

文章编号 1004-924X(2021)04-0877-19

联合多尺度和注意力-残差的深度监督 乳腺癌分割

秦传波¹, 宋子玉¹, 曾军英^{1*}, 田联房², 李 芳³

- (1. 五邑大学 智能制造学部, 广东 江门 529020;
2. 华南理工大学 自动化科学与工程学院, 广东 广州 510640;
3. 江门市妇幼保健院, 广东 江门 529020)

摘要:针对 DCE-MRI 乳腺癌病变区的浸润范围勾画精度低、结构形态变化大、强度不均和边界对比度低等原因,导致乳腺癌病变区自动化分割存在准确率低和错分割的问题,为此,本文构建了一个二阶段乳腺癌病变区分割框架,提出一种乳腺癌病变区分割模型 UTB-net,分别在编码路径和末端整合多尺度和 Non-local,在解码路径构建注意力-残差模块。首先,利用基准 U-net 网络模型实现对乳房区域的粗略勾画,消除影像中胸肌肉、脂肪、心脏等不相关组织对乳腺癌分割的影响。然后,基于提取的 ROI 结果,在模型的编码路径嵌入了多尺度信息融合和 Non-local 模块。最后,在解码路径构建了一种注意力-残差混合解码模块,并引入深度监督机制,以提高乳腺癌病灶的分割精度。实验结果表明:相较于 U-Net 基准模型,乳腺癌分割指标 DICE, IOU, SEN, PPV 分别提升了 4%, 4.78%, 5.92% 和 3.94%。所提模型在提高了乳腺癌分割结果的同时,减少了小面积误分割和钙化分割。

关键词: DCE-MRI; 分割; 乳腺癌; U-net; 注意力; 残差网络

中图分类号: TP391 文献标识码: A doi: 10.37188/OPE.20212904.0877

Deeply supervised breast cancer segmentation combined with multi-scale and attention-residuals

QIN Chuan-bo¹, SONG Zi-yu¹, ZENG Jun-ying^{1*}, TIAN Lian-fang², LI Fang³

- (1. Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen 529020, China;
2. School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China;
3. Jiangmen Maternal and Child Healthcare Hospital, Jiangmen 529020, China)

* Corresponding author, E-mail: zengjunying@126.com

Abstract: Given the low accuracy of delineation of the infiltration area of breast cancer lesions in DCE-MRI, variable structure and shape, large intensity heterogeneity changes, and low boundary contrast, the automatic segmentation of breast cancer lesions has the problems of low accuracy and mis-segmentation.

收稿日期: 2020-08-28; 修订日期: 2020-10-16.

基金项目: 国家自然科学基金资助项目 (No. 61771347); 广东省基础与应用基础研究基金资助项目 (No. 2019A1515010716); 广东普通高校人工智能重点领域专项资助项目 (No. 2019KZDZX1017); 广东省数字信号与图像处理技术重点实验室开放基金资助项目 (No. 2020GDDSIPL-03; No. 2019GDDSIPL-03)

For this reason, a two-stage breast cancer image segmentation framework is constructed, and a breast cancer lesion segmentation model UTB-net is proposed to integrate multi-scale and non-local at the encoding path and the end, respectively, which constructs attention-residuals in the decoding module. First, the benchmark U-net network model is used to achieve a rough delineation of the breast area, eliminating the influence of unrelated tissues, such as chest muscle, fat, and the heart, on breast tumor segmentation in the image. Then, based on the extracted ROI results, a multi-scale information fused and non-local module is constructed in the coding path of the model. Finally, an attention-residual hybrid decoding module is constructed in the decoding path, and a deep supervision mechanism is introduced to improve the segmentation accuracy of breast tumor lesions. Experimental results show that breast tumor segmentation indexes DICE, IOU, SEN, and PPV increase by 4%, 4.78%, 5.92%, and 3.94% respectively, in comparison with the U-Net benchmark model. The proposed model not only improves the segmentation results of breast cancer but also reduces the small-area mis-segmentation and calcification segmentation.

Key words: Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI); segmentation; breast tumor; U-net; attention; residual network

1 引言

乳腺癌是众多严重影响女性患者生活质量和身心健康的疾病之一。相较于钼靶、超声、X-ray、MRI、Thermogram 等类型影像^[1], 动态对比度增强 MRI 影像 (Dynamic Contrast-Enhanced Magnetic Resonance Imaging, DCE-MRI) 对高密度腺体组织的乳腺癌诊断价值大, 是一种以病变区和正常组织中的微血管系统为生理基础的无辐射磁共振功能成像技术。DCE-MRI 对于早期乳腺癌诊断敏感性强、乳腺癌的分期、分子分型的确诊和病变范围的一致性表现良好。乳腺癌病变区的自动化分割在早期诊断、预防、术中引导和预后具有重大临床研究价值, 是实现人工智能计算机辅助诊断系统的关键基础。

DCE-MRI 乳腺癌主要分为肿块和非肿块两种类型。乳腺癌病变区主要沿腺体分布, 肿块形乳腺癌相对容易辨识和分割, 而非肿块乳腺癌病变区的形态变化不一, 且病变区强度不一, 存在空洞、弱边界和较低的对亮度, 容易造成放射科医生的漏检和错判。然而, 手动勾画乳腺癌病变区要求放射科医生具备丰富的临床经验和技巧, 并且是一种劳动密集型的工作, 易受主观、心理波动和生理疲劳的影响, 导致乳腺病变勾画结果

存在一定的随机性。综上, DCE-MRI 乳腺癌自动化准确分割是一个具有挑战性的课题, 极具临床应用价值和研究意义。

当前乳腺癌影像的研究主要集中在钼靶、超声、X 射线和组织病理图像分析^[2-4], 但对 DCE-MRI 乳腺癌影像的研究目前尚没有统一的公开数据库可以使用, 且临床乳腺癌图像病例收集工作量大, 样本相对少, 且病灶区面积小, 背景复杂, 易造成漏分割和钙化区误分割, 分割精度尚有较大提升空间。

目前, 乳腺癌分割算法研究可以大致化分为经典机器学习方法和当前流行的深度学习方法。乳腺癌分割经典机器学习方法主要包括模糊聚类 (Fuzzy C-Means, FCM)、活动轮廓模型 (Active Contour Model, ACM)、条件随机场、Graph Cut 等方法^[5-8]。文献[5]选取冠状位影像, 采用基于 FCM 和体积增长实现 DCE-MRI 乳腺癌图像分割, 但需要先手动选定感兴趣区域 (Region of Interest, ROI), 对于大部分病灶分割结果需手动调整参数; 文献[6]根据肿瘤的时间、空间和形态属性, 使用图割 (Graph Cuts) 进行分割, 构建一个癌症诊断的框架; 文献[7]利用 MRF 和条件独立性构建了动力学观测模型, 将 DCE-MRI 乳腺癌动态特征图作为 MRF 肿瘤分割的观察通道。并将分割结果应用于乳腺癌复发预测。基于偏微

分方程的ACM方法在医学影像分割上得到广泛应用;文献[8]针对非肿块乳腺癌病变区分割,通过MRF能量区分肿瘤与背景,再使用能量项后验概率作为ACM区域项,而后根据纹理、强度信息构建ACM速度函数,构建了一个复杂的ACM分割目标函数。但需要事先手动分割出ROI,在小范围病例内有较好分割效果。且计算成本高,耗时长,需要人为调整参数以适应变化的乳腺癌影像。

对于基于传统机器学习方法而言,由于乳腺癌病变区具有对比度低、边界模糊及亮度不均匀等特点,通常需要手动选取ROI,再进行分割处理,或者对分割后结果进行后处理,才能得到平滑和连续的分割边界。通常需要手动调整参数才能适应形态各异的病灶区,显然不是一套完备的自动化分割方案。不适合在大量病例数据上实现自动化分割。

近年来,基于深度学习的图像分割、检测和分类方法取得很大进展^[4,9-10]。杜克大学提出了一种掩模引导分层式FCN分割框架^[11],通过两个级联FCN粗分割和FCN细分割模型,在前后两个阶段分别完成乳房区域粗略勾画和乳腺癌病变区的精确分割。在其自建的数据库上进行训练和测试,分割关键指标平均Dice值达到0.72。然而,对于体素比较小的乳腺癌病变区存在漏分割、错分割和分割不准确的问题,并且模型架构单一,仅使用FCN级联架构,没有引入更多有效的措施提升分割精度,诸如注意力和监督机制。这也为本文进行下一步深入研究乳腺癌病变区的精确分割提供了思路。

注意力机制^[12]、多尺度和监督学习^[13-14]表明能有效改善深度学习模型效果。深圳研究院提出基于深度表示的跨模态乳腺MRI影像分割模型^[15]。受知识蒸馏和注意力转移策略启发,采用有监督跨模态学习方式对T1c和T2W两种序列进行融合,同时引入注意力机制实现有效的信息选择和抑制噪声。

由于DCE-MRI乳腺癌背景复杂,胸腔、乳房等组织占据影像的很大比例的面积,而乳腺癌病变区所占面积相对较小,在乳腺癌分割任务中,先勾画出乳房区域ROI,可进一步提高分割精

度^[11]。文献[16-17]采用经典U-net^[18]为基础,将乳房区域与胸壁线分离。证明了使用U形网络进行乳房ROI提取的可行性。

2 所提出的方法

乳腺癌病变区形态各异且边界模糊,所占面积小,是个典型的分类不均衡问题,并且周围背景复杂,存在心脏、胸壁、皮肤、肌肉和脂肪等组织,严重影响和干扰乳腺癌图像的自动化分割^[11]。由于乳腺癌区域相对整张影像所占面积小,为充分挖掘其语义信息,本文提出了一个二阶段的乳腺癌分割框架。

在第一个阶段,重新训练了U-Net模型来提取一个粗略的乳房区域ROI,消除乳房周边胸肌肉、脂肪、心脏等不相关组织对乳房区域内乳腺癌分割的影响。

在第二个阶段,基于阶段一粗分割乳房ROI,提出了一种UTB-Net模型,实现乳腺肿瘤分割,模型如图1所示。

所提模型以改进的U-Net为基准,设计并融合多尺度并行卷积融合模块、Non-local模块^[19]、注意力-残差混合解码模块,提出了融合深度监督的UTB-Net模型。采用深度监督的UTB-Net模型可以提取更加丰富的细节信息,结合全局注意力和局部注意力对细小的乳腺癌病灶进行更准确的定位,而深度监督机制则能加强网络对前景的响应。

本文主要所做的工作和贡献如下:

