

史健婷, 李雪瑶, 李志军. 基于 SRL-UNet 的肾脏病变图像分割算法[J]. 智能计算机与应用, 2024, 14(11): 80-87. DOI: 10.20169/j. issn. 2095-2163. 24042801

## 基于 SRL-UNet 的肾脏病变图像分割算法

史健婷<sup>1</sup>, 李雪瑶<sup>1,2</sup>, 李志军<sup>3</sup>

(1 黑龙江科技大学 计算机与信息工程学院, 哈尔滨 150022; 2 黑龙江科技大学 研究生学院, 哈尔滨 150022;  
3 梧州学院 广西机器视觉与智能控制重点实验室, 广西 梧州 543002)

**摘要:** 肾脏的病变通常隐匿无明显症状, 随着时间推移, 还可能导致并发症的出现。由于肾脏产生病变原因复杂多样, 且造成病变因素众多, 因此针对病变肾脏的准确诊断至关重要。精确的病变肾脏图像分割对临床诊断具有重要价值, 为此提出一种基于改进 UNet 的肾脏病变图像分割算法 SRL-UNet, 用于自动检测和分割肾脏中的病变区域。通过在 UNet 模型中引入 SE 注意力机制, 并且使用 ResNet50 作为 UNet 的主干网络, 然后将 UNet 和水平集方法结合生成混合损失函数, 实现网络端端的训练来提高模型泛化性能, 增强图像边缘的分割能力。仿真结果表明, SRL-UNet 算法得到的准确率为 0.993 3, *MIoU* 指数为 0.812 7, 与 UNet、DeepLabv3 和 HRNet 三种方法比较, 准确率分别提高了 0.61%、0.35% 和 0.24%; *MIoU* 指数分别提高了 18.21%、4.04% 和 10.02%。同时, 也分别解决了 3 种方法产生的空洞问题、分割缺失问题以及边缘不光滑问题, 达到了良好的分割效果, 体现出所提方法在分割算法性能上的优势。

**关键词:** 图像分割; UNet; 深度学习; 水平集方法; 病变肾脏

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2024)11-0080-08

## Renal lesion image segmentation algorithm based on SRL-UNet

SHI Jianting<sup>1</sup>, LI Xueyao<sup>1,2</sup>, LI Zhijun<sup>3</sup>

(1 School of Computer and Information Engineering, Heilongjiang University of Science and Technology, Harbin 150022, China;  
2 Graduate College, Heilongjiang University of Science and Technology, Harbin 150022, China;  
3 Guangxi Key Laboratory of Machine Vision and Intelligent Control, Wuzhou University, Wuzhou 543002, Guangxi, China)

**Abstract:** Renal lesions are usually asymptomatic and may lead to complications over time. Due to the complex and diverse causes of kidney disease, as well as the numerous factors that cause it, accurate diagnosis of diseased kidneys is crucial. Accurate segmentation of diseased kidney images is of great value for clinical diagnosis. Therefore, a kidney disease image segmentation algorithm SRL-UNet based on improved UNet is proposed for automatic detection and segmentation of diseased areas in the kidney. By introducing SE attention mechanism into the UNet model and using ResNet50 as the backbone network, and combining UNet and level set methods to generate a mixed loss function, end-to-end training of the network is achieved to improve the model's generalization performance and enhance the segmentation ability of image edges. The results show that the accuracy obtained by the SRL-UNet algorithm is 0.993 3, with a *MIoU* index of 0.812 7. Compared with UNet, DeepLabv3, and HRNet methods, the accuracy has been improved by 0.61%, 0.35%, and 0.24%, respectively; The *MIoU* index increased by 18.21%, 4.04%, and 10.02%, respectively. At the same time, the three methods have also solved the problems of voids, missing segmentation, and uneven edges, achieving good segmentation results, reflecting the advantages of the proposed method in segmentation algorithm performance.

**Key words:** image segmentation; UNet; deep learning; level set method; pathological kidney

## 0 引言

肾脏是人类主要脏器之一, 当肾脏出现疾病时,

会对人的身体健康造成严重的不良影响。在肾脏病变初期与肺癌初期的肺结节相似, 通常没有明显症状。随着时间的推移, 异常细胞逐渐积聚, 可能导致

基金项目: 2023 年度黑龙江省属高校基本科研业务费项目(2023-KYYWF-0547)。

作者简介: 史健婷(1981—), 女, 博士, 教授, 主要研究方向: 图像处理与模式识别, 多维数据分析与可视化; 李雪瑶(1999—), 女, 硕士研究生, 主要研究方向: 图像处理, 模式识别。

通信作者: 李志军(1973—), 男, 副教授/高级工程师, 主要研究方向: 物联网技术, 人工智能。Email: 292905400@qq.com。

收稿日期: 2024-04-28

炎症和其他病症。因此,肾脏病变的确切诊断对其治疗和预后至关重要。

肾脏图像分割效果可以提升肾脏病变的确诊率,通过医学影像来观察患者肾脏的位置和形态,判断是否发生病变。对于放射科医生,手动分割病变肾脏是一个非常繁琐的过程,这些情况可能会导致放射治疗手术不及时而使病人存活率低等情况的发生<sup>[1]</sup>。手动分割不仅会耗费医生大量的时间和精力,还会过度依赖医生的经验和技巧,因此手动分割得到的结果在准确性和有效性上存在不确定性。自动分割技术能够提供有效且更精确的分割结果,提升分割效率,节省时间用来优化放射治疗计划,进而降低放射治疗的副作用。传统的自动分割技术通常根据图像的灰度、纹理和梯度等浅层特征分割目标,常用方法有阈值法<sup>[2-3]</sup>、图谱法<sup>[4]</sup>和区域生长法<sup>[5]</sup>等。在进行分割操作时,需要把图像中的每个像素划分为背景和前景<sup>[6]</sup>。当图像的背景与目标之间的灰度值差异不明显时,例如在血管分割中,要准确有效地进行图像分割就极具挑战性。此外,由于肾脏结构复杂,病变区域可能较小且形状不规则,传统的图像分割方法受到一定限制,可能导致模型的分割精度不高,甚至在深层次网络模型中丢失关键信息,影响病变肾脏的准确诊断。

近年来,深度学习在医学图像处理中取得了显著的成绩。在各种基于深度学习网络模型的病变肾脏自动分割技术中,卷积神经网络(Convolutional Neural Network, CNN)表现出了出色的图像特征提取能力。与此同时, CNN 具有相对较小的网络参数量,但却展现出强大的泛化能力。例如, Bianconi 等学者<sup>[7]</sup>对目前主流的传统分割方法和深度学习方法进行了对比和分析。Gui 等学者<sup>[8]</sup>提出了一种基于微分几何的等周不等式模型,将等周约束融入水平集框架中。可以有效地分割医学图像中那些边界模糊或部分缺失的病灶。深度学习架构的子集 UNet 神经网络,特别设计用于图像分割,已经在各种医学图像分割任务中得到广泛应用。例如:陈铭等学者<sup>[9]</sup>通过改进 UNet 模型对肺结节图像分割做出了进一步的研究,提高了分割的准确率。Ronneberger 等学者<sup>[10]</sup>研究了 UNet 模型在生物医学图像分割上得到了良好的实验结果并且赢得了细胞追踪挑战赛。Kim 等学者<sup>[11]</sup>提到水平集方法是一种数学工具,用于图像分割,具有处理复杂边界和区域的能力。

深度学习方法 UNet 神经网络在医学图像分割中能够自动学习特征和模式,提高分割的准确性。为了提高模型在肾脏图像中学习和识别病变区域的能力,提高分割的准确性和稳定性,提出了 SRL-UNet (Squeeze - and - excitation Resnet Level - loss UNet) 模型算法,其中 SE (Squeeze - and - Excitation) 注意力机制可以提高网络的判别能力和泛化能力。ResNet 主干网络能够提高网络的深度,使模型更加有效地捕获关键特征并提高模型的性能,加速模型的训练和优化;结合水平集方法和 UNet 神经网络,可以根据图像中的像素强度来确定病变的位置,然后计算分割掩码中病变像素的数量估计病变的大小,融合 2 种方法提高对病变肾脏区域分割的准确率、边界分割的清晰度。Litjens 等学者<sup>[12]</sup>表明自动化图像分割技术可以显著降低医生的误诊率和漏诊率,为后续的肿瘤的认识和疾病诊断打下基础,同时将 UNet 深度学习模型与水平集方法相结合具备了良好的可解释性,便于医生理解,降低医学诊断的工作负担。

## 1 SRL-UNet 算法

### 1.1 SRL-UNet 网络整体结构

UNet 是一种常用于图像分割的卷积神经网络(CNN)结构,特别适用于医学图像分割任务。主要由编码器(Encoder)、解码器(Decoder)中间层和跳跃连接结构组成,输入的图像通过编码器产生特征图,逐步提取图像的特征信息,特征图通过跳跃连接传递给解码器,保留了编码器中不同层次的特征信息,解码器利用这些特征图进行上采样和卷积操作,逐步恢复图像的空间尺寸并生成最终的分割结果。

改进后的 SRL-UNet 网络整体结构如图 1 所示。在 UNet 的解码器引入 SE (Squeeze and Excitation) 注意力机制,通过学习特征通道之间的关系,动态调整特征图的权重,使得模型能够更加有效地学习和利用图像中的重要信息,从而抑制对不重要信息的过度关注。同时把 ResNet 网络结构作为 UNet 编码器的部分,由于 ResNet 具有较深的网络结构,可以帮助缓解梯度消失和网络退化问题,使得模型更容易训练和优化。SRL-UNet 算法结构扩展了对重要特征的关注和对高级特征的学习能力,从而提高了病变肾脏分割任务的准确性和稳定性。

