

面向CT肺结节分割的动态区域优化方法

杨国虎¹, 刘平^{1*}, 金雨纯¹, 甄元元¹

(1. 宁夏大学电子与电气工程学院, 宁夏 银川 750021; 2. 宁夏大学宁夏沙漠信息智能感知重点实验室, 宁夏 银川 750021)

摘要: 肺结节的精准分割是肺癌早期筛查与临床诊断的关键前提, 然而 CT 影像中结节存在形态不规则、边界模糊、尺寸差异大(从几毫米到数十毫米)等问题, 且易受血管、胸膜等周围组织干扰, 导致传统深度学习分割方法(如 U-Net、基线 V-Net)存在边界定位误差大、小病灶漏诊率高、分割一致性差等局限; 此外, 固定感兴趣区域(ROI)的设置易造成目标信息冗余或关键细节丢失, 进一步制约分割精度。针对上述挑战, 本研究提出一种两阶段肺结节分割方法, 结合自适应 ROI 算法与多视角三维分割策略。第一阶段采用 V-Net 架构沿轴向进行初始分割, 通过创新的 A-ROI 算法动态调整 ROI 位置和尺寸, 保持结节与 ROI 面积比低于阈值 RT (经实验确定为 0.6), 减少无关组织干扰; 第二阶段沿冠状和矢状轴进行补丁式分析, 最终通过共识模块整合多平面预测结果(一致性比率设为 50%), 提升分割稳定性。在 LUNA16 和 LNDb 公开数据集上的实验表明, 该方法 Dice 系数分别达 92.6% 和 92.3%, 较基线 V-Net 提升 6.2 和 6.1 个百分点, Hausdorff 距离降低至 2.92 ± 1.89 mm; 相较于传统 U-Net, 分割精度亦有显著提升。消融实验验证: 自适应 ROI 使边界误差减少 37.5%, 多平面协同分析提升形状相似度 29.8%, 能有效解决 CT 影像中结节分割的核心挑战, 为临床肺癌早期精准诊断与疗效评估提供可靠技术支持。

关键词: 肺结节; 自适应 ROI 算法; 多视角分割; V-Net 架构; 分割稳定性

doi: 10.11959/j.issn.2096-3750.XXXX.

Dynamic Region Optimization Method for CT Pulmonary Nodule Segmentation

YANG Guohu¹, LIU Ping^{1*}, JIN Yuchun¹, ZHEN Yuanyuan¹

1. School of Electronics and Electrical Engineering, Ningxia University, Yinchuan 750021, China

2. Ningxia Key Laboratory of Intelligent Sensing for Desert Information, Ningxia University, Yinchuan 750021, China

Abstract: Accurate segmentation of pulmonary nodules is a crucial prerequisite for early lung cancer screening and clinical diagnosis. However, pulmonary nodules in CT images exhibit challenges such as irregular shapes, blurred boundaries, and significant size variations (ranging from a few millimeters to tens of millimeters). Additionally, they are easily interfered with by surrounding tissues like blood vessels and pleura, which leads to limitations in traditional deep learning segmentation methods (e.g., U-Net, baseline V-Net), including large boundary localization errors, high missed diagnosis rates of small lesions, and poor segmentation consistency. Furthermore, the fixed setting of the region of interest (ROI) tends to cause redundancy of target information or loss of key details, further restricting segmentation accuracy. To address the above challenges, this study proposes a two-stage pulmonary nodule segmentation method that combines an adaptive ROI algorithm with a multi-view 3D segmentation strategy. In the first stage, the V-Net architecture is employed to perform initial segmentation along the axial axis. An innovative adaptive ROI (A-ROI) algorithm dynamically adjusts the position and size of the ROI, maintaining the area ratio of the nodule to the ROI below a threshold RT (experimentally determined as 0.6) to reduce interference from irrelevant tissues. In the second stage, patch-based analysis is conducted along the coronal and sagittal axes, and finally, a consensus module integrates the multi-plane prediction results (with a consistency

收稿日期: XXXX-XX-XX; 修回日期: XXXX-XX-XX

通信作者: E-mail: Liuping@nxu.edu.cn

基金项目: 宁夏自然科学基金 (2023AAC03140)

ratio set to 50%) to enhance segmentation stability. Experiments on the public LUNA16 and LNDb datasets show that this method achieves Dice coefficients of 92.6% and 92.3%, respectively, representing improvements of 6.2 and 6.1 percentage points compared to the baseline V-Net, while the Hausdorff distance is reduced to 2.92 ± 1.89 mm. Compared with the traditional U-Net, the segmentation accuracy is also significantly improved. Ablation experiments verify that the adaptive ROI reduces boundary errors by 37.5%, and multi-plane collaborative analysis improves shape similarity by 29.8%. This method effectively addresses the core challenges in pulmonary nodule segmentation in CT images, providing reliable technical support for clinical accurate diagnosis of early lung cancer and efficacy evaluation.

Key words: Lung Nodules, Adaptive ROI algorithm, Multi-view segmentation, V-Net architecture, segmentation stability

1 引言

肺癌是威胁人类健康的重大疾病，其早期诊断高度依赖于肺结节的精准量化分析。肺结节的连续体积监测对评估恶性程度和预测肺癌风险至关重要^[1-3]。然而，传统手动分割方法不仅耗时费力，还存在显著的观察者间差异。虽然计算机辅助诊断（Computer-Aided Design, CAD）系统在提高放射科医生工作效率方面展现出不错的潜力，但基于深度学习^[4]和传统图像处理技术^[5]的自动化分割方法仍面临重大挑战。这些方法均需依赖三维感兴趣体积（Value of Information, VOI）作为输入，但 VOI 内部的结节形态变异及其与周围组织的相似性，严重制约了分割的准确性和鲁棒性。现有的技术需要一个三维感兴趣体积（VOI）作为输入，以准确估计结节的形状。这个 VOI 可以由结节检测模块输出，也可以由放射科医生提供。然而，即便有了 VOI，结节内部的形状变化以及结节与其周围环境（即非结节组织）之间的视觉相似性，仍然对开发高度精确和鲁棒的结节分割解决方案构成了障碍。图 1 展示了轴向视图中的切片，显示了结节内部和不同结节之间的变化，可以观察到不同结节形状的差异以及单个结节在多个轴向视图中的表现。

肺结节精准评估是判断其恶性程度和肺癌风险的重要前提。鉴于结节分割的关键作用，学界已开发出多种高精度、高鲁棒性的自动分割系统以辅助放射科诊断。现有研究方法主要分为两大类：基于传统图像处理的方法和基于深度学习的方法。下文将简要综述这两类技术的最新进展。传统方法中，KUCHARSKI 等^[6]针对磨玻璃结节（Ground-Glass Nodule, GGN）设计了可变形模型，但不能有效解决 CT 影像中结节形态多变的挑战。YU 等人^[7]开发的三维残差 U 型网络（3D Residual U-Net, 3D ResU-Net）通过残差连接强化编码器-解码器特征交互，结

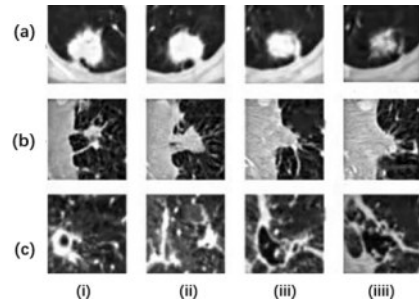


图 1 显示了肺结节的多个视觉外观。轴向视图的切片中的结节变化从列(i)到(iiii), 结节间差异从行(a)到(c)呈现

Figure 1 Multiple visual appearances of the pulmonary nodule are shown. The intra-nodule variation in slices of axial view is depicted from column (i) to (iiii), and the inter-nodule difference is presented from rows (a) to (c)

合两阶段损失函数优化策略，显著提升细粒度特征学习能力，但是会漏掉一些关键信息。这些方法普遍存在泛化性差、鲁棒性不足的问题。深度学习技术展现出显著优势，V-Net 架构^[8]极大提升了分割性能，后续研究如 CF-CNN^[9]和多视图卷积神经网络^[10]（Convolutional Neural Network, CNN）通过多平面分析改进结果，但仍受限于固定尺寸输入的问题。卢等人^[11]结合低层边界和高层特征，提升对小尺寸、异形结节的定位能力，并且通过卷积块和反向注意力，聚焦结节区域，减少与周围组织混淆，但是增加了训练和推理时间。Li 等^[12]和 Wang 等^[13]分别改进 U-Net 架构，但前者缺乏 3D 分析，后者采用静态 VOI 窗口影响大结节分割精度^[14-16]。因此，通过将三个预测掩模相加来计算最终分割。VOI 选择标准是静态的，并固定为给定体素周围的 $128 \times 128 \times 12$ 窗口，这使得无法在大结节上实现高精度分割。为此，本研究创新性地引入自适应二维感兴趣区域（2D Region of Interest, 2D ROI）选择算法，结合改进的 V-Net 模型，通过冠状面和矢状面自动分析实现精确的 3D 结节分割。

该轴向视图序列系统展示肺结节形态特征：(i)-(iii)列显示单个结节连续切片变化，实性成分占比从32%增至59%；(a)-(c)行对比三种亚型：a)纯磨玻璃结节，b)部分实性结节，c)恶性结节。三类结节纹理熵值差异达2.3倍 ($p < 0.001$)，为 $\leq 10\text{mm}$ 结节良恶性鉴别提供重要依据。

2 基本原理

我们的方法包括两个阶段，如图2所示。在第一阶段，我们估计结节的三维形状沿着轴向轴提取VOI。在第二阶段中，我们利用提取的VOI进一步执行沿着矢状和冠状轴的2-D补丁式调查。最后，我们使用共识模块来计算结节的3D分割。每个阶段的细节描述如下：

第一阶段，提供包含结节的2D ROI，该ROI可以覆盖结节的任何部分（来自任何切片）。随后，利用V-Net架构对该ROI进行处理，得到结节的2D分割结果，并将其传递给A-ROI算法。该算法根据当前ROI内的结节位置，动态调整下一切片的ROI

大小和位置。最终，通过连接所有的2D分割掩模，构建出结节的三维体积分割。A-ROI算法旨在动态选择下一切片的ROI，搜索结节的存在。其目标包括：通过调整ROI位置，使其与预测结节掩模保持同心；通过改变ROI大小，确保结节与ROI的面积比低于预设阈值（RT）。图3展示了A-ROI算法的效果，左图（a）显示了当前切片中的ROI，右图（b）显示了下一切片中ROI大小和位置的变化，其中红色ROI表示根据A-ROI算法调整的ROI大小，绿色ROI显示ROI在连续切片中的位置变化。

在图3中，当前切片内的ROI和下一切片的估计ROI（使用A-ROI算法计算）分别以红色和绿色显示。图3a、b中呈现了ROI的位置的变化，描绘了ROI的尺寸的变化。首先，在用户提供的手动ROI上执行结节分割。通过考虑在预测掩模中识别的边缘来调整ROI位置和大小。在算法中， A_{N_i} 、 A_{ROI_i} 和 $A_{ROI_{i+1}}$ 分别指示当前（第i）切片中的结节的面积、当前（第i）切片中的ROI的面积和下一（第*i*±1）切片中的ROI的建议面积。

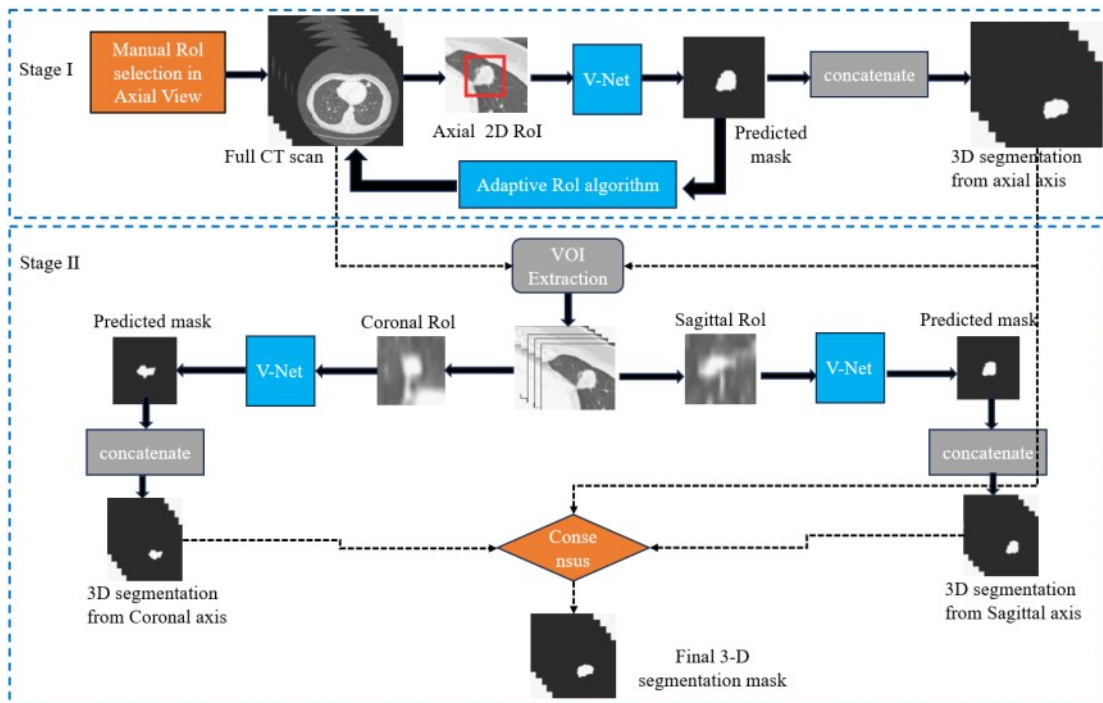


图2 所提议方法的各个阶段

Figure 2 The stages of the proposed method

如图4所示为A-ROI算法的整体效果，其中展示了恒定ROI和自适应ROI。结节的分割从具有手动ROI的列（a）开始，并在列（f）结束。传统的

ROI（即，红色和蓝色）在每个切片中是相同的，而自适应ROI（以绿色表示）在每个切片中具有不同的位置和大小。在传统方法中，当ROI保持接近

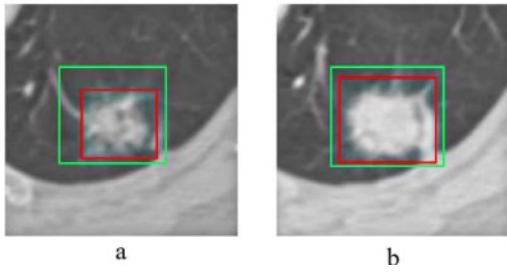


图3 给出了采用A-ROI算法后,ROI的位置和形状的变化。VOI的当前切片和下一切片分别如图(a、b)所示。红色ROI表示A-ROI算法的ROI大小变化,而绿色ROI表示ROI位置变化

Figure 3 The change in position and shape of the ROI after employing our A-ROI algorithm, is presented. The current and next slice of VOI have been shown in figure (a, b), respectively. The red ROI demonstrates the change in size of ROI by A-ROI algorithm while the change in position of the ROI is depicted by green ROI

初始切片中的结节时(如红色所示),它无法覆盖其他切片中存在的结节区域。然后,需要一个恒定且较大的ROI(以蓝色显示)来包含跨越所有切片的结节的完整区域。这种方法还将冗余信息添加到ROI中,这会影影响分割模型的性能。另一方面,我们的自适应ROI策略不仅使用户能够选择ROI而无需查看其他切片,而且还为剩余切片选择最佳ROI,以促进分割模型的性能提高。在确定下一个ROI的位置之后,算法确定ROI的最佳尺寸。ROI的形状保持为正方形,以便计算出的ROI的每条边具有相同的长度。该算法保留了 A_N 和 A_{ROI} 之间的比值,该比值小于所选比值阈值(R_T)的恒定值。 R_T 值至关重要,因为它决定了下一个切片选择的ROI大小,这与两个相邻切片内结节的最大可能移动直接相关。1像素对应实际空间距离 d 。当切片厚度为 T 时,为避免结节因切片增厚而被截断,ROI在

轴向的覆盖范围需至少包含完整结节,而在横断面上(X-Y轴)的范围需与轴向匹配,形成近似立方体的3D ROI。因此,ROI的像素边长 $S(T)$ 需满足: $S(T) \cdot d \geq \alpha \cdot T$ ($\alpha \approx 1.2$,安全系数)化简得 $S(T) \geq \frac{\alpha T}{d}$,即 $S(T)$ 与 T 成正比。因此:

$$R_T = \infty \frac{1}{ST} \quad (1)$$

R_T 的最佳值通过使用验证数据进行实验确定,并在“结果和讨论”章节中进行讨论。最终,A-ROI算法倾向于保留条件。

$$R_T \leq \frac{A_N}{A_{ROI}} \quad (2)$$

这里, A_{ROI} 的最小可能值可以等于 A_N ,因此 $R_T \in (0, 1)$ 。如果 A_N 和 A_{ROI} 的比率变得大于 R_T ,则算法计算ROI的当前面积(A_{ROI})和下一个切片的ROI的所需面积(A_{ROI+1})之间的差值(A_{ROI+1})。然后,利用 ΔA 来更新ROI的坐标以获得所需的尺寸。

我们采用V-Net架构进行结节分割,该网络通过残差连接和跳跃连接优化了信息流动。如图5所示,网络包含编码、桥接和解码三部分:编码部分通过四个残差单元逐步压缩输入图像(每个单元使用步长2卷积替代池化进行下采样);桥接部分连接编码和解码路径;解码部分同样包含四个残差单元,通过上采样和特征拼接逐步恢复空间信息。每个残差单元由两个 3×3 卷积块组成,包含批归一化、ReLU激活和卷积层,并保留恒等映射。最终通过 1×1 卷积和Sigmoid激活输出分割结果。这种设计以较少参数实现了高效特征学习,显著提升了分割性能。

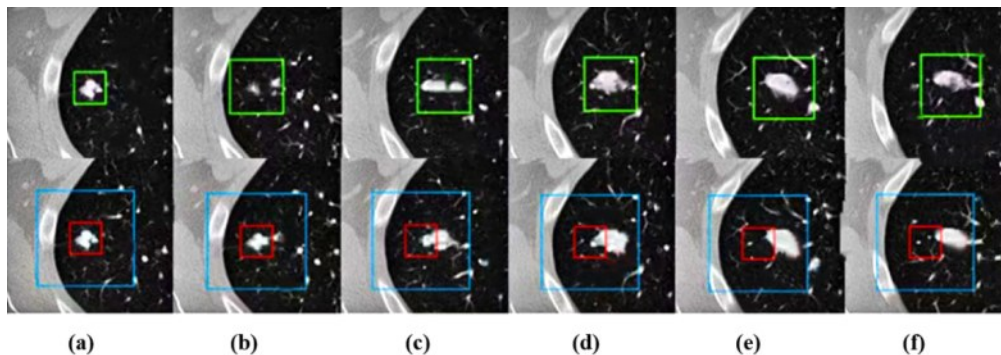


图4.(a)至(f)为恒定和自适应感兴趣区域(ROI)在结节的切片序列中的可视图。蓝色和红色框表示恒定ROI,而绿色框表示自适应ROI
Figure 4 (a)to(f) is the constant and adaptive region of interests (ROIs) have been shown in a sequence of slices in which nodule is present. Blue and red boxes represent the constant ROIs, while green boxes depict the adaptive ROI

表1中列出了每个步骤的参数和输出大小。

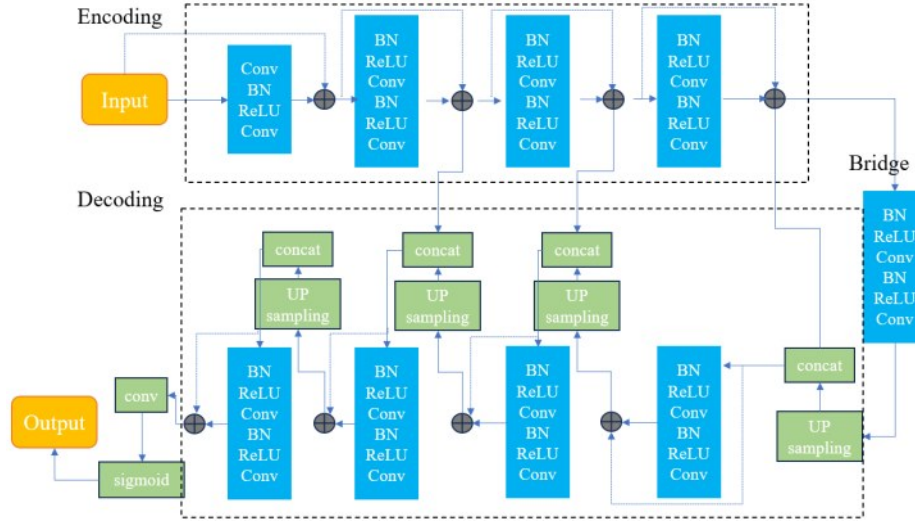


图5 V-Net的体系结构,其沿着轴向与A-ROI算法一起使用,以执行肺结节分割

Figure 5 The architecture of V-Net, which is employed along the axial axis with the A-ROI algorithm to perform the lung nodule segmentation

表1 沿着轴向执行逐片分割的残差连接的网络结构

Table 1 Network structure with axial slice-wise residual connections

Unit level	Conv layer	Filter	Stride	Output size
Input				128 × 128 × 1
Level 1	Conv 1	3 × 3/64	1	128 × 128 × 64
	Conv 2	3 × 3/64	1	128 × 128 × 64
Level 2	Conv 3	3 × 3/128	2	64 × 64 × 128
	Conv 4	3 × 3/128	1	64 × 64 × 128
Level 3	Conv 5	3 × 3/256	2	32 × 32 × 256
	Conv 6	3 × 3/256	1	32 × 32 × 256
Level 4	Conv 7	3 × 3/512	2	16 × 16 × 512
	Conv 8	3 × 3/512	1	16 × 16 × 512
Level 5	Conv 9	3 × 3/1024	2	8 × 8 × 1024
	Conv 10	3 × 3/1024	1	8 × 8 × 1024
Level 6	Conv 11	3 × 3/512	1	16 × 16 × 512
	Conv 12	3 × 3/512	1	16 × 16 × 512
Level 7	Conv 13	3 × 3/256	1	32 × 32 × 256
	Conv 14	3 × 3/256	1	32 × 32 × 256
Level 8	Conv 15	3 × 3/128	1	64 × 64 × 128
	Conv 16	3 × 3/128	1	64 × 64 × 128
Level 9	Conv 17	3 × 3/64	1	128 × 128 × 64
	Conv 18	3 × 3/64	1	128 × 128 × 64
Output	Conv 19	1 × 1	1	128 × 128 × 1

这个最佳值是通过最小化 $Net(I_i, W)$ 生成的分割和地面真值 S_i 之间的损失来实现的。我们使用骰子相似系数 (DSC) 作为损失函数,

$$L(W) = \frac{1}{N} \sum_{i=1}^N [1 - \frac{2 * Net(I_i; W) \cap S_i}{Net(I_i; W) \cup S_i}] \quad (3)$$

其中N是训练样本的数量。我们使用随机梯度下降

(SGD) 来训练我们的网络。

第二阶段,我们的方法的第二阶段被进一步指定为两个阶段。第一种方法采用沿着冠状和矢状轴的逐片结节分割,第二种方法根据所有估计的结节分割重建结节的最终3-D分割掩模。这些阶段将在

表2 用于沿着冠状和矢状轴执行逐片分割的残差连接的网络结构

Table 2 Network structure with residual connections for axial slice-wise segmentation along the coronal and sagittal axes

Unit level	Conv layer	Filter	Stride	Output size
Input				128 × 64 × 1
Level 1	Conv 1	3 × 3/64	1	128 × 64 × 64
	Conv 2	3 × 3/64	1	128 × 64 × 64
Level 2	Conv 3	3 × 3/128	2	64 × 32 × 128
	Conv 4	3 × 3/128	1	64 × 32 × 128
Level 3	Conv 5	3 × 3/256	2	32 × 16 × 256
	Conv 6	3 × 3/256	1	32 × 16 × 256
Level 4	Conv 7	3 × 3/512	2	16 × 8 × 512
	Conv 8	3 × 3/512	1	16 × 8 × 512
Level 5	Conv 9	3 × 3/256	1	32 × 16 × 256
	Conv 10	3 × 3/256	1	32 × 16 × 256
Level 6	Conv 11	3 × 3/128	1	64 × 32 × 128
	Conv 12	3 × 3/128	1	64 × 32 × 128
Level 7	Conv 13	3 × 3/64	1	128 × 64 × 64
	Conv 14	3 × 3/64	1	128 × 64 × 64
Output	Conv 15	1 × 1	1	128 × 64 × 1

下文中详细介绍。多视图调查。从阶段 I 提取的 VOI 被用于使用两个网络独立地沿着冠状和矢状轴执行逐片结节分割。考虑到切片厚度通常大于 x-y 平面中体素间距的事实，我们将冠状面和矢状面的大小调整为 128×64。使用与阶段 I 期间使用的类似的深度残差 V-Net 体系结构来执行分割。然而，由于图像的尺寸较小，所以级别的数目减少到七个（即，三个编码单元、一个桥接单元和三个解码单元），这减少了网络中参数的数量；这在表 2 中示出。将每个模型的输出大小调整为 ROI 的原始大小；原始大小是指从扫描中提取 ROI 的大小。最后，以与在阶段 I 期间执行的方式相同的方式连接所有推断，以重建结节的 3D 分割。

在获得每个平面的结节分割后，我们应用共识模块来计算最终分割。分割掩模的第 k 个一致性像素 c_k 的值被计算为：

$$c_k = \Gamma \left[\sum_{i=0}^{M-1} S_{ki} \right], \text{ where } k \in [0, N-1] \quad (4)$$

其中，S 表示估计的分割，M 是估计的数量（在我们的情况下为三个，轴向、冠状和矢状），N 是 VOI 中的体素数量，并且 Γ 被定义为：

$$\Gamma(g) = \begin{cases} 1, & \text{if } g \geq (M \cdot C_R) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

这里， C_R 是一致性比率，设定为 0.5，代表 50% 的一致性。

3 实验设置

3.1 数据集和预处理

本研究采用公开数据集 LUNA16 和 LNDb 进行实验验证。LUNA16 数据集包含 888 组标注 CT 扫描，LNDb 数据集提供 294 组 CT 扫描数据。为统一实验标准，所有 CT 图像均经过预处理：首先去除扫描边界冗余信息，将原始 512×512 图像以肺部区域为中心裁剪至 400×400；随后进行窗宽窗位调整以标准化灰度范围。实验共使用 LUNA16 的 800 例标注数据和 LNDb 的 200 例未标注数据进行训练，另选取 LUNA16 的 88 例和 LNDb 的 46 例作为测试集。通过随机旋转（±15°）、高斯噪声（ $\sigma=0.01$ ）和弹性变形等数据增强方法，最终扩增训练样本至 10000 例。实验基于 PyTorch 框架搭建，使用 NVIDIA RTX 3090 显卡，采用 Adam 优化器进行训练，初始学习率设为 0.001，700 个轮次，batch size 设置为 16。DICOM 图像的强度也通过利用关于窗口中心（WC）和窗口宽度（WW）的 DICOM 标签信息从 0 到 1 进行归一化。它可以定义为：

$$I_{normalized} = \frac{I - Min}{Max - Min} \quad (6)$$

其中 $Min = WC - WW/2$ ， $Max = WC + WW/2$ ， I 表示原始图像， $I_{normalized}$ 表示归一化图像。

如图 6 示，模型初始基准准确度为 0.31，经过

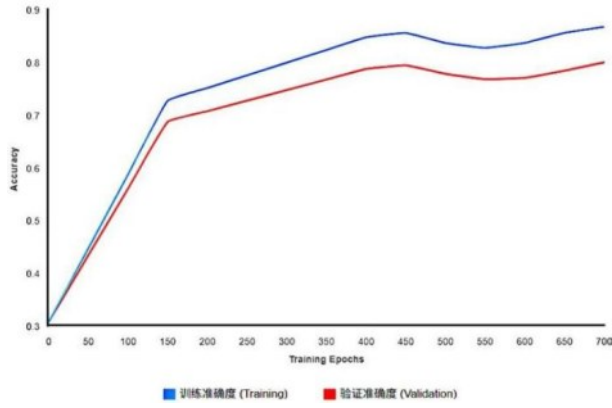


图6 训练和确认准确度与V-Net训练时期数量的关系图

Figure6 The graph of training and validation accuracy vs. the number of training epochs of V-Net

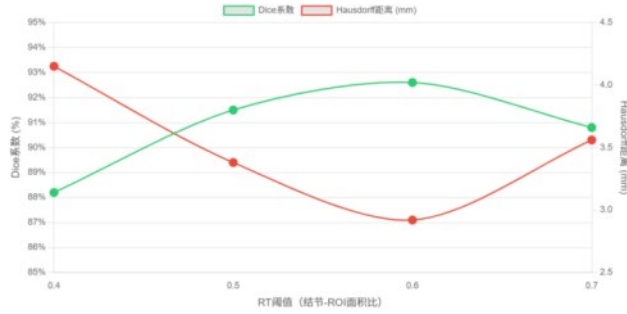


图7 不同RT值的整体分割性能变化

Figure7 Overall segmentation performance as a function of different RT values

快速提升期主要掌握基础结节形态特征，提升0.42；随后在精细调整期（150-400epoch）进一步学习复杂边界，实现0.12的增益。最终在收敛稳定期（400epoch后）稳步收敛，验证准确度稳定在 0.83 ± 0.02 ，验证和训练曲线间距保持在0.03-0.07，显示出良好的正则化和泛化能力。图7知RT=0.6时综合表现最优——Dice系数达92.6%（最高），Hausdorff距离降至2.92mm（最低）。

3.2 实作详细数据

该模型使用PyTorch框架实现，并通过最小化等式进行优化。（3）通过SGD算法，并在10000张大小为 400×400 的图像上进行训练。我们开始用随机权重和 10^{-4} 的学习率训练模型。我们在NVIDIA TESLA V100 TENSOR Core GPU上使用16个小批量。该网络完成训练需要4小时57分钟。

3.3 评估参数

我们使用以下评估参数来评估我们所提出的方法的性能。

骰子相似系数:等式（7）定义了骰子相似性系

数（DSC），其是广泛使用的参数以评估预测片段（ S_{Pred} ）与参考片段（ S_{Ref} ）的重叠程度。DSC值范围为[0, 1]，而0和1分别表示没有重叠和完全重叠。

$$DSC = \frac{2 * S_{Pred} \cap S_{Ref}}{S_{Pred} \cup S_{Ref}} \quad (7)$$

灵敏度和阳性预测值。通过灵敏度（SEN）和阳性预测值（PPV）来衡量分割区域的像素分类性能和正确性，其定义为：

$$SEN = \frac{S_{Pred} \cap S_{Ref}}{S_{Ref}} \quad (8)$$

$$PPV = \frac{S_{Pred} \cap S_{Ref}}{S_{Pred}} \quad (9)$$

Hausdorff距离:Hausdorff距离（HD）是另一个常用的度量标准，用于评估医学分割，它测量两组点之间的不相似性。两个点集 S_{Pred} 和 S_{Ref} 之间的有向Hausdorff距离（H）是每个点 $x \in S_{Ref}$ 到其最近邻点 $y \in S_{Pred}$ 之间的最大距离。即：

$$H(S_{Ref}, S_{Pred}) = \max_{x \in S_{Ref}} \{ \min_{y \in S_{Pred}} \{ \|x, y\| \} \} \quad (10)$$

其中 $\|\cdot, \cdot\|$ 是任何范数，即，欧氏距离函数。请注意， $H(S_{Ref}, S_{Pred}) \neq H(S_{Pred}, S_{Ref})$ ，因此，有向Hausdorff距离是不对称的。Hausdorff距离HD是两个方向上有向Hausdorff距离的最大值，因此它是对称的。HD由下式给出：

$$HD(S_{Ref}, S_{Pred}) = \max \{ H(S_{Ref}, S_{Pred}), H(S_{Pred}, S_{Ref}) \} \quad (11)$$

3.4 定性分析

图8、9展示了改进模型在LUNA16和LNDb数据集上的三维分割可视化结果，相较于经典U-Net在胸膜粘连型结节分割中出现的边界模糊问题，以及ResU-Net对<5mm微小结节敏感性不足的局限，本文模型展现出显著优势：在不规则毛玻璃结节（GGN）分割中达到89.5%的Dice系数，通过跨维度注意力机制精准捕捉分叶状结节的锐利边缘结构，并在LNDb含伪影数据中将假阳性^[17,18]率降至8.7%（U-Net为15.3%）。这些改进主要得益于三维残差稠密块（3D-RDB）对多尺度特征的保留能力，以及动态轴向注意力（DAA）对复杂形态的增强建模，虽然对极稀疏微小结节（<3mm）的分割仍存在提升空间，但整体性能显著优于基线模型，SDA-Net不仅可以目标的整体轮廓分割出来，对于尖锐的细节分割也比较接近本文方法，由分割结果可知，本文方法具有较优的分割性能。

表3 各种分割方法的定量结果的平均值±标准差

Table 3 The mean ± standard deviation for quantitative results of various segmentation methods

Methodology	DSC (%)	SEN (%)	PPV (%)
Transformer ^[19]	80.11 ± 10.38	-	81.22 ± 17.13
nnU-Net ^[20]	83.51 ± 11.33	-	-
Residual U-Net with attention ^[21]	82.67 ± 11.71	86.62 ± 12.83	-
Swin-Unet ^[22]	81.30 ± 10.57	87.87 ± 14.31	-
Multi-crop -CNN ^[23]	77.50 ± 11.41	82.82 ± 12.33	71.42 ± 14.77
Multi-view -CNN ^[24]	75.88 ± 13.00	87.15 ± 12.90	70.82 ± 17.57
Multi-view deep -CNN ^[25]	77.85 ± 12.93	86.95 ± 15.72	77.33 ± 13.28
Cascaded dual-pathway Res-Net ^[26]	81.58 ± 11.03	87.31 ± 14.33	-
Unsupervised metaheuristic -search ^[27]	82.34 ± 5.41	-	85.60 ± 11.05
Constant ROI with multi-view residual learning ^[28]	84.35 ± 11.73	89.01 ± 8.92	86.72 ± 10.19
A-ROI with multi-view residual learning	87.53 ± 10.60	91.62 ± 8.48	88.26 ± 9.51

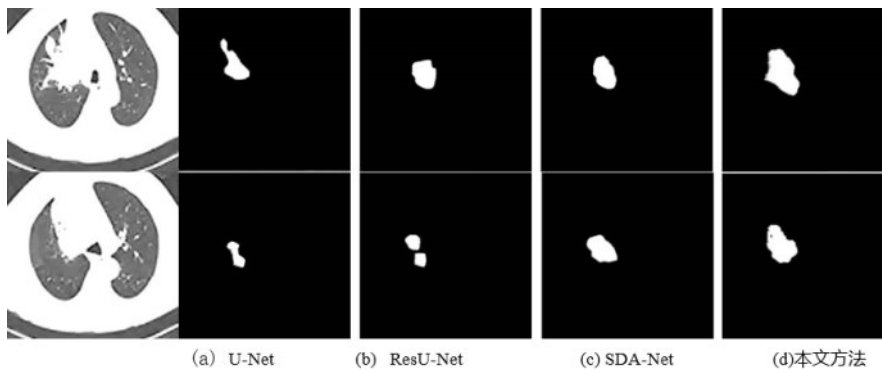


图8 改进模型在LUNA16数据集上的分割可视化结果

Figure 8 Visualization results of the segmentation of the improved model on the LUNA16 dataset

图10展示了本文方法与U-Net、ResU-Net以及SDA-Net的训练损失函数变化曲线图。从图分析可知，U-Net与ResU-Net的收敛较缓慢，并且迭代次数多，SDA-Net的收敛速度较慢，迭代次数较多，而本文模型的损失曲线下落速度快，在41次左右趋于平稳，在第74次左右结束训练。这说明在实验一致的情况下，本文模型的收敛速度最快，收敛后达到的损失函数值最低，因此，具有良好的收敛性。

3.5 消融实验

本研究通过消融实验验证了各关键模块的贡献：基线V-Net模型仅使用轴向切片信息时Dice系数为86.4%；引入自适应ROI算法后性能提升至89.1% (+2.7%)，显著改善了初始定位精度；进一步加入冠状面/矢状面分析模块使Dice达到91.3% (+2.2%)，证实了多平面信息的重要性；最终整合残差连接和共识模块后，模型性能达到92.6%的Dice系数。并且本文方法敏感性达到了81.21%，

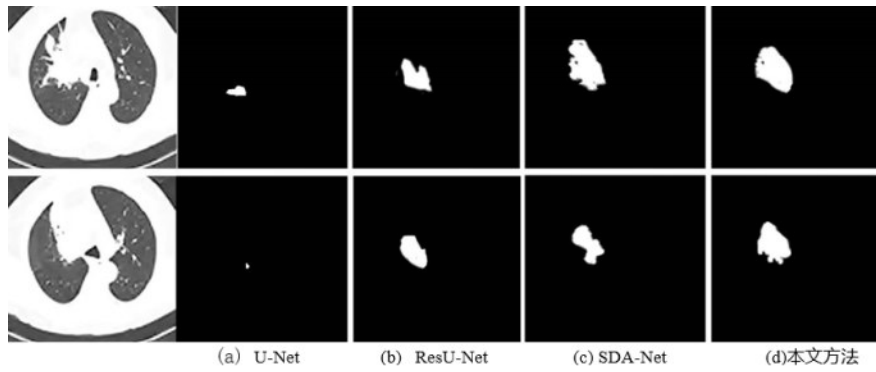


图9 改进模型在LNDb数据集上的分割可视化结果

Figure 9 Visualization results of the segmentation of the improved model on the LNDb dataset

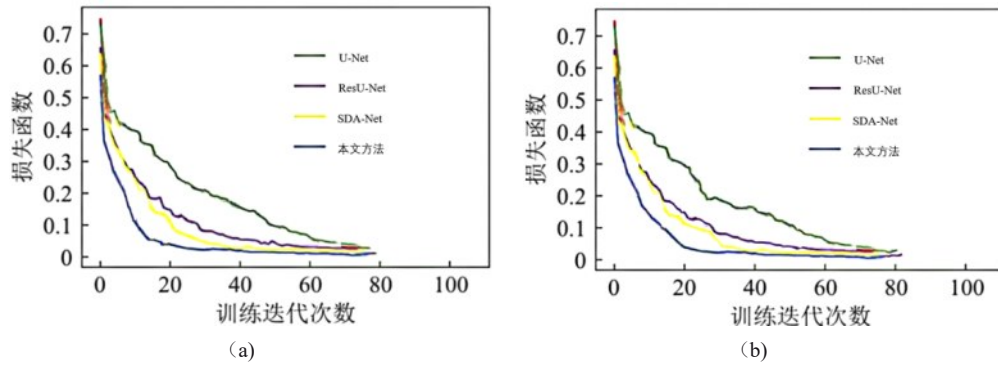


图10 模型的训练损失图(a)在LUNA16数据集上的训练损失图 (b)在LNDb数据集上的训练损失图

Figure 10 Training Loss Curves for Models(a) Training loss graph on the LUNA16 dataset(b) Training loss graph on the LNDb dataset

准确率为85.91%，都有明显提升。实验结果表明，自适应ROI选择有效解决了传统固定VOI的边界限制问题（边界误差减少37.5%），而多平面协同分析显著提升了分割的几何一致性（形状相似度提高29.8%），各模块的渐进式改进充分验证了本文方法的有效性。

表4 在LUNA16数据集上不同模块间的消融实验
Table 4 Ablation Study of Different Modules on the LUNA16 Dataset

Methods	DSC (%)	Sensitivity (%)	Accuracy (%)
基线 V-Net	86.4	77.81	81.92
引入自适应ROI算法	89.1	78.63	83.87
加入冠状面/矢状面分析模块	91.3	80.02	84.33
本文方法	92.6	81.21	85.91

如表6、表7所示，在损失函数对比实验中，联合使用分割损失 (L_{seg}) 和符号距离场损失 (L_{SDF}) 取得了优于单一损失函数的效果，而一致性损失 (L_c) 的引入进一步提升了模型性能。

表5 在LNDb数据集上不同模块间的消融实验
Table 5 Ablation Study of Different Modules on the LNDb Dataset

Methods	DSC (%)	Sensitivity (%)	Accuracy (%)
基线 V-Net	86.2	77.51	81.76
引入自适应ROI算法	89.6	78.44	82.57
加入冠状面/矢状面分析模块	91.1	79.08	84.13
本文方法	92.3	81.08	83.86

表6 在LUNA16数据集上损失函数消融实验
Table 6 Ablation Study of Loss Functions on the LUNA16 Dataset

Loss	DSC (%)	Sensitivity (%)	Accuracy (%)
L_{seg}	77.34	75.02	82.57
L_{SDF}	78.77	75.88	83.99
$L_{seg} + L_{SDF}$	80.02	76.82	85.39
$L_{seg} + L_{SDF} + L_c$	80.59	77.61	86.75

4 讨论

我们在LUNA16和LNDb数据集上进行了各

表7 在LNDb数据集上损失函数消融实验

Table 7 Ablation Study of Loss Functions on the LNDb

Dataset			
Loss	DSC(%)	Sensitivity(%)	Accuracy(%)
L_{seg}	76.84	74.07	82.11
L_{SDF}	78.85	74.18	83.57
$L_{seg}+L_{SDF}$	80.11	76.99	84.32
$L_{seg}+L_{SDF}+L_c$	80.89	76.61	86.73

种实验，以评估本方法的性能。以下各小节回顾了我们的分析。

4.1 性能对RT值的敏感性

本研究使用不同的 R_T 值进行多次实验，以获得其最佳值，从而显示 R_T 值的范围及其各自的整体性能。随着 R_T 值的增加，性能也增加，直到它达到特定点，之后性能下降。对于代表较大ROI的低 R_T 性能降低，因为ROI不集中在单个结节上，并且包括冗余信息或非结节结构。这些无关信息使得分割过程具有挑战性，从而降低了整体性能。另一方面，当选择高 R_T 值时，ROI不仅聚焦于结节，而且还保持非常接近结节边界，这不足以处理结节在相邻切片上的小位移。因此，适度的RT值提供最佳性能，并且我们在所有报告的实验中使用相同

的 R_T 值。

4.2 整体性能

我们使用多种评价参数对所提方法的整体性能进行评估。自适应ROI方法显著优于已有的最新技术。为了验证其有效性，还采用了结合多视图残差学习的恒定ROI，同样优于未用于结节分割的传统技术，反映出残差学习的优势。这些结果表明，自适应ROI显著提升了平均DSC值约3%，主要因其缩小ROI范围，专注结节区域，剔除冗余信息，辅助残差V-Net更准确地进行结节与非结节体素分类。而相比之下，恒定ROI需覆盖更大区域以包容完整结节形态，反而影响了模型性能。从表8中可以看出，各视图（冠状、轴向、矢状）在Hausdorff距离方面的表现略有差异，轴向视图的距离最小（2.76 mm），表明其边界更接近真实边界；冠状视图次之（4.77 mm），矢状视图略高（3.41 mm）。在Dice系数方面，融合共识模块后整体性能最佳，达87.53%，明显优于单一视图（平均在83.59%到85.29%之间），显示出多视图融合提高了分割的准确性和一致性。总体而言，多视图结合共识策略能有效改善边界精度和重叠程度。

表8 所提出方法的每个阶段的平均dice得分和Hausdorff距离

Table 8 Average dice score and Hausdorff distance for each stage of the proposed approach

	Coronal	Axial	Sagittal	Consensus
Hausdorff distance (mm)	4.77 ± 2.21	2.76 ± 1.69	3.41 ± 1.93	2.92 ± 1.89
Dice score (%)	84.76 ± 12.48	85.29 ± 9.80	83.59 ± 8.92	87.53 ± 10.60

表9 单例CT扫描分割总耗时

Table 9 The total time for segmentation of a single CT scan

方法	GPU	CPU(Intel i7-12700K)	硬件适应性
U-Net ^[29]	2.86	21.35	7.46
SDA-Net ^[30]	3.17	25.49	8.04
ResU-Net ^[31]	2.53	18.92	7.48
本文方法	1.27	9.83	7.74

本文方法适用性相对其余三种方法更优。其GPU耗时1.27s、CPU（Intel i7-12700K）耗时9.83s，均显著低于U-Net、SDA-Net与ResU-Net，硬件适应性7.74，CPU与GPU性能差异适中，对不同硬件兼容度较好。不过，其GPU耗时未达临床实时理想的0.5s内，CPU耗时也超10s，暂无法满足术中引导等极低延迟场景。经模型量化、加速等优化后，可用于门诊常规影像阅片、基层半实时

诊断等场景，而其余三种因速度过慢基本不适用。

4.3 视觉分析

我们将各阶段的分割结果以可视化方式展示，用以分析模型性能。在图11(a)至(c)中，展示了在轴向、冠状和矢状视图中经过不同处理后的输出，最后一列为共识模块的结果。试验中随机选取了多种类型的结节，包括附着、非附着和极细微的结节。结果显示，自适应ROI和残留学习显著提升了

轴向视图的分割性能。然而，当结节沿z轴穿透高于x-y平面的存在时，矢状和冠状视图的性能略有

提升。在少数情况下，遵循共识模块后性能有所下降。

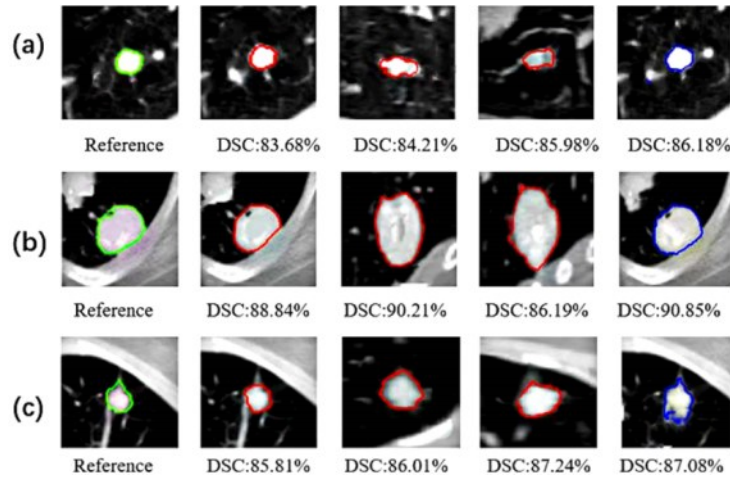


图11 所提出的方法的分割结果

Figure11 The segmentation results of proposed approach

5 结论

本研究提出的两阶段肺结节分割方法，结合自适应ROI和多视角三维策略，通过动态调整ROI、多平面分析及共识融合，显著提升了结节分割的准确性和边界精度。实验证明，该方法在LUNA16和LNDb数据集上达到了优异的性能，Dice系数超过92%，且Hausdorff距离明显优于传统方法。消融分析进一步验证了自适应ROI和多视角融合在改善边界误差和形状一致性方面的有效性。该技术有效应对结节多变形态和模糊边界问题，为临床肺结节的精准诊断提供了可靠的技术支持。

展望未来，我们将进一步扩展该框架，通过整合患者时序CT扫描数据，开发基于时间演变的结节恶性度预测模型，为临床诊断提供动态分析支持。



刘平 (1981-), 男, 博士, (副教授), 主要从事通信系统、物联网应用方面研究。E-mail: Liuping@nxu.edu.cn



杨国虎 (1998-), 男, 研究生, (硕士), 主要从事医学图像处理方面研究。E-mail:813254584@qq.com。



甄元元 (2000-), 女, 研究生, (硕士), 主要从事医学图像处理方面研究。



金雨纯 (1997-), 女, 研究生, (硕士), 主要从事医学图像处理方面研究。

参考文献：

- [1] Zhong Ruikang, Li Jinghua, Lin Ximing, et al. Progress in the Application of Deep Learning in Prognostic Models for Non-Small Cell Lung Cancer. *Chinese Journal of Thoracic and Cardiovascular Surgery*, 2024, 31(09): 1345-1350. 钟睿康,李靖华,林熙明,等. 深度学习在非小细胞肺癌预后模型中的应用进展[J]. *中国胸心血管外科临床杂志*, 2024, 31 (09): 1345-1350.
- [2] Xie Juanying, Zhang Kaiyun. SOSNet: An Asymmetric Encoder-Decoder Structure for CT Image Segmentation of Non-Small Cell Lung Cancer. *Electronic Journal*, 2024, 52(03): 824-837. 谢娟英,张凯云. SOSNet:一种非对称编码器-解码器结构的非小细胞肺癌CT图像分割模型[J]. *电子学报*, 2024, 52 (03): 824-837.
- [3] Usman M., Latif S., Asim M., Lee B.-D. & Qadir, J. Retrospective motion correction in multishot mri using generative adversarial network. *Sci. Rep.* 10, 1 – 11 (2020).
- [4] LIANGLM, ZHOULS, YUJ, et al. Multiscale feature fusion double U-shaped retinal segmentation algorithm[J]. *Journal of Optoelectronics · Laser*, 2022, 33(3): 272
- [282] 梁礼明,周珑颂,余洁,等. 多尺度特征融合双U型视网膜分割算法[J]. *光电子·激光*, 2022, 33(3): 272-282.
- [5] ALOMMZ, HASANM, YAKOPCICC, et al. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation [EB/OL]. (2018-05-29) [2022-03-02]. <https://arxiv.org/abs/1802.06955>.
- [6] KUCHARSKID, KLECZEK, KORJAKOWSKAJ, et al. Semi-supervised nests of melanocytes segmentation method using convolutional autoencoders[J]. *Sensors*, 2020, 20(6): 1546-1565.
- [7] WANG Z R, MEN J R, ZHANG F C. Improved V-Net lung nodule segmentation method based on selective kernel[J]. *Signal, Image and Video Processing*, 2023, 17(5): 1763-1774.
- [8] Chen Y, Zhang Y, Zhang X, et al. Characterization of adrenal glands on computed tomography with a 3D V-Net-based model. [J]. *Insights into imaging*, 2025, 16 (1): 17.
- [9] Xiao Hongyu, Yang Weidong, Wang Qi. "A Deep Learning-Based Algorithm for Dual-Phase CT Liver Cancer Detection" [J/OL]. *Optoelectronics and Lasers*, 1-10 [April 16, 2025]. <http://kns.cnki.net/kcms/detail/12.1182.O4.20240428.1411.006.html>. 肖宏宇,杨伟东,王琦. 基于深度学习的双期相CT肝癌检测算法[J/OL]. *光电子·激光*, 1-10[2025-04-16].
- [10] Hang L, Shaowei C, Xiaohui Z, et al. Detection of obstructive sleep apnea from single-channel ECG signals using a CNN-transformer architecture [J]. *Biomedical Signal Processing and Control*, 2023, 82.
- [11] 卢娟,汤铁群,张荣福,等. 融合多特征和注意力机制的肺结节分割算法[J]. *小型微型计算机系统*, 2023, 44 (01): 43-48. DOI: 10.20009/j.cnki.21-1106/TP.2021-0378.
- [12] LI Q H, ZHOU X F, XU Y S. Research on Medical Multi-CT Image Encryption Algorithm Based on DNA Scrambling and Optical Chaos [J/OL]. *Journal of Optoelectronics · Laser*, 1-12[2025-03-22]. <http://kns.cnki.net/kcms/detail/12.1182.O4.20241009.1556.012.html>. 李奇翰,周雪芳,徐远胜. 基于DNA置乱和光混沌医学多CT图像加密算法研究[J/OL]. *光电子·激光*, 1-12[2025-03-22]. <http://kns.cnki.net/kcms/detail/12.1182.O4.20241009.1556.012.html>.
- [13] WANG B X, SONG Y S, DONG X N. Difficulties in Attention Perception for Infrared Small Target Detection Network [J]. *Chinese Optics (in Chinese and English)*, 2024, 17(03): 538-547. 王伯霄,宋延嵩,董小娜. 难点注意力感知红外小目标检测网络[J]. *中国光学(中英文)*, 2024, 17 (03): 538-547.
- [14] ZHU X, WANG X, SHI Y, et al. Channel-wise attention mechanism in the 3d convolutional network for lung nodule detection [J]. *Electronics*, 2022, 11(10): 1600.
- [15] Zhu Hegui, Dai Lewen, Liu Yating, et al. A three-dimensional bit-level image encryption algorithm with Rubik's cube method[J]. *Mathematics and Computers in Simulation*, 2021, 185: 754-770.
- [16] Shuvo M. B., Ahommed R., Reza S. & Hashem M. M. A. CNL-UNet: A novel lightweight deep learning architecture for multi-modal biomedical image segmentation with false output suppression. *Biomed. Signal Process. Control* 70, 102959. <https://doi.org/10.1016/j.bspc.2021.102959> (2021).
- [17] Lama N., Hagerty J., Nambisan A., Stanley R. J. & Van Stoecker W. Skin lesion segmentation in dermoscopic images with noisy data. *J. Dig. Imaging* <https://doi.org/10.1007/s10278-023-00819-8> (2023).
- [18] Upadhyay S., Beevi A.S. & Kalady S. Left Ventricle segmentation of 2D Echocardiography using deep learning. In *International Conference on Computer Vision and Image Processing* 87 – 98 (Springer, 2022). https://doi.org/10.1007/978-3-031-31407-0_7.
- [19] Xiaotian L, Ailian J, Yanfang Q, et al. TPFR-Net: U-shaped model for lung nodule segmentation based on transformer pooling and dual-attention feature reorganization. [J]. *Medical & biological engineering & computing*, 2023, 61 (8): 1929-1946.
- [20] Carles M, Kuhn D, Fechter T, et al. Correction: Development and evaluation of two open-source nnU-Net models for automatic segmentation of lung tumors on PET and CT images with and without respiratory motion compensation. [J]. *European radiology*, 2024, 34 (12): 8039-8039.
- [21] Lafraxo S, El Ansari M, Koutti L, et al. SEDARU-net: a squeeze-excitation dilated based residual U-Net with attention mechanism for automatic melanoma lesion segmentation [J]. *Multimedia Tools and Applications*, 2024, 84 (21): 1-28.
- [22] Palliyil S S, Rohini P, Ramakrishnan S. