作为输出模型。此外, 为了提高模型的泛化性能, 设置权重衰减 (Weight Decay) 系数为 0.001。

2.2.6 模型评估

为评估模型的表现, 本实验使用了和文献[7]相同的方式, 并基于该工作公开的评估代码评估模型在测试集上的结果。即基于不同阈值来计算相关指标, 然后求取均值, 作为最终的评估

结果。评估指标包括常用的参数,如 Dice 相似系数、敏感性、特异性和准确性。另外,本文模型引入了目标检测领域的相关评估方法,包括结构度量 (Structure Measure)、配准增强度量 (Enhanced-Alignment Measure) 和平均绝对误差 (Mean Absolute Error)。

结构度量 S_a 用于评估模型预测结果和真实分割之间的相似性,这种度量方式更加符合人类认知,具体如公式(6)所示:

$$S_a = (1 - \alpha) \times S_o(\mathbf{S}_p, \mathbf{G}) + \alpha \times S_r(\mathbf{S}_p, \mathbf{G}) \quad (6)$$

其中, α 为平衡系数,本实验设置 α 为 0.5,和原始文献[14]相同; S_o 为对象感知相似性; S_r 为区域感知相似性。 \mathbf{S}_p 为输出的预测结果; \mathbf{G} 为真实的分割。

配准增强度量 E_ϕ 可以度量两个二值矩阵的相似性,该方法可以同时度量局部特征和整体特征,具体如公式(7)所示:

$$E_\phi = \frac{1}{w \times h} \sum_x \sum_y \phi[\mathbf{S}_p(x, y), \mathbf{G}(x, y)] \quad (7)$$

其中, w 和 h 为图像的宽和高; (x, y) 为图像中的像素坐标; ϕ 为增强配准矩阵。首先,通过 0~1 的 256 个等分的阈值分别计算 E_ϕ ; 然后,把所有的 E_ϕ 均值 E_ϕ^{mean} 作为最终的度量指标。该指标不是仅考虑某一个阈值下模型的表现,而是综合所有阈值结果,因此对阈值具有鲁棒性。

平均绝对误差 (MAE) 度量模型的预测结果和真实分割之间的误差,具体如公式(8)所示:

$$MAE = \frac{1}{w \times h} \sum_x \sum_y |\mathbf{S}_p(x, y) - \mathbf{G}(x, y)| \quad (8)$$

3 结果

模型在 500 个训练轮次上的损失函数值变化和 IOU 变化如图 2 所示。随着模型的训练过程,

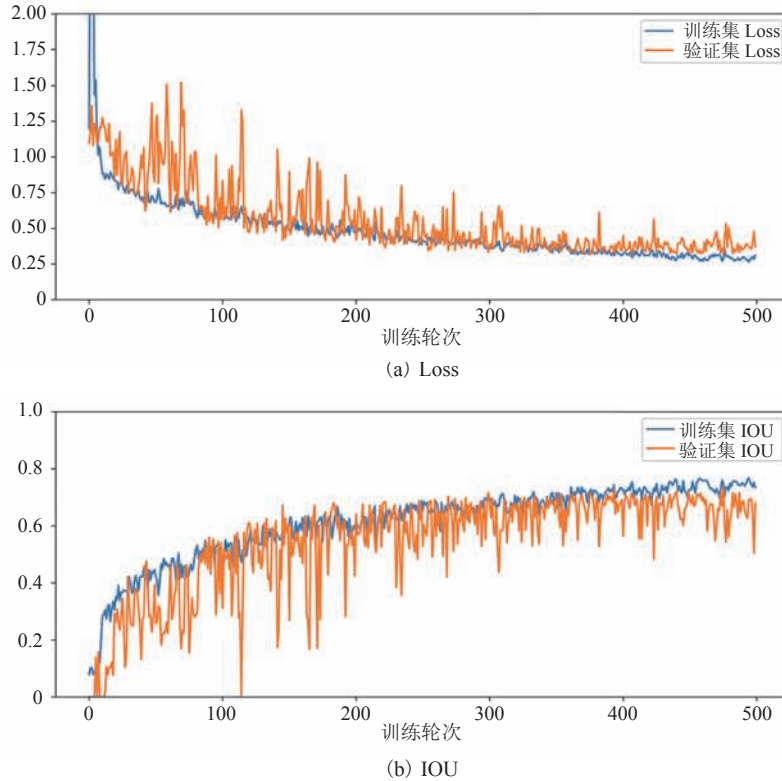


图 2 模型在训练集和测试集上 Loss 和 IOU 随训练轮次的变化曲线

Fig. 2 Curves of the Loss and IOU with training epochs on the training and test sets

训练集的 Loss 不断下降、IOU 不断上升。在验证集上没有出现明显的过拟合现象, 说明数据增强策略有效提高了模型的泛化能力。最终, 选择了在验证集上 IOU 最高的模型, 即第 475 个训练轮次的模型。该模型在训练集、验证集和测试集上以 0.5 位阈值二值化后的分割结果的 IOU 达到了 0.75、0.73 和 0.66。

模型的分割结果如图 3 所示。对模型的输出结果以 0.5 为阈值进行二值化, 可以得到最终预

测的分割结果。在该阈值下, 模型的 Dice 系数达到了 0.734 ± 0.129 。

对于 0~1 的 256 个等间隔的阈值, 本模型在测试集上的 Dice 系数达到了 0.715, 超过了仅用 COVID-19 CT 分割数据集^[10]训练的 Inf-Net 模型。此外, 基于同样的数据集和同样的测量方法, 本模型优于其他传统分割模型。同时, 模型的敏感性、特异性、 E_{ϕ}^{mean} 和 MAE 均有更好的表现, 尤其是敏感性, 相比之前的模型有较大的提

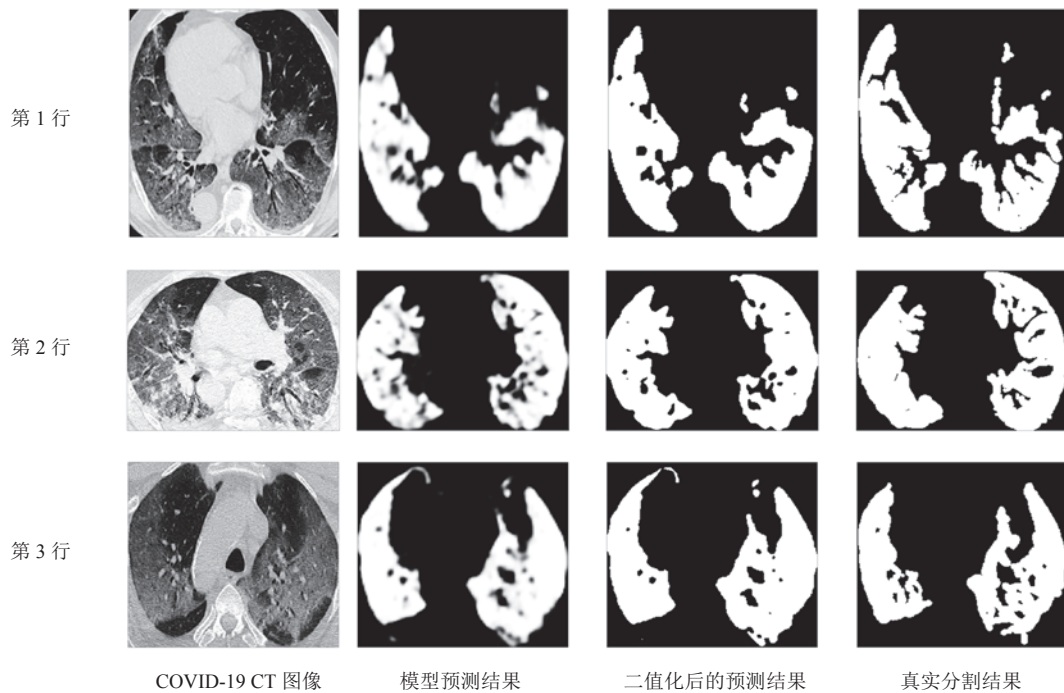


图 3 模型的分割结果示意图

Fig. 3 The segmentation results of our model

表 1 不同深度学习模型在测试集上的指标对比

Table 1 The comparison of the performance with other deep learning models on the test sets

模型	Dice 系数	敏感性	特异性	S_o	E_{ϕ}^{mean}	MAE
U-Net ^[15] *	0.439	0.534	0.858	0.622	0.625	0.186
Attention-UNet ^[16] *	0.583	0.637	0.921	0.744	0.739	0.112
Gated-UNet ^[17] *	0.623	0.658	0.926	0.725	0.814	0.102
Dense-UNet ^[18] *	0.515	0.594	0.840	0.655	0.662	0.184
U-Net++ ^[9] *	0.581	0.672	0.902	0.722	0.720	0.120
Inf-Net ^[7]	0.682	0.692	0.943	0.781	0.838	0.082
本文模型	0.715	0.754	0.952	0.769	0.884	0.073

注: *为来自文献[7]的复现结果

升。尽管 S_a 略低于 Inf-Net, 但优于其他分割模型(表 1)。

4 讨论与分析

基于 CT 图像的 COVID-19 的病灶分割在临床诊断和治疗的过程中具有非常重要的应用价值。通过计算机辅助诊断技术实现新 COVID-19 病灶分割, 可以减轻医生勾画病灶边界的工作, 提高工作效率, 同时提高了勾画的一致性和可重复性。COVID-19 病灶的分割结果, 可助力于临床诊断、定量分析和影像组学分析, 具有较高的临床应用价值。

本模型基于 Unet++ 网络^[9], 增加了注意力机制模块, 对空间重要性和通道重要性进行分析。同时, 在残差模块引入了跃层连接, 提高了模型的表达能力和梯度的传播。此外, 本实验是基于深度监督的方式来训练模型。最终, 模型对 CT 图像上 COVID-19 病灶的分割表现有所提高。

Inf-Net^[7]提出了新的网络结构和 Loss 形式, 通过挖掘更多病灶的边界信息, 提高模型的分割能力。相比于 Inf-Net, 本模型的敏感性、特异性、 E_{ϕ}^{mean} 和 MAE 均有更好的表现。尤其是敏感性从 0.692 提高到了 0.754。基于 Inf-Net 的网络架构^[7]和 1 600 张无标注图像, 可以训练半监督模型(Semi-Inf-Net)。而本实验没有进行半监督模型的构造, 可以预见的是, 引入半监督的训练策略后, 本模型结构的表现会有进一步的提升。Zhang 等^[8]构造的 COVID-19 分割模型是基于 4 965 张逐像素人工标注的 CT 图像进行训练的, 最终达到了 0.59 的 Dice 系数。尽管该模型可用于分割更多种类的病灶, 但模型训练集远大于本文所提出的模型。对于较小的数据集上的训练, 本模型提供了一种行之有效的方式。相比于基于 UNet++ 模型^[7]和其他传统分割模型^[15-18], 本文

所提出模型的性能有较大提升。

值得一提的是, 本研究所用的数据集为公开数据集, 仅由一位放射科医生进行标注。因此, 在应用到临床任务之前, 仍需要多位医生进行严谨的标注和审核, 并基于多中心对泛化性能进行广泛的验证。

5 结论

本文所提出的模型通过引入残差模块和卷积块注意力模块改进了 Unet++, 是一种基于深度监督方式的训练模型。在 COVID-19 病灶纹理、大小和位置变化较大且与正常组织间差异较小的条件下, 本方法通过对训练数据和人体解剖信息的充分利用, 提高了模型对 COVID-19 肺炎病灶的分割效果。

该模型可用于 CT 图像上的 COVID-19 病灶分割, 从而减少医生勾画肺炎病灶的时间, 提高勾画效率和勾画的一致性、客观性。达到输入一组胸部 CT 图像, 即可输出一组分割 CT 图像的效果, 实现了全自动分割 COVID-19 病灶区域的目的。

参考文献

- [1] Ai T, Yang ZL, Hou HY, et al. Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: areport of 1014 cases [J]. Radiology, 2020, 296(2): E32-E40.
- [2] Fang YC, Zhang HQ, Xie JC, et al. Sensitivity of chest CT for COVID-19: comparison to RT-PCR [J]. Radiology, 2020, 296(2): E115-E117.
- [3] Wu QX, Wang S, Li L, et al. Radiomics analysis of computed tomography helps predict poor prognostic outcome in COVID-19 [J]. Theranostics, 2020, 10(16): 7231-7244.
- [4] Wei W, Hu XW, Cheng Q, et al. Identification of common and severe COVID-19: the value of

- CT texture analysis and correlation with clinical characteristics [J]. *European Radiology*, 2020.
- [5] Chaganti S, Balachandran A, Chabin G, et al. Quantification of tomographic patterns associated with COVID-19 from chest CT [Z/OL]. arXiv Print, arXiv: 2004.012792020, 2004.
- [6] Fu H, Fan DP, Chen G, et al. COVID-19 imaging-based AI research collection [EB/OL]. [2020-10-27]. https://github.com/HzFu/COVID19_imaging_AI_paper_list.
- [7] Fan DP, Zhou T, Ji DP, et al. Inf-Net: automatic COVID-19 lung infection segmentation from CT images [J]. *IEEE Transactions on Medical Imaging*, 2020, 39(8): 2626-2637.
- [8] Zhang K, Liu XH, Shen J, et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements and prognosis of COVID-19 pneumonia using computed tomography [J]. *Cell*, 2020, 181(6): 1423-1433.
- [9] Zhou ZW, Siddiquee MMR, Tajbakhsh N, et al. UNet++: anested U-Net architecture for medical image segmentation [C] // *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018, 11045: 3-11.
- [10] COVID-19 CT segmentation dataset [EB/OL]. [2020-10-27]. <https://medicalsegmentation.com/covid19>.
- [11] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition [C] // *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770-778.
- [12] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module [C] // *European Conference on Computer Vision*, 2018: 3-19.
- [13] Kingma D, Ba J. Adam: a method for stochastic optimization [Z/OL]. arXiv Print, arXiv: 1412.6980, 2014.
- [14] Fan DP, Cheng MM, Liu Y, et al. Structure-measure: anew way to evaluate foreground maps [C] // *2017 IEEE International Conference on Computer Vision*, 2017.
- [15] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [C] // *Medical Image Computing and Computer-Assisted Intervention*, 2015: 234-241.
- [16] Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: learning where to look for the pancreas [Z/OL]. arXiv Print, arXiv: 1804.03999, 2018.
- [17] Schlemper J, Oktay O, Schaap M, et al. Attention gated networks: learning to leverage salient regions in medical images [J]. *Medical Image Analysis*, 2019, 53: 197-207.
- [18] Li XM, Chen H, Qi XJ, et al. H-DenseUNet: hybrid densely connected UNet for liver and liver tumor segmentation from CT volumes [J]. *IEEE Transactions on Medical Imaging*, 2017, 37(12): 2663-2674.