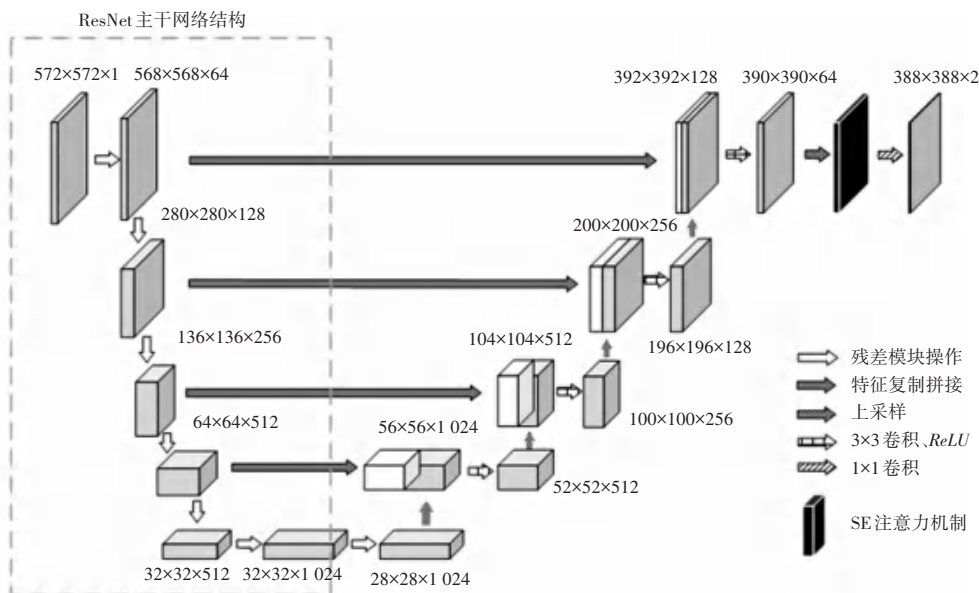


图1 SLR-UNet 网络结构

Fig. 1 SLR-UNet network architecture

## 1.2 SE 注意力机制

引入SE注意力机制在UNet网络解码器的最后一个上采样层的输出之后。具体包括全局平均池化层(AdaptiveAvgPool2d)和2个全连接层(Linear),最后使用Sigmoid激活函数。压缩(Squeeze)和激励(Excitation)是模块中的2个关键操作,压缩操作对尺寸为 $H \times W \times C$ 的特征图进行全局平均池化,压缩为 $1 \times 1 \times C$ 向量。其中,激励操作使用一个全连接神经网络,对压缩之后的结果做一个非线性变换,得到的结果作为每个通道的权重。SE注意力机制结构如图2所示。通过引入SE注意力机制,可以提升UNet模型的特征学习能力和预测性能,从而对病变肾脏图像进行更好分割。

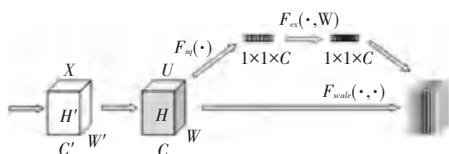


图2 SE网络结构

Fig. 2 SE network structure

## 1.3 Res50-UNet 网络结构

本文提出将ResNet50作为UNet主干网络,形成Res50-UNet网络结构,用该模型来实现对病变肾脏图像的语义分割任务。由于ResNet50网络结构较深并且引入了残差连接,可以更好地学习图像的高级特征。ResNet50根据VGG19网络的形式,进一步采用全局平均池化替代全连接层,并引入了快捷连接的机制。在ResNet50中,特征图由残差

图和恒等映射组合而成,输出表示为:

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

这种输出被转化为残差函数:

$$F(x) = H(x) - x \quad (2)$$

该转化便于网络学习到残差函数并进行相应调整,ResNet50网络结构如图3所示,其中包含3个归一化层(Batch Normalization, BN)、2个激活函数层(Rectified Linear Unit, ReLU)和3个卷积层(Convolutional layer, Conv)标识块。Identity Block主要用于网络加深,包含1个卷积层(Conv)和1个归一化层(BN),此结构与原始特征提取网络相比学习能力更强。训练时,特征提取网络的残差会出现趋近于零的情况,实际上是网络变成了一个恒等映射 $H(x) = x$ ,这表明网络可能已经达到了最佳状态。虽然网络可能会继续加深,但由于网络仅进行映射,其性能不会随着深度增加而下降,因此避免了因深度增加而导致的性能降级问题。

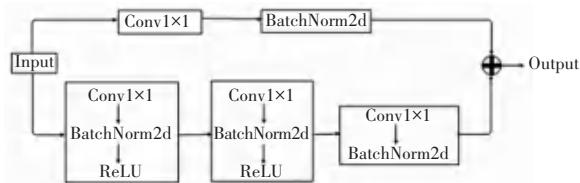


图3 ResNet50 残差模块

Fig. 3 ResNet50 residual module

将UNet的主干网络结构替换为ResNet50网络首先需要输入图片并对图像进行预处理,将图像作为输入传递到ResNet50的前向传播过程中,包括一

系列的卷积、池化、残差块等操作, 逐渐提取图像中的特征信息。由于 ResNet50 的主干网络结构会逐层提取图像的特征, 使其从低级到高级的特征逐渐加深, 所以在 UNet 中 ResNet50 提取的特征与 UNet 的解码器中的特征通过跳跃连接的方式进行融合, 会将高级和低级特征结合起来, 使得模型可以同时关注全局和局部的信息, UNet 的解码器部分也会进一步处理特征, 恢复图像的空间信息, 最后生成输出的分割结果。具体流程如图 4 所示。

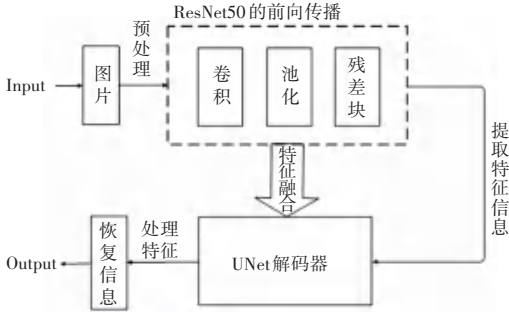


图 4 Res50-UNet 分割流程

Fig. 4 Res50-UNet segmentation process

#### 1.4 水平集方法

水平集方法<sup>[13]</sup> (Level-Set Method) 是一种用于描述和模拟物体轮廓、形状演化及界面追踪的数学和计算机科学方法, 该方法的核心思想是将物体的轮廓或界面表示为一个动态的、随时间演化的隐式函数, 该函数的零级集合定义了物体的边界。通过适当设计能量函数, 可以实现自动或半自动的图像分割, 从而降低了医务人员的操作压力, 并大大提高分析的精确度<sup>[14]</sup>。

##### 1.4.1 基于区域的水平集方法

水平集模型大致分为区域模型和边缘模型。CHAN 和 VESE 简化了分段常值 Mumford-Shah 模型<sup>[15]</sup>, 结合水平集方法提出了 CV (Chan-Vese)<sup>[16-17]</sup> 模型, 是典型的区域模型。其核心思想是通过计算动态轮廓分割区域的平均值生成能量函数<sup>[18]</sup>, 然后利用该能量函数最小化的方法来实现对患者肾脏病变图像的分割。此模型的应用使得肾脏病变边缘获得更好的分割结果。CV 模型的能量函数表示如下:

$$\begin{aligned}
 E^{CV}(c_1, c_2, C) = & \mu \cdot \text{Length}(C) + v \cdot \text{Area}(\varphi) + \\
 & \lambda_1 \cdot \int_{\text{inside}(C)} |v_0(x, y) - c_1|^2 dx dy + \\
 & \lambda_2 \cdot \int_{\text{outside}(C)} |v_0(x, y) - c_2|^2 dx dy \\
 & c_1 = av(v_0(x, y)), \text{inside}(C) \\
 & c_2 = av(v_0(x, y)), \text{outside}(C)
 \end{aligned}
 \quad (3)$$

其中,  $\mu, \lambda_1, \lambda_2$  为正常数, 且令  $\lambda_1 = \lambda_2 = 1$ 。 $v_0(x, y)$  表示  $(x, y)$  的像素值;  $c_1$  和  $c_2$  分别表示图像上曲线内和曲线外像素的平均值<sup>[19]</sup>。可由式如下公式计算求出:

$$c_1(\varphi) = \frac{\int_{\Omega} v_0(x, y) H(\varphi(x, y)) dx dy}{\int_{\Omega} H(\varphi(x, y)) dx dy} \quad (4)$$

$$c_2(\varphi) = \frac{\int_{\Omega} v_0(x, y) (1 - H(\varphi(x, y))) dx dy}{\int_{\Omega} (1 - H(\varphi(x, y))) dx dy} \quad (5)$$

能量泛函的首项式通常用作长度约束, 以确保能量演化曲线  $C$  足够小, 或确保得到足够小的曲线; 式(3)的第 3 项和第 4 项一起构成了保真项, 并将演化曲线  $C$  导引于目标轮廓上<sup>[20]</sup>。演化曲线  $C$  使用水平集函数  $\varphi(x, y)$  表示, 如果坐标  $(x, y)$  处于演化曲线  $C$  的内部, 那么水平集函数  $\varphi(x, y) > 0$ ; 如果坐标点  $(x, y)$  处于演化曲线  $C$  的外部, 那么水平集函数  $\varphi(x, y) < 0$ ; 而如果坐标点  $(x, y)$  恰好位于演化曲线  $C$  上, 那么水平集函数  $\varphi(x, y) = 0$ 。由此得到:

$$\begin{aligned}
 E^{CV}(c_1, c_2, \varphi) = & \mu \cdot \int_{\Omega} \delta_{\varepsilon}(\varphi(x, y)) |\nabla \varphi(x, y)| dx dy + \\
 & \lambda_1 \cdot \int_{\Omega} |v_0(x, y) - c_1|^2 H_{\varepsilon}(\varphi(x, y)) dx dy + \\
 & \lambda_2 \cdot \int_{\Omega} |v_0(x, y) - c_2|^2 (1 - H_{\varepsilon}(\varphi(x, y))) dx dy \quad (6)
 \end{aligned}$$

