ContextSE-UNet：基于上下文信息的肝脏肿瘤分割算法研究

肖 露[1,2] 王亚茹[1,2] 刘锦博[1,2] 邓 锴[1,2] 黄星帆[1]
XIAO Lu WANG Yaru LIU Jinbo DENG Kai HUANG Xingfan

摘 要 肝癌是全球发病率与致死率均居前列的恶性肿瘤，腹部CT影像下的肝脏肿瘤精准分割是肝癌早期诊断、手术切除规划、放疗靶区勾画及预后评估的核心环节，对提升临床诊疗精度与患者生存率具有重要价值。然而，肝脏肿瘤自动化分割面临类间灰度相似性高、类内形态异质性强、病灶边界模糊及尺度差异显著等核心难题。文章提出一种肝脏肿瘤分割模型ContextSE-UNet，通过局部上下文模块与全局上下文模块分别捕获同层多尺度与跨层级的长程上下文依赖关系，有效缓解类间高相似性、类内高差异性及多尺度建模困难；同时引入SCR模块建模前景与背景的结构约束关系，并通过EA模块显式增强边界特征表达，进一步细化分割边缘。在公开LiTS肝脏肿瘤分割数据集上的实验结果表明，所提算法的mIoU与Dice系数分别可达79.3%、73.7%，与主流对比算法相比，该方法能有效提升肝脏肿瘤分割精度，尤其在小病灶与模糊边界场景下表现优异。

关键词 肝脏肿瘤；上下文信息；图像分割

doi：10.3969/j.issn.1672-9528.2026.08.005

0 引言

在临床诊断中，腹部CT图像是肝脏肿瘤筛查、诊断的核心影像学依据，精准的肝脏肿瘤分割是肝癌早期诊断、手术切除规划以及预后评估的关键前提，对实现精准诊疗、提升患者生存率具有重要临床价值[1]。然而，肝脏肿瘤的人工分割高度依赖医师的临床经验，难以满足大规模临床诊疗与定量分析的需求。因此，研究高效、精准的自动化肝脏肿瘤分割方法是当前医学影像处理领域的重要研究方向。肝脏肿瘤在CT影像中表现出高度的复杂性，使得自动化分割面临诸多核心挑战[2-3]：一是类间高相似性与类内高差异性并存，肿瘤组织与正常肝实质、周边血管及邻近器官的灰度对比度较低，易造成误分割；同时不同病理类型、不同分期的肿瘤在大小、形态、密度上差异显著，微小转移灶与巨块型肿瘤的尺度跨度极大，单一感受野难以完整捕捉病灶特征；二是病灶边界模糊不清，部分浸润性肿瘤与正常肝组织无明确分界，且易受图像噪声、伪影及部分容积效应的干扰，导致边缘分割精度不足；三是多尺度特征建模困难，肝脏肿瘤从毫米级微小病灶到占据肝叶的大范围病变均有分布，网络下采样过程中小尺度病灶的细节信息易丢失，大病灶的全局语义特征也难以充分建模。

1 相关工作

在医学图像领域，U-Net凭借其出色的性能表现被广泛应用于肝脏肿瘤分割任务中。Seo等[4]针对U-Net方法对肝肿瘤图像中大目标边缘分割效果差以及小目标全局信息丢失问题，提出mU-Net网络，在跳跃连接部分加入反卷积+激活函数模块，避免低分辨率信息的重复利用，有效提升了肝肿瘤分割精度。钟思华等[5]将密集连接嵌入U-Net网络的卷积层中，增强了特征的复用和传播，该思路被广泛迁移至肝脏肿瘤分割任务中，有效缓解了梯度消失问题，提升了小病灶的特征提取能力。

针对肝脏肿瘤的多尺度与上下文建模问题，学者们开展了诸多针对性研究。Li等[6]提出SPA-UNet，通过空间金字塔卷积模块融合不同扩张率的卷积特征，实现多尺度病灶信息的提取，提升了对不同大小肝肿瘤的分割效果。Ruan等[7]提出EGE-UNet，通过分组特征处理与哈达玛积注意力机制从多维度提取病理信息，增强了模型对异质性肿瘤的特征表达能力。Wang等[8]提出AGCAF-Net，结合注意力引导模块与上下文融合模块，进一步强化了模型对全局上下文信息的捕捉能力。

1. 郑州西亚斯学院工学部 河南郑州 451150
2. 河南省智能制造数字孪生工程研究中心 河南郑州 451150
[基金项目]2025年河南省科技发展计划（科技攻关）项目（252102411005）；郑州西亚斯学院2025年度科研资助项目（2025XKE104）

尽管上述方法已取得一定进展，但仍存在三方面核心不足：一是多数方法仅在同层内进行多尺度特征融合，缺乏跨层级的长程上下文依赖建模，难以充分缓解类间相似性与类内异质性带来的挑战；二是对边界信息的显式建模不足，多依赖像素级损失间接优化边界，导致弱边界场景下分割精度有限；三是现有方法较少从结构一致性角度对分割结果施加区域级约束，病灶的形态完整性难以保证。针对上述问题，本文从上下文信息融合与边界结构增强两个维度改进 U-Net 架构，提出 ContextSE-UNet 模型，以提升肝脏肿瘤的分割精度与鲁棒性。

