



DOI: 10.13364/j.issn.1672-6510.20230148

数字出版日期: 2024-04-17; 数字出版网址: <http://link.cnki.net/urlid/12.1355.N.20240416.1630.001>

## 基于选择性自校正卷积 U-Net 的肺部 X 射线 图像肺实质分割

王 怡, 李 昆

(天津科技大学人工智能学院, 天津 300457)

**摘要:** 针对 U-Net 分割算法无法提取多尺度特征、易受到伪影和噪声干扰而导致在肺部 X 射线图像中肺实质分割不精确的问题, 提出一种基于选择性自校正卷积的 U-Net 改进算法。改进后的 U-Net 算法将普通卷积模块替换为选择性自校正卷积模块, 该模块采用多分支结构提取多尺度特征信息, 使用 Sigmoid 函数和 Softmax 函数对多尺度特征信息进行选择性校正, 使校正后的特征信息聚焦于肺实质区域, 输出特征更加具有针对性。实验表明, 该方法对骰子系数、交并比、 $F_1$  评分结果以及对肺实质分割结果都有一定程度的提升。

**关键词:** 肺部 X 射线图像; 肺实质分割; U-Net 模型; 选择性自校正卷积

中图分类号: TP183; TP391.4

文献标志码: A

文章编号: 1672-6510(2024)04-0073-08

## Lung Parenchyma Segmentation of Lung X-ray Images Based on Selective Self-Calibration Convolution U-Net

WANG Yi, LI Kun

(College of Artificial Intelligence, Tianjin University of Science & Technology, Tianjin 300457, China)

**Abstract:** Aiming at the problem that U-Net segmentation algorithm cannot extract multi-scale features and is susceptible to artifacts and noise, which leads to imprecise segmentation of lung parenchymal in lung X-ray images, an improved U-Net algorithm based on Selective Self-Calibration convolution is proposed in this article. The improved U-Net algorithm replaces the common convolutional module with the Selective Self-Calibration convolution module, which adopts a multi-branch structure to extract multi-scale feature information, and uses Sigmoid function and Softmax function to selectively correct the multi-scale feature information, so that the corrected feature information focuses on the lung parenchyma region and the output features are more targeted. Experiments showed that this method brought some improvement on dice coefficient, intersection over union,  $F_1$  score and improved the segmentation accuracy of the lung parenchymal.

**Key words:** lung X-ray images; lung parenchymal segmentation; U-Net model; Selective Self-Calibration convolution

### 引文格式:

王怡, 李昆. 基于选择性自校正卷积 U-Net 的肺部 X 射线图像肺实质分割[J]. 天津科技大学学报, 2024, 39(4): 73-80.

WANG Y, LI K. Lung parenchyma segmentation of lung X-ray images based on Selective Self-Calibration convolution U-Net[J]. Journal of Tianjin university of science & technology, 2024, 39(4): 73-80.

医学图像分割在临床医生的诊断、治疗和疾病评估中有着不可忽视的作用。利用医学图像分割技术可以辅助医生做出定性和诊断, 从而提高医生诊断病情的正确率和有效性<sup>[1-3]</sup>。精确的肺实质分割是肺部

疾病检测的关键, 放射科医生在工作中往往需要费很大精力分析肺部 X 射线图像并作出诊断。如果在同一个肺部区域存在多种不同的病变模式时, 快速准确辨识病灶就更加不容易。加之获取的图像一般会受

收稿日期: 2023-07-29; 修回日期: 2023-11-26

作者简介: 王 怡(1971—), 男, 天津人, 副教授, wy@tust.edu.cn

噪声、伪影的影响,快速准确诊断就更加困难。

为了克服这些困难,研究者提出了基于卷积神经网络(CNN)的分割算法,卷积神经网络中的卷积核不仅可以从大量数据中提取出感受野范围内的关系特征,还具有良好的泛化能力。U-Net 模型是广泛应用于医学图像分割方面的模型架构,它是由 Ronneberger 等<sup>[4]</sup>提出的一种 U 型的编码器和解码器架构。U-Net 模型通过编码器对特征信息进行提取,通过解码器对提取出来的特征信息进行优化处理。编码器和解码器中间采用了跳跃连接修复解码过程中损失的特征信息。但是,U-Net 模型对一些细微特征的分割效果不够精确。为了解决 U-Net 模型的不足,研究人员提出了诸多以 U-Net 模型为基础的改良方法,Zhou 等<sup>[5]</sup>提出的 U-Net++通过捕捉多个层次的特征,并以特征叠加的方式进行融合,可以减小融合时特征图之间的尺度差异,但是这种方式会引入很多参数,导致内存占用增加。Wu 等<sup>[6]</sup>在编码路径和解码路径之间新增跳跃连接,并组合基础 U-Net,提出了 SCU-Net 模型,新增的跳跃连接由卷积和裁剪两个操作组成,这种方法可以提高网络的分割性能并加快收敛速度,但是裁剪和拼接后组成的特征图可能会导致特征图差异变大。Oktay 等<sup>[7]</sup>提出的 Attention U-Net 模型,通过在编码器过渡到解码器间添加 Attention 模块,逐步加强了特征图中感兴趣区域的权重,有效抑制无关区域的特征,并减少部分冗余的跳跃连接。但是,使用大量的  $1 \times 1$  卷积会导致内存访问量增加,并且在注意力模块的输出阶段只对解码区的特征值进行了比例分配,未对编码区的特征值进行比例分配,导致跳跃连接只能使用解码区的特征值。Gu 等<sup>[8]</sup>提出的 CE-Net 模型在编码部分采用了残差模块作为固定特征提取器,同时使用了空洞卷积和多尺度残差池化层作为上下文信息提取器,解码部分采用了可以学习自适应映射的转置卷积恢复高分辨率特征。虽然采用空洞卷积可以扩大感受野,但是使用空洞卷积进行提取特征值时,特征图中的所有像素不会都得到计算,从而造成卷积核不连续的情况。

为了在每个空间位置周围自适应地建立远程空间和通道间的相关性,Liu 等<sup>[9]</sup>提出了一种自校正卷积,在不同模型中应用自校正卷积可以增强分割效果,但是自校正卷积不能在校正处理阶段选择多尺度信息,该方法存在一定的局限性。本文受到自校正卷积和注意力机制等相关应用<sup>[10-15]</sup>的启发,提出了一种新型的选择性自校正卷积,通过不同的卷积核生成

不同尺度的特征信息,然后将特征信息融合,再使用自校正卷积有选择性地对不同尺度的特征进行校正。在 U-Net 算法的基础上,构建了基于选择性自校正卷积的 U-Net 算法。在 U-Net 网络中加入本文方法,可以让 U-Net 网络学习不同尺度的重要特征信息,并且不同尺度的特征信息是动态聚合的,可以根据输入的图像进行动态校正操作,使网络对肺部 X 射线图像肺实质的分割更具有鲁棒性,并且提升了网络的分割精度。

## 1 基础工作

### 1.1 传统卷积

在传统的 2D 卷积中主要有两个操作步骤:第一步是在输入的特征图中利用卷积核提取出特征值,其具体操作就是将特征图  $X$  的每个像素值和局部感受野区域的权重相乘;第二步是将其加权求和和汇聚总特征值。因此,提取出的特征图  $Y$  可以表示为

$$Y(K) = \sum_{K_n \in R \times R} W(K_n) \times X(K + K_n) \quad (1)$$

式中: $Y$  表示输出, $X$  表示输入, $K$  表示输入和输出特征图上的当前位置,表示的是局部感受野区域的第  $n$  个位置, $W(K_n)$  表示卷积核中第  $n$  个权重。每个通道的输出特征图都是由所有通道求和而得,所有的特征图都是经过多次重复式(1)生成。通过这种方式提取特征值,然后由反向传播对卷积核中的权重进行修正。但是,这种学习模式具有相似性,并且卷积层组成的网络缺乏大的感受野而不能提取更高级的特征信息,导致学习效果低下。

### 1.2 U-Net 网络

在医学图像分割领域中,U-Net 网络在医学图像分割中效果最好,它可以在数据集较小的情况下获得良好的分割结果。

U-Net 网络主干主要包括编码器、解码器和跳跃连接 3 个部分。编码器对输入的特征图进行编码操作,每轮编码需要执行 2 次  $3 \times 3$  卷积操作和 1 次  $2 \times 2$  的最大池化操作,每轮编码都使图像通道数增加 1 倍,共需 4 轮编码。解码器对特征值执行 1 次  $2 \times 2$  的转置卷积和 2 次  $3 \times 3$  的卷积操作,每次都图像通道数减半,共执行 4 轮,每轮转置卷积后需要和编码器卷积操作后获得的特征值进行跳跃连接,获取加强的特征图,经过 4 轮解码和跳跃连接,通过激活函数对特征中的所有像素点进行分割操作,最终输出分割效果图。U-Net 的网络模型结构如图 1 所示。

