



基于物理显著性特征感知的医学小病灶分割网络

任向阳 焦博 杨震 贾育衡 郭伟峰 梁静

Medical Small Lesion Segmentation Network Based on Physical Saliency Feature Perception

REN Xiang-Yang, JIAO Bo-Yang, YANG Zhen, JIA Yu-Heng, GUO Wei-Feng, LIANG Jing

在线阅读 View online: <https://doi.org/10.16383/j.aas.c250623>

您可能感兴趣的其他文章

基于循环显著性校准网络的胰腺分割方法

Pancreas Segmentation Based on Recurrent Saliency Calibration Network

自动化学报. 2022, 48(11): 2703–2717 <https://doi.org/10.16383/j.aas.c210865>

基于语义引导特征聚合的显著性目标检测网络

Semantic Guided Feature Aggregation Network for Salient Object Detection

自动化学报. 2023, 49(11): 2386–2395 <https://doi.org/10.16383/j.aas.c210425>

一种迭代边界优化的医学图像小样本分割网络

A Few-shot Medical Image Segmentation Network With Iterative Boundary Refinement

自动化学报. 2024, 50(10): 1988–2001 <https://doi.org/10.16383/j.aas.c220994>

基于双高斯分布混合的可解释自适应鲁棒神经网络建模方法

An Interpretable and Adaptive Robust Neural Network Modeling Method Based on Dual Gaussian Mixture Distribution

自动化学报. 2026, 52(3): 463–480 <https://doi.org/10.16383/j.aas.c250602>

基于多模型融合的肺部CT新冠肺炎病灶区域自动分割

Automatic Segmentation of Covid-19 Infected Regions in Chest CT Images Based on 2D/3D Model Ensembling

自动化学报. 2023, 49(2): 317–328 <https://doi.org/10.16383/j.aas.c210400>

基于误差回传机制的多尺度去雾网络

Multi-scale Dehazing Network Based on Error-backward Mechanism

自动化学报. 2023, 49(9): 1857–1867 <https://doi.org/10.16383/j.aas.c210264>

基于物理显著性特征感知的医学小病灶分割网络

任向阳¹ 焦博旸² 杨震³ 贾育衡⁴ 郭伟峰⁵ 梁静⁵

摘要 精准分割早期小病灶对疾病诊疗至关重要,但现有方法面对特征稀疏、易受背景干扰的小病灶时性能显著下降。受计算机断层扫描成像中 X 射线衰减物理原理启发,发现小病灶在影像中呈现出中心亮、边缘渐弱的强度分布,其轮廓与二维高斯分布高度吻合。为此,提出一种辐射衰减原理引导的显著性感知分割网络 (RAP-Net)。RAP-Net 将射线衰减导致的二维高斯分布作为相关滤波卷积核融入深度学习架构,并设计专用于小病灶分割的多尺度感知特征网络。该网络的核心显著特征感知提取模块利用多尺度高斯空洞卷积建模衰减特性,实现稀疏特征的深度挖掘与背景抑制。实验表明, RAP-Net 在小肾结石与小肝脏钙化灶分割任务中的 Dice 系数和 IoU 提升至少 18.05% 和 19.55%, 显著超越现有主流方法。

关键词 医学小病灶; 辐射衰减; 多尺度感知特征网络; 高斯分布

引用格式 任向阳, 焦博旸, 杨震, 贾育衡, 郭伟峰, 梁静. 基于物理显著性特征感知的医学小病灶分割网络. 自动化学报, 2026, 52(7): 1387-1400

DOI 10.16383/j.aas.c250623 **CSTR** 32138.14.j.aas.c250623

Medical Small Lesion Segmentation Network Based on Physical Saliency Feature Perception

REN Xiang-Yang¹ JIAO Bo-Yang² YANG Zhen³ JIA Yu-Heng⁴ GUO Wei-Feng⁵ LIANG Jing⁵

Abstract Accurate segmentation of early-stage small lesions is crucial for disease diagnosis and treatment. However, existing methods suffer from a significant drop in performance when dealing with small lesions that are feature-sparse and susceptible to background interference. Inspired by the physical principles of X-ray attenuation in computed tomography imaging, we observe that small lesions exhibit an intensity distribution characterized by a bright center and gradually fading edges, with their contours closely matching a two-dimensional Gaussian distribution. To address this, we propose a radiation attenuation principle-guided saliency perception segmentation network (RAP-Net). RAP-Net integrates the two-dimensional Gaussian distribution induced by radiation attenuation into a deep learning architecture as a correlation-filtering convolution kernel, and designs a multi-scale perceptual feature network specifically for small lesions segmentation. Its core perceptual saliency feature extraction module employs multi-scale Gaussian dilated convolutions to model attenuation characteristics, thereby enabling deep mining of sparse features and effective background suppression. Experiments demonstrate that RAP-Net achieves at least 18.05% and 19.55% improvements in Dice coefficient and IoU, respectively, for the segmentation tasks of small kidney stones and small liver calcifications, significantly outperforming existing mainstream methods.

Keywords medical small lesion; radiation attenuation; multi-scale perceptual feature network; Gaussian distribution

Citation Ren Xiang-Yang, Jiao Bo-Yang, Yang Zhen, Jia Yu-Heng, Guo Wei-Feng, Liang Jing. Medical small lesion segmentation network based on physical saliency feature perception. *Acta Automatica Sinica*, 2026, 52(7): 1387-1400

收稿日期 2025-11-11 录用日期 2026-03-19

Manuscript received November 11, 2025; accepted March 19, 2026

国家自然科学基金 (62206251, 62476253, 62576094, U23A20340, 625B2066), 河南省自然科学基金 (262300421218, 262300421635), 河南省科技研发联合基金 (产业类) 重大项目 (245101610001), 河南省医学科技人才海外研修项目 (HNMOT2025063) 资助

Supported by National Natural Science Foundation of China (62206251, 62476253, 62576094, U23A20340, 625B2066), Natural Science Foundation of Henan (262300421218, 262300421635), Henan Province Science and Technology Research Joint Fund (Industrial Category) Major Project (245101610001), and Henan Medical Researcher Overseas Training Program (HNMOT2025063)

本文责任编辑 罗彪

Recommended by Associate Editor LUO Biao

1. 郑州大学第一附属医院医学 3D 打印中心 郑州 450003 2. 郑州大学第一临床医学院 郑州 450003 3. 湖南大学人工智能与机器人学院 长沙 410012 4. 东南大学计算机科学与工程学院 南京 211189 5. 郑州大学电气与信息工程学院 郑州 450001

医学图像中小病灶的精准分割对疾病早期诊断、治疗规划与疗效评估具有重要意义^[1-2]。近年来,深度学习技术,尤其是卷积神经网络 (convolutional neural networks, CNN), 在医学图像分割领域取得显著进展。CNN 能够逐层提取从纹理到语义的多尺度特征,有效抑制背景噪声,在肝脏^[3-6]、肾脏^[7-9]、血管^[10-12]及左心房^[13-14]等特定解剖结构的分割任

1. Medical 3D Printing Center, the First Affiliated Hospital of Zhengzhou University, Zhengzhou 450003 2. The First Clinical Medical College, Zhengzhou University, Zhengzhou 450003 3. School of Artificial Intelligence and Robotics, Hunan University, Changsha 410012 4. School of Computer Science and Engineering, Southeast University, Nanjing 211189 5. School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001

务中表现优异. 代表性工作包括肺肿瘤特征融合^[15]、视网膜血管多尺度监督^[16-17]、肺肿瘤 PET-CT 模态特异性网络^[18]、视网膜血管多级注意力机制^[19]、冠状动脉数据增强^[20]、视网膜血管数据集构建^[21] 以及黄斑中央凹无血管区与视网膜血管联合分割策略^[22] 等. 然而, 小病灶分割在医学图像中仍面临诸多挑战. 相较于大尺寸病灶, 小病灶的特征更为稀疏, 极易被复杂背景淹没, 导致检测精度下降. 值得注意的是, 小病灶的准确分割在临床实践中往往比大病灶的分割更具决定性意义, 例如通过精准识别多器官小病灶, 为疾病早期预防、精准治疗与临床决策提供关键技术支持, 助力器官保护与医疗资源优化 (见图 1). 以小肝脏钙化灶分割为例, 精确勾画其位置、大小、形态及数量, 可为肝脏相关疾病的诊断提供关键依据. 类似地, 小肾结石的分割结果直接影响治疗策略的选择: 较小结石可采取保守观察, 而较大结石则需积极干预 (如体外碎石). 此外, 小病灶分割还可用于评估组织受累情况及并发症风险, 为保护肾功能提供及时预警.

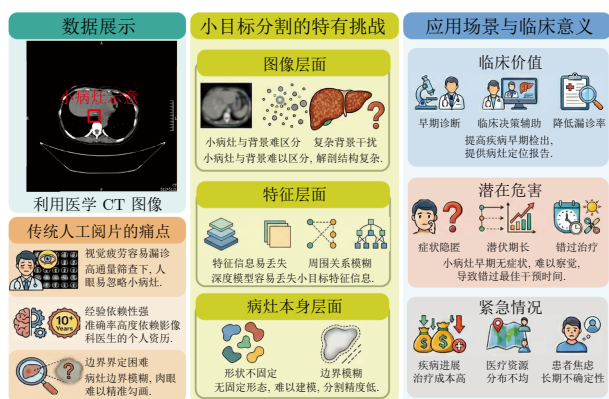


图 1 本文提出的小病灶分割任务及研究动机

Fig.1 The small lesion segmentation task proposed in this paper and its research motivation

针对上述问题, 本文提出辐射衰减原理引导的显著性感知分割网络 (radiation attenuation principle-guided saliency perception segmentation network, RAP-Net), 一种受计算机断层扫描 (computed tomography, CT) 成像 X 射线衰减原理^[23] 启发的小病灶分割网络. 该方法将辐射衰减特性与相关滤波理论相结合, 融入 CNN 架构, 显著提升小病灶的分割性能. 在 CT 成像中, X 射线穿过生物组织时会发生强度衰减, 其衰减程度由组织的密度与厚度共同决定. 从成像原理来看, 组织对 X 射线的衰减能力越强, 探测器接收到的信号就越弱, 但系统通过重建算法所计算出的对应像素 CT 值越高, 最终在图像中表现为更高的灰度值. 因此, 在重

建后的 CT 图像中, 病灶中心具有更高的 CT 值与灰度值, 向外逐渐递减, 形成中心亮、边缘渐弱的强度分布, 该分布形态与二维高斯分布高度吻合 (见图 2). 基于这一原理, 为实现对小病灶显著性特征的精准感知, RAP-Net 引入多尺度感知特征网络 (multi-scale perceptual feature network, MSPF-Net), 通过引入相关滤波机制强化显著特征表达, 能够有效应对小病灶特征稀疏的挑战. MSPF-Net 中的相关滤波器同时具备背景抑制能力, 成功将 CT 成像物理机制嵌入深度学习架构中. 实验结果表明, RAP-Net 在医学小病灶图像数据集中取得当前最优性能, 显著提升小病灶的检测精度, 为疾病的早期干预提供可靠支持. 本文将 CT 成像的原理嵌入深度学习网络中, 利用物理原理约束引导网络在低对比度和模糊边界下进行更精准的特征复原, 并有效解决了通用分割网络在小病灶任务中参数冗余且细节丢失的痛点问题. 本文的主要贡献如下:

1) 提出一种基于 X 射线衰减原理的小病灶分割网络 RAP-Net, 有效解决了医学图像中小病灶分割的难题. 该方法融合 CT 成像原理, 提升分割方法的鲁棒性, 为疾病早期发现与治疗提供技术支撑.