其中,  $H_{\varepsilon}(z)$  和  $\delta_{\varepsilon}(z)$  分别表示经过正则化的海氏 (Heaviside) 函数和狄拉克 (Dirac) 函数的形式。其公式如下:

$$H(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z \leq 0 \end{cases} \quad (7)$$

$$\delta(z) = \frac{d}{dz} H(z) \quad (8)$$

通过对上述公式对应的 Euler-Lagrange 方程进行求解, 可得到如下的演化方程:

$$\frac{\partial \varphi}{\partial t} = \delta_{\varepsilon}(\varphi) \text{div} \left( \frac{\nabla \varphi}{|\nabla \varphi|} \right) - \lambda_1 (v_0 - c_1)^2 + \lambda_2 (v_0 - c_2)^2 \quad (9)$$

其中,  $\delta_{\varepsilon}$  表示 Heavisid 函数的导数。且当  $\lambda_1 = \lambda_2 = 0, \mu = v = 0$ , 在 CV 模型中, 每个点的法线方向上的速度函数可以表示为:

$$F = 2\delta_\varepsilon(\varphi)(c_1 - c_2)\left(v_0(x, y) - \frac{c_1 + c_2}{2}\right) \quad (10)$$

#### 1.4.2 基于深度学习的水平集方法

水平集方法的主要目的是根据图像中的特征分割物体。深度学习框架,如卷积神经网络(CNN)可以用于提取图像特征,可以将这些特征结合水平集方法来改进分割结果。研究人员已经提出了深度学习网络,例如深度水平集网络(Deep Level Set)<sup>[21]</sup>将水平集方法与深度学习结合在一起。通过联合训练来学习分割任务中的水平集演化和特征提取。Lin等学者<sup>[22]</sup>建议使用焦点损失来解决一级目标检测中的极端前景、背景不平衡问题。研究通过将包含图像全局灰度信息的水平集能量泛函引入卷积神经网络(CNN),为图像中的每个目标类别生成单独的

水平集能量损失,增强网络对小目标和多目标的边缘细节信息的捕获能力。其中,水平集能量项表示为:

$$E_{LS}(\varphi, G) = \sum_{l \in L} \left( \int_{\Omega_l} |G_l(x, y) - c_{l,1}| H_\varepsilon^*(\varphi_l(x, y)) dx dy + \int_{\Omega} |G_l(x, y) - c_{l,2}| (1 - H_\varepsilon^*(\varphi_l(x, y))) dx dy \right) \quad (11)$$

其中,  $\Omega_l$  表示  $G_l$  的整个区域,水平集函数  $\varphi$  是一个移位的密集概率。图5为病变肾脏分割图的水平集损失流程护照原理。输出概率映射变换作为水平集函数  $\varphi$  工作。将基本真理分解为每类的二元映射(白色为1,黑色为0)。对于每个类进行水平集能量的计算,能量之和作为训练分割网络的损失函数。

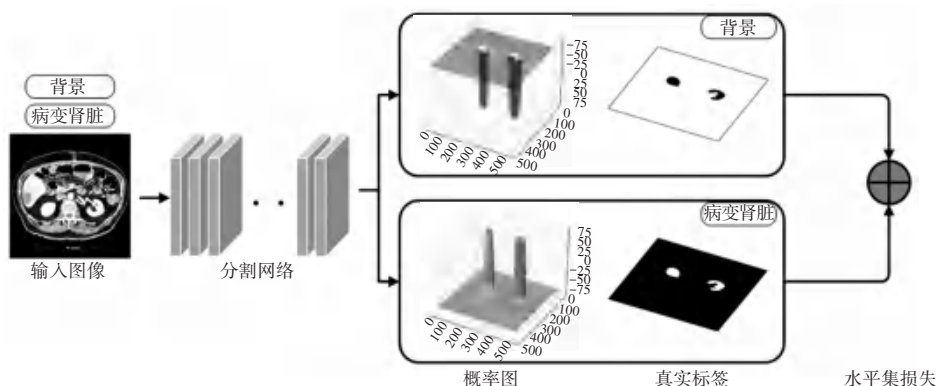


图5 病变肾脏分割图的水平集损失流程护照原理

Fig. 5 Level set loss process passport principle for renal lesion segmentation map