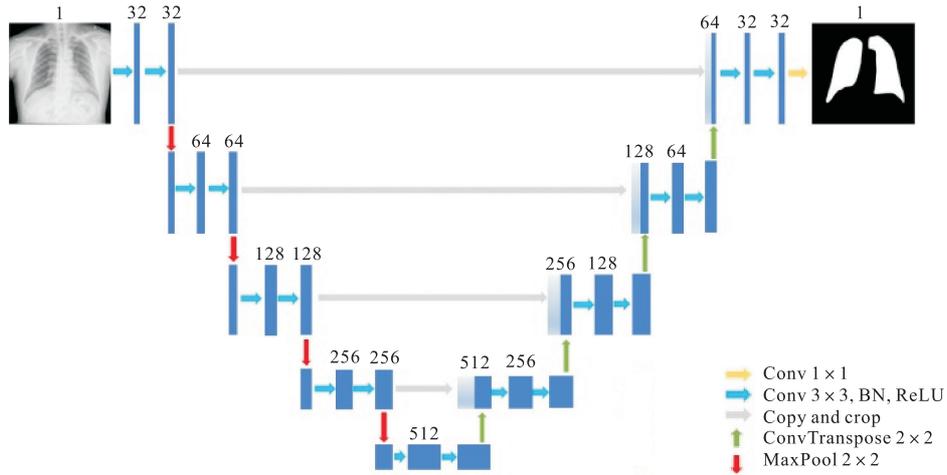


图1 U-Net网络模型结构

Fig. 1 Structure of U-Net network model

从图1中可以看出, U-Net网络模型在卷积过程中采用的是单一的卷积核, 因此无法提取不同尺度的特征信息, 而医学图像中的噪声、伪影和医疗器械线材等干扰信息在较多情况下会导致预分割区域特征提取困难, 会影响分割的准确率。针对上述问题, 本文提出基于选择性自校正卷积算法, 并与U-Net模型结合以分割肺部X射线图像中的肺实质区域。

## 2 基于选择性自校正卷积的U-Net算法

### 2.1 概述

为了解决U-Net网络模型易受到伪影和噪声干扰、无法获得不同尺度特征信息, 而导致的肺实质分割不精确的问题, 提出了选择性自校正卷积。该卷积模型采用多分支结构提取多尺度特征信息, 使用Softmax函数和Sigmoid函数对多尺度特征信息进行选择性校正, 模型结构如图2所示。

训练过程中,  $I$  的输入通道和  $Y$  的输出通道都为  $C$ , 其中  $\{W_1, W_3, W_5, W_7\}$  表示大小为  $3 \times 3$  的卷积核,  $\{W_2, W_4, W_6, W_8\}$  表示大小为  $5 \times 5$  的卷积核。输入  $I$  经过  $W_1$  和  $W_2$  卷积得到  $I_1$  和  $I_2$  两部分, 每部分都经过3个分支, 提取不同尺度下的特征信息。 $W_3$ 、 $W_5$  和 Select 模块对  $I_1$  进行选择性自校正操作生成  $Y_1$ ,  $W_4$ 、 $W_6$  和 Select 模块对  $I_2$  进行选择性自校正操作生成  $Y_2$ 。 $I_1$ 、 $I_2$  分别拼接  $Y_1$ 、 $Y_2$  生成  $Y'_1$  和  $Y'_2$ ,  $I_1$  和  $I_2$  的作用是保留原始的特征信息。最后将  $Y'_1$  和  $Y'_2$  融合得到输出  $Y$ 。

### 2.2 选择性自校正卷积

选择性自校正卷积由 Split 和 Selective Self-

Calibration 两部分组成。

**Split:** 对输入的特征图  $I$  分别进行2次卷积核为  $3 \times 3$  和  $5 \times 5$  的卷积操作, 生成  $I_1$  和  $I_2$  两种不同尺度的特征图, 丰富图像中肺实质特征信息。

$$I_1 = \mathcal{F}_1(I) = \sum_{K_n \in 3 \times 3} W_1(K_n) \times I(K_n) \quad (2)$$

$$I_2 = \mathcal{F}_2(I) = \sum_{K_n \in 5 \times 5} W_2(K_n) \times I(K_n) \quad (3)$$

式中:  $\mathcal{F}_1$  表示由卷积核大小为  $3 \times 3$  卷积操作、归一化和 ReLU 激活函数构成的网络层。 $\mathcal{F}_2$  表示由卷积核大小为  $5 \times 5$  的卷积操作、归一化和 ReLU 激活函数构成的网络层。在模型中使用 ReLU 激活函数增加网络的稀疏性, 可以缓解过拟合情况并缩短模型的训练时长。

**Selective Self-Calibration:** 对  $I_1$  和  $I_2$  融合的特征图进行卷积核尺寸为  $r \times r$ , 步长为  $r$  的平均池化, 这一步的目的是强化不同尺度的特征信息间共有的特征, 并且让下一层的神经元学习不同尺度的特征。下采样操作会扩充感受野的特性, 通过较小内在空间变换的插值操作可以指导原始特征空间中的特征变换过程, 可表示为

$$I' = \text{AvgPool}_r(I_p) = \text{AvgPool}_r(I_1 + I_2) \quad (4)$$

对  $I'$  分别使用卷积核大小为  $3 \times 3$  和  $5 \times 5$  的卷积操作, 并对其进行特征变换, 表示为

$$I'_1 = \text{Up}(\mathcal{F}_3(I')) = \text{Up}\left(\sum_{K_n \in 3 \times 3} W_3(K_n) \times I'(K_n)\right) \quad (5)$$

$$I'_2 = \text{Up}(\mathcal{F}_4(I')) = \text{Up}\left(\sum_{K_n \in 5 \times 5} W_4(K_n) \times I'(K_n)\right) \quad (6)$$

其中:  $\text{Up}()$  为双线性插值操作, 目的是将多尺度卷积提取的特征信息从小尺度空间映射到原始特征空间。

Selective Self-Calibration 中的 Select 模块: 该模块负责选择多尺度特征信息校正的规模, 其基本思想是使用 Softmax 函数控制特征信息传入 Sigmoid 函数的权重比。

首先, 通过元素求和融合多分支的结果, 表示为

$$I_a = I'_1 + I_p \quad (7)$$

$$I_b = I'_2 + I_p \quad (8)$$

式中  $I_p$  为残差项, 解决梯度消失和网络退化的问题。

然后, 增加新的维度对  $I_a$  和  $I_b$  进行堆叠, 通过 Softmax 函数自适应选择  $I_a$  和  $I_b$  传入 Sigmoid 函数的权重比。

$$r = \sigma(\eta(I_a, I_b)) \quad (9)$$

$$r_1 = \frac{e^{I_a}}{e^{I_a} + e^{I_b}} \quad (10)$$

$$r_2 = \frac{e^{I_b}}{e^{I_a} + e^{I_b}} \quad (11)$$

$$r_1 + r_2 = 1 \quad (12)$$

式中:  $\eta$  表示扩充维度后堆叠特征信息的操作, 得到形状为  $(2, B, C, W, H)$  的特征信息, 其中 2 为扩充维度后堆叠特征信息的个数,  $B$  表示样本数量,  $C$  表示特征信息通道,  $W$  和  $H$  表示特征信息的宽和高;  $\sigma$  表示 Softmax 操作;  $r$  表示 Softmax 函数的输出结果, 其中  $r_1 \in r$  表示  $r$  结果集中第 0 维度的第 1 个, 即  $I_a$  的权重比,  $r_2 \in r$  表示  $r$  结果集中第 0 维度的第 2 个,

即  $I_b$  的权重比。

因此, Select 模块的输出结果可以表示为

$$I'_a = I_a \times r_1 \quad (13)$$

$$I'_b = I_b \times r_2 \quad (14)$$

将  $I'_a$  和  $I'_b$  传入 Sigmoid 函数, 进行选择性自校正操作, 为

$$Y_1 = \mathcal{F}_7(\varphi(I'_a) \times \mathcal{F}_5(I_1)) \quad (15)$$

$$Y_2 = \mathcal{F}_8(\varphi(I'_b) \times \mathcal{F}_6(I_2)) \quad (16)$$

式中:  $\varphi$  表示 Sigmoid 操作。

拼接  $Y_1$ 、 $I_1$  和  $Y_2$ 、 $I_2$  结果为

$$Y'_1 = \varepsilon(Y_1, I_1) \quad (17)$$

$$Y'_2 = \varepsilon(Y_2, I_2) \quad (18)$$

式中:  $\varepsilon$  表示拼接操作,  $I_1$  和  $I_2$  为残差项。

通过式 (15) 和式 (16) 的选择性自校正操作, 可以聚焦每个空间位置周围的特征信息, 并自适应地融合到高维度的内在空间, 将其作为原始比例空间相应的输入, 因此可以对通道间的依赖关系进行建模。

最后, 融合经过选择性自校正卷积操作的多尺度特征图, 计算公式为

$$Y = Y'_1 + Y'_2 \quad (19)$$

选择性自校正操作只关注相关区域感兴趣的特征信息, 提取每个感兴趣区域周围的上下文信息, 通过多通道间的依赖关系, 可以扩充融合了多尺度信息并且具有选择性自校正卷积层的感受野。

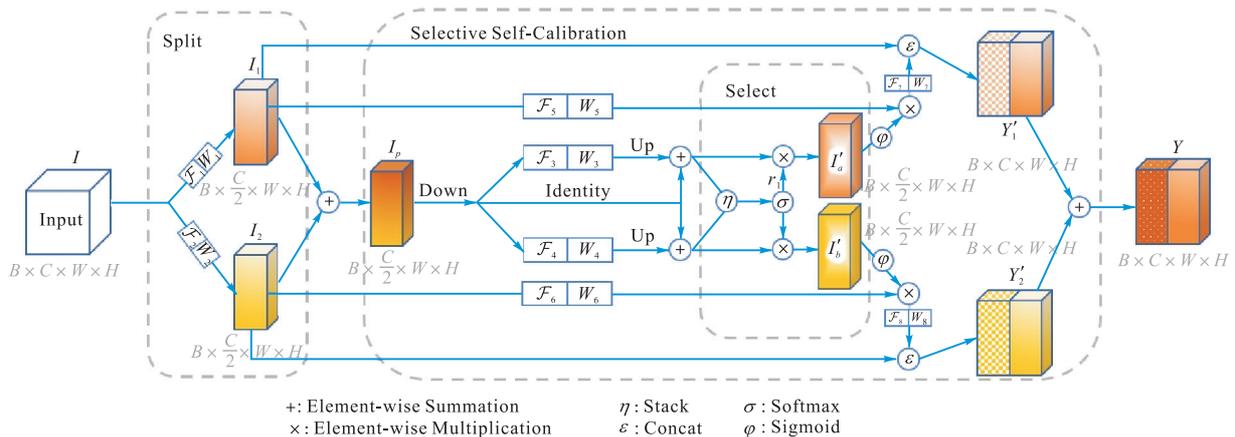


图2 选择性自校正卷积结构

Fig. 2 Selective Self-Calibration convolution structure

### 2.3 基于选择性自校正卷积的 U-Net 算法

本文提出的基于选择性自校正卷积的 U-Net 算法保留了 U-Net 模型上的跳跃连接, 并将编码部分和解码部分的第 2 次卷积操作更改为选择性自校正卷积, 使用选择性自校正卷积可以有选择性地对图像中不同尺度的特征信息进行校正, 并且重点关注肺实质