2 ContextSE-UNet

结合上文提到的各种问题，本文提出的 ContextSE-UNet 网络主要通过结合多尺度的上下文信息，克服 U-Net 上下文短距离依赖的限制，提取更加丰富的语义和细节特征，其整体框架如图 1 所示。局部上下文模块和全局上下文模块用来缓解类间的高相似性和类内的高差异性问题：局部上下文模块通过逐渐变化的感受野提取特征信息，先是向感受野逐步扩大的方向扫描生成特征图，然后向感受野缩小的方向生成特征图，最后形成同层更为“清晰”的特征表示；全局上下文模块应用全局增强机制，从各种“角度”（不同层间的多尺度信息）观察感染区域，并得出“全局”特征图。同时，为解决低对比度和模糊边界问题，加入 SCR 模块强调区域之间的相关性，以提供区域级约束，另外加入 EA 模块来显式地提升目标区域的边界特征的表达。

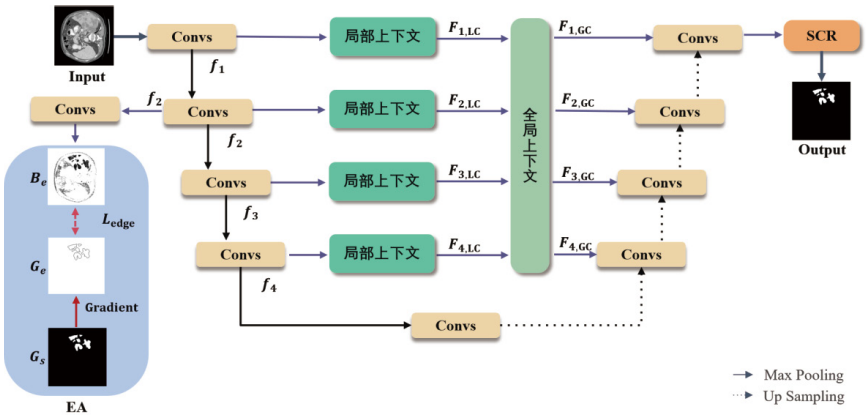


图 1 ContextSE-UNet 网络结构

3 局部上下文模块

在分割任务中，由于类间的高相似性和类内的高差异性，模型需要对各种感染病变做出反应，单个固定大小的局部感受野无法满足对不同大小病变的精准分割。具体来说，不同的感受野对各种大小的病变有不同的反应；大的感受野可以捕捉更多的宏观特征并描绘不同病变之间的关系，而小的感受野则更加关注局部信息。为了解决这一问题，采用摄影中的自动对焦思想对模型进行改进。相机镜头不断缩小和放大是摄影中常见的技术，通过不断地变换焦距确定具有最佳对焦距离的场景。本文方法运用自动对焦机制思想将变换操作与逐渐扩大和缩小的感受野结合，观察不同大小感受野下的特征表示，最终将其聚合成“清晰”的表示。图 2 为局部上下文模块的原理图，主要包括两个过程：生成连续观察和整合聚焦。

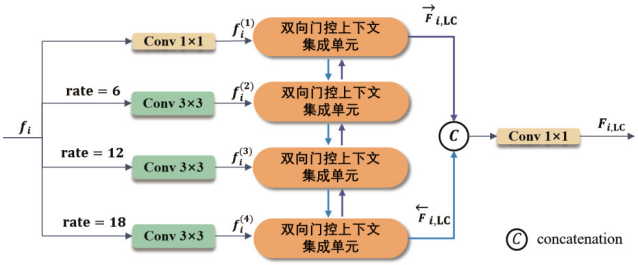


图 2 局部上下文模块原理图

（1）生成连续观察

模仿镜头变焦过程，将一系列伴随逐渐扩大感受野的变换操作记作 $T^{(t)}$。采用多个卷积运算来进一步获取具有不同尺寸的细节特征。具体而言，在编码器子网的每个阶段，将初始特征图 $\boldsymbol{f}_i$ 馈送到 $T^{(t)}$ 中，获得结果特征图 $\boldsymbol{f}_i^{(t)}$，其中记录了由不同尺寸的感受野产生的病变表示，具体表示为：

$$\boldsymbol{f}_i^{(t)} = T^{(t)}(\boldsymbol{f}_i;\theta_{T^{(t)}}) \tag{1}$$

（2）整合聚焦

通过双向门控上下文集成单元来实现在同一层自动整合这些特征图，此过程是模拟相机的自动对焦过程。图 3 是双向门控上下文集成单元的核心集成过程。其中，$\mathrm{CA}(\boldsymbol{M}_i^{(t-1)})$ 是一个通道注意力单元，它找到特征图上的关键通道，并重新为每个像素生成一个权重；$\mathrm{SA}(\boldsymbol{M}_i^{(t-1)})$ 是一个空间注意力单元，找到在特征图上关键信息最多的位置，并重新为其赋予权重，是对通道注意力的补充，简单来说，通道注意力单元是为了找到有重要信息的通道，而空间注意力则是在此基础上，基于通道方向找到信息聚集最多的位置，其过程如图 3 所示，输入的特征表示为：

$$\mathrm{CA}(\boldsymbol{M}_i^{(t)}) = \sigma(\boldsymbol{F}(\mathrm{AvgPool}(\boldsymbol{M}_i^{(t-1)});\theta_{i,\mathrm{CA}})) \tag{2}$$

$$\boldsymbol{M}_{i,\mathrm{CA}}^{(t-1)} = \mathrm{CA}(\boldsymbol{M}_i^{(t)}) \otimes \boldsymbol{M}_i^{(t-1)} \tag{3}$$

$$\mathrm{SA}(\boldsymbol{M}_i^{(t)}) = \sigma(\boldsymbol{F}(\mathrm{AvgPool}(\boldsymbol{M}_{i,\mathrm{CA}}^{(t-1)});\theta_{i,\mathrm{SA}})) \tag{4}$$

式中：σ 表示用于产生过滤注意力的 sigmoid 函数，AvgPool() 表示全局平均池化，$\otimes$ 表示像素级乘积。

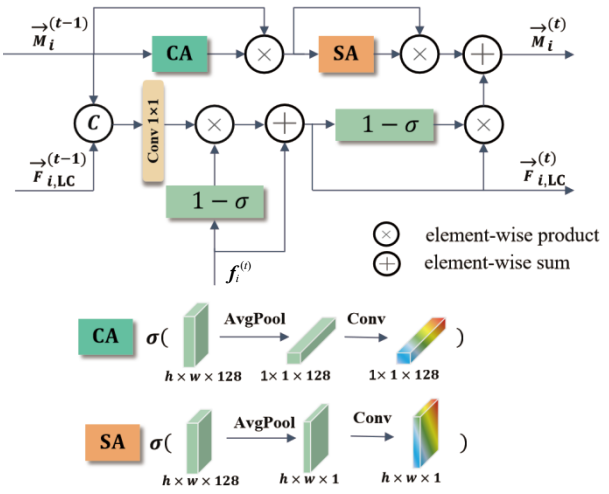


图 3 双向门控上下文集成单元的核心过程

双向门控上下文集成单元先以逐渐扩大的感受野整合（向前方向），然后反向以逐渐缩小的感受野整合 $\boldsymbol{f}_i^{(t)}$（向后方向）。以这种双向的方式，可以记录下不同大小感受野产生的丰富的多尺度上下文信息，在此过程中各个感受野产生的信息都被无差别地对待。每个双向门控上下文集成单元的内部结构主要由两个部分组成：上下文控制器 $\boldsymbol{F}_{i,\mathrm{LC}}^{(t)}$ 和存储器 $\boldsymbol{M}_i^{(t)}$。$\boldsymbol{F}_{i,\mathrm{LC}}^{(t)}$ 负责合成流入信息，$\boldsymbol{M}_i^{(t)}$ 负责更新和恢复来自先前各自字段的信息流，用公式表示为：

$$\boldsymbol{M}_i^{(t)} = \boldsymbol{r}_{F_{i,\mathrm{LC}}}^{(t)} \otimes \boldsymbol{F}_{i,\mathrm{LC}}^{(t)} + \boldsymbol{M}_{i,\mathrm{CA}}^{(t\text{-}1)} \otimes \mathrm{SA}(\boldsymbol{M}_i^{(t)}) \tag{5}$$

$$\boldsymbol{F}_{i,\mathrm{LC}}^{(t)} = \boldsymbol{f}_i^{(t)} + \boldsymbol{r}_{f_i}^{(t)} \otimes \boldsymbol{F}(C(\boldsymbol{M}_i^{(t-1)}, \boldsymbol{F}_{i,\mathrm{LC}}^{(t-1)}); \theta_{i,\mathrm{LC}}^{(t)}) \tag{6}$$

式中：C 表示特征图的拼接运算，$\theta_{i,\mathrm{LC}}^{(t)}$ 表示参数，$\boldsymbol{r}_{F_{i,\mathrm{LC}}}^{(t)}$ 和 $\boldsymbol{r}_{f_i}^{(t)}$ 分别表示自适应选择流信息的两个门，分别定义为：

$$\boldsymbol{r}_{F_{i,\mathrm{LC}}}^{(t)} = 1 - \sigma(\boldsymbol{F}_{i,\mathrm{LC}}^{(t)}) \tag{7}$$

$$\boldsymbol{r}_{f_i}^{(t)} = 1 - \sigma(\boldsymbol{f}_i^{(t)}) \tag{8}$$

另外，对于前向整合的初始值，令 $\boldsymbol{F}_{i,\mathrm{LC}}^{(1)} = \boldsymbol{f}_i^{(1)}$，$\boldsymbol{M}_i^{(1)} = \boldsymbol{r}_{F_{i,\mathrm{LC}}}^{(1)} \otimes \boldsymbol{F}_{i,\mathrm{LC}}^{(1)}$，分别将前向整合和后向整合定义为 $\Phi_{\mathrm{LC}}(\boldsymbol{f}_i^{(1)}, \boldsymbol{f}_i^{(2)}, ..., \boldsymbol{f}_i^{(T)})$ 和 $\Phi_{\mathrm{LC}}(\boldsymbol{f}_i^{(T)}, \boldsymbol{f}_i^{(T-1)}, ..., \boldsymbol{f}_i^{(1)})$。最后，将这两个双向整合过程的输出结合起来，得到一个输入特征图 $\boldsymbol{f}_i$ 的自适应特征表示 $\boldsymbol{F}_{i,\mathrm{LC}}$：

$$\boldsymbol{F}_{i,\mathrm{LC}} = \boldsymbol{F}(C(\boldsymbol{F}_{i,\mathrm{LC}}, \boldsymbol{F}_{i,\mathrm{LC}}); \theta_{i,\mathrm{LC}}) \tag{9}$$

$$\vec{\boldsymbol{F}}_{i,\mathrm{LC}} = \Phi_{\mathrm{LC}}(\boldsymbol{f}_i^{(1)}, \boldsymbol{f}_i^{(2)}, ..., \boldsymbol{f}_i^{(T)}) \tag{10}$$

$$\overleftarrow{\boldsymbol{F}}_{i,\mathrm{LC}} = \Phi_{\mathrm{LC}}(\boldsymbol{f}_i^{(T)}, \boldsymbol{f}_i^{(T-1)}, ..., \boldsymbol{f}_i^{(1)}) \tag{11}$$

式中：$\vec{\boldsymbol{F}}_{i,\mathrm{LC}}$ 和 $\overleftarrow{\boldsymbol{F}}_{i,\mathrm{LC}}$ 表示前向整合和后向整合过程的输出。

4 全局上下文模块

许多基于 CNN 的模型都会遇到受上下文信息短距离依赖性限制的问题，并且由于重复的采样操作，无法同时保持不同尺寸的细节信息及丰富的语义信息。参考全景拍摄技术，通过全维方式观察主体并形成全景图像，以全局视角描绘主体。接收到来自局部上下文模块从不同层学习到的特征表示 $\boldsymbol{F}_{i,\mathrm{LC}}$，$i \in [1, 2, 3, 4]$ 后，通过全局上下文集成单元将其转换为经过全局增强的特征表示。其主要包括两个过程：生成全局特征和跨层上下文集成，如图 4 所示。

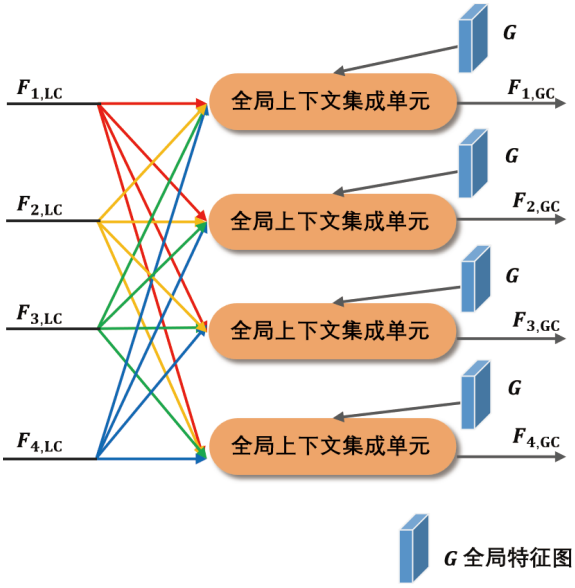


图 4 全局上下文模块原理图

（1）生成全局特征

将局部上下文模块产生的特征图 $\boldsymbol{F}_{i,\mathrm{LC}}$ 进行汇总，生成全局特征图 $\boldsymbol{G}$，用于从不同阶段辅助下一个特征聚合，$\boldsymbol{G}$ 用公式表示为：

$$\boldsymbol{G} = \boldsymbol{F}(C(\boldsymbol{F}_{1,\mathrm{LC}}, \boldsymbol{F}_{2,\mathrm{LC}}, \boldsymbol{F}_{3,\mathrm{LC}}, \boldsymbol{F}_{4,\mathrm{LC}}); \theta_G) \tag{12}$$