2) 提出一种专为医学图像小病灶特征提取而构建的主干网络 MSPF-Net. 该网络通过强化显著特征提取与抑制背景噪声, 丰富特征表达, 有效解决了强背景干扰下的稀疏特征提取难题, 提升小病灶分割性能.

3) 在本文构建的医学小病灶数据集 MSL 上评估所提网络的分割性能. 本文提出的方法与现有先进的医学图像处理相比, 在多个分割任务上的 Dice 系数提升至少 18.05%, IoU 提升至少 19.55%, 表现出优异的小病灶分割性能.

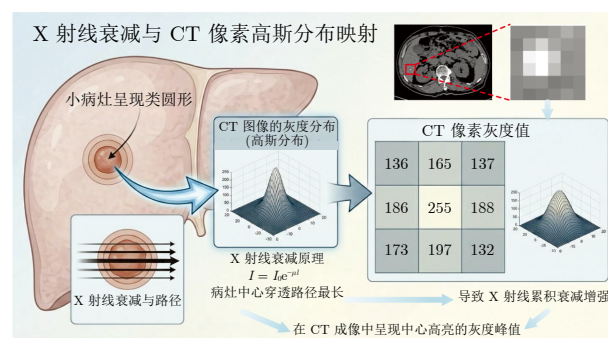


图 2 CT 成像中的 X 射线衰减原理示意图

Fig.2 Schematic diagram of X-ray attenuation principles in CT imaging

1 相关工作

医学图像分割的发展历程可以概括为从全卷积

网络的结构优化到注意力机制的引入,再到近期基于状态空间模型 (state space model, SSM) 与基础大模型的范式转变。

1.1 CNN 架构的演进与多尺度特征增强

全卷积网络^[24]和 U-Net^[25]的提出奠定了医学图像分割的范式。为解决 U-Net 在下采样过程中对小病灶空间信息丢失的问题,研究者在编码器—解码器结构上进行广泛的改进。UNet++^[26]通过嵌套的密集跳跃连接减少语义鸿沟; R2U-Net^[27]利用循环残差单元增强深层特征的表达能力; DoubleU-Net^[28]则通过堆叠两个 U-Net 结构来捕捉更丰富的语义信息。针对特定解剖结构的分割难点, Xiao 等^[29]提出加权 Res-UNet 以优化视网膜血管的连通性。Isensee 等^[30]提出的 nnU-Net 代表另一条优化路径,它不依赖复杂的架构堆叠,而是通过自适应的数据预处理和网络配置策略,建立医学图像分割的性能标杆。这些基于 CNN 的改进工作^[24-30]有效提升特征提取的鲁棒性,但受限于卷积核的局部感受野,在处理具有长距离依赖的复杂背景时仍存在局限。

1.2 注意力机制与混合建模策略

为突破局部感受野的限制,注意力机制被广泛引入以强化关键区域的特征响应。Attention U-Net^[31]利用门控机制抑制背景噪声; PraNet^[32]通过并行反向注意力模块显著提升了息肉等不规则目标的边界分割精度; FANet^[33]则设计反馈注意力网络以实现特征的迭代细化。

随着 Transformer 在视觉领域的兴起,混合架构成为主流。FTransCNN^[34]创新性地结合模糊逻辑与 Transformer,以处理医学图像中的不确定性; UCTransNet^[35]和 TransUNet^[35]通过在 CNN 瓶颈处引入 Transformer 模块来实现全局上下文建模; Swin-UNet^[36]基于移动窗口机制实现纯 Transformer 的医学分割; STC-UNet^[37]提出一种结合选择性核 (selective kernel) 与 Vision Transformer 的混合架构,通过在不同网络层级进行针对性的特征增强,显著提升了模型对复杂病灶的解析能力。尽管这些方法^[31-37]极大地增强了全局建模能力,但 Transformer 二次复杂度的计算开销限制了其在高分辨率小病灶任务中的应用。

1.3 状态空间模型、基础模型与小目标挑战

近年来,以 Mamba 为代表的 SSM 因其线性复杂度与强大的序列建模能力成为新热点。Ruan 等^[38]提出的 VM-UNet 首次构建了基于 Visual Mamba 的非对称 U 形网络,在保持低显存占用的同时实现

优异的分割性能; Chen 等^[39]进一步提出基于 Zigzag 扫描机制的 Zig-RiR,利用 RWKV 架构高效处理医学图像的多向特征。与此同时,分割大模型 (segment anything model, SAM) 的基础模型研究也迅速发展。MedSAM^[40]和最新的 MedCLIP-SAMv2^[2]通过适配器微调和多模态对齐,展示了强大的泛化能力。然而,针对肾结石、小钙化等像素占比极低的目标,通用大模型往往因下采样率过高而丢失细节。现有的小目标专用方法如 UoloNet^[1]或 RECISTSurv^[41]多侧重于多任务联合学习或临床生存分析,较少利用医学成像原理。本文提出的 RAP-Net 正是为了填补这一空白,通过引入 X 射线衰减原理^[23],在特征提取源头增强小病灶的显著性。

以 U-Net^[25]及其变体^[26-30]为代表的深度学习方法,虽在诸多医学图像分割任务中表现突出,但其架构多针对尺寸较大、特征明显的病灶或器官进行优化。一旦应用于小病灶,深层网络中连续的下采样操作容易使本就微弱的病灶特征在传递过程中逐渐丢失;同时,通用的卷积核难以在复杂背景中聚焦于病灶的显著性特征,导致模型在极端正负样本不平衡的情况下性能下降明显。近年来,虽有研究尝试通过注意力机制、多尺度融合等策略提升小目标检测能力,但大多仍集中于模型内部的结构优化,未能从根本上解决小病灶因其尺度有限、特征稀疏及复杂背景干扰所带来的分割难题^[42-43]。

2 方法

尽管基于 CNN 的医学分割方法在提取大病灶丰富特征方面表现卓越^[44-45],但是当其处理小病灶时,由于特征极度稀疏且背景干扰强烈,其性能显著下降。这些挑战不仅增加了精准定位与分割的难度,也制约了网络对有效特征的表达与学习能力。受 X 射线衰减原理及 CT 图像的典型灰度模式启发,本文基于 Beer-Lambert 定律对病灶的灰度分布特性进行建模。X 射线在介质中的衰减可表示为:

$$I = I_0 e^{-\mu l} \quad (1)$$

其中, I_0 为初始射线强度; I 为衰减后的射线强度; μ 为线性衰减系数; l 为射线穿过目标的路径长度。由于小病灶通常呈球状或类球状,路径长度 l 从中心向边缘逐渐减小,从而导致观测强度 I 呈现出符合二维高斯分布的衰减特性 (如图 2 所示)。

基于这一原理,本文针对医学小病灶分割任务设计全新架构 RAP-Net。如图 3 所示,图中 (a) 为 RAP-Net 的 MSPF-Net 连接图, (b) 为 RAP-Net 的特征映射网络架构, (c) 为用于提取小病灶显著特征的特征映射网络模块, (d) 为 MSPF-Net

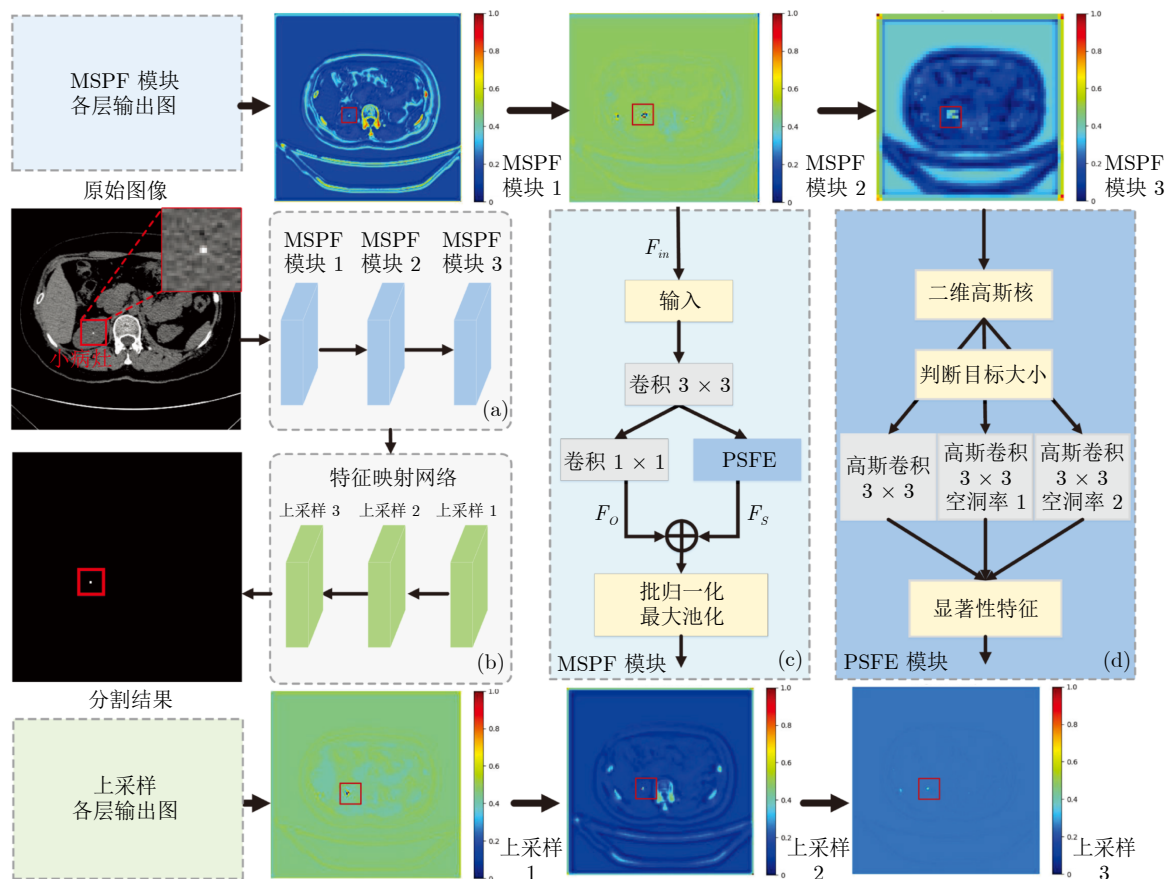


图 3 本文所提出的 RAP-Net 架构示意

Fig.3 The architecture diagram of the proposed RAP-Net

的显著特征感知提取 (perceptual saliency feature extraction, PSFE) 模块. 该网络引入 MSPF-Net 作为主干, 其核心思想源于 X 射线在 CT 成像中的辐射衰减原理^[23]. MSPF-Net 通过提出的 PSFE 模块显著增强了小病灶的特征表达能力. PSFE 模块融合高斯空洞卷积, 并借鉴传统图像处理中的相关滤波理论, 可在提取稀疏小病灶显著特征的同时有效抑制背景噪声. 通过将小病灶的 CT 图像显著性特征直接嵌入特征提取过程, RAP-Net 极大提升了网络学习判别性特征的能力, 实现对挑战性医学小病灶的精准分割. 后续内容将系统阐述高斯空洞卷积 (Gaussian dilated convolution)、MSPF-Net 整体架构的设计细节和模型训练设置.

2.1 高斯空洞卷积

医学 CT 成像利用 X 射线衰减原理反映组织密度差异, 即病灶厚度越大, 对射线的吸收越强, 对应像素值越高, 从而在空间上形成中心灰度峰值、向外逐渐衰减的分布, 可精确近似为二维高斯分布^[23]. 然而, 小病灶尺寸有限、特征稀疏, 给现有分

割方法带来严峻挑战. 受衰减原理及其典型灰度模式启发, 本文提出高斯空洞卷积, 用于小病灶显著特征的提取. 如图 4 所示, 二维高斯卷积在小病灶

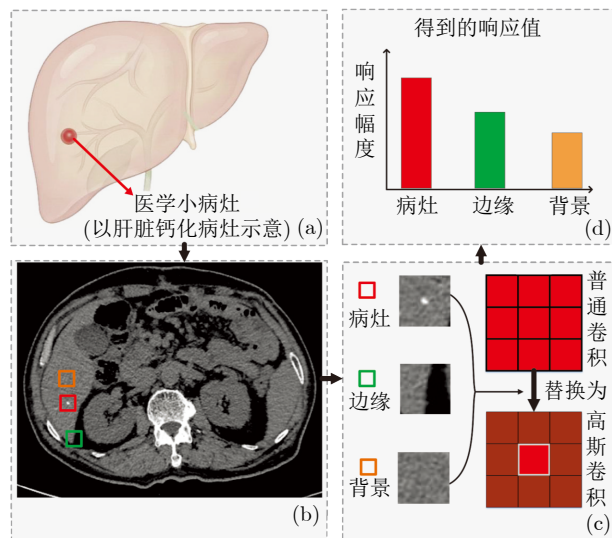


图 4 显著特征提取原理图

Fig.4 Schematic diagram of saliency feature extraction

区域产生的响应值远高于边缘或平滑背景. 其根本原因在于, 小病灶的像素灰度值呈现出由中心向外周径向递减的分布特性, 与高斯核形态高度匹配, 边缘仅沿单一方向快速衰减, 背景则几乎无灰度变化. 依据相关滤波理论, 该过程可同步实现病灶增强与背景抑制. 图 4 中 (a) ~ (d) 以肝脏 CT 为例直观展示了这一效果: 经二维高斯核处理后, 病灶区域表现为最大响应值, 而边缘和背景区域的响应次之. 图 4 中 (d) 的纵坐标表示响应幅度, 在数学上定义为二维高斯卷积核权重矩阵 K 与局部图像像素矩阵 X 的内积运算结果, 其计算公式可表示为:

$$R(i, j) = \sum_{m=-k}^k \sum_{n=-k}^k K(m, n) \times X(i+m, j+n) \quad (2)$$

其中, (i, j) 代表像素坐标; (m, n) 代表卷积核偏移量. 从物理意义上讲, 该响应值量化了特征提取器与解剖结构之间的相似度. 病灶区域因其灰度分布与高斯核 K 高度契合, 从而产生峰值响应, 而边缘和背景区域因形态不匹配, 其响应幅度被显著抑制.

然而, 在实际医学成像中, 小病灶的尺寸存在差异. 为解决这一问题, 本文提出一种高斯空洞卷积来提取不同尺寸小病灶的显著特征. 具体而言, 高斯卷积核的大小被设定为固定的较小尺度 (例如 3×3), 高斯卷积核作为一个通用的空间滤波器, 旨在有效抑制背景噪声的同时, 突出小病灶的核心显著性. 实验结果表明, 该设置在特征增强与边缘保留之间取得了最佳平衡, 从而确保模型针对不同形态类型的小病灶均具备稳定且出色的特征提取能力. 高斯卷积的示意图如图 4 所示. 通过将空洞率引入高斯卷积, 该卷积能够自适应地提取小病灶的显著特征, 从而实现对不同尺寸病灶的鲁棒检测. 本文选取空洞率 $d = 1, 2, 3$ 的组合, 该设置基于对数据集中病灶尺寸分布的统计分析 & 多组消融实验. 实验表明, $d = 1$ 与 $d = 2$ 能有效覆盖绝大多数直径小于 3 mm 病灶的核心区域, 而 $d = 3$ 的加入显著提升了对较大病灶边缘衰减特征的建模能力. 在尝试其他组合 (如 $[1, 2]$ 、 $[1, 3]$ 、 $[1, 2, 4]$) 后的结果显示, $[1, 2, 3]$ 在保持模型轻量的同时, 取得了最优的多尺度特征提取效果, 过大空洞率则易引入无关背景信息. 高斯空洞卷积的数学表示如下:

$$(x * d\omega)(i, j) = \sum_{m, n} x(i + d \times m, j + d \times n) \times \omega(m, n) \quad (3)$$

$$G(m, n) = \exp \left(-\frac{1}{2} \left(\left(\frac{m - m_c}{\sigma_m} \right)^2 + \left(\frac{n - n_c}{\sigma_n} \right)^2 \right) \right) \quad (4)$$

其中, $(x * d\omega)(i, j)$ 表示在输出位置 (i, j) 处的卷积响应输出; 符号 “*” 代表卷积算子; $x \in \mathbf{R}^{C \times W \times H}$ 为输入特征图, 其中的物理量 C 、 W 和 H 分别代表特征图的通道数、宽度和高度; $d = 1, 2, 3$ 为空洞率, 用于调节卷积核的采样步长; $\omega(m, n)$ 为卷积核在局部偏移量 (m, n) 处的权重参数; $G(m, n)$ 是基于二维高斯分布生成的卷积核权重; m_c 与 n_c 分别表示高斯核在水平和垂直方向上的中心坐标. $0 < \sigma_m < 1$ 和 $0 < \sigma_n < 1$ 是高斯卷积核函数的尺度参数和控制核函数形状的尺度参数, 这两个符号作为可学习的参数在训练过程中自适应更新, 确保卷积核始终保持高斯分布形态, 从而引导网络精准捕获小病灶的物理显著性特征. 结合 $(x * d\omega)(i, j)$ 和 $G(m, n)$, 高斯空洞卷积表示为:

$$GD(i, j) = (x * dG)(i, j) = \sum_{m, n} x(i + d \times m, j + d \times n) \times G(m, n) \quad (5)$$

其中, $GD(i, j) = (x * dG)(i, j)$ 为高斯空洞卷积的数学表达式, (i, j) 为输出特征图中的位置. 基于高斯空洞卷积的小病灶显著特征提取, 可视为最大化高斯空洞卷积核与输入局部特征之间内积的过程. 其数学表示如下:

$$\max \left(\sum_{m, n} x(i + d \times m, j + d \times n) \times G(m, n) \right) \quad (6)$$

在所提出的 MSPF-Net 中, 本文利用三种不同空洞率的高斯空洞卷积构建了 PSFE 模块, 以提取小病灶的多尺度显著特征. 通过该 PSFE 模块, 医学小病灶的显著特征被有效提取, 这解决了现有医学图像处理无法有效提取医学小病灶稀疏特征的问题, 并极大丰富了医学小病灶的特征表达. 同时, 该模块显著抑制了强背景杂波的干扰, 有效解决了医学小病灶分割网络训练中正负样本分布不平衡的问题, 大幅提升了训练效率, 并提高了检测精度.