分割区域的特征信息。基于选择性自校正卷积的 U-Net 模型结构如图 3 所示。主干网络的编码部分由 5 次卷积、选择性自校正卷积和池化层组成, 每次编码时都通过卷积提升通道数, 卷积后都添加了一个归一化层和 ReLU 激活函数。归一化层和 ReLU 激活函数在反向传播时能够产生更明显的梯度, 使模型更容易

收敛, 并避免梯度消失的问题。使用选择性自校正卷积对其不同尺度的特征信息进行选择校正, 选择性自校正卷积分别采用了卷积核大小为  $3 \times 3$  和  $5 \times 5$  的卷积操作提取图像中不同尺度的特征信息。选择校正操作后经过池化层后进入下一轮编码操作。解码器部分由 4 次的转置卷积、选择性自校正卷积和卷积操

作组成, 并在转置卷积前使用了跳跃连接, 使高级语义信息可以更好地融合低级语义信息, 使用卷积操作对融合了跳跃连接特征信息的通道数缩减为原来的  $1/2$ , 然后通过选择性自校正卷积对融合的语义信息进行校正操作, 在解码的最后一轮中选择性自校正卷积操作后使用分类函数输出分割图。

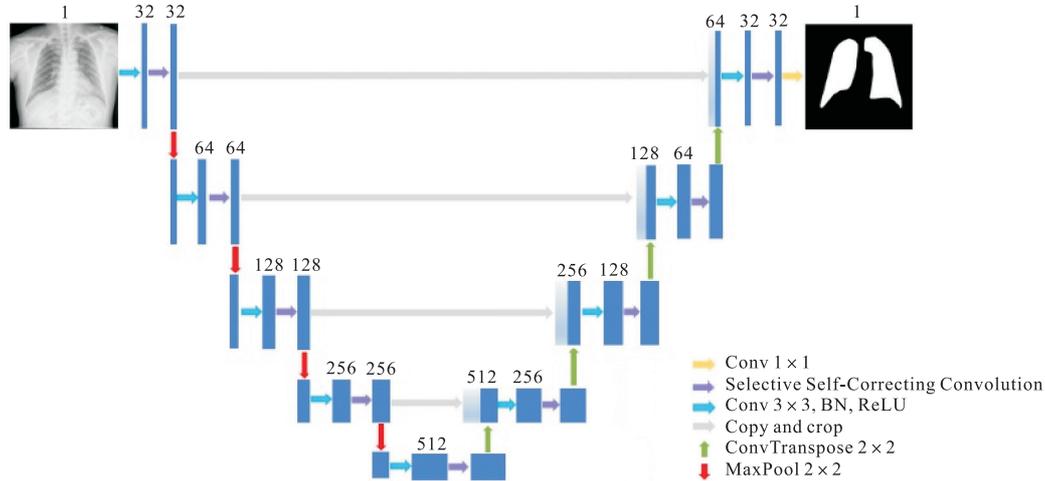


图 3 基于选择性自校正的 U-Net 模型结构

Fig. 3 Structure of U-Net model based on Selective Self-Calibration

基于选择性自校正卷积的 U-Net 算法通过选择性自校正卷积操作对原始图像生成不同尺度的特征信息, 有选择性地校正特征信息, 并重点关注预分割区域的特征信息, 从而达到模型对重要特征的提取, 提高模型对于肺部 X 射线图像肺实质的分割效果。

### 3 实验

#### 3.1 数据集

实验所用 2 个数据集为 COVID-19 Radiography Database 和 Pulmonary Chest X-Ray Abnormalities。COVID-19 Radiography Database 数据集是 Kaggle 数据集奖获得者, 它由 21 165 张肺部 X 射线图像和 21 165 张对应的分割掩码组成, 图像的分辨率为  $299 \times 299$ 。Pulmonary Chest X-Ray Abnormalities 数据集由 800 张肺部 X 射线图像和 704 张分割掩码组成, 其中 658 张肺部 X 射线图像由广东医学院和深圳市第三人民医院提供, 142 张图片由美国马里兰州蒙哥马利县卫生与公共服务部提供, 图像的分辨率为  $2900 \times 2900$ 。数据集中的部分图像如图 4 所示, 其中图 4(a)—4(c) 为患者肺部 X 射线图像, 红色箭头指向的是医疗设施线材的干扰, 图 4(d)—4(f) 中黑色为背景像素, 白色表示分割出的肺实质区域。