式中：$\boldsymbol{F}$ 表示参数为 θ_G 的 1×1 卷积。

（2）跨层上下文集成

在跨层特征整合过程中，采用全局上下文集成单元将互补的上下文信息 $\boldsymbol{F}_{j,\mathrm{LC}}$，$\forall j$，$j \neq i$ 集成到 $\boldsymbol{F}_{i,\mathrm{LC}}$ 中。图 5 显示了全局上下文集成单元的结构。为了保持特征大小的统一性，实践中，在第 i 阶段，首先使用双线性插值将其他层次的特征图转换为与当前特征图相同的大小。全局增强 $\boldsymbol{F}_{i,\mathrm{GC}}$ 可表示为：

$$\begin{aligned}\boldsymbol{F}_{i,\mathrm{GC}} &= \Psi_{\mathrm{GC}}(\boldsymbol{F}_{1,\mathrm{LC}}, \boldsymbol{F}_{2,\mathrm{LC}}, \boldsymbol{F}_{3,\mathrm{LC}}, \boldsymbol{F}_{4,\mathrm{LC}}) \\ &= \boldsymbol{F}_{i,\mathrm{LC}} + r_i \otimes (\sum_{j \neq i} ((\mathrm{CA}(\boldsymbol{G}, \boldsymbol{F}_{j,\mathrm{LC}}) \otimes \boldsymbol{F}_{j,\mathrm{LC}}) \otimes \mathrm{SA}(G, \boldsymbol{F}_{j,\mathrm{LC}})))\end{aligned} \tag{13}$$

式中：$\mathrm{CA}(\boldsymbol{G}, \boldsymbol{F}_{j,\mathrm{LC}})$ 将全局信息导入局部特征，根据其对全局上下文的贡献为每个像素生成一个权重；$\mathrm{SA}(\boldsymbol{G}, \boldsymbol{F}_{j,\mathrm{LC}})$ 是空间注意力单元。将输入 CA 的特征表示为 $\boldsymbol{F}_{j,G}$，输入 SA 的特征表示为 $\boldsymbol{F}_{j,\mathrm{CA}}$：

$$\boldsymbol{F}_{j,G} = \boldsymbol{F}(C(\boldsymbol{G}, \boldsymbol{F}_{j,\mathrm{LC}}); \theta_{i,G}) \tag{14}$$

$$\mathrm{CA}(\boldsymbol{G}, \boldsymbol{F}_{j,\mathrm{LC}}) = \sigma(\boldsymbol{F}(\mathrm{AvgPool}(\boldsymbol{F}_{j,G}); \theta_{i,\mathrm{CA}})) \tag{15}$$

$$\mathrm{SA}(\boldsymbol{G}, \boldsymbol{F}_{j,\mathrm{LC}}) = \sigma(\boldsymbol{F}(\mathrm{AvgPool}(\boldsymbol{F}_{j,\mathrm{CA}}); \theta_{i,\mathrm{SA}})) \tag{16}$$

r_i 是一个用于过滤信息流的门，可以表示为：

$$r_i = 1-\sigma(F_{i,\mathrm{LC}}) \tag{17}$$

利用全局上下文模块，实现了跨层上下文信息的互补，细化了 U-Net 编码器阶段的深浅层特征图。同时，由于全局视角的机制，可以减轻类间的高相似性和类内的高差异性的影响，并解决多尺度问题。

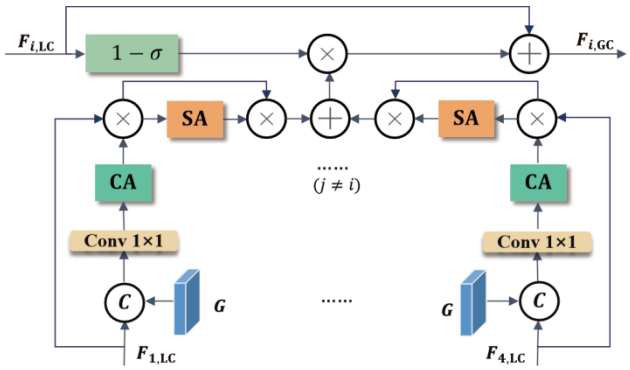


图 5 全局上下文集成单元结构图

5 SCR 模块

在传统的 U-Net 网络中，编码过程是直接转换成有像素级监督的二值分割图像。然而，这种方式只考虑了每个单独的像素与其对应的 ground-truth 值之间的关系，缺乏区域间的关联性。由于感染区域形状尺寸的多样性，因而考虑在优化过程中引入更多的结构信息，以进一步提高分割精度。尤其是在复杂的边界区域，通过 SCR 模块来进一步整合像素之间的结构关系，使其边缘分割更加精细。

对于二值图像，每个像素可以分为两组：前景（即感染区域）和背景 I_{bg}，利用距离变换技术（DT）对像素分类进行校正。DT 算法是描述二值图像形态学运算的经典图像处理算法。DT 映射是反映区域间相对结构关系的矩阵。具体来说，如果像素属于感染区域，那么 DT 映射中对应的输出值是指其与背景中距离最近的像素 k 之间的距离，而不是一个只描述其存在的离散数字。否则，如果像素 l 属于背景，则将其值设为 0。另外，感染区域中的像素都属于前景，但是这些像素与位于背景中与其最接近的像素之间的距离是完全不同的。因此，这种结构信息可以很好地用于校正预测结果。在形式上，DT 可以表示为：

$$\mathrm{dt}(l) = \begin{cases} \min_{k \in I_{\mathrm{bg}}} \{\mathrm{dist}(l,k)\}, l \in I_{\mathrm{fg}} \\ 0, l \in I_{\mathrm{bg}} \end{cases} \tag{18}$$

式中：dt(l) 表示任意像素 l 对应的 DT 值，dist(l, k) 表示像素 l 与 k 之间的加权欧几里得度量。

$$\mathrm{dist}(l,k) = s_l \sqrt{(l_x - k_x)^2 + (l_y - k_y)^2} \tag{19}$$

式中：l_x 和 l_y 表示像素的空间位置坐标；s_l 表示像素被预测为前景的概率密度，反映了像素 l 属于前景的置信度。通过上式可以得到任意二值图像的 DT 值。

最后，期望预测结果图 P 的 DT 映射 $I_{D(P)}$ 能与 Ground-truth G_S 的 $I_{D(G_S)}$ 保持相似的结构，通过（kullback-leibler, KL）散度来计算其之间的结构关系：

$$D_{\mathrm{SCR}} = \mathrm{KL}(I_{D(P)} \| I_{D(G_S)}) \tag{20}$$

6 EA 模块

边缘信息可以在指导分割特征的过程中提供有效的约束关系，因此，考虑到低层特征（例如：本文网络模型中的 f_2）保留了一些富含边缘的信息，将具有合适分辨率的低层特征 f_2 输入 EA（edge attention）模块，从而显式地学习关注于边缘的特征表示。具体地说，f_2 特征通过一个卷积层得到边缘映射图 B_e。然后，使用标准的二值交叉熵损失函数来度量生成的边缘映射图 B_e 与由真值图（ground-truth）导出的边缘真值图 G_e 的差异性：

$$L_{\mathrm{edge}} = -\sum_{x=1}^{w}\sum_{y=1}^{h}[G_e \log(B_e) + (1-G_e)\log(1-B_e)] \tag{21}$$

式中：(x, y) 代表所预测的边缘图 B_e 和边缘真值图 G_e 中每一个像素点的坐标，w 和 h 分别代表对应的特征映射图的宽和高。

7 损失函数

通过引入 SCR 模块及 EA 模块，同时整合了 DT 映射的结构关系、损失函数 L_{edge} 和图像的标准二值交叉熵损失 L_{BCE} 来对最终的预测图进行约束。总损失函数为：

$$L = \alpha L_{\mathrm{edge}} + \beta L_{\mathrm{BCE}} + D_{\mathrm{SCR}} \tag{22}$$

式中：标准二值交叉熵损失 L_{BCE} 表示为：

$$L_{\mathrm{BCE}} = -\sum_{j}^{N}(G_j \log P_j + (1-G_j)\log(1-P_j)) \tag{23}$$

式中：N 表示像素的总个数，P_j 和 G_j 分别表示预测图和真实标记 Ground-truth 的第 j 个像素点。

8 实验测试

本文采用 LiTS2017 肝脏肿瘤分割挑战赛数据集（Liver Tumor Segmentation Challenge 2017）进行实验验证。本数据集共包含 201 例腹部增强 CT 扫描数据，其中训练集 131 例、测试集 70 例。数据集中 CT 图像分辨率统一为 512 px×512 px，层厚范围为 0.7~5.0 mm，具有较强的临床代表性。本实验从 LiTS 训练集的 131 例 CT 序列中提取所有包含肿瘤区域的二维轴向切片，共得到 3 216 张有效图像。实验随机选取 80% 的图像作为训练集，剩余 20% 作为测试集。实验基于深度学习框架 PyTorch 实现，硬件环境为 NVIDIA GeForce RTX 3090 显卡（显存 24 GB），采用 Adam 优化器，Adam 算法可以在训练时自适应地调节学习率，且有更快的收敛速度，初始学习率设置为 0.000 1。通过实验验证当训练迭代次数超过 100 时，验证集的损失值趋于稳定。因此，本文实验中的训练迭代次数设置为 100。

8.1 定量分析结果

为充分验证本文方法的性能优势，选取 5 种医学图像分割领域的经典主流算法作为对比，所有方法均采用公开代码实现，实验数据预处理与超参数配置保持一致。表 1 为各算法在 LiTS 数据集上肝脏肿瘤分割的定量评价结果。从表中可以看出，本文提出的 ContextSE-UNet 在所有评价指标上均优于对比算法，验证了方法的有效性。从数值对比来看，本文方法的核心指标 Dice 系数与 mIoU 较 U-Net 分别提升了 4.4% 和 8.5%，性能提升显著；与当前表现优异的 TransUNet 相比，Dice 系数提升 1.9%，mIoU 提升 1.5%。其中灵敏度提升最为明显，较 TransUNet 提升 2.6%，说明本文方法能够更完整地检出肿瘤区域，有效降低漏诊率。

表 1 LiTS 数据集上肝脏肿瘤分割定量结果

单位：%

网络模型	mIoU	Dice	ACC	Sen	Spe
U-Net	70.8	66.3	92.2	54.2	86.8
U-Net++	72.9	68.4	95.4	55.9	87.2
Attention-UNet	75.4	69.2	97.5	56.1	89.6
ResUNet	76.4	70.8	98.3	58.5	90.8
TransUNet	77.8	71.8	99.5	60.3	91.4
本文方法	**79.3**	**73.7**	**99.8**	**63.5**	**92.4**