(1)在编码路径,提出多尺度并行卷积融合模块。采用多尺度的并行卷积结构,通过不同核大小的卷积操作提取不同尺度的特征信息,丰富模型提取边缘信息与精细信息的能力。

(2)对基准U-Net网络模型进行微调。在编码路径采用PReLU激活函数,减少网络非线性表达的信息丢失,将BatchNorm换成Group-Norm^[20],以此来减少计算量不足对分割精度的影响。并在编码路径的末端增加了Non-local,从而高效地获取全局相关信息,以此形成一个全局的注意力机制。

(3)在解码端,提出结合注意力和残差的混合解码模块,将注意力模块的非线性整流单元与残差结合起来构建解码模块,以此来改善网络信息流向,提高特征提取能力。

实验表明,所提出的模型不但对于肿块和非肿块有较好的分割结果,提高分割精度的同时对于小体素、小面积的乳腺癌图像也具有较好的分割效果。

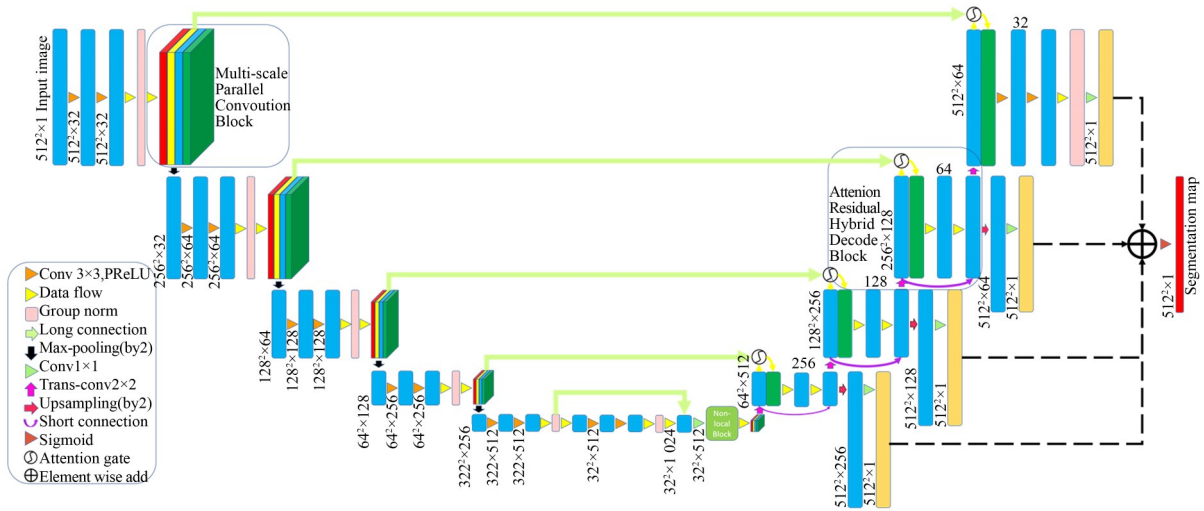


图1 所提深度监督 UTB-Net 架构

Fig. 1 Proposed deep supervision UTB-Net 架构

2.1 乳房 ROI 提取

乳腺肿瘤病变区通常出现在乳房区域,极少出现在胸壁线的位置或者乳房边缘。事先提取一个粗糙的乳房区域 ROI,有利于下一步乳腺癌的准确分割^[11]。在第一个阶段,根据临床医生勾画的乳房区域标签,重新设计了医学影像分割任务中常用的基准 U-Net 模型,实现乳房 ROI 粗略

勾画,如图 2 所示(彩图见期刊电子版)。图 2(a)是 DCE-MRI 乳腺癌水平位影像,乳房区域 ROI 用绿色标签标记,乳腺癌区域用黄色区域标记。图 2(b)是局部放大图,可见乳腺癌有边界不规则、空洞等特征。图 2(c)是用 ITK-Snap 工具重建 3D 效果图。图 2(d)从空间角度展示了乳房组织和乳腺癌之间空间关系。

2.2 微调基准 U-Net 模型

为了充分挖掘乳腺肿瘤区域的语义信息,对基准 U-net 模型分别做了微调和改进。微调 U-Net 架构细节如图 3 所示。

首先,编码区与解码区特征提取单元采用的激活函数用 PReLU 替换掉 ReLU,以此来减少信息的丢失;其次,又因医学影像计算量大,采用 GroupNorm 作为归一化层,以缓解显卡内存不足对分割精度的影响;特征提取单元基本结构为 Conv→PReLU→Conv→PReLU→GroupNorm;最后,在编码路径的最后一层,进行两次 Conv→PReLU→Conv→PReLU→GroupNorm 操作,而后将信息融合压缩至与

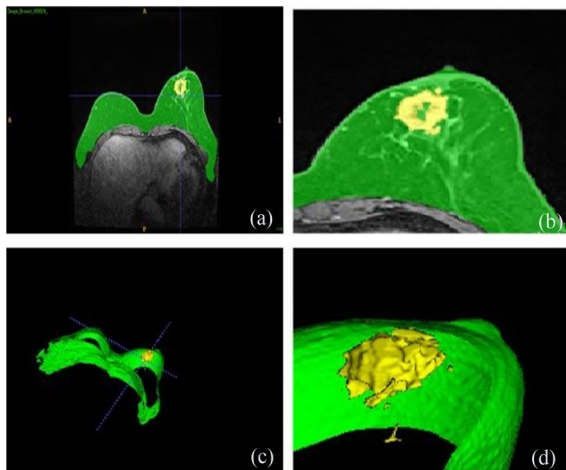


图 2 DCE-MRI 乳腺癌示意图

Fig. 2 Illustration of DCE-MRI breast tumor

原通道数一致,目的是保证不增加参数的同时提取更多高级特征信息。在后续实验分析部分,其实验结果表明,相较于基准 U-net 模型,经微调的 U-net 也可以取得较好分割效果。

其次,基于所提模型的微调 U-Net 模型结构基础,在其编码路径上,进一步整合本文所

构建的多尺度并行卷积模块;在编码路径末端加入 Non-local 模块;在解码路径上加入了所提出的注意力-残差混合解码模块,并对各层解码输出使用多尺度深度监督机制进行融合。即分别在微调 U-Net 模型的编码路径、编码末端和解码路径进行改进,构建所提出的 UTB-Net 模型。

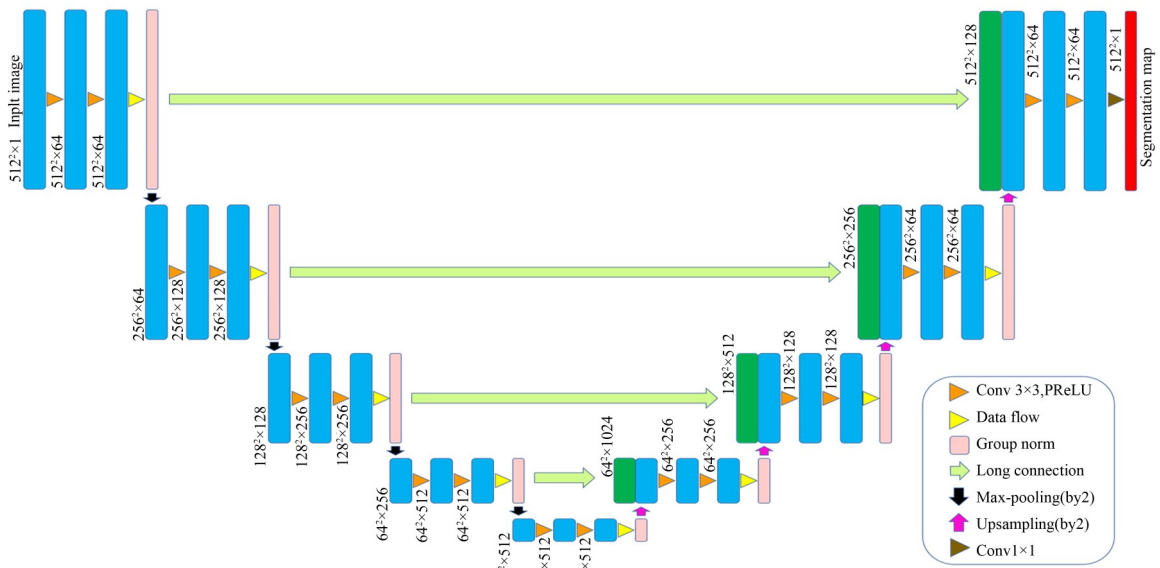


图 3 微调的 U-net 模型

Fig. 3 Illustration of Fine-tuned U-net model

2.3 所提出的 UTB-Net 分割模型构建方法

下面将详尽阐述 UTB-Net 的构建过程,即如何将多尺度并行卷积模块、Non-local 模块和注意力、残差混合解码模块和多尺度监督分别嵌入至微调的 U-Net 模型的编码路径和解码路径。

2.3.1 多尺度并行卷积融合模块

因乳腺癌区域大小不定、形状各异、灰度不均且边界模糊,在模型引入多尺度并行卷积融合 (Multi-Scale Parallel Convolution Fusion, MSPCF) 模块,如图 4 所示。并行卷积结构能够在同一层级获得不同感受野的特征,提高模糊边缘和小面积肿块的乳腺癌分割精度,然后将不同尺度的信息融合传递至下一层,可以在提高精度的同时灵活地平衡计算量和模型能力。

MSPCF 模块包含 3 个并行卷积层和 1 个全

局平均池化层、1 个 1×1 卷积层。3 个并行卷积操作的卷积核尺寸分别为 1×1, 3×3 和 5×5,其中 5×5 的卷积核经过卷积分解处理^[21],用 2 个 3×3 的卷积核替代。卷积分解使用小的卷积核级联替代大的卷积核,可以获得相同的感受野,同时在一定程度上减少模型的参数,提高模型的非线性表达能力。

MSPCF 模块采用不同大小的卷积核获取多尺度的上下文特征信息,其中 1×1 的卷积层则保留当前尺度的特征信息。MSPCF 模块还使用了全局平均池化层,获取像素级的全局信息,而后通过双线性插值上采样得到所需的维度,再通过 1×1 卷积将全局平均池化层信息与三个并行卷积层信息融合并压缩。然后将输出特征图传递至下一层。