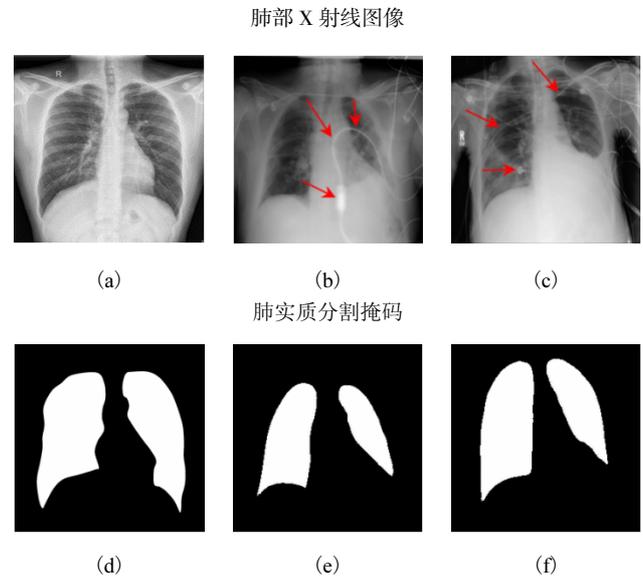


图 4 部分数据集图像

Fig. 4 Partial data set images

#### 3.2 实验设置

实验在 PyTorch 1.10 环境下进行, RTX3090 (24 GB) 训练加速, Adam 优化器<sup>[16]</sup>优化训练, 余弦退火 (cosine annealing)<sup>[17]</sup>学习率调节策略, 损失函数为混合损失函数, 分别是二元交叉熵损失函数 (binary cross entropy, BCE)<sup>[18]</sup>和 Dice 相似系数 (dice similarity coefficient, DSC) 损失函数<sup>[19]</sup>。BCE 损失函数是二

分类的常用损失函数,为

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N t \times \ln p + (1-t) \times \ln(1-p) \quad (20)$$

式中:  $N$  为样本数量,  $p$  为模型预测值,  $t$  为真实值。

由于 BCE 损失函数在计算前景像素和背景像素不均衡的图像数据时,会导致训练时学习的重心会更加偏向于图像中前景或背景中像素较多的那一类。由于在肺部 X 射线图像中存在肺实质的像素面积小于背景像素面积的情况,导致前景像素和背景像素分布不均。如果只采用 BCE 损失函数,就会导致训练模型更加倾向于背景的分割而忽略了肺实质分割,从而降低模型的分割效果。

DSC 损失函数的数值越小就代表预测值和真实值越相似,因此 DSC 损失函数可以表示为

$$L_{\text{DSC}} = 1 - 2 \frac{|X \cap Y|}{|X| + |Y|} \quad (21)$$

式中:  $|X \cap Y|$  表示的是样本和之间像素值取交集的数量,  $|X|$  和  $|Y|$  分别表示样本  $X$  和  $Y$  中像素的数量。

采用 DSC 损失函数可以缓解样本中前景和背景不均衡的情况,但是只采用 DSC 损失函数又会导致损失饱和问题,所以让 DSC 损失函数和 BCE 函数相融合,生成混合损失函数 (Mixed Loss),使两函数互补,增强分割效果,混合损失函数公式为

$$L_{\text{Mixed}} = L_{\text{DSC}} + L_{\text{BCE}} \quad (22)$$

针对 COVID-19 Radiography Database 数据集选取其中的 8 000 张图像作为训练集,1 000 张图像作为测试集。在 Pulmonary Chest X-Ray Abnormalities 数据集中采用其中 600 张图像作为训练集,100 张图像作为测试集。

### 3.3 评分标准

本文采用了 4 种在医学图像分割领域中常用的分割性能指标:骰子系数(dice)、交并比(IoU)、 $F_1$ 、特异性(specificity)。

骰子系数表示预测分割图像和真实掩码图像相交的面积,与真实掩码图像面积的比值,取值范围在  $[0,1]$  之间,骰子系数计算公式可表示为

$$S_{\text{dice}} = \frac{2N_{\text{TP}}}{2N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}} \quad (23)$$

交并比是计算真实值和分割值两个集合的交集和并集的比值,交并比的计算公式为

$$S_{\text{IoU}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}} \quad (24)$$

$F_1$  是综合考虑精确率评分和召回率评分的结果计算出的衡量函数。

$$P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \quad (25)$$

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (26)$$

$$S_{F_1} = \frac{2P \times R}{P + R} \quad (27)$$

式中:  $P$  为精确率,表示在被预测为正的样本中实际为正的比例;  $R$  为召回率,表示在实际为正的样本中预测为正的比例;  $S_{F_1}$  为  $F_1$  评分指标。

特异性指的是预测为负的样本数在所有负样本中的占比,计算公式为

$$S_{\text{specificity}} = \frac{N_{\text{TN}}}{N_{\text{TN}} + N_{\text{FP}}} \quad (28)$$

式中:  $N_{\text{TP}}$  表示真正样本数,即预测为正,真实为正的样本数量;  $N_{\text{TN}}$  表示真负样本数,即预测为负,真实为负的样本数量;  $N_{\text{FN}}$  表示假负样本数,即预测为负,真实为正的样本数量;  $N_{\text{FP}}$  表示假正样本数,即预测为正,真实为负的样本数量。