性能提升的核心原因在于：局部与全局上下文模块的结合，既在同层捕获了多尺度病灶特征，又通过跨层级信息互补增强了全局语义理解，有效缓解了肿瘤类内异质性与尺度差异带来的挑战；SCR 模块的结构约束与 EA 模块的边界增强，则进一步修正了边缘分割误差，提升了边界区域的分割精度。

8.2 定性分析结果

图 6 为本文方法与各对比算法的肝脏肿瘤分割结果可视化对比。从图中可以观察到：一是对于体积较小的微小结节病灶，U-Net、U-Net++ 等基础模型易出现漏分割或分割不完整的现象，假阴性较高；而本文方法能够更完整地检出小病灶，对小尺度目标的感知能力更强；二是对于边界模糊、呈浸润性生长的肿瘤，对比算法普遍存在边界过度平滑、病灶形态失真的问题，易将正常肝组织误判为肿瘤或将肿瘤区域遗漏；而本文方法的分割边界更贴合真实病灶轮廓，边缘细节更丰富；三是对于多发肿瘤场景，基础模型易出现邻近病灶粘连、小病灶被忽略的问题；本文方法能够更清晰地区分相邻病灶，保持病灶的独立形态。整体而言，本文方法在多尺度病灶、弱边界病灶、多发病灶等困难场景下均表现出更优的分割效果，与定量分析的结论一致。

9 结语

本文提出 ContextSE-UNet 用于 CT 图像中肝脏肿瘤的精准分割。针对肝脏肿瘤分割中类间相似性高、类内异质性强、边界模糊、多尺度差异显著等核心难题，本文从上下文信息融合与边界结构增强两个维度对 U-Net 进行改进。为捕获长程上下文依赖、提升多尺度病灶建模能力，本文构建了局部上下文模块与全局上下文模块：局部上下文模块通过变焦式多感受野特征提取与双向门控聚合，在同层内整合多尺度局部语义信息；全局上下文模块通过跨层特征互补机制，选择性融合不同层级的特征信息，实现全局语义增强。为提升边界分割精度，通过 SCR 模块引入区域级结构约束，建模前景与背景的相对位置关系；同时通过 EA 模块显式学习边缘特征，优化边界损失，强化病灶边缘的表达能力。在公开 LiTS 数据集上的大量实验结果表明，本文提出的 ContextSE-UNet 在多项评价指标上均优于主流对比算法，能够有效提升肝脏肿瘤的分割精度，尤其在小病灶、模糊边界等困难场景下优势明显，具备一定的临床应用潜力。

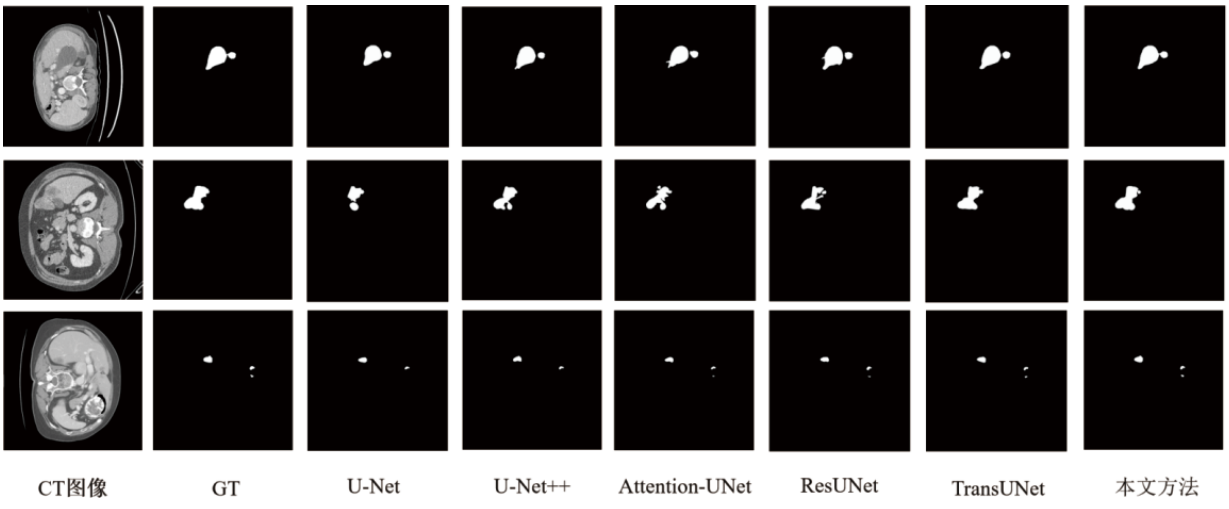


图 6 分割结果对比

（下转第 33 页）

[9] Qi Wen, Ovur S E, Li Zhijun, et al. Multi-sensor guided hand gesture recognition for a teleoperated robot using a recurrent neural network[J]. IEEE Robotics and Automation Letters, 2021, 6(3): 6039-6045.