2.2 显著特征提取主干网络

传统的医学分割网络通常采用深层架构来提取病灶特征, 这些方法对于较大病灶有效, 但会导致参数量大、实时性差; 更为关键的是, 深层网络中的池化层会造成小病灶固有稀疏特征的丢失. 为解决上述问题, 本文提出专为医学小病灶分割设计的轻量级 MSPF-Net. 如图 3 中的 (a) 所示, MSPF-Net 采用极浅的架构, 仅包含三个网络模块, 这是考虑到小病灶特征有限, 无需采用过深的网络结构. 每个 MSPF 网络块 (其结构详见图 3 中的 (c)) 处理输入特征图 $F_{in} \in \mathbf{R}^{C \times W \times H}$ 时, 首先应用一个 $3 \times$

3 卷积, 其输出随后进行并行处理: 一路输入核心的 PSFE 模块, 该模块采用多尺度高斯空洞卷积来提取小病灶在不同尺度上的显著特征, 从而丰富特征表达; 同时, 另一路通过一个 1×1 卷积以保留原始特征. 最后, 将 PSFE 模块提取的多尺度显著特征与 1×1 卷积路径的原始特征进行融合, 融合后的特征图被传递至后续的 MSPF 网络块. 此轻量级设计使得 MSPF-Net 能够快速、准确地提取医学小病灶分割所必需的稀疏但关键的特征. MSPF 网络块的计算过程表示如下:

$$F_{Out} = \text{BN}(\text{Maxpool}(F_O + F_S)) \quad (7)$$

$$F_c = \text{Conv}_{3 \times 3}(F_{in}), F_O = \text{Conv}_{1 \times 1}(F_c) \quad (8)$$

其中, $F_{Out} \in \mathbf{R}^{C \times W \times H}$ 是 MSPF 网络块的输出结果表达, 它融合了医学小病灶的显著特征 $F_S \in \mathbf{R}^{C \times W \times H}$ 与原始特征 F_O ; $\text{BN}(\cdot)$ 表示整体的批归一化 (batch normalization); $\text{Maxpool}(\cdot)$ 代表最大化池化 (max pooling); $F_{in} \in \mathbf{R}^{C \times W \times H}$ 是输入特征; $\text{Conv}_{3 \times 3}(\cdot)$ 是卷积核大小为 3 的普通卷积. 式中 F_c 通过普通卷积和池化获得, F_S 通过 PSFE 模块获得. PSFE 模块的计算过程表示如下:

$$F_S = \text{PSFE}(F_c) = G(F_c) + GD_{d=1}(F_c) + GD_{d=2}(F_c) \quad (9)$$

其中, $\text{PSFE}(\cdot)$ 代表该核心功能模块的映射运算; $G(\cdot)$ 是卷积核大小为 3 的二维高斯卷积; $GD_{d=1}(\cdot)$ 和 $GD_{d=2}(\cdot)$ 分别是空洞率为 1 和 2 的卷积核大小为 3 的高斯空洞卷积. 通过 $G(\cdot)$ 、 $GD_{d=1}(\cdot)$ 和 $GD_{d=2}(\cdot)$, 医学小病灶的多尺度显著特征被有效提取并最终融合, 输出显著特征 $F_S \in \mathbf{R}^{C \times W \times H}$.

2.3 模型训练

如前述, 本文提出的 RAP-Net 借鉴了传统相关滤波方法的原理来实现病灶多尺度显著特征的提取, 即通过计算特征图之间的响应来定位显著区域. 为实现小病灶的精准定位与分割, 本文采用均方误差 (mean squared error, MSE) 作为损失函数. 这一选择主要源于 RAP-Net 独特的射线衰减原理与小病灶的特征稀疏特性. 由于网络引入了符合 X 射线衰减原理的高斯核, 其核心任务是通过相关滤波原理回归出反映病灶“中心强、边缘弱”特征的高斯响应图, 而非单纯的二值掩膜分类. MSE 作为一种回归损失, 能够迫使网络预测的特征图在数值分布上逼近理想的二维高斯分布, 从而精确学习病灶的形态与强度衰减原理, 这与相关滤波中的响应图回归目标高度一致. 此外, 考虑到小病灶在图像中像素占比极低, 基于重叠率的 Dice 损失容易由于前

景区域过小导致梯度剧烈波动, 而 MSE 提供了严格的逐像素强度约束, 在有效抑制背景噪声干扰的同时, 为稀疏目标提供了更平滑、稳定的优化方向, 确保网络在极不平衡样本分布下的收敛性能. 本文的损失函数可表示如下:

$$\text{MSE}(P, I_{\text{Label}}) = \frac{\sum_{j=1}^H \sum_{i=1}^W (I_{\text{Label}}(i, j) - P(i, j))^2}{N} \quad (10)$$

其中, $\text{MSE}(P, I_{\text{Label}})$ 为 RAP-Net 的损失函数值; $I_{\text{Label}}(i, j) \in \mathbf{R}^{1 \times W \times H}$ 表示标签图像在坐标 (i, j) 处的像素值; $P(i, j) \in \mathbf{R}^{1 \times W \times H}$ 为网络在相同位置的预测输出响应; N 表示图像的总像素数, 满足 $N = W \times H$.

3 实验结果

3.1 实施细节

所有实验均在同一台计算机上进行, 该计算机硬件配备 Intel Core Ultra 9 275HX (2.70 GHz) 处理器、32 GB 系统内存, 并搭载 NVIDIA GeForce RTX 5080 GPU 进行计算加速. 提出的 RAP-Net 模型采用 Python 3.7.0 与 PyTorch 1.7.0 实现, 在 Windows10 系统计算机上运行. 模型训练耗时 24 小时, 卷积运算采用 PyTorch 后端, 并使用 Adam 优化器 (学习率设为 0.001). 训练方案包含 200 个 epoch, 每批次包含 16 张图像. 每批次量设置为 16 是基于 GPU 显存限制与训练稳定性之间的平衡考量; 初始学习率设定为 0.001 是经多轮实验验证, 在 Adam 优化器下能够实现模型最优收敛速度与性能表现的设定.

3.2 数据集

为评估所提出的 RAP-Net 模型, 本文基于郑州大学第一附属医院临床病例的腹部 CT 扫描构建了医学小病灶 (medical small lesion, MSL) 数据集. 本研究经过了郑州大学第一附属医院伦理委员会批准 (2024-KY-1887), 因为本文属于回顾性研究, 故免除了患者的知情同意. 本文回顾性收集了 2024 年 1 月至 2024 年 12 月期间在郑州大学第一附属医院接受腹部 CT 扫描的患者数据. MSL 数据集包含 6420 例小肾结石与小肝脏钙化灶的 CT 扫描病例. 从图 5 中可以清晰看出, 小病灶尺寸极小且特征不明显, 这对现有的医学病灶检测方法构成了巨大挑战. 病例的纳入标准包括: 1) 年龄大于等于 18 岁; 2) 确诊患有肾结石或肝脏钙化灶; 3) 肾结石最大直径小于等于 5 mm, 单发或多发, 位于肾实

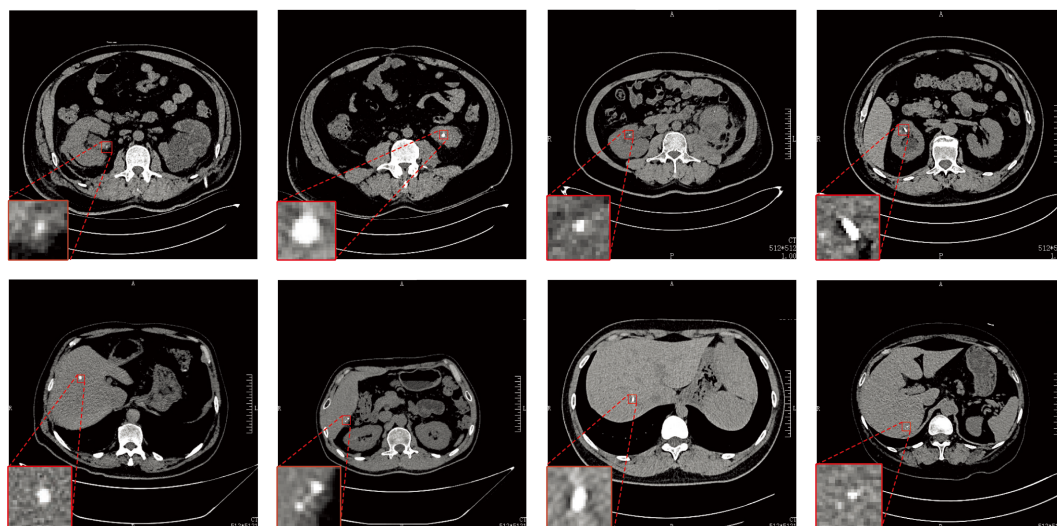


图 5 MSL 数据集样本示例图

Fig.5 Sample diagram of the MSL dataset

质或集合系统, 肝脏钙化灶最大直径小于等于 10 mm, 孤立分布, 无伴随占位性病变; 4) 无其他占位性病变, 无肝肾手术史; 5) 无运动伪影或金属植入物. 排除标准包括: 1) 合并其他肝肾占位性病变; 2) 图像质量较差; 3) 既往有肝肾手术、介入治疗或放疗史; 4) 腰腹部有金属植入物; 5) 结石体积过大.

最终, 本文纳入了 5 203 例小型肾结石和 1 217 例小型肝脏钙化灶. 所有数据集的训练集、验证集和测试集划分比例为 6:2:2. 所有图像均由两名具有十年以上放射学经验的临床医生手动标注, 并由资历更高的医师确认和校正. 本研究使用的 CT 扫描设备包括 Discovery 750 HD CT (GE HealthCare,

美国)、SOMATOM Force CT (Siemens, 德国)、Brilliance iCT (Philips, 荷兰) 以及 Aquilion ONE 640 (Toshiba, 日本). 图像尺寸为 512×512 像素, 窗宽设置为 220 HU, 窗位设置为 40 HU, 扫描层厚设置为 1 ~ 5 mm.

3.3 有效性实验

为评估所提 RAP-Net 的有效性, 本文在 MSL 数据集上分别进行了小肾结石与小肝脏钙化灶的分割实验. 如图 6、图 7 所示, RAP-Net 在这两类小病灶分割任务中均表现出优良性能. 如图 8 所示, RAP-Net 能够准确分割小病灶, 并在轮廓保持与边缘定

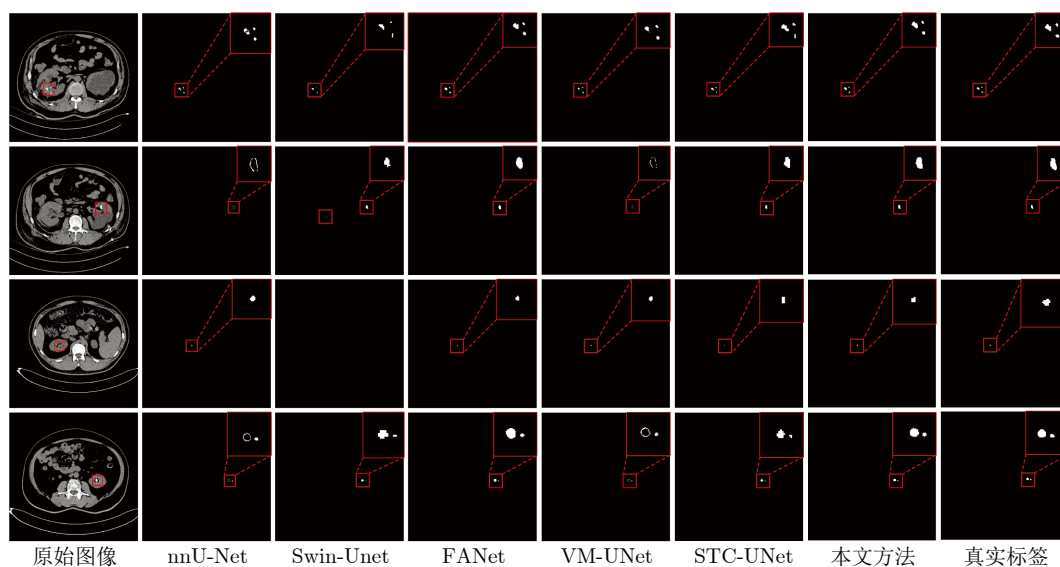


图 6 采用不同方法对 MSL 数据集中小肾结石进行分割的结果

Fig.6 Results of segmenting small kidney stones in the MSL dataset by using different methods