### 3.4 实验结果与分析

为了证实本文提出的基于选择性自校正卷积的 U-Net 算法模型的有效性,在 COVID-19 Radiography Database 和 Pulmonary Chest X-Ray Abnormalities 两个数据集上进行实验,评分结果对比见表 1。在 COVID-19 Radiography Database 数据集中,本文方法比其他算法在骰子系数、交并比和  $F_1$  评分的最优值分别高 1.57%、3.03% 和 0.22%,在特异性上的评分比最优值低 0.21%。在 Pulmonary Chest X-Ray Abnormalities 数据集中,本文方法同样比其他算法在骰子系数、交并比和  $F_1$  评分的最优值分别高 3.19%、5.25% 和 0.77%,在特异性上的评分比最优值低 0.06%。

从图 5 展示的结果对比来看,本文方法比其他网络具有更好的特征提取能力,例如图 5 中 (b-3) 和 (c-3),在图像中有干扰的情况下 U-Net 分割的肺实质主体不够理想。在图 5 (a-4)、(b-4)、(c-4)、(a-5)、(b-5)、(c-5) 中,U-Net++ 和 Attention U-Net 虽然可以分割出肺实质主体部分,但是针对肺实质肺尖和肺底等细节分割不够精细,而本文方法分割的肺实质主体、边界、肺尖和肺底等细节信息均比 U-Net、U-Net++ 以及 Attention U-Net 网络清晰精确,如图 5 (a-6)、(b-6)、(c-6) 所示。

评分结果和分割结果表明,本文方法在两个数据集中的骰子系数、交并比和  $F_1$  评分比其他算法均有所提升,虽然特异性的得分稍逊色于其他模型,但通

过分割结果对比可知,本文方法对肺部X射线图像肺实质的分割效果比其余3种算法在抗干扰能力、肺实质主体分割以及肺实质细节分割能力优秀,证明选择性自校正卷积的有效性。本文方法通过选择性自

校正卷积提取多尺度特征并进行选择性校正,增强了网络对于特征的提取能力,充分考虑每个空间位置周围上下文信息,避免来自不相关区域无用特征信息的干扰,使肺实质细节等信息处理更加清晰、精准。

表1 评分结果对比

Tab. 1 Comparison of scoring results

模型	COVID-19 Radiography Database				Pulmonary Chest X-Ray Abnormalities			
	骰子系数/%	交并比/%	$F_1$ /%	特异性/%	骰子系数/%	交并比/%	$F_1$ /%	特异性/%
U-Net	94.23	89.01	98.13	99.61	86.93	76.88	95.81	98.91
U-Net++	95.99	92.29	97.82	99.62	88.27	79.01	95.47	98.68
Attention U-Net	95.31	91.04	98.11	99.63	87.34	77.53	96.05	98.79
本文方法	97.56	95.32	98.35	99.42	91.46	84.26	96.82	98.85