UNet 通常使用交叉熵损失函数来进行训练,对于 UNet 的二进制分割任务,这部分损失用于度量模型生成的分割掩码与实际目标分割掩码之间的差异。但是传统的 UNet 二进制交叉熵损失函数可能存在着目标边缘不清晰的情况,导致模型难以准确捕捉边缘信息,也会倾向于生成锯齿状的分割轮廓,影响分割结果的平滑性,不符合实际要求。Ronneberger 等学者<sup>[10]</sup>通过改进 UNet 的变体,来提升边缘的平滑性。Zhao 等学者<sup>[23]</sup>使用多层次深度监督信息,将损失函数分布在网络的不同层次,以帮助模型更好地捕获图像的细节信息和边缘信息;Krähenbühl 等学者<sup>[24]</sup>将条件随机场(CRF)作为损失函数的一部分,并将其作为后处理步骤进行使用,可以改善分割结果的平滑性和准确性。通过上述方法来分割图片时,结果的精准性和平滑性之间的平衡存在不足。为改善两者之间的平衡,对水平集进行初始化,如图6所示。将多类水平集能量损失和

UNet 网络的二进制交叉熵损失进行有机融合,生成混合损失函数。水平集能量损失通常有助于改善分割结果的平滑性和形状一致性,特别是对目标边缘不清晰的情况。二进制交叉熵损失可以帮助提高分割的精确性,两者的有机融合可以在精确性和平滑性之间找到平衡,实现网络端到端的训练,增强对病变肾脏尖锐或凹凸目标边缘的分割能力。

对于上文中引入的作为水平集标记函数的 Heavisid 函数,使用  $H_\varepsilon(z)$  函数作为 Heaviside 的近似,若  $\varepsilon \rightarrow 0$  时,可以得出  $H_\varepsilon(z) \rightarrow H(z)$ <sup>[25]</sup>。根据文献[11]可以得出以下结论:

$$H_\varepsilon(z) = \frac{1}{2} \left( 1 + \tanh \frac{z}{\varepsilon} \right) \quad (12)$$

其中,使用  $\tanh$  作为激活函数能够展示深度学习结构的卓越性能<sup>[26-27]</sup>。对于水平集能量损失函数与交叉熵损失的结合如下:

$$Loss = E_{CE}(P, G) + \lambda \cdot E_{LS}(\varphi, G) \quad (13)$$

其中,  $\lambda$  表示加权水平集损失的参数。图 7 为水平集方法训练过程中的损失函数。函数图像在训练过程中逐步收敛, 表明模型在学习过程中逐渐提高性能。

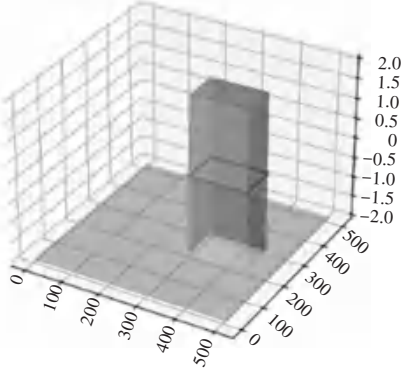


图 6 初始水平集

Fig. 6 Initial level set

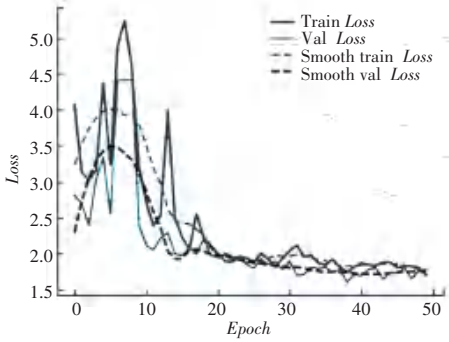


图 7 训练过程中的损失函数

Fig. 7 Loss function during training

## 2 实验配置

### 2.1 数据集介绍

本文所用的数据集<sup>[28]</sup>来自医学图像数据集 3D-IRCADB, 图片格式为 DICOM, 这种格式的图像不适用于论文实验所使用的软件进行相关实验操作。因此, 需要对图像格式进行转换, 并对数据集进行标准化处理<sup>[9]</sup>。同时, UNet 网络要求使用掩膜作为原始图像的标签, 将其处理为 8 位灰度图。数据集共有 662 张病变肾脏的实验图片, 划分训练集 486 张, 测试集 176 张。

表 2 消融实验

Table 2 Ablation experiment

模型				评价指标			
UNet	SE	ResNet50	Level-Set	Acc/%	mPA/%	P/%	MIoU/%
✓	-	-	-	98.72	68.38	75.07	63.06
✓	✓	-	-	98.81	70.38	77.53	64.94
✓	-	✓	-	98.87	68.03	80.08	64.19
✓	-	-	✓	99.01	88.40	79.45	74.76
✓	✓	✓	-	98.92	73.78	80.46	67.82
✓	✓	-	✓	99.20	89.01	82.90	77.65
✓	✓	✓	✓	99.33	94.21	84.23	81.27