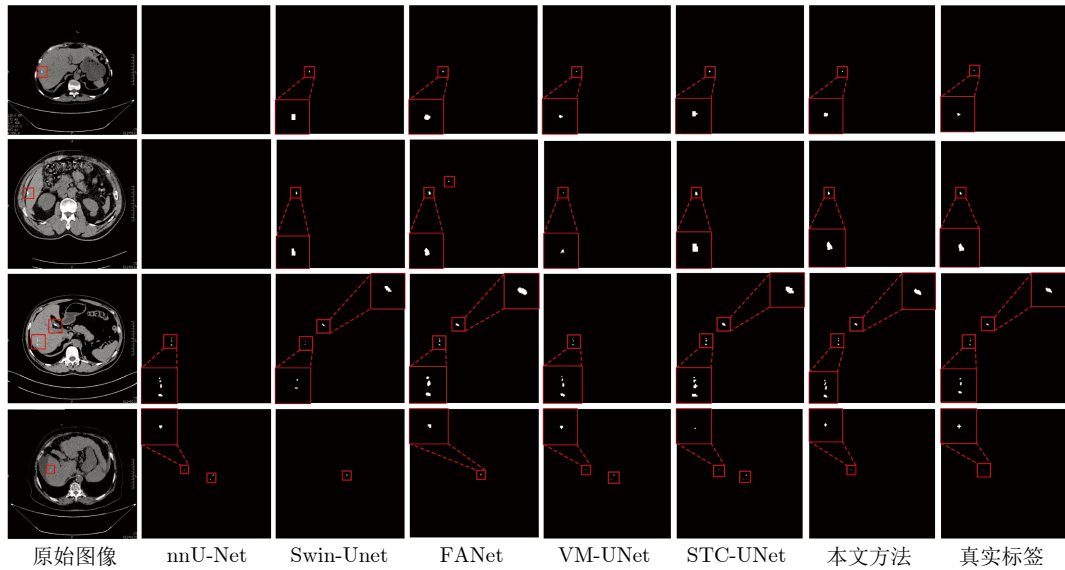


图 7 采用不同方法对 MSL 数据集中小肝脏钙化灶进行分割的结果

Fig.7 Results of segmenting small liver calcifications in the MSL dataset by using different methods

位方面表现出较高精度. 图 9 中的混淆矩阵进一步展示了 RAP-Net 在像素层面的分类情况. 整体来看, 模型在背景类上具有极高的准确率, 表明其能有效抑制复杂背景的干扰. 假阳性像素主要出现在与病灶灰度特征相似的结构边缘, 如小血管横截面或组织交界处; 假阴性则多集中于病灶边缘弱对比区域及部分小病灶的中心. 这类错误模式反映出, 尽管高斯衰减能有效增强病灶的整体响应, 但对形态偏离典型高斯分布或信噪比极低的局部区域仍存

在一定局限. 该分析也印证了 PSFE 模块在增强显著特征的同时, 对背景抑制的有效性, 为后续优化提供了明确方向. 此外, 为更直观地体现 PSFE 模块的贡献, 图 3 中对 RAP-Net 各层输出特征图进行了可视化分析. 可见, 经 PSFE 模块处理后, 小病灶特征得到显著增强与保留, 同时复杂背景干扰被有效抑制, 这充分表明基于 X 射线衰减原理设计的 RAP-Net 对医学小病灶具备强鲁棒性与良好适应性.

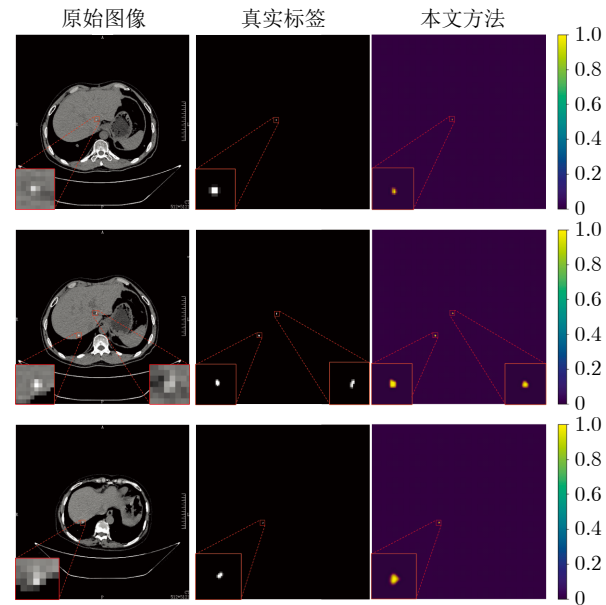


图 8 本文方法在复杂 CT 图像中对小病灶的分割结果

Fig.8 The segmentation results for small lesions in complex CT images by using the method described in this paper

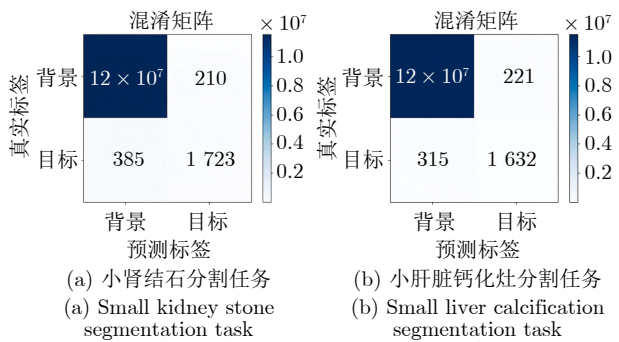


图 9 RAP-Net 在 MSL 数据集两个医学小病灶分割任务中的像素级混淆矩阵

Fig.9 Pixel-level confusion matrices of RAP-Net in the MSL dataset for two medical small lesion segmentation tasks

3.4 比较实验结果

在医学分割领域, Dice 系数与 IoU 是评估方法性能的常用指标^[46-53]. 因此, 为更直观有效地验证所提方法的性能, 本文选择这两个指标进行验证实验.

为验证 RAP-Net 在医学小病灶图像分割任务中的有效性, 本文在小肾结石分割任务和小肝脏钙化灶分割任务 (即 MSL 数据集) 中, 与现有行业内具有代表性的医学病灶分割深度学习方法进行了广泛的对比实验. 为确保对比的全面性与先进性, 本文选取七种涵盖不同主流架构范式的方法, 包括: 1) 医学图像分割领域的自适应标杆方法 nnU-Net^[30]; 2) 通过迭代反馈机制优化特征提取的网络 FANet^[33]; 3) 基于纯 Transformer 架构的代表性网络 Swin-Unet^[36]; 4) 结合选择性核与 ViT 进行多层次特征增强的混合架构 STC-UNet^[37]; 5) 引入 SSM 以捕捉长距离依赖的新型 SOTA 方法 VM-UNet^[38] 和 Zig-RiR^[39]; 6) 基于 SAM 进行医学适配的架构 MedSAM^[40]. 上述对比方法涵盖了 CNN、Transformer、SSM 以及 SAM 等多种主流架构范式, 代表了当前医学图像分割领域不同技术流派的典型方案.

表 1 和表 2 分别汇总了所有方法在 MSL 数据集上的小病灶分割性能. 结果表明, 所提方法在性能指标 Dice 和 IoU 上均优于现有先进深度学习方法. 与现有方法相比, 本文方法在小肾结石分割任务中将 Dice 系数提高至少 20.96%, IoU 提高至少 22.76%. 在小肾结石分割任务中, 小结石的图像病灶小而微弱. 所提方法具有较强的特征表达能力, 能有效处理此类不显著的小病灶分割任务. 表 1 清晰表明, 与现有 SOTA 方法 Zig-RiR^[39] 相比, RAP-Net 在小肾结石分割任务中表现更优. 与同为 CNN 检测模型的 STC-UNet^[37] 相比, 本文方法在所有性能指标上均具有显著优势. 为更好地展示所提方法的检测性能, 本文对检测结果进行了可视化. 如图 6 所示, RAP-Net 网络能有效、准确地检测肾脏中尺寸小且信号弱的结石病灶, 这充分证明 RAP-Net 在决策过程中选择和增强显著性特征的卓越能力.

表 1 不同方法在小肾结石分割任务中的 Dice 系数与 IoU
Table 1 The Dice coefficient and IoU of different methods in the small kidney stone segmentation task

方法	Dice (%)	IoU (%)	参数量 ($\times 10^6$)
nnU-Net ^[30]	63.85	51.34	30.60
FANet ^[33]	61.35	46.19	7.72
Swin-Unet ^[36]	64.79	49.32	27.20
STC-UNet ^[37]	61.47	52.29	41.06
VM-UNet ^[38]	58.90	44.12	27.43
Zig-RiR ^[39]	65.48	55.15	24.58
MedSAM ^[40]	59.64	47.71	93.74
本文方法 (3-layer MSPF)	86.44	77.91	0.34
本文方法 (5-layer MSPF)	86.65	78.05	—
本文方法 (7-layer MSPF)	87.43	78.51	—

图 6 中红框内的检测结果清晰表明所提方法的优异检测性能, 该方法能有效分割更精细的病灶区域, 在多病灶的场景下仍能保持良好的分割性能.

表 2 不同方法在小肝脏钙化灶分割任务中的 Dice 系数与 IoU

Table 2 The Dice coefficient and IoU of different methods in the small liver calcification segmentation task

方法	Dice (%)	IoU (%)	参数量 ($\times 10^6$)
nnU-Net ^[30]	53.03	44.45	30.60
FANet ^[33]	44.12	35.70	7.72
Swin-Unet ^[36]	37.81	38.44	27.20
STC-UNet ^[37]	43.12	31.85	41.06
VM-UNet ^[38]	27.81	20.19	27.43
Zig-RiR ^[39]	66.59	56.02	24.58
MedSAM ^[40]	55.13	45.06	93.74
本文方法 (3-layer MSPF)	84.64	75.57	0.34
本文方法 (5-layer MSPF)	84.96	75.84	—
本文方法 (7-layer MSPF)	85.15	76.02	—

除小肾结石分割任务外, 小肝脏钙化灶分割任务是另一项富有挑战性的小病灶分割任务, 由于病灶与肝实质之间对比度低且病灶尺寸极小, 二者在图像中难以区分, 导致分割难度显著增加. 本文在 MSL 数据集上将 RAP-Net 与现有行业内具有代表性的多类医学图像分割方法进行了比较, 结果如表 2 所示. 所提方法将小肝脏钙化灶分割的 Dice 系数提高至少 18.05%, IoU 提高至少 19.55%. 如图 7 所示, 图中红框标注的钙化病灶尽管与背景组织边界模糊, 但仍实现了精确分割, 突显其对低对比度小病灶的优异特征提取能力. 值得注意的是, 通过不同深度的 MSPF 模块消融实验 (表 1、表 2) 发现, 更深的网络结构并未显著提升性能. 这是因为小病灶特征稀疏, 无需深层网络进行拟合. 因此, RAP-Net 采用 3 层 MSPF 模块进行医学小病灶特征提取. 由于本文方法采用 CT 成像原理引导, RAP-Net 仅需少量模型参数即可有效拟合小病灶特征, 使得网络非常轻量化. 在表 1、表 2 中, 本文列举了现有先进方法用于本文数据时的参数量. 从表中可以清晰看出, RAP-Net 的参数量非常小, 极大提升了 RAP-Net 的部署便利性. 图 8 展示了在复杂图像场景中本文方法的网络输出结果图, 最终分割结果仅为阈值分割结果, 此图清晰直观地展示了网络对病灶边缘和特征的表达能力.

本文提出的 RAP-Net 能在众多强基线方法中取得显著优势, 关键在于衰减原理的有效引入与显著性特征提取专用网络结构的设计相辅相成. 具体

而言, 衰减的作用主要体现在: 通过将 CT 影像中 X 射线的衰减原理建模为高斯衰减, RAP-Net 在特征提取初期可实现类似“显著性滤波”的效果, 从而增强病灶区域的响应, 以解决小病灶特征稀疏、对比度低的难题. 另一方面, 所设计的轻量主干网络 MSPF-Net 则侧重于高效的多尺度特征融合与判别, 该网络结构精简, 借助多尺度高斯空洞卷积, 能够在不过度增加参数量的前提下, 有效聚合不同感受野下的病灶特征, 从而减少深层网络中可能出现的细节丢失与计算冗余. 正因这两方面紧密结合, RAP-Net 才得以在小病灶分割任务中同时实现高精度与高效率.

3.5 大病灶的分割性能

为探究所提方法是否具备更广泛的泛化能力, 本节将 RAP-Net 延伸应用于肝肿瘤与肾肿瘤等相对大尺寸病灶的分割任务中. 需要强调的是, 本研究的核心目标是解决医学图像中小病灶的分割难题, 大病灶实验旨在作为方法泛化能力的一种佐证, 验证基于衰减原理的特征提取机制在不同尺度病灶上的适应性, 而非主要应用场景. 尽管所提出的 RAP-Net 主要针对医学小病灶设计, 但其对较大尺寸的医学病灶仍展现出良好的分割性能. 为验证模型在此类目标上的泛化能力, 本文在肝肿瘤分割任务^[41, 54–55]和肾肿瘤分割任务^[56–58]上进行了系统评估, 结果如图 10 所示. 实验表明, 基于其卓越的显著性特征提取机制, RAP-Net 能够有效处理肝、肾肿瘤等较大尺寸病灶的分割问题. 这是由于在这两

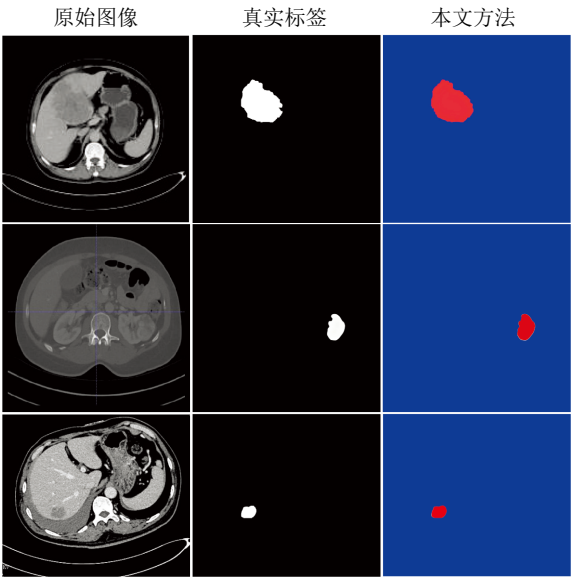


图 10 RAP-Net 在肝/肾肿瘤的分割结果
Fig.10 RAP-Net segmentation results for liver/kidney tumors

个分割任务中, 肿瘤区域与周围组织仍有较为明显的边缘特征, PSFE 模块能够准确识别并强化此类边缘区域的显著性特征, 从而获得精确的分割结果. 如图 4 所示, 尽管大尺寸病灶边缘的显著性特征响应强度不如小病灶强烈, 但仍显著高于背景组织 (即大面积的正常组织), 故 PSFE 模块能够有效捕捉大尺寸病灶的结构边界信息, 显著提升分割边界的一致性. 这一特性使得 RAP-Net 在保持对小病灶高灵敏度的同时, 对大尺寸病灶同样具备可靠的分割能力.

3.6 消融实验

为验证所提出的 PSFE 模块的有效性, 本文将其替换为常规卷积层进行对比实验. 如表 3、表 4 和图 11 所示, 实验结果表明, 相较于普通卷积, 集成了 CT 成像物理原理的 PSFE 模块能够精准捕获微小病灶的多尺度显著特征, 显著提升了在复杂背景干扰下对特征稀疏目标的分割性能. 这充分证明, 基于 CT 成像物理原理驱动的 PSFE 模块相较于传统 CNN 卷积算子, 对医学小病灶具有更强的特征表达能力与鲁棒性. 此外, 为验证 RAP-Net 架构中各核心组件的有效性, 本文进行了系统的消融实验. 首先, 针对骨干网络 MSPF-Net 的优越性, 本文将与其与三种经典的特征提取网络 (VGG-16^[59]、ResNet-101^[60] 及 DenseNet-121^[61]) 进行对比分析, 测试结果如表 3、表 4 和图 12 所示. 实验分析表明, VGG-16 与 ResNet-101 网络在频繁下采样过程中极易导致微小目标的细节丢失与信息稀释, 而 DenseNet-121 因其密集拼接方式在处理小病灶时容易产生大量冗余信息. 相比之下, MSPF-Net 结合物理原理引导实现了对稀疏特征的精准捕获, 克服了传统骨干网络在空间分辨率保持与特征辨识力方面的不足. 从定量结果可以清晰看出, 尽管对比网络具有更深的模型深度或更复杂的特征复用机制, 但在面对尺寸极小、对比度低的病灶时, 其分割精度远低

表 3 小肾结石分割任务中消融实验的 Dice 系数与 IoU (%)
Table 3 The Dice coefficient and IoU of ablation experiment in the small kidney stone segmentation task (%)

方法	Dice	IoU
VGG-16 ^[59]	44.25	31.40
ResNet-101 ^[60]	51.02	42.67
DenseNet-121 ^[61]	51.49	39.81
常规卷积	40.97	28.13
本文方法	85.35	76.86

表 4 小肝脏钙化灶分割任务中消融实验的 Dice 系数与 IoU (%)
Table 4 The Dice coefficient and IoU of ablation experiment in the small liver calcification segmentation task (%)

方法	Dice	IoU
VGG-16 ^[50]	41.32	30.13
ResNet-101 ^[60]	51.12	40.13
DenseNet-121 ^[61]	54.73	41.51
常规卷积	39.97	27.23
本文方法	83.81	74.46

于本文提出的 MSPF-Net. 分析其原因, 传统大参数网络在频繁的下采样过程中极易丢失小病灶的细节信息, 并引入冗余的背景噪声. 相比之下, MSPF-Net 能够深度融合辐射衰减原理, 引导网络在浅层即可实现对稀疏特征的精准捕获, 在大幅降低参数量的同时, 有效消除了背景杂波的干扰, 实现了性能与效率的双重突破.

3.7 局限性分析

尽管本文提出的 RAP-Net 能够有效分割出复杂背景中的医学小病灶, 但其性能仍在一定程度上受与小病灶形态相似的高斯噪声影响. 该局限性的根源在于传统相关滤波方法对高斯噪声固有的敏感性. 虽然引入 CNN 结构能在一定程度上抑制这类干扰, 但高斯噪声的特征可能与呈二维高斯分布的医学小病灶存在相似性, 这会对 PSFE 模块的效果产生一定影响. 虽然本文方法对高斯噪声存在一定的敏感性, 但这主要体现在噪声水平异常高的情境

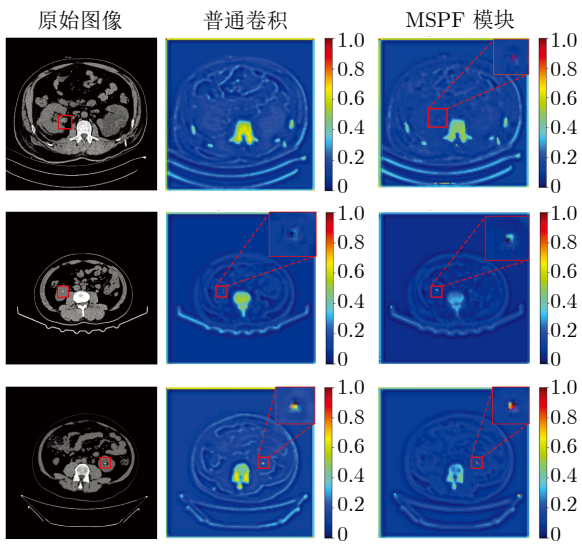


图 11 消融实验的可视化结果
Fig.11 The visualization results of ablation experiments

下. 对于常规临床图像, 模型通过训练中的数据增强已具备抗噪能力, 无需强制进行额外的去噪预处理. 然而, 为确保在低剂量或高噪声采集场景下的最佳性能, 本文建议在推理前应用一次轻量级的去噪预处理步骤, 以提高输入特征的信噪比, 从而最大化 RAP-Net 的分割精度和鲁棒性. 在未来研究中可致力于构建多元高斯分布模型, 以更精确地拟合医学小病灶的灰度分布特征, 从而有效抑制高斯噪声干扰. 此外, 本研究构建的 MSL 数据集及实验主要针对小肾结石与小肝脏钙化灶两类特定病灶. 本文所提方法基于一个核心假设: 小病灶在 CT 图

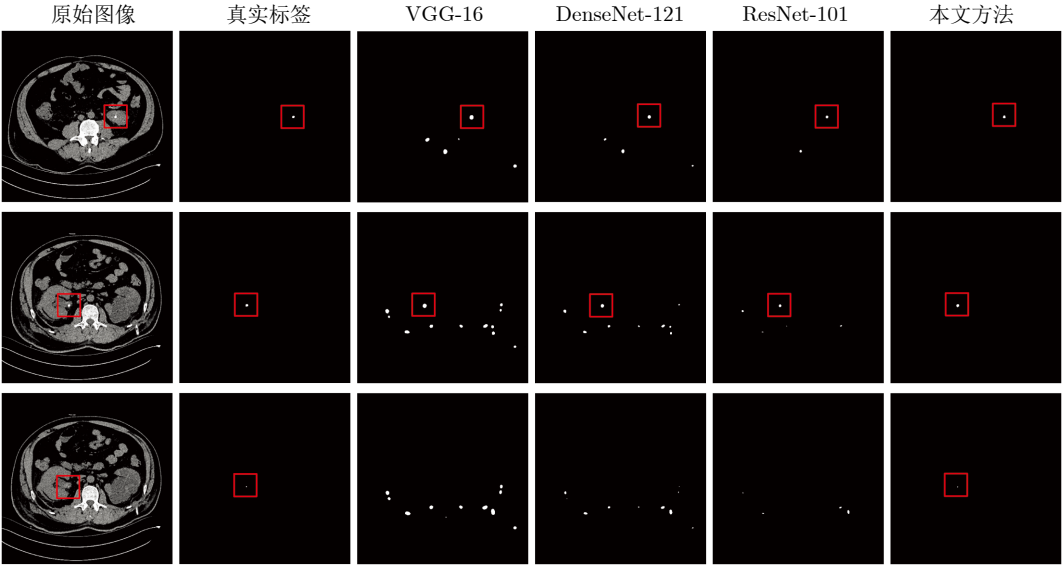


图 12 消融实验的可视化结果
Fig.12 The visualization results of ablation experiments

像中的灰度分布近似二维高斯型. 这一假设在多数钙化灶、小结石中成立, 但面对形态明显偏离该规律的病灶 (如条索状、分叶状或不均匀强化病灶) 时, 其分割性能可能下降. 同时, 方法的性能可能受到不同 CT 扫描设备、成像协议所引入的图像纹理与噪声特性差异的影响. 尽管所提出的衰减原理源自 CT 成像的一般特性, 理论上对具有类似灰度分布的其他器官小病灶 (如肺小结节、胰腺小囊肿等) 也具备适用潜力, 但其实际分割性能可能受病灶形态、尺寸变异度及周围组织对比度等因素影响. 未来工作将进一步采集多器官、多类型小病灶数据, 以系统验证方法的泛化能力, 并探索如何自适应地调整高斯衰减参数以应对不同成像协议与病灶形态的差异.

4 结束语

本文从 CT 成像物理机制与小病灶显著性特征建模的角度出发, 提出一种用于医学小病灶分割的网络 RAP-Net. 该网络的创新在于构建一种物理驱动的特征感知机制, 将传统相关性滤波理论与深度学习架构深度融合, 以解决小病灶因特征稀疏和背景干扰而导致的分割难题. 具体而言, 本文基于小病灶在 CT 图像中呈现出的符合物理规律的近似二维高斯分布的强度特性, 在 RAP-Net 中构建 MSPF-Net 作为主干特征提取器, 并通过其核心组件 PSFE 模块引入多尺度高斯空洞卷积, 实现了对病灶物理显著性特征的有效感知, 同时有效抑制了背景噪声. 在本文构建的 MSL 数据集上的大量实验表明, RAP-Net 的性能显著优于现有主流方法. 此外, 作为泛化能力的一种验证, 该方法在大病灶分割任务中也展现出良好的适应性, 进一步体现了基于物理感知的网络架构的鲁棒性与推广潜力. 本研究工作表明, 成像物理原理转化为深度网络中的显著性感知能力, 是提升医学小病灶分割模型鲁棒性的一条有效路径. 展望未来, 本文提出的 RAP-Net 展现出良好的临床部署潜力. 模型轻量化、无需复杂预处理的特点, 有利于集成至现有影像分析系统. 未来研究需进一步在多中心、多设备数据上验证其泛化能力, 并探索如何与临床 workflow 深度融合, 以推动其在早期精准诊疗中的实际应用.

参考文献

- Zhang K J, Zhang L, Pan H W. UoloNet: Based on multi-tasking enhanced small target medical segmentation model. *Artificial Intelligence Review*, 2024, **57**(2): Article No. 31
- Koleilat T, Asgariandehkordi H, Rivaz H, Xiao Y M. MedCLIP-SAMv2: Towards universal text-driven medical image segmentation. *Medical Image Analysis*, 2025, **106**: Article No. 103749
- He A L, Wang K, Li T, Du C K, Xia S, Fu H Z. H2Former: An efficient hierarchical hybrid transformer for medical image segmentation. *IEEE Transactions on Medical Imaging*, 2023, **42**(9): 2763–2775
- Zhou X D, Chen T X. BSBP-RWKV: Background suppression with boundary preservation for efficient medical image segmentation. In: *Proceedings of the 32nd ACM International Conference on Multimedia*. Melbourne, Australia: Association for Computing Machinery, 2024. 4938–4946
- Mou L, Zhao Y T, Fu H Z, Liu Y H, Cheng J, Zheng Y L, et al. CS²-Net: Deep learning segmentation of curvilinear structures in medical imaging. *Medical Image Analysis*, 2021, **67**: Article No. 101874
- Gu Z W, Cheng J, Fu H Z, Zhou K, Hao H Y, Zhao Y T, et al. CE-Net: Context encoder network for 2D medical image segmentation. *IEEE Transactions on Medical Imaging*, 2019, **38**(10): 2281–2292
- Zhang N, Yu L, Zhang D Z, Wu W D, Tian S W, Kang X J, et al. CT-Net: Asymmetric compound branch Transformer for medical image segmentation. *Neural Networks*, 2024, **170**: 298–311
- Lei Y J, Wu Z J, Li Z Y, Yang Y E, Liang Z M. BP-CapsNet: An image-based deep learning method for medical diagnosis. *Applied Soft Computing*, 2023, **146**: Article No. 110683
- Patro K K, Allam J P, Neelapu B C, Tadeusiewicz R, Acharya U R, Hammad M, et al. Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images. *Information Sciences*, 2023, **640**: Article No. 119005
- Lin J, Huang X R, Zhou H Y, Wang Y Q, Zhang Q N. Stimulus-guided adaptive transformer network for retinal blood vessel segmentation in fundus images. *Medical Image Analysis*, 2023, **89**: Article No. 102929
- Shaker A, Maaz M, Rasheed H, Khan S, Yang M H, Khan F S. UNETR++: Delving into efficient and accurate 3D medical image segmentation. *IEEE Transactions on Medical Imaging*, 2024, **43**(9): 3377–3390
- Wei J H, Zhu G J, Fan Z, Liu J C, Rong Y B, Mo J J, et al. Genetic U-Net: Automatically designed deep networks for retinal vessel segmentation using a genetic algorithm. *IEEE Transactions on Medical Imaging*, 2022, **41**(2): 292–307
- Fang X, Yan P K. Multi-organ segmentation over partially labeled datasets with multi-scale feature abstraction. *IEEE Transactions on Medical Imaging*, 2020, **39**(11): 3619–3629
- Zhang T L, Wang X Z. Anchorwise fuzziness modeling in convolution-Transformer neural network for left atrium image segmentation. *IEEE Transactions on Fuzzy Systems*, 2024, **32**(2): 398–408
- Kumar A, Fulham M, Feng D G, Kim J. Co-learning feature fusion maps from PET-CT images of lung cancer. *IEEE Transactions on Medical Imaging*, 2020, **39**(1): 204–217
- Zhou Y, Ahmed T S, Wang M, Newman E A, Schmetterer L, Fu H Z, et al. Masked vascular structure segmentation and completion in retinal images. *IEEE Transactions on Medical Imaging*, 2025, **44**(6): 2492–2503
- Wu Y C, Xia Y, Song Y, Zhang D H, Liu D N, Zhang C Y, et al. Vessel-Net: Retinal vessel segmentation under multi-path supervision. In: *Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Shenzhen, China: Springer, 2019. 264–272
- Xiang D H, Zhang B, Lu Y X, Deng S M. Modality-specific segmentation network for lung tumor segmentation in PET-CT images. *IEEE Journal of Biomedical and Health Informatics*, 2023, **27**(3): 1237–1248
- Yuan Y C, Zhang L, Wang L T, Huang H Y. Multi-level attention network for retinal vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2022, **26**(1): 312–323
- Ma Y X, Wang S, Hua Y, Ma R H, Song T, Xue Z G, et al. Perceptual data augmentation for biomedical coronary vessel segmentation. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2023, **20**(4): 2494–2505