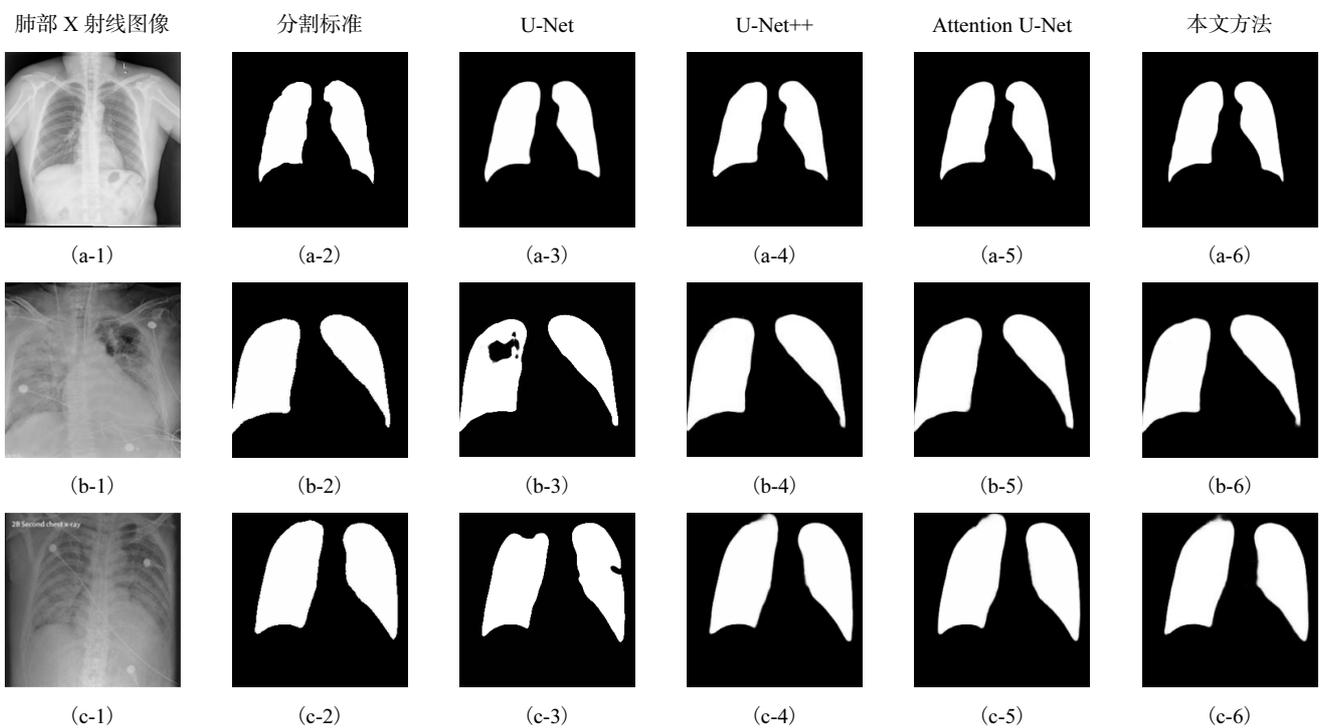


图5 分割结果对比

Fig. 5 Comparison of segmentation results

## 4 结 语

本文提出了一种基于选择性自校正卷积的U-Net网络模型,该模型采用选择性自校正卷积扩大感受野范围,并对输入图像进行不同尺度的特征提取,然后进行选择性自校正处理,通过校正后的特征会更加关注预分割区域,实验过程中采用BCE损失函数和DSC损失函数作为混合损失函数,采用余弦退火学习率策略调节学习率。通过在COVID-19 Radiography Database和Pulmonary Chest X-Ray Abnormalities数据集上进行评分实验和分割实验,结果表明本文方法在肺部X射线图像中肺实质分割的有效性。

## 参考文献:

- [1] 黄晓鸣,何富运,唐晓虎,等. U-Net及其变体在医学图像分割中的应用研究综述[J]. 中国生物医学工程学报, 2022, 41(5): 567-576.
- [2] 王国力,孙宇,魏本征. 医学图像图深度学习分割算法综述[J]. 计算机工程与应用, 2022, 58(12): 37-50.
- [3] 曹玉红,徐海,刘荪傲,等. 基于深度学习的医学影像分割研究综述[J]. 计算机应用, 2021, 41(8): 2273-2287.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and

- Computer-Assisted Intervention-MICCAI 2015 : 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Berlin : Springer International Publishing, 2015: 234-241.
- [5] ZHOU Z, RAHMAN S M M, TAJBAKHS N, et al. U-Net++: a nested U-Net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Berlin : Springer International Publishing, 2018: 3-11.
- [6] WU J, ZHANG Y, WANG K, et al. Skip connection U-Net for white matter hyperintensities segmentation from MRI[J]. IEEE Access, 2019, 7: 155194-155202.
- [7] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: learning where to look for the pancreas [EB/OL]. [2023-06-01]. <https://arxiv.org/pdf/1804.03999.pdf>.
- [8] GU Z, CHENG J, FU H, et al. Ce-Net: context encoder network for 2D medical image segmentation[J]. IEEE Transactions on medical imaging, 2019, 38(10): 2281-2292.
- [9] LIU J J, HOU Q, CHENG M M, et al. Improving convolutional networks with self-calibrated convolutions [C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 10096-10105.
- [10] LI X, WANG W, HU X, et al. Selective kernel networks[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2019: 510-519.
- [11] ZHANG H, WU C, ZHANG Z, et al. ResNeSt: Split-Attention networks[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2022: 2736-2746.
- [12] YU Z, ZHAO C, WANG Z, et al. Searching central difference convolutional networks for face anti-spoofing [C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 5295-5305.
- [13] CHEN Y, DAI X, LIU M, et al. Dynamic convolution: attention over convolution kernels[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 11030-11039.
- [14] HUANG Z, WANG X, HUANG L, et al. CCNet: criss-cross attention for semantic segmentation[C]//IEEE. Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2019: 603-612.
- [15] ZHU Q, PENG S, ZHANG R, et al. An improved U-Net based on self-adaptive difference convolution[C]//IEEE. 2022 7th International Conference on Computer and Communication Systems (ICCCS). New York: IEEE, 2022: 17-21.
- [16] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL]. [2023-06-01]. <http://arxiv.org/pdf/1412.6980.pdf>.
- [17] HUANG G, Li Y, PLEISS G, et al. Snapshot ensembles: train 1, get  $M$  for free [EB/OL]. [2023-06-01]. <https://arxiv.org/pdf/1704.00109.pdf>.
- [18] DE BOER P T, KROESE D P, MANNOR S, et al. A tutorial on the cross-entropy method[J]. Annals of operations research, 2005, 134: 19-67.
- [19] DE SOUZA REBOUÇAS E, DE MEDEIROS F N S, MARQUES R C P, et al. Level set approach based on Parzen Window and floor of log for edge computing object segmentation in digital images[J]. Applied soft computing, 2021, 105: 107273.

责任编辑: 郎婧