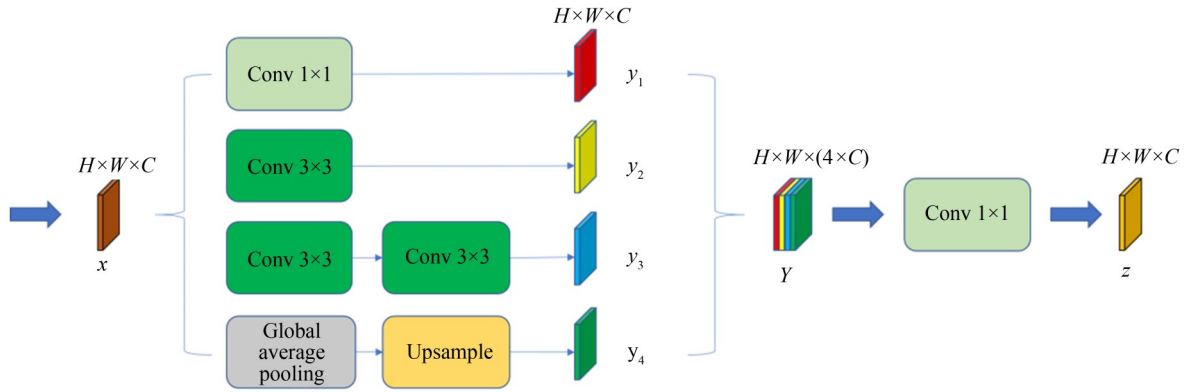


图 4 多尺度并行卷积融合模块
Fig. 4 MSPCF Block

在分割任务中,将此 MSPCF 模块融合至所提出的网络的编码部分。在图 4 中, y_4 的计算公式如式(1)所示:

$$y_4 = F_{bi} \left(\frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x(i, j) \right), \quad (1)$$

其中: $x=[x_1, x_2, \dots, x_c]$, $y_4=[y_{4_1}, y_{4_2}, \dots, y_{4_c}]$, $F_{bi}(x)$ 表示双线性插值,然后将所有分支的特征图整合为 Y , 其中 $Y=[y_1, y_2, y_3, y_4]$ 。最后,通过一个 $[1 \times 1, C]$ 的卷积操作将多尺度信息聚集到一起得到输出特征图 z 。

2. 3. 2 Non-local 模块

众所周知,卷积操作通常关注局部的感受野。如果要增大感受野,传统的方法是堆叠卷积层和池化层来实现,但这样会增加模型的参数量和计算量。Wang 等人^[19]借鉴图像去噪的思想而提出 Non-local 模块,用来获取长距离之间的关联信息,最终实现每个像素的全局感受野,并通过不同的操作,可以获得不同空间、时间、通道的像素之间的信息相关性,按照非局部均值定义,具体结构如图 5 所示。其中, $C(x)$ 是归一化参数,在实验中设置为 x 的位置数量,用 N 来表示。 x 表示输入的信号。 y 表示输出的信号,其面积和输入信号相同。 $g(x_i)$ 用于计算输入信号在 j 位置的特征值。 $f(x_i, x_j)$ 则表示点乘相似度,在本方法中采用点积表示。其中:

$$y_i = \frac{1}{c(x)} \sum_{v_j} f(x_i, x_j) g(x_j). \quad (2)$$

将 Non-local 架构定义为一个残差形式,使其具有灵活性,不需要循环渐进的训练,可以十分

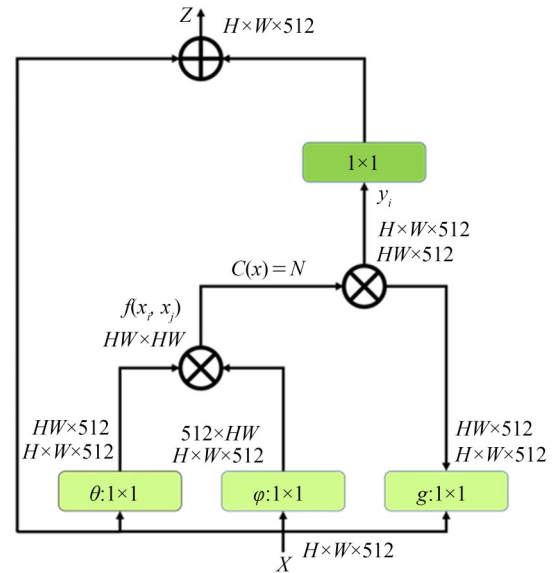


图 5 Non-local 模块架构^[19]
Fig. 5 Non-local block architecture^[19]

方便的整合到一个标准 CNN 架构中,以此来改善模型的敏感性和准确性,并且不需要改变原有结构,输出信号 $z_i = W_z y_i + x_i$, x_i 表示残差, W_z 表示一系列的卷积操作, z_i 则表示下层的输入信号。可灵活地将 Non-local 模块整合到本文模型中。

2. 3. 3 注意力-残差混合解码模块

本文提出了一种注意力-残差混合解码模块 (Attention-Residual Hybrid Decoding, ARHD), 并将其融入本网络的解码路径当中, ARHD 架构细节如图 6 所示。He 等人^[21]通过短连接构建残差模块,在一定程度上缓解了梯度消失和梯度爆

炸问题,使得在训练更深的网络的同时,不增加参数量同时且能保证良好的性能。文献[12]根据软自注意力机制提出附加注意门(Additive attention Gate, AG)模块,在自动学习聚焦于不同形状和大小的目标结构同时,隐式的学习抑制输入图像中不相关区域,以此来突出特定任务中有效的显著特征,从而有效地提高分割精度。根据上述两种方法,构建了 ARHD 模块。

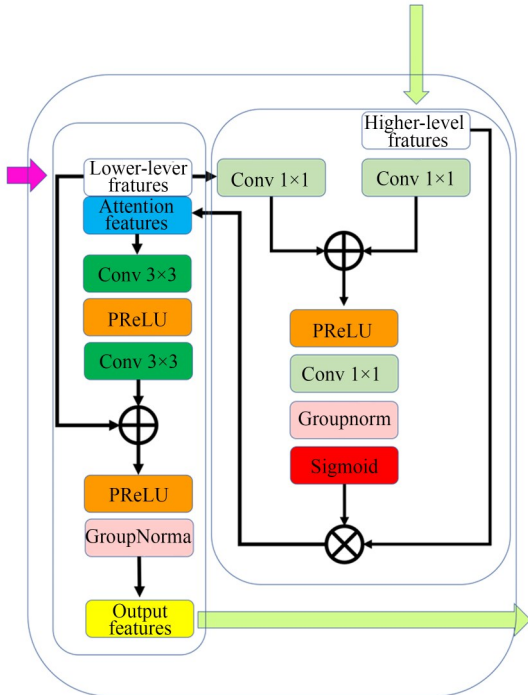


图 6 注意力、残差混合解码模块

Fig. 6 Block of attention and residual hybrid decoding

如图 6 所示, AG 模块将深层特征信息与来自浅层的特征信息进行矩阵加法运算,再经 PReLU^[20]得到激活特征,然后通过 sigmoid 得到注意力系数。在计算注意力系数 $\alpha_i^l \in [0, 1]$ 时,由于 ReLU 激活函数会造成信息的丢失,所以在模块中将使用 PReLU^[22] 函数,公式如式(3)所示:

$$\alpha_i^l = \text{sigmoid}(q_{\text{att}}^l(x_i^l, g_i; \theta_{\text{att}})), \quad (3)$$

其中:

$$q_{\text{att}}^l = \text{GroupNorm}(\varphi^T(\text{PReLU}(\mathbf{W}_x^T x_i^l + \mathbf{W}_g^T g_i)))$$

$$\text{PReLU}(x) = \begin{cases} x, & x > 0 \\ \alpha x, & x < 0 \end{cases}$$

其中:GroupNorm^[19]是针对 batch size 较小时的一种规范化方式, α 是可学习参数。 x_i^l, g_i 分别代表

浅层特征和深层特征信息, $\varphi^T, \mathbf{W}_x^T, \mathbf{W}_g^T$ 是卷积操作的线性变换, $\text{sigmoid}(x) = 1/(1 + e^x)$ 。

将 AG 模块输出的注意力系数与深层特征信息进行矩阵相乘运算,消除不相关特征与噪声,突出显著特征,再将结果和浅层信息加入残差连接模块,改善解码部分的梯度流动,缓解梯度弥散和网络退化问题,使深层的网络更容易训练,最后输出注意力特征图。为了方便表达,用 $z_i^l = \alpha_i^l \otimes x_i^l$, \otimes 代表矩阵相乘运算, z_i^l 表示最后输出的注意力特征图,实验结果分析见后续 3.5 节内容。

在 ARHD 模块中,将 AG 模块嵌入到解码层单元中,再加入残差连接,改善解码部分的梯度流动,缓解梯度弥散和网络退化问题,使深层的网络更容易训练。用 cat 来表示通道拼接操作, d 表示通道拼接后的特征图, $F(x)$ 表示一系列的卷积操作,那么有:

$$d = \text{cat}\{z_i^l, g_i\}, \quad (4)$$

$$y = \text{GroupNorm}(\text{PReLU}(F(d, \{W_i\}) + z_i^l)),$$

其中: y 表示最后的输出结果。最后的输出结果将经过上采样,作为下一层的输入 x_{i+1}^l 。

2.3.4 深度监督(Deep Supervision, DS)

为解决深度模型训练时梯度消失或爆炸的优化问题, Lee 等^[13]提出深度监督,加快网络收敛速度、提高识别网络的识别能力。这种机制通过导出一个直接联系上下层训练的目标函数而实现,以便在训练过程中抵消不稳定的梯度变化产生的负面影响。

在本文所提出的网络中,如图 1 右侧所示,最后输出的结果将来自不同解码层输出的特征图输出结果作为深度监督的输入。对四个输出特征图使用同一种损失函数,最后将所有分支的输出结果求和取平均作为最终损失。具体表述公式如式(5)所示:

$$L(y, \hat{y}_j) = -\frac{1}{N} \sum_{i=1}^N f(y, \hat{y}_j), \quad (5)$$

其中: y_j, \hat{y}_j 分别表示不同解码层输出的概率值、真实标签值, N 表示分支的总数。

2.3.5 改进过程

将上述三种模块:多尺度并行卷积模块、Non-local 模块和注意力、残差混合解码模块和深度监督机制,分别嵌入至微调的 U-net 模型的编

码区和解码区。对于三种模块的加入方式,依次研究了如下 3 种改进:

首先,提出基于 ARHD 模块的 ARH-UNet 模型。仅在解码部分进行改进。在每个特征提取单元前嵌入 ARHD 模块,在乳腺肿瘤分割中,ARHD 模块可以将编码区的浅层特征信息和解码区的深层特征信息进行融合,再通过注意力系数得到所要的局部显著特征,同时,因为短连接的存在使解码单元的语义信息更加丰富,改善解码部分的梯度流动,缓解梯度弥散和网络退化问题,使深层的网络更容易训练,从而简化网络的训练,有效的提高乳腺肿瘤分割精度,在实验中可以取得一定效果。

其次,提出基 ARHD 模块和 MSPCF 模块的 AMH-UNet。基于上面的 ARH-UNet 模型,在编码部分进行改进。在每个编码部分的特征提取单元后嵌入 MSPCF 模块,与 ARH-UNet 网络融合提出 AMH-UNet。用于提取位于编码部分不同层级的不同尺度的特征信息,将其融合压缩后传递至一下层或通过长连接传递至对应的解码层,以此来丰富网络提取的特征信息,在实验中对于乳腺癌病灶的分割取得较好的效果。

最后,基于上面的 ARH-UNet 模型,引入具有良好可移植性的 Non-local 模块,从而高效地获取全局相关信息,以此形成一个全局的注意力机制。又因为 Non-local 模块计算开销大,所以我们将其放置在编码部分最后一层,从而提出 UTB-Net 网络。并在解码路径的各层输出引入深度监督,构建基于深度监督的 UTB-Net,模型具体细节如图 1 所示,其实验效果表现最佳,详见后续章节 3 实验对比分析。

UTB-Net 可以提取丰富的浅层、深层特征信息,跟已有的乳腺病灶分割网络相比,通过全局注意力与局部注意力相辅相成,实现准确地病灶定位、分割。最后,为了避免长连接的浅层语义信息影响病灶像素级分割,采用深度监督机制使解码单元间特征在每个图像尺度上具有语义识别能力,以此来确保不同尺度的注意力单元能够对特征中前景信息产生最大响应。

2.3.6 Loss Function

在医学图像分割中的一个重要挑战就是数据不平衡。一般情况下,大部分数据集中的病灶

体素的数量远低于非病变体素的数量,乳腺癌数据集亦如此,并且乳腺病灶的面积远小于乳房区域面积。数据集训练出的模型会对分割精度产生影响。为此,针对该种情况采用 Tversky^[23] 损失函数。它是一种基于 Tversky 指数的广义损失函数,可以有效解决数据不平衡的问题,在精度和召回率之间达到更好的平衡。公式如式(6)所示:

$$T(\alpha, \beta) = \frac{\sum_{i=1}^N p_{0i} g_{0i}}{\sum_{i=1}^N p_{0i} g_{0i} + \alpha \sum_{i=1}^N p_{0i} g_{1i} + \beta \sum_{i=1}^N p_{1i} g_{0i}}, \quad (6)$$

其中: p, g 为网络最后一层 sigmoid 输出的结果。 p_{0i} 为病变体素的输出概率, p_{1i} 非病变体素的输出概率。类似地,对于 $g, 0$ 表示病变体素输出概率, 1 表示非病变体素输出的概率。 α, β 则分别控制假阴性和假阳性,可以通过调整 α, β 来控制假阳性和假阴性之间的权衡,在实验中 $\alpha = 0.3, \beta = 0.7$ 。

3 实验与分析

在这项研究应用中,服务器硬件配置为: Intel(R) Core(TM) i9-9900K CPU @ 3.60 GHz $\times 16$; GPU 显卡配置为 GeForce RTX 2080 Ti; 操作系统为 Ubuntu 18.04; 采用 Python3.6 编程语言; 使用基于 Pytorch1.3 的深度学习框架。

3.1 数据库及预处理

本文实验所用 DCE-MRI 影像数据为 160 例自建临床数据库,采集乳腺癌影像的设备型号是西门子 MAGNETO 1.5T MRA,配备乳腺专用 4 通道相控阵表面线圈。在采集过程中病人的姿态呈俯卧位。筛选后最终入库 160 例 T2 期 DCE-MRI 临床数据,影像为横切面方向,图像大小 512×512 。乳房 ROI 区域的标签制作由临床医生借助 LabelMe 软件手动进行标注。乳腺癌的标签制作使用了医学影像软件 3D slicer。对于临床 160 例影像数据,随机划分为训练集 128 例,测试集 32 例,在训练时对数据进行了镜像、缩放、弹性形变等数据增强操作。在训练过程中又对数据进行有效的预处理,首先删除前 0.1% 与后 0.1% 的像素值进行去噪,然后进行灰度归一化

操作。对于乳房ROI区域提取与乳腺癌病灶分割采用同一批训练集和测试集,二者在整体框架上是一种先后顺序的二阶段结构,在训练过程中分别对两个模型进行先后训练。首先使用微调U-net模型进行乳房ROI区域提取,其次对于乳腺癌病灶分割在训练过程中使用He^[22]归一化算法进行初始化,Adam优化器^[25]进行梯度下降,参数设置为: $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 1e^{-8}$ 、学习率为 $1e^{-3}$ 。Batch size设置为2,每个模型迭代次数为200,最后通过sigmoid函数得到输出结果,因为其输出值域为(0,1),所以最后输出结果通过设置阈值($threshold = 0.3$)来判断。

3.2 分割评价标准

评估分割结果借鉴五种常用的医学影像分割评估指标: Dice coefficient (DSC)、Intersection over Union (IoU)、敏感性 (Sensitivity)、特异性 (Specificity)、阳性预测值 (Positive predictive value),公式如式(7)所示。

$$\begin{aligned} DSC &= \frac{2TP}{2TP + FN + FP} \\ IOU &= \frac{TP}{TP + FP + FN} \\ SEN &= \frac{TP}{TP + FN} \\ SPE &= \frac{TN}{TN + FP} \\ PPV &= \frac{TP}{TP + FP} \end{aligned}, \quad (7)$$

其中:TP是像素级分割中前景分割正确的像素点数,TN是像素级分割中背景分割正确的像素点数,FP是像素级分割中背景分割错误的像素点数,FN是像素级分割中前景分割错误的像素点数。

3.3 乳房ROI分割

重新设计基准U-Net模型实现乳房ROI粗略勾画。分割示意如图7所示(彩图见期刊电子版),第1行为3例含有乳腺癌的影像;第2行为医生手动分割乳房区域标签;第3行红色区域为U-Net乳房区ROI粗分割结果。乳房区域可以延深至胸部两侧,以确保没有乳腺癌区域漏掉。同时实现即使较小的乳房区域也得以正确分割。

分割结果的实验统计表明,其DICE值可以达到0.92,ACC达到0.98,如表1所示。满足下一步乳腺癌精确分割的需求。基本不影响后续乳腺癌病灶分割。

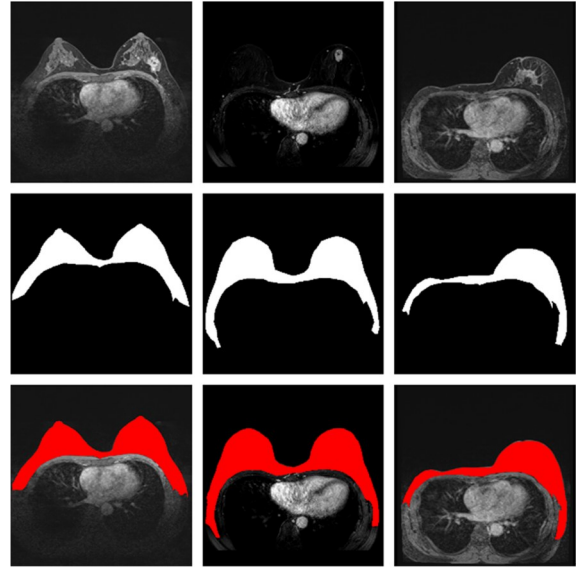


图7 乳房区域粗分割示意图

Fig. 7 Illustration of breast area coarse segmentation

表1 乳房区域分割结果

Tab. 1 breast area segmentation results

Type	Dice	IOU	SEN	SPE	PPV	ACC
U-net	91.98	85.34	92.89	98.78	91.24	98.07

3.4 不同方法实验对比分析

在有限的相同数据集上,将所提出模型与现有表现优异的主流分割模型进行对比测试。我们的微调模型有:U-Net, ARH-UNet, AMH-UNet, AMH-UNet+DS, UTB-Net+DS。而选择用于比对的6个表现优异主流医学影像分割模型:U-Net^[18], 2D-VNet^[26], ResUNet^[28], 2D-DenseU-Net^[27], Pix2pix^[29], Attention-UNet^[12]进行对比分析。实验结果分别如图8~图10所示。

图8展示了3例病人在6个不同表现优异的主流分割模型的乳腺癌病变区分割结果。第1行为3例病人原始影像。第2行分别是3例病人相应的金标准。从第3行到第7行分别对应U-Net^[18], 2D-VNet^[26], ResUNet^[28], 2D-DenseU-Net^[27], Pix2pix^[29]的5种模型相应的分割结果。由图可知,第3行U-net分割结果的边缘粗糙,不能准确区分正常腺体与乳腺癌灶区域,分割结果主要集中在灰度值较高的区域,未能完整地分割出乳腺癌灶,分割结果小于金标准区域。第4行是2D-VNet的分割的结果,可以看出受乳房边缘和周围腺体的影响,甚至无法分割出乳腺

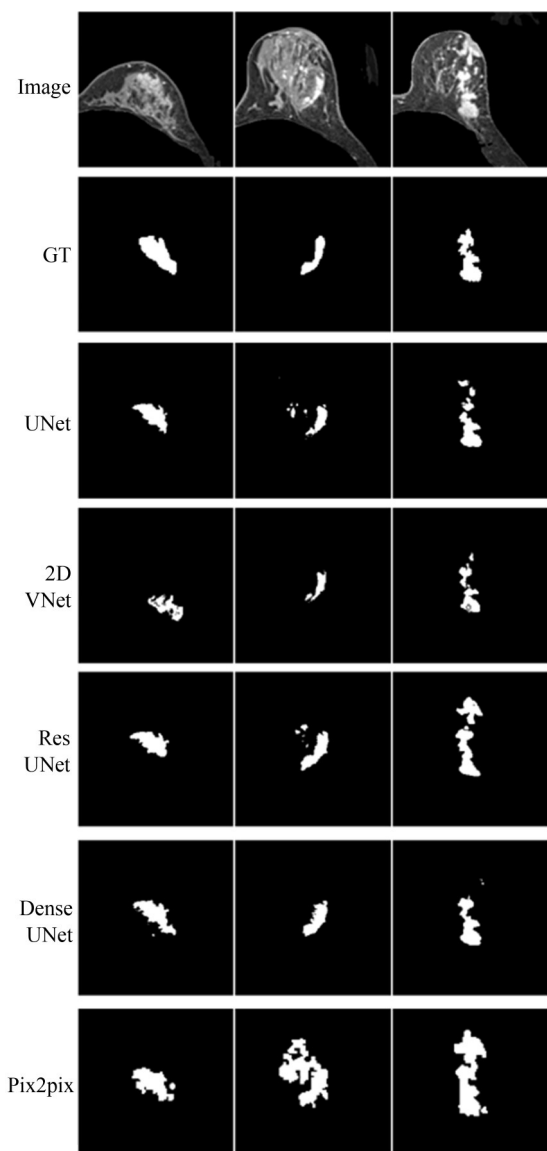


图8 乳腺癌病变区分割结果局部示图

Fig. 8 Partial segmentation of breast tumor lesions

癌灶。

在图8中,第5行是ResUNet^[28]的分割结果,可以看出受周围正常腺体的影响,将正常腺体分割成乳腺癌灶区域。第6行是2D-DenseUNet^[27]的分割结果,由于乳腺癌灶边界具有模糊特性,分割存在过度分割的问题。第7行是Pix2pix^[29]分割结果,作为可用于分割的生成对抗网络,由于乳腺癌灶存在亮度不均,语义信息复杂,最终影响了分割的准确性。Attention-UNet^[12]分割结果在图9中第1行,主观上可以看出对乳腺癌灶的定位较为准确,但是对于边缘和细节的部分还有欠缺。

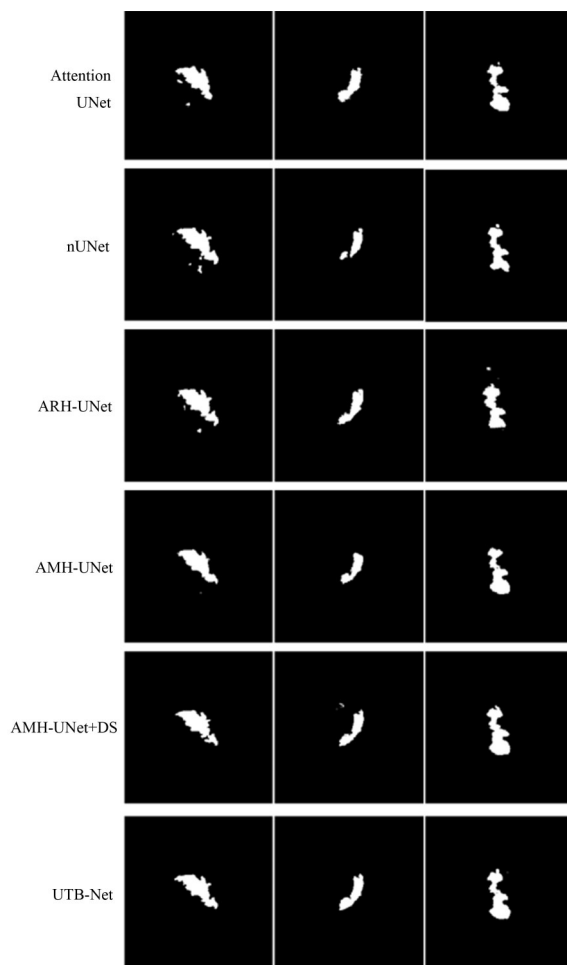


图9 乳腺癌病变区分割结果局部示图

Fig. 9 Partial segmentation of breast tumor lesions

对于本文所提出的5个改进模型分割结果,主要展示在图9的第2行至第6行。

图9第2行是微调U-Net模型的分割结果,在编码部分的最后一层,进行两次Conv→PreLU→Conv→PreLU→GroupNorm操作,在保证不增加参数的同时提取更多高级特征信息。相较于图8中第三行基准U-net分割结果,乳腺癌结果在面积上有所增加。图9中第3行是ARH-UNet分割结果,ARH-UNet与基准U-Net相比不仅继承了微调U-Net优点,全局上观察则对分割的病灶进行了较为准确的定位,局部上展示出获取了更多细节信息。而图9中第4行是AMH-UNet分割结果,其在全局定位与局部细节都得到更进一步提升,分割结果噪点更少。

图9中第5行是AMH-UNet+DS分割结果,模型在局部边缘和信息的识别上展示得更优异。

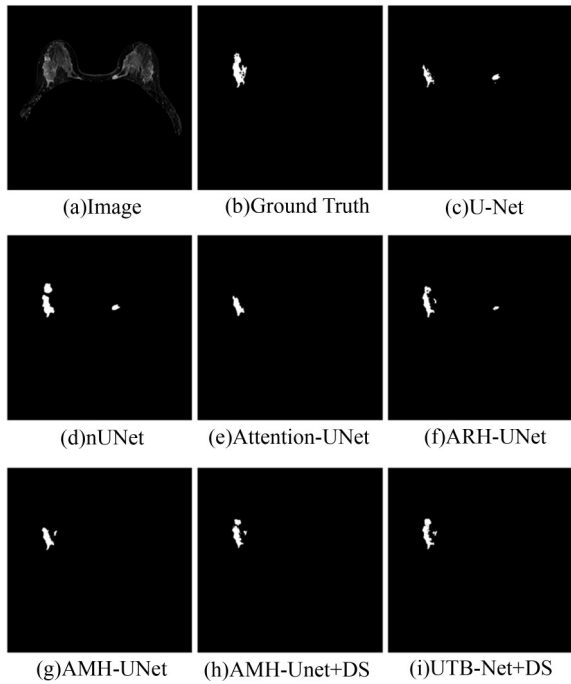


图 10 乳腺癌分割结果全局展示图

Fig. 10 Global illustration of breast tumor segmentation

图 9 中最后一行是 UTB-Net+DS 模型分割结果,与所有实验网络相比表现出最高的分割准确性。这是因为引入了多尺度注意力混合机制,相对其他模型而言,更接近于医生标注的金标准结果,分割结果优于其他方法。另外,结合图 8 和图 9 可以看出微调 U-Net 与基准 U-Net 相比无论全局还是局部都具有更高的分割准确性,最明显的优点是保留了更多图片信息。分割不准确的主要原因主要由于非肿块的乳腺肿瘤区域边界模糊,形态变化大,纹理不规则,受周围正常腺体组织的影响较大。

图 10 进一步使用单个病例从全局分析分割结果。ARH-UNet 与基准 U-Net 相比不仅继承了微调 U-Net 优点,全局上观察则对分割的病灶进行了较为准确的定位,局部上展示出获取了更多细节信息。而 AMH-UNet 则在全局定位与局部细节都得到更进一步提升。使得分割结果的噪点更少。AMH-UNet+DS 则在局部边缘和信息识别上展示的更优异。最后,UTB-Net+DS 与所有实验网络相比表现出最高的分割准确性。

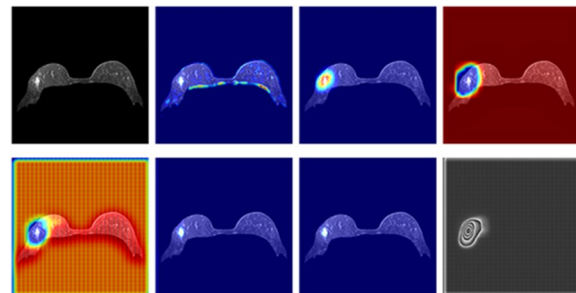
通过分析局部细节分割结果图 8~图 9 和全

局分割图 10,可以看出 UTB-Net+DS 在全局与局部的分割区域定位、边缘信息识别和细节信息分割方面都取得了最好的表现。最终得出在分割过程中对分割区域进行全局与局部精准定位、保留更多图像信息与增强信息识别能力对像素级分割结果至关重要,这也证实了 UTB-Net+DS 网络存在一定的鲁棒性与合理性。

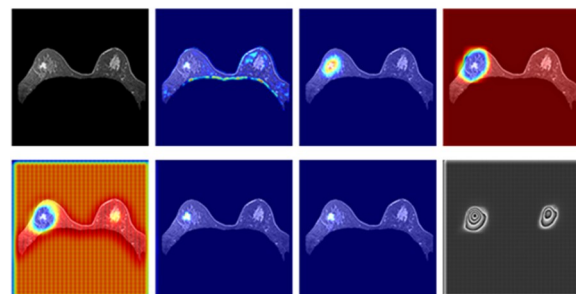
3.5 注意力图和特征分析

如图 11 所示,为两例乳腺癌灶注意力的热力图和特征图。网络初始化时注意力点在乳房全局区域服从均匀分布。随着数据在网络中流动,网络的焦点逐渐聚焦于乳腺癌灶的边缘,最后从粗糙的定位和大概轮廓逐渐聚焦于乳腺癌灶,从而进行精确地分割。即使当数据流经网络过程中产生错误的聚焦也可以被逐渐修正。从中可以观察到,残差与注意力单元 ARHD 与 Non-local 能够帮助进行病灶定位与纠正,以及获取目标的重要特征。

图 11 为两例病人注意力图采用 Grad-CAM^[30]方法产生热力图和特征图作为注意力的仿真结



(a) 病例 1 热力图与特征图
(a) Case 1 at tentional heat map and feature map



(b) 病例 2 热力图与特征图
(b) Case 2 at tentional heat map and feature map

图 11 乳腺癌灶分割注意力热力图与特征图

Fig. 11 Breast tumor segmentation Attention and feature map

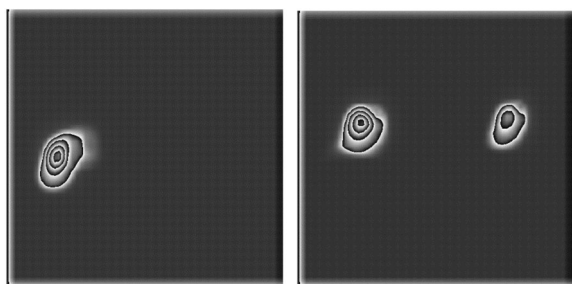


图 12 乳腺癌灶分割特征图

Fig. 12 Breast tumor segmentation Attention and feature map

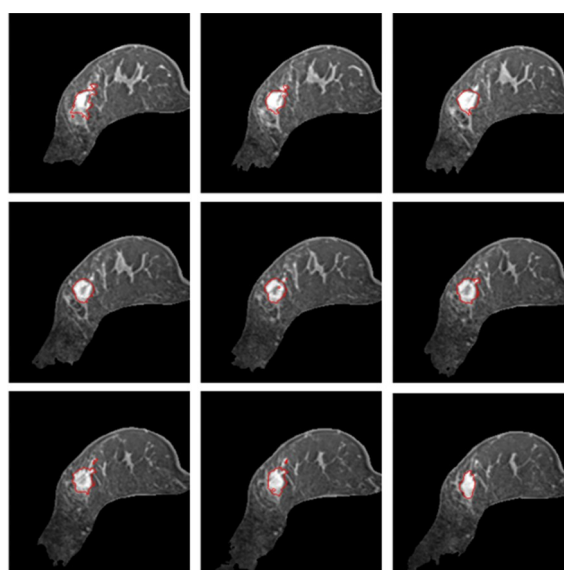
果,通过病例 1 可以观察到模型使用注意力机制后可以稳定的定位到病灶部位。病例 2 可以观察到,图中左乳区有乳腺癌病灶,图 2 中右乳区存在虚假疑似病灶区域,所提模型的注意力机制配合残差与注意力单元(ARHD)与 Non-local 能够帮助进行疑似病灶定位与纠正,以及获取目标的重要特征。所获得的注意力特征图如图 12 所示,从所提模型的编码路径与解码路径的连接处提取注意力特征图,从中可以看到所提模型稳定的定位到病灶区和疑似病灶区,而对于图 2 中右乳区虚假病灶区,所提模型可以在解码部分进行抑制,能够有效的减少虚假病灶区错分割。

3.6 模型可视化与消融实验分析

图 13 展示深度监督 UTB-Net 对单例病灶的乳腺癌灶分割结果的 2D 和 3D 拟合效果。图 13 (a) 预测结果轮廓基本拟合乳腺癌灶区。由于 MSPCF 能够在同一层级获得不同感受野的特征,可提高网络获取模糊边缘和小面积肿块信息的能力。为了更加直观展现所提模型预测结果,分别对金标准和预测结果进行 3D 构建,结果分别如图 13(b) 和 13(c)。由图可知,所提模型的分割结果虽然接近金标准的大体轮廓,但在乳腺癌病变边缘像素级分割上仍有待加强。

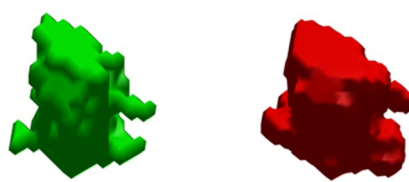
为了进一步验证注意力-残差混合解码模块、多尺度并行卷积融合模块、深度监督机制和 Non-local 模块对模型的贡献,进行了模型的消融实验仿真和分析。如图 14 所示,第 1 列为 3 例病患的原始影像;第 2 列为金标准在 3D 空间中的重构图;第 3 列为微调 U-net 加入注意力-残差混合解码(ARHD)模块后,分割结果在 3D 空间中的重

构图;第 4 列为模型加入多尺度并行卷积融合(MSPCF)模块后,分割结果在 3D 空间中的重构图;第 5 列为加入深度监督机制(Depth supervision)后,分割结果在 3D 空间中的重构图;最后 1 列为加入 Non-local 后,所提出模型的分割结果在 3D 空间中的重构图。在各个模块添加的过程中,在表面细节、颗粒度、中间的空洞处理、病灶边缘上分割效果逐渐提升,对于综合后的模型表现为最优。



(a)乳腺癌预测效果图

(a) Prediction of breast tumor lesions



(b)金标准3D构建图

(b) Ground truth 3D reconstruction

(c)预测结果3D构建图

(c) Prediction result 3D reconstruction

图 13 单病例分割与 3D 重构

Fig. 13 Segmentation and 3D reconstruction of single case

对于模型中的卷积层特征进行了提取和再现,如图 15 所示,分别对编码层 1~4 进行特征可视化,从图中可知所提模型在编码层主要学习纹理、边缘、灰度等浅层特征信息,在解码路径上分别对解码层 1~4 进行特征可视化,从图中可知解码层我们的模型主要学习深层的语义信息。

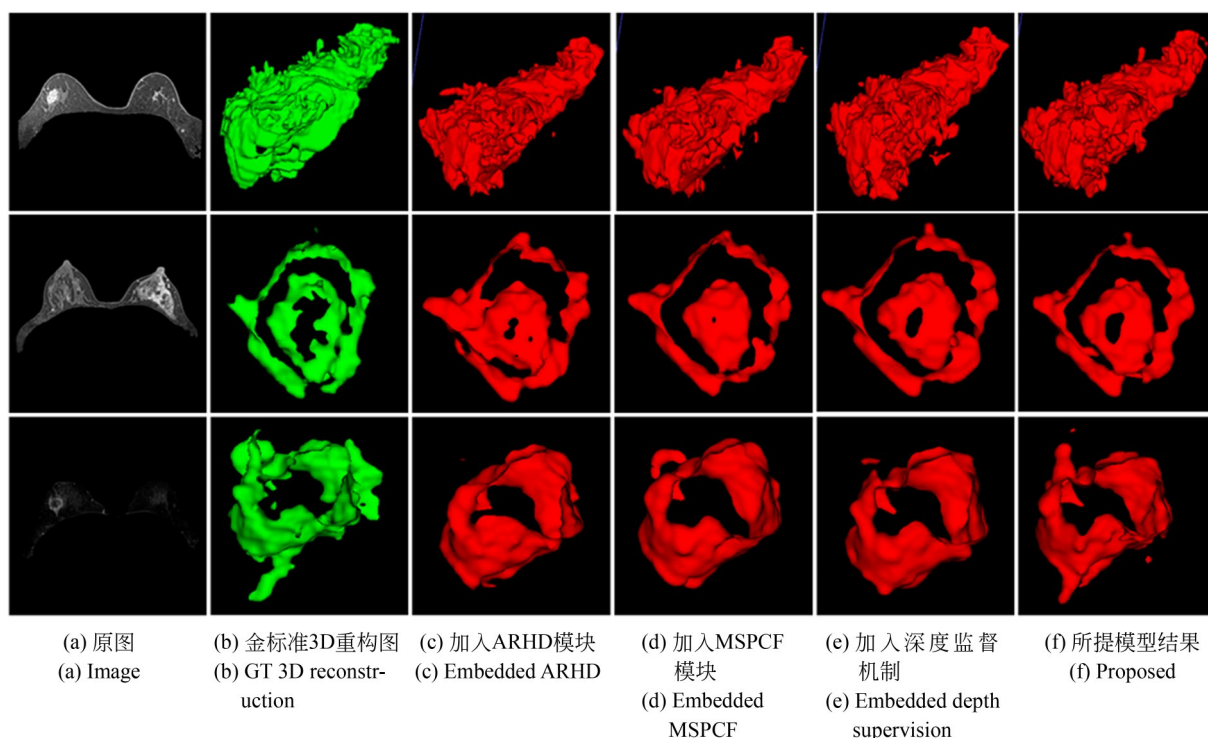


图 14 消融实验过程展示

Fig. 14 Diagram of ablation experiment process

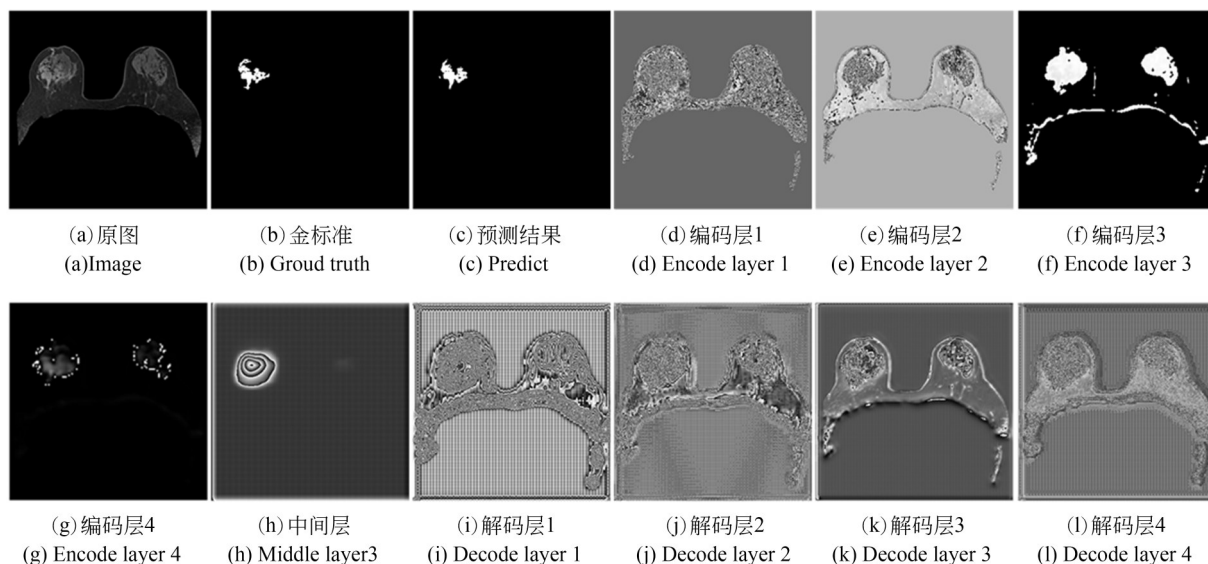


图 15 卷积层仿真结果展示

Fig. 15 Diagram of ablation experiment process

3.7 分割结果统计分析

为了更进一步验证深度监督 UTB-Net 对于乳腺癌灶的分割性能,使用 Dice coefficient (DSC), Intersection over Union (IoU), 敏感性 (Sensitivity), 特异性 (Specificity), 阳性预测值

(Positive predictive value) 作为该数据集上的评估指标,对测试结果进行统计分析。

如表 2 所示,在该数据集深度监督 UTB-Net 在所有评估指标上取得了很好的结果。但由于乳腺 DCE-MRI 临床数据集包含大量噪声,背景

表 2 不同模型分割结果统计分析

Tab. 2 Compare of segmentation results

Model Type	Dice	IOU	SEN	SPE	PPV
U-Net[18]	73.14	60.18	80.20	99.88	73.10
2D-VNet[26]	74.16	61.59	83.90	99.85	72.33
ResUNet[28]	75.55	63.22	82.19	99.87	75.26
DenseUNet[27]	76.67	65.29	80.31	99.90	77.81
Pix2pix [37]	58.25	43.57	78.37	99.82	59.00
Att-UNet[12]	75.94	63.85	80.12	99.90	77.86
微调 U-Net	75.68	63.54	81.12	99.88	76.71
ARH-UNet	77.06	65.23	82.05	99.89	77.26
AMH-UNet	77.21	65.10	83.83	99.89	77.02
AMH-UNet+DS	77.32	65.37	84.32	99.88	75.99
UTB-Net+DS	77.92	66.10	84.14	99.88	77.53

区域与前景区域不平衡,所以造成模型表现出敏感性指标低特异性指标高这一特点。深度监督 UTB-Net 的 DICE 值超出对比基准 U-Net 4.78%, IOU 上超出 5.92%, SEN 超出 3.94%, 在 SPE 上持平,在 PPV 上超出 4.43%。可以看出,深度监督 UTB-Net 结果与基准 U-Net^[18]相比更具有稳定性,且测试结果平均水平也高于它。所提最终深度监督 UTB-Net 是经过一系列模型不断创新改进而来。

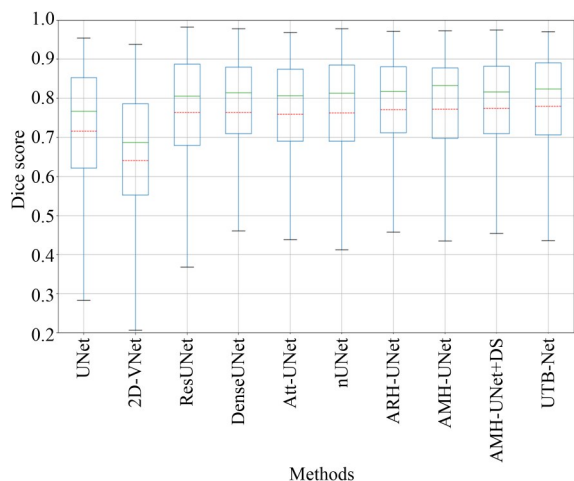


图 16 所有实验模型的箱线图展示

Fig. 16 BOX diagrams of all comparative experiments

图 16 为各个算法的 BOX 箱线图统计对比(彩图见期刊电子版),其横轴表述前面实验对比算法中所用到的不同模型,纵轴表示 DICE 系数,

本文所提微调 U-Net、ARH-UNet、AMH-UNet、AMH-UNet+DS、UTB-Net+DS 模型,随着改进过程,所有指标值、稳定性均逐步升高,进一步验证了所提模块的有效性和对模型改进与创新方向的正确性。BOX 图中的绿色线和红色线分别表示测试数据的平均值和中值。

3.8 模型参数分析

由表 3 可知,所提模型参数约 23.7 MB,所提出模型使用微调 U-net 作为基础的编码与解码结构,即 Conv→PReLU→Conv→PReLU→Group-Norm,在编码路径加入多尺度并行卷积融合(MSPCF)模块构成编码层的一部分。UTB-Net 包含四层编码层和四个最大池化运算,每个运算之后特征图面积减半。除 MSPCF 外,所有卷积层都有 3×3 的卷积核,通道数如上表所示,第一个编码层输出有 32 个通道,随着网络的加深,每个编码层和最大池化运算之后通道数会加倍,直到通道数为 512,在之后的卷积层与 Non-local 模块输出通道数不变。

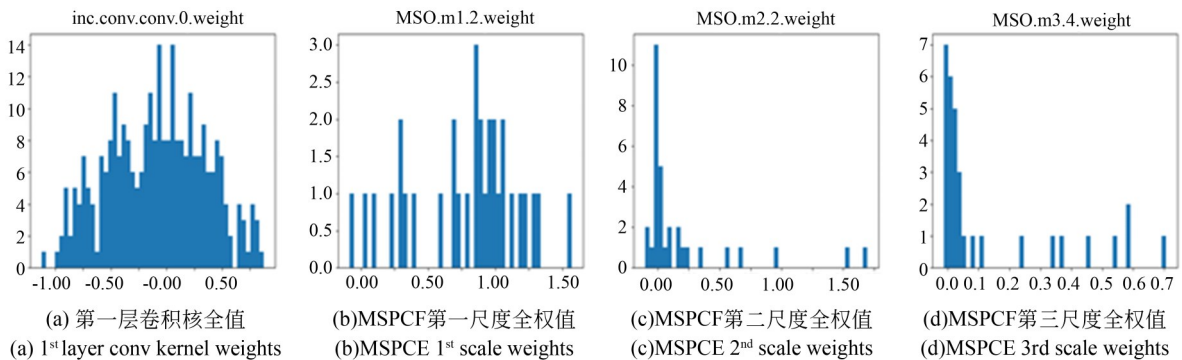
在解码部分,包含注意力-残差混合解码(ARHD)模块构建的三层解码层、一个基础解码结构和四个转置卷积运算,每个运算后特征图面积增大两倍,同时通道参数减少一半。转置卷积的输出连接着编码器对应部分的输出,以此作为下层的输入。由此得到的特征图通过卷积运算来处理,以保持通道数量与对称编码部分相同。最后将每个解码层的输出的特征图,通过上采样与 1×1 卷积操作,还原成 512×512 大小的单通道特征图,经过深度监督机制求和取平均获得最终输出结果。

对所提出模型的关键卷积核权值与偏置进行了可视化分析。如图 17 所示,从图 17(a)可以看出编码层第一层的卷积核权值服从正态分布,这一层卷积核主要学习一些粗糙的浅层特征。图 17(b),(c),(d),(i)是编码层第一层多尺度并行卷积融合(MSPCF)模块不同尺度的卷积核权值,从中可以看到不同尺度的卷积核所学习的特征是不相同,这进一步验证所提出 MSPCF 模块的合理性与有效性。图 12(k),(l),(o)是解码层 1~3 的卷积核权值,从图中可

表 3 UTB-Net主要模块参数
Tab.3 Main module parameters of UTB net

Layer	Type	Parameters	Size/Stride	Output
Encode1	Conv.	1→32 (Filters)	3×3/1	512×512
	PReLU	32	/	/
	Conv.	32(Filters)	3×3/1	512×512
	PReLU	32	/	/
	GN	32 (Channels)	/	32(Groups)
Parallel structure				
Multi-Scale Parallel Conv. Fusion 1	Conv(1,1)	Conv(3,3)	Conv(3,3)	AvgPool(256)
	PReLU(32)	PReLU(32)	PReLU(32)	Upsample (2.0, bilinear)
	GN (32,32)	Conv(3,3)	Conv(3,3)	PReLU(32)
		GN (32,32)	PReLU(32)	GN(32,32)
		GN(32,32)		
	Conv.	128→32 (Filters)	1×1/1	512×512
Down-Sampling1	MaxPool	/	2×2/2	256×256
Connect+ Non_local		256→512(Filters)		32×32
UpSampling3	Conv. T.	128→64(Filters)	2×2/2	256×256
ARHD Block3	/	128→64(Filters)	/	256×256
UpSampling4	Conv. T.	64→32(Filters)	2×2/2	512×512
Decode	Conv.	64→32(Filters)	3×3/1	512×512
	PReLU	32	/	/
	Conv.	32→32(Filters)	3×3/1	512×512
	PReLU	32	/	/
	GN	32 (Channels)	/	32 (Groups)
Out	Conv.	32→1 (Filters)	1×1/1	512×512
Deep Supervisor		Sum to average		512×512

Number of trainable parameters 23.7 MB in Model



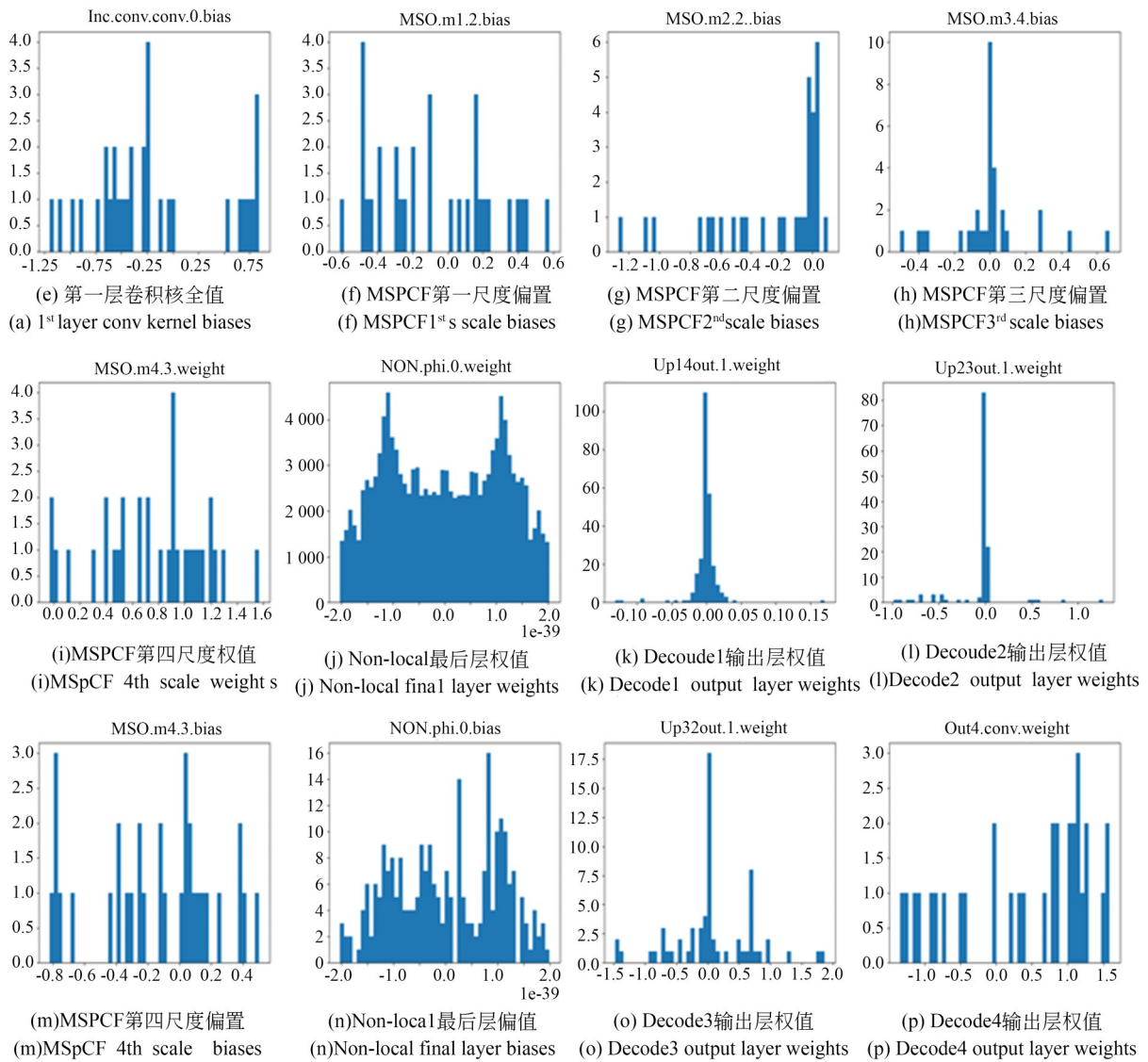


图 17 UTB-Net 关键卷积核权值与偏置

Fig. 17 BOX diagrams of all comparative experiments

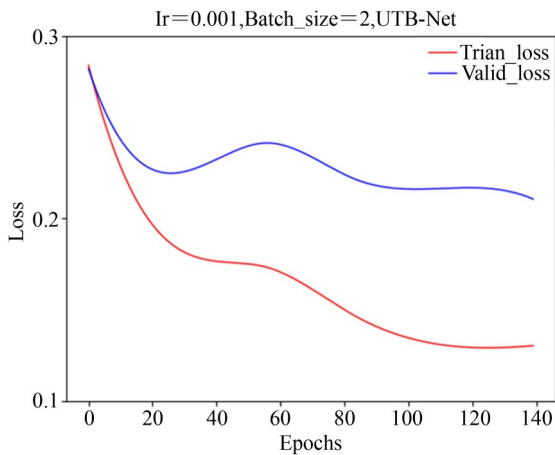


图 18 UTB-Net 训练与验证 Loss 曲线

Fig. 18 UTB-N train and valid loss

以看出注意力-残差混合解码(ARHD)模块的卷积核所学习特征分布基本一致,进一步表明所提 ARHD 模块能稳定识别与定位病灶区域,能有效学习到深层语义信息。图 12(p)为解码层最后一层的卷积核参数,由于长连接的存在使其不仅能保留上层所学习的深层特征信息,也能保留对应编码层的浅层特征信息。综上,通过卷积核权值的分析,所提模型具有一定的合理性与正确性。

3.9 分割模型训练与验证

如图 18 可知,本文所提出的模型在训练过程中,训练 loss 在迭代 0~25 次之间收敛速度较快,

在经历 25~140 次迭代后曲线收敛逐渐趋于平缓。验证 loss 在 0~25 次迭代之间收敛速度也较快,在经历 25~140 次迭代后曲线趋于平缓。综上所述,本文所提的模型在训练与验证过程中具有一定的稳定性与鲁棒性。

4 结 论

在有限的数据集上,针对具有复杂背景的 DCE-MRI 乳腺癌图像分割问题。本文提出了一种分阶段乳腺癌图像分割模型。首先,借鉴 U-

net 模型完成乳房区域的粗略勾画。其次,提出一种 UTB-Net 模型,在解码路径整合注意力-残差混合解码结构,在编码路径和末端分别引入多尺度并行卷积模块、Non-local 模块。实验结果表明所提出的模型实验相较于 U-Net 基准模型,乳腺癌分割指标 DICE, IOU, SEN, PPV 分别提升了 4%。4.78%, 5.92%, 3.94%。不足之处:对于网孔状较小的病变区细节分割上需进一步提升。下一步将研究乳腺癌图像的 3D 分割和定位,结合影像组学,建立预测模型,提高乳腺癌的诊断、治疗和预后。

参考文献:

- [1] LITJENS G, KOOI T, BEJNORDI B E, *et al.* A survey on deep learning in medical image analysis [J]. *Medical Image Analysis*, 2017, 42(9):60-88.
- [2] LI S, DONG M, DU G, *et al.* Attention dense-unet for automatic breast mass segmentation in digital mammogram[J]. *IEEE Access*, 2019, (99):1-1.
- [3] 周国华,蒋晖,顾晓清,等.多视角判别度量学习的乳腺影像检索方法[J].*液晶与显示*, 2020, 35(6): 619-630.
ZHOU G H, JIANG H, GU X Q, *et al.* Multi-view metric learning with Fisher discriminant analysis and its applications for breast image retrieval[J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(6): 619-630. (in Chinese)
- [4] LIU L L, CHENG J H, QUAN Q, *et al.* A survey on u-shaped networks in medical image segmentations[J]. *Neurocomputing*, 2020, 409: 224-258.
- [5] JANA M, OLGA C, MAJA M. Automated breast-region segmentation in the axial breast MR images [J]. *Computers in Biology and Medicine*, 2015, 62:55-64.
- [6] ZHENG Y, BALOCH S, ENGLANDER S, *et al.* Segmentation and Classification of breast tumors Using Dynamic Contrast-Enhanced MR Images. In: Ayache N., Ourselin S., Maeder A. (eds) *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2007* [J]. *Lecture Notes in Computer Science*, 4792: 393-401. Springer, Berlin, Heidberg.
- [7] ASHRAF A, GAVENONIS S, DAYE D, *et al.* A multi-channel markov random field framework for tumor segmentation with an application to classification of gene expression based breast cancer recurrence risk[J]. *IEEE Transactions on Medical Imaging*, 2013, 32(4):637-648.
- [8] 冯宝,陈业航,刘壮盛,等.结合MRF能量和模糊速度的乳腺癌图像分割方法[J].*自动化学报*, 2019. 2020, 46(6): 1188-1199.
FENG B, CHEN Y HA, LIU ZH SH, *et al.* Segmentation of breast cancer on DCE-MRI images with MRF energy and fuzzy speed function[J]. *Acta Automatica Sinica*, 2020, 46(6): 1188-1199. (in Chinese)
- [9] 范丽丽,赵宏伟,赵浩宇,等.基于深度卷积神经网络的目标检测研究综述[J].*光学精密工程*, 2020, 28(5):1152-1164.
FAN L L, ZHAO H W, ZHAO H Y, *et al.* Survey of target detection based on deep convolutional neural networks[J]. *Opt. Precision Eng.*, 2020, 28(5): 1152 -1164. (in Chinese)
- [10] 陈彦彤,陈伟楠,张献中,等.基于深度卷积神经网络的蝇类面部识别[J].*光学精密工程*, 2020, 28(7): 1558-1567.
CHEN Y T, CHEN W N, ZHANG X ZH, *et al.* Fly facial recognition based on deep convolutional neural network[J]. *Opt. Precision Eng.*, 2020, 28(7): 1558-1567. (in Chinese)

- [11] ZHANG J, SAHA A, ZHU Z, *et al.* Hierarchical Convolutional Neural Networks for Segmentation of breast tumors in MRI with Application to Radiogenomics[J]. *IEEE Transactions on Medical Imaging*, 2018, 38(2):435-447.
- [12] OKTAY O, SCHLEMPER J, FOLGOC L, *et al.* Attention u-net: Learning where to look for the pancreas [J]. *arXiv preprint arXiv: 1804.03999*, 2018.
- [13] LEE C Y, XIE S N, GALLAGHER P, *et al.* Deeply-supervised nets[C]. 2015. *arXiv preprint arXiv:1409.5185v2*.
- [14] 徐胜军, 欧阳朴衍, 郭学源, 等. 多尺度特征融合空洞卷积 ResNet 遥感图像建筑物分割[J]. *光学精密工程*, 2020, 28(7):1588-1599.
- XU SH J, OU Y P, GUO X Y, *et al.* Building segmentation in remote sensing image based on multiscale-feature fusion dilated convolution resnet [J]. *Opt. Precision Eng.*, 2020, 28(7):1588-1599. (in Chinese)
- [15] LI C, SUN H, LIU Z, *et al.* Learning cross-modal deep representations for multi-modal mr image segmentation[C]. 2019 *Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. MICCAI 2019*, LNCS 11765:57-65.
- [16] WEI D, WEINSTEIN S, HSIEH M K, *et al.* Three-dimensional whole breast segmentation in sagittal and axial breast mri with dense depth field modeling and localized self-adaptation for chest-wall line detection[J]. *IEEE Transactions on Biomedical Engineering*, 2019, 66(6):1567-1579.
- [17] GABRIELE P, MARIO S, ROBERTA F, *et al.* Multi-planar 3D breast segmentation in MRI via deep convolutional neural networks[J]. *Artificial Intelligence in Medicine*, 2019, 103:101781.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for biomedical image segmentation[J]. *International Conference on Medical Image Computing and Computer-Assisted Intervention-MICCAI*, 2015. Cham: Springer; 2015: 234-241.
- [19] WANG X, GIRSHICK R, GUPTA A, *et al.* Non-local neural networks[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018: 7794-7803.
- [20] WU Y, HE K. Group normalization[C]. *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3-19.
- [21] SZEGEDY C, VANHOUCKE V, IOFFE S, *et al.* Rethinking the inception architecture for computer vision[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 2818-2826.
- [22] HE K, ZHANG X, REN S, *et al.* Deep residual learning for image recognition [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 770-778.
- [23] HE K, ZHANG X, REN S, *et al.* Delving deep into rectifiers: Surpassing human-level performance on image net classification[C]. *Proceedings of the IEEE International Conference on Computer vision*, 2015: 1026-1034.
- [24] SALEHI S S M, ERDOGMUS D, GHOLIPOUR A. Tversky loss function for image segmentation using 3D fully convolutional deep networks [C]. *International Workshop on Machine Learning in Medical Imaging, Springer, Cham*, 2017: 379-387.
- [25] KINGMA D P, BA J. Adam: A method for stochastic optimization [J]. *arXiv preprint arXiv: 1412.6980*, 2014.
- [26] MILLETARI F, NAVAB N, AHMADI S. V-net: fully convolutional neural networks for volumetric medical image segmentation [C]. 2016 *Fourth International Conference on 3D Vision (3DV)*, Stanford, CA, 2016:565-571.
- [27] LI X, CHEN H, QI X, *et al.* H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes[J]. *IEEE transactions on medical imaging*, 2018, 37(12):2663-2674.
- [28] HE K, ZHANG X, REN S, *et al.* Deep residual learning for image recognition [C] *IEEE Conference on Computer Vision and Pattern Recognition*

(CVPR), 2016: 770-778.

- [29] PHILLIP I, ZHU J Y, ZHO T H, *et al.* Efras. *Image-to-Image Translation with Conditional Adversarial Networks* [J]. *arXiv preprint, arXiv: 1611.07004v3*, 2016.

- [30] SELVARAJU R R, COGSWELL M, DAS A, *et al.* Grad-cam: Visual explanations from deep networks via gradient-based localization[C]. *Proceedings of the IEEE International Conference on Computer Vision(CVPR)*, 2017: 618-626.

作者简介:



秦传波(1982—),男,安徽宿州人,博士,讲师,硕士生导师,分别于2004年、2008年在五邑大学取得学士、硕士学位,2015年在华南理工大学获得博士学位,主要研究方向为医学影像处理、生物特征识别。E-mail: tenround@163.com



曾军英(1977—),男,江西赣州人,博士,副教授,硕士生导师,2005年在云南大学取得硕士学位,2008年在北京邮电大学获得博士学位,主要研究方向为图像处理、生物特征识别。E-mail: zengjunying@126.com