### 2.2 实验环境与参数配置

本文实验 PC 机的配置为: 处理器 Intel (R) Core (TM) i7-12700H CPU @ 2.30 GHz 和 NVIDIA GeForce RTX 3060 Laptop GPU, 运行环境为 Python3. 10. 9, Win11 x64, Pytorch2. 0. 1。具体参数见表 1。

表 1 实验参数

Table 1 Experimental parameters

参数名称	含义	参数值
<i>epochs</i>	训练轮次	50
<i>batch_size</i>	批量大小	2
<i>num_classes</i>	分类个数	2
<i>optimizer</i>	优化器	SGD
<i>backbone</i>	主干网络	ResNet50
<i>momentum</i>	SGD 动量	0.9

### 2.3 评价指标

本文将选择准确率 *Acc* (Accuracy)、类别平均像素准确率 *mPA* (mean Pixel Accuracy)、精确率 *P* (Precision) 和平均交并比 *MIoU* (Mean Intersection over Union) 指标来评价分割效果。各指标公式具体如下:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

$$mPA = \frac{\sum_{i=1}^n (P_i)}{n}, \quad n \text{ 为类别数} \quad (15)$$

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

其中, *TP*、*TN*、*FP*、*FN* 分别表示真阳性、真阴性、假阳性、假阴性, *P<sub>i</sub>* 表示每个类别像素准确率。

## 3 实验结果与分析

### 3.1 消融实验

对于改进后的 UNet 网络通过消融实验进行验证, 分别讨论 SE 注意力机制、ResNet50 主干网络和水平集方法对病变肾脏分割精度的影响。对病变肾脏 CT 图像数据集进行的消融实验结果见表 2。

由表2可以看出,UNet网络分别融入SE注意力机制、使用ResNet50作为主干网络以及加入水平集方法生成混合损失函数之后,与原始UNet模型相比平均交并比分别提升18.21%、4.04%、10.02%,同时准确率、精准率和类别平均像素准确率也分别提升0.61%、9.16%和25.83%,表明SE注意力机制、ResNet50主干网络和水平集方法对网络性能的提升均有所帮助。

### 3.2 对比实验

为了更好地了解本文使用分割方法的性能,将模型SRL-UNet与原模型UNet和传统模型DeepLabv3以及HRNet模型进行对比实验。不同模型的评价指标对比见表3。

表3 不同模型的评价指标

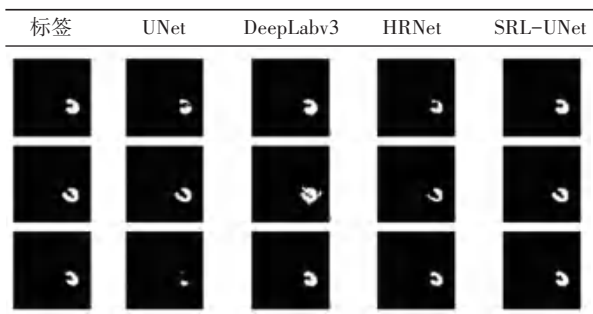
Table 3 Evaluation indicators for different models

模型	Acc/%	mPA/%	P/%	MIoU/%
UNet	98.72	68.38	75.07	63.06
DeepLabv3	98.98	98.16	78.16	77.23
HRNet	99.09	76.50	80.42	71.25
SRL-UNet	99.33	94.21	84.23	81.27

由表3可以看出,本文方法对比其他3种模型方法在MIoU、Acc和Pre指标数值上都有所提升,mPA指标对比UNet模型和HRNet模型上也有所提高。对于DeepLabv3模型中的mPA指标较高,考虑到mPA是用来衡量模型正确分类每个像素的能力且DeepLabv3模型中运用了空洞卷积、空间金字塔池化,还使用了尺度的特征图以在不同层次捕获对象的特征,所以在mPA指标中表现尤为突出。不同模型下分割效果见表4。

表4 不同模型的分割效果

Table 4 Segmentation effect of different models



由表4可以看出,UNet模型方法得出的分割结果中出现空洞,分割结果不理想,可能与模型的复杂度不足有关。DeepLabv3模型方法分割将不属于病变肾脏部分分割出来,并且边缘不光滑,没有分割出细小边缘。HRNet模型方法没有将病变肾脏分割完

整。而SRL-UNet模型能够较准确完整地分割出病变肾脏且对于边缘的平滑性也有良好的分割效果。

## 4 结束语

本文采用对UNet网络改进后的SRL-UNet模型算法对病变肾脏图像进行分割,以此来提高对肾脏病变的准确识别和定位。融合SE注意力机制可以增强网络对关键特征的关注,将ResNet50作为UNet的主干提取特征网络,可在训练过程中因其具有深层网络而避免梯度消失或梯度爆炸的问题,并且通过结合UNet神经网络的图像特征学习和水平集方法的轮廓演化,得到了更加有效的分割方法。同时解决了其他3种模型分割结果产生的空洞、分割缺失以及边缘不光滑的问题。研究还展示了深度学习技术在医学图像处理中的巨大潜力,为未来的疾病诊断和治疗提供了更多可能性。

尽管研究取得了较好的结果,但仍存在挑战和改进的空间,例如数据集的质量和数量、模型的泛化能力等方面。后续研究可以全面助推网络架构优化,探索新的损失函数和数据增强方法,以进一步提高病变肾脏图像分割的性能。

## 参考文献

- [1] 沈小波,张波,李文豪,等. U-Net在肺结节分割中的应用进展[J]. 计算机应用, 2023, 43(S1): 250-257.
- [2] BEVERIDGE J R, GRIFFITH J, KOHLER R R, et al. Segmenting images using localized histograms and region merging[J]. International Journal of Computer Vision, 1989, 2(3): 311-347.
- [3] PALN R, PAL S K. A review on image segmentation techniques[J]. Pattern Recognition, 1993, 26(9): 1277-1294.
- [4] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [5] VO A V, TRUONG-HONG L, LAEFER D F, et al. Octree-based region growing for point cloud segmentation[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 104: 88-100.
- [6] 赵麒然. 基于改进参数核图割的OCT眼底图像分割方法研究[D]. 长春: 吉林大学, 2016.
- [7] BIANCONI F, FRAVOLINI M L, PIZZOLI S, et al. Comparative evaluation of conventional and deep learning methods for semiautomated segmentation of pulmonary nodules on CT[J]. Quantitative Imaging in Medicine and Surgery, 2021, 11(7): 3286-3305.
- [8] GUI Luying, LI Chunming, YANG Xiaoping. Medical image segmentation based on level set and isoperimetric constraint[J]. Physica Medica, 2017, 42: 162-173.
- [9] 陈铭,梅雪,朱文俊,等. 一种新型Mobile-Unet网络的肺结节图像分割方法[J]. 南京工业大学学报(自然科学版), 2022, 44

- (1):76-81,91.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//Proceedings of the 18<sup>th</sup> International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham:Springer, 2015: 234-241.
- [11] KIM Y, KIM S, KIM T, et al. CNN-based semantic segmentation using level set loss [C]//2019 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway NJ:IEEE, 2019: 1752-1760.
- [12] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis, 2017, 42:60-88.
- [13] CHAN T, VESE L. An active contour model without edges [M]//NIELSEN M. JOHANSEN P, OLSEN O F, et al. Scale-space theories in Computer Vision. Scale-space1999. Lecture Notes in Computer Science. Cham:Springer, 1999: 141-151.
- [14] 吴磊磊. 基于数学形态学的细胞图像分割算法研究[D]. 成都:电子科技大学,2009.
- [15] ESEDOG S, TSAI Y H R. Threshold dynamics for the piecewise constant Mumford-Shah function[J]. Journal of Computational Physics, 2006, 211(1): 367-384.
- [16] CHAN T F, VESE L A. Active contours without edges[J]. IEEE Transactions on Image Processing, 2001, 10(2): 266-277.
- [17] 黄扬. 基于熵的变分水平集图像分割[D]. 宁波:宁波大学,2016.
- [18] 朴春赫,赵海,朱宏博,等. 基于水平集方法的交互式目标分割算法[J]. 计算机工程,2016,42(2):229-235.
- [19] 兰红,韩纪东. 基于 CV 模型改进的磁共振成像图像分割方法[J]. 科学技术与工程,2018,18(28):229-234.
- [20] 王啸伟,张兴媛,史红标,等. 基于 C-V 模型和形态学的道路图像处理方法[J]. 计算机时代,2019(3):49-52.
- [21] HU Ping, SHUAI Bing, LIU Jun, et al. Deep level sets for salient object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE, 2017: 2300-2309.
- [22] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2017: 2980-2988.
- [23] ZHAO Hengshuang, SHI Jianping, QI Xiaojun, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE,2017: 2881-2890.
- [24] KRÄHENBÜHL P, KOLTUN V. Efficient inference in fully connected crfs with gaussian edge potentials[J]. Advances in Neural Information Processing Systems, 2011, 24:244-216.
- [25] 赵瑞雪,潘振宽,侯国家,等. 基于一个水平集函数的多相图像分割方法[J]. 计算机工程,2018,44(4):317-321.
- [26] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradientbased learning applied to document recognition[J]. Proceedings of the IEEE,1998, 86(11):2278-2324.
- [27] LECUN Y A, BOTTOU L, ORR G B, et al. Efficient backprop [M]//MONTAVON G. Neural networks: Tricks of the trade. Cham:Springer, 2012:9-48.
- [28] SOLER L, HOSTETTLER A, AGNUS V, et al. 3D image reconstruction for comparison of algorithm database: A patient specific anatomical and medical image database [Z]. Strasbourg, France: IRCAD, 2010.