[10] 薛阳 , 李清 , 卢秋红 , 等 . 基于改进 YOLOv5-Conv 手势识别研究 [J]. 信息技术与信息化 , 2024(7): 207-211.

[11] Kapitanov A, Makhliarchuk A, Kvanchiani K. HaGRID–HAnd gesture recognition image dataset[PP/OL]. (2022-06-14) [2026-07-29]. https://arxiv.org/pdf/2206.08219v2.

[12] Materzynska J, Berger G, Bax I,et al. The jester dataset: a large-scale video dataset of human gestures [C/OL]// 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Piscataway: IEEE, 2019[2026-07-29]. https://ieeexplore.ieee.org/document/9022297.DOI: 10.1109/ICCVW.2019.00349.

[13] 张硕 . 基于 KNN 算法的空间手势识别研究与应用 [D]. 长春 : 吉林大学 ,2017.

[14] 陈嘉伟 , 韩晶 , 郝瑞玲 , 等 . 基于改进 KNN 算法的动态手势识别研究 [J]. 中北大学学报 (自然科学版), 2020, 41 (3): 232-237.

[15] Zhou Kai, WanYi, Huang Hailong. Research on recognition and application of hand gesture based on skin color and SVM[J]. Journal of Computational Methods in Sciences and Engineering, 2019, 20(1): 269-278.

[16] 王庆鑫 . 基于轻量级 CNN 的手势识别算法研究 [D]. 哈尔滨 : 哈尔滨理工大学 ,2023.

[17] 王路遥 . 基于改进 YOLOv5s 的静态 / 动态手势识别算法探究 [J]. 信息系统工程 ,2024(12):138-141.

[18] 张常斌 , 孙连山 , 唐景琰 . 基于改进 YOLOv5 的轻量级手势识别算法 [J]. 智能计算机与应用 ,2025,15(6):184-189.

【作者简介】

刘姝怡（2004—），女，江苏苏州人，本科，研究方向：数字媒体技术，email:2133993942@qq.com。

徐梦溪（1983—），通信作者（email: xumx@njit.edu.cn），女，江苏南京人，博士，教授，研究方向：图像处理、智能信息处理、多媒体系统与应用软件。

刘姝悦（2004—），女，江苏苏州人，本科，研究方向：数字媒体技术，email:3137631284@qq.com。

刘梓莹（2005—），女，广东广州人，本科，研究方向：数字媒体技术，email:3375788023@qq.com。

（收稿日期：2026-06-24 修回日期：2026-08-03）

（上接第 27 页）

参考文献：

[1] 张欢 , 刘静 , 冯毅博 , 等 . U-Net 及其在肝脏和肝脏肿瘤分割中的应用综述 [J]. 计算机工程与应用 , 2022, 58(2): 41-52.

[2]Khan S U R, Lee H. Deep learning-based liver tumor segmentation: a scoping review [J]. Convergence Hepatology, 2026, 2 (1): 29-55.

[3]Qi Yingyi, Xiong Wei, Leow K, et al. Semi-automatic segmentation of liver tumors from CT scans using bayesian rulebased 3D region growing[J/OL].The MIDAS Journal, 2008[2026-02-04].https://doi.org/10.54294/rofibw.

[4]Seo H, Huang C, Bassenne M, et al. Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images [J]. IEEE Transactions on Medical Imaging, 2020, 39 (5): 1316-1325.

[5] 钟思华 , 郭兴明 , 郑伊能 . 改进 U-Net 网络的肺结节分割方法 [J]. 计算机工程与应用， 2020, 56 (17): 203-209.

[6]Li Wenkun, Jia Maoning, Yang Chen, et al. SPA-UNet: A liver tumor segmentation network based on fused multi-scale features [J]. Open Life Sciences, 2023,18 (1): 1245-1256.

[7]Ruan Jiacheng, Xie Mingye, Gao Jingsheng, et al. EGE-UNet: an efficient group enhanced UNet for skin lesion segmentation[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2023. Cham: Springer, 2023: 481-490.

[8]Wang Fang, Cheng Xueli, Luo Ningbin, et al.Attention-guided context asymmetric fusion networks for the liver tumor segmentation of computed tomography images[J]. Quantitative Imaging in Medicine and Surgery, 2024, 14(7): 4825-4841.

【作者简介】

肖露（1997—），女，河南南阳人，硕士，助教，研究方向：计算机视觉、图像分割。

王亚茹（1994—），女，河南信阳人，硕士，助教，研究方向：机器学习、大数据分析。

刘锦博（1999—），男，河南郑州人，硕士，助教，研究方向：计算机视觉、人工智能。

邓锴（1997—），男，河南南阳人，硕士，助教，研究方向：算力网络、未来网络。

黄星帆（2005—），女，河南商丘人，本科在读，研究方向：计算机视觉。

（收稿日期：2026-07-14 修回日期：2026-08-12）