- 21 Hatamizadeh A, Hosseini H, Patel N, Choi J, Pole C C, Hoeflerlin C M, et al. RAVIR: A dataset and methodology for the semantic segmentation and quantitative analysis of retinal arteries and veins in infrared reflectance imaging. *IEEE Journal of Biomedical and Health Informatics*, 2022, **26**(7): 3272–3283
- 22 Hu K, Jiang S, Zhang Y, Li X Y, Gao X P. Joint-Seg: Treat foveal avascular zone and retinal vessel segmentation in OCTA images as a joint task. *IEEE Transactions on Instrumentation and Measurement*, 2022, **71**: Article No. 4007113
- 23 McKetty M H. The AAPM/RSNA physics tutorial for residents. X-ray attenuation. *RadioGraphics*, 1998, **18**(1): 151–163
- 24 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 3431–3440
- 25 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention (MICCAI). Munich, Germany: Springer, 2015. 234–241
- 26 Zhou Z W, Siddiquee M M R, Tajbakhsh N, Liang J M. UNet++: A nested U-Net architecture for medical image segmentation. In: Proceedings of the 4th International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada, Spain: Springer, 2018. 3–11
- 27 Alom M Z, Yakopcic C, Taha T M, Asari V K. Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net). In: Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON). Dayton, USA: IEEE, 2018. 228–233
- 28 Jha D, Riegler M A, Johansen D, Halvorsen P, Johansen H D. DoubleU-Net: A deep convolutional neural network for medical image segmentation. In: Proceedings of the 33rd International Symposium on Computer-based Medical Systems (CBMS). Rochester, USA: IEEE, 2020. 558–564
- 29 Xiao X, Lian S, Luo Z M, Li S Z. Weighted Res-UNet for high-quality retina vessel segmentation. In: Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME). Hangzhou, China: IEEE, 2018. 327–331
- 30 Isensee F, Jaeger P F, Kohl S A A, Petersen J, Maier-Hein K H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 2021, **18**(2): 203–211
- 31 Oktay O, Schlemper J, Folgoc L L, Lee M, Heinrich M, et al. Attention U-Net: Learning where to look for the pancreas. arXiv preprint arXiv: 1804.03999, 2018.
- 32 Fan D P, Ji G P, Zhou T, Chen G, Fu H Z, Shen J B, et al. PraNet: Parallel reverse attention network for polyp segmentation. In: Proceedings of the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Lima, Peru: Springer, 2020. 263–273
- 33 Tomar N K, Jha D, Riegler M A, Johansen H D, Johansen D, Rittscher J, et al. FANet: A feedback attention network for improved biomedical image segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, **34**(11): 9375–9388
- 34 Ding W P, Wang H P, Huang J S, Ju H R, Geng Y, Lin C T, et al. FTransCNN: Fusing Transformer and a CNN based on fuzzy logic for uncertain medical image segmentation. *Information Fusion*, 2023, **99**: Article No. 101880
- 35 Wang H N, Cao P, Wang J Q, Zaiane O R. UCTransNet: Rethinking the skip connections in U-Net from a channel-wise perspective with Transformer. In: Proceedings of the 36th AAAI Conference on Artificial Intelligence. Virtual Event: AAAI Press, 2022. 2441–2449
- 36 Cao H, Wang Y Y, Chen J, Jiang D S, Zhang X P, Tian Q, et al. Swin-Unet: Unet-like pure Transformer for medical image segmentation. In: Proceedings of the European Conference on Computer Vision. Tel Aviv, Israel: Springer, 2022. 205–218
- 37 Hu W, Yang S Y, Guo W F, Xiao N, Yang X P, Ren X Y. STC-UNet: Renal tumor segmentation based on enhanced feature extraction at different network levels. *BMC Medical Imaging*, 2024, **24**(1): Article No. 179
- 38 Ruan J C, Li J C, Xiang S C. VM-UNet: Vision mamba UNet for medical image segmentation. arXiv preprint arXiv: 2402.02491, 2024.
- 39 Chen T X, Zhou X D, Tan Z T, Wu Y, Wang Z Y, Ye Z, et al. Zig-RiR: Zigzag RWKV-in-RWKV for efficient medical image segmentation. *IEEE Transactions on Medical Imaging*, 2025, **44**(8): 3245–3257
- 40 Ma J, He Y T, Li F F, Han L, You C Y, Wang B. Segment anything in medical images. *Nature Communications*, 2024, **15**(1): Article No. 654
- 41 Jiao R S, Liu Q P, Zhang Y, Pu B Z, Xue B S, Cheng Y, et al. RECIST^{Surv}: Hybrid multi-task Transformer for hepatocellular carcinoma response and survival evaluation. *IEEE Transactions on Image Processing*, 2025, **34**: 3873–3888
- 42 Zhao L X, Wang T, Chen Y B, Zhang X L, Tang H, Lin F X, et al. A novel framework for segmentation of small targets in medical images. *Scientific Reports*, 2025, **15**(1): Article No. 9924
- 43 Ahmadi M, Biswas D, Lin M H, Vrionis F D, Hashemi J, Tang Y F. Physics-informed machine learning for advancing computational medical imaging: Integrating data-driven approaches with fundamental physical principles. *Artificial Intelligence Review*, 2025, **58**(10): Article No. 297
- 44 Wang X L, Girshick R, Gupta A, He K M. Non-local neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 7794–7803
- 45 Huang Z L, Wang X G, Huang L C, Huang C, Wei Y C, Liu W Y. CCNet: Criss-cross attention for semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, South Korea: IEEE, 2019. 603–612
- 46 Ruan J C, Xiang S C, Xie M Y, Liu T, Fu Y Z. MALUNet: A multi-attention and light-weight UNet for skin lesion segmentation. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Las Vegas, USA: IEEE, 2022. 1150–1156
- 47 Wu Y X, Liao K L, Chen J T, Wang J H, Chen D Z, Gao H H, et al. D-Former: A U-shaped dilated Transformer for 3D medical image segmentation. *Neural Computing and Applications*, 2023, **35**(2): 1931–1944
- 48 Intiaz T, Fattah S A, Kung S Y. BAWGNet: Boundary aware wavelet guided network for the nuclei segmentation in histopathology images. *Computers in Biology and Medicine*, 2023, **165**: Article No. 107378
- 49 Bernard O, Lalonde A, Zotti C, Cervenansky F, Yang X, Heng P A, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Transactions on Medical Imaging*, 2018, **37**(11): 2514–2525
- 50 Xu Y S, Tang J Q, Men A D, Chen Q C. EviPrompt: A training-free evidential prompt generation method for adapting segment anything model in medical images. *IEEE Transactions on Image Processing*, 2024, **33**: 6204–6215
- 51 Zhou S H, Nie D, Adeli E, Yin J P, Lian J, Shen D G. High-resolution encoder-decoder networks for low-contrast medical image segmentation. *IEEE Transactions on Image Processing*, 2020, **29**: 461–475
- 52 Qiu Z X, Hu Y, Chen X S, Zeng D, Hu Q Y, Liu J. Rethinking dual-stream super-resolution semantic learning in medical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, **46**(1): 451–464
- 53 Zhang T Y, Zheng S M, Cheng J, Jia X, Bartlett J, Cheng X X, et al. Structure and intensity unbiased translation for 2D medical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, **46**(12): 10060–10075

- 54 Bilic P, Christ P, Li H B, Vorontsov E, Ben-Cohen A, Kaissis G, et al. The liver tumor segmentation benchmark (LiTS). *Medical Image Analysis*, 2023, **84**: Article No. 102680
- 55 Lyu F, Ye M, Ma A J, Yip T C F, Wong G L H, Yuen P C. Learning from synthetic CT images via test-time training for liver tumor segmentation. *IEEE Transactions on Medical Imaging*, 2022, **41**(9): 2510–2520
- 56 Heller N, Sathianathan N, Kalapara A, Walczak E, Moore K, Kaluzniak H, et al. The KiTS19 challenge data: 300 kidney tumor cases with clinical context, CT semantic segmentations, and surgical outcomes. arXiv preprint arXiv: 1904.00445, 2019.
- 57 Liao Miao, Yang Rui-Xin, Zhao Yu-Qian, Di Shuan-Hu, Yang Zhen. Multi-organ segmentation from abdominal CT images based on CE-TransNet. *Acta Automatica Sinica*, 2025, **51**(6): 1371–1387
(廖苗, 杨睿新, 赵于前, 邸拴虎, 杨振. 基于 CE-TransNet 的腹部 CT 图像多器官分割. *自动化学报*, 2025, **51**(6): 1371–1387)
- 58 Jia Xi-Bin, Guo Xiong, Wang Luo, Yang Da-Wei, Yang Zheng-Han. A few-shot medical image segmentation network with iterative boundary refinement. *Acta Automatica Sinica*, 2024, **50**(10): 1988–2001
(贾熹滨, 郭雄, 王璐, 杨大为, 杨正汉. 一种迭代边界优化的医学图像小样本分割网络. *自动化学报*, 2024, **50**(10): 1988–2001)
- 59 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. San Diego, USA: ICLR, 2015.
- 60 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE, 2016. 770–778
- 61 Huang G, Liu Z, van der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE, 2017. 4700–4708



任向阳 郑州大学第一附属医院高级工程师. 主要研究方向为医学图像处理, 计算机视觉, 模式识别和目标检测.
E-mail: xyren@zzu.edu.cn
(**REN Xiang-Yang** Senior engineer at the First Affiliated Hospital of Zhengzhou University. His research interests include medical image processing, computer vision, pattern recognition, and object detection.)



焦博昶 郑州大学第一临床医学院硕士研究生. 主要研究方向为医学图像处理, 计算机视觉和模式识别.
E-mail: jiaoboyang@gs.zzu.edu.cn
(**JIAO Bo-Yang** Master student at the First Clinical Medical College, Zhengzhou University. His research interests include medical image processing, computer

vision, and pattern recognition.)



杨 震 湖南大学人工智能与机器人学院博士研究生. 主要研究方向为计算机视觉, 模式识别和目标检测.
E-mail: zhenyangssy@gmail.com
(**YANG Zhen** Ph.D. candidate at the School of Artificial Intelligence and Robotics, Hunan University. His research interests include computer vision, pattern recognition, and object detection.)



贾育衡 东南大学计算机科学与工程学院副教授. 主要研究方向为机器学习, 数据表示.
E-mail: yhjia@seu.edu.cn
(**JIA Yu-Heng** Associate professor at the School of Computer Science and Engineering, Southeast University. His research interests include machine learning and data representation.)



郭伟峰 郑州大学电气与信息工程学院特聘教授. 主要研究方向为计算智能, 图信号机器学习和生物医学大数据挖掘. E-mail: guowf@zzu.edu.cn
(**GUO Wei-Feng** Distinguished professor at the School of Electrical and Information Engineering, Zhengzhou University. His research interests include computational intelligence, machine learning for graph signals, and biomedical big data mining.)



梁 静 郑州大学电气与信息工程学院教授. 主要研究方向为进化计算, 群集智能, 机器学习和集成学习. 本文通信作者.
E-mail: liangjing@zzu.edu.cn
(**LIANG Jing** Professor at the School of Electrical and Information Engineering, Zhengzhou University. Her research interests include evolutionary computation, swarm intelligence, machine learning, and ensemble learning. Corresponding author of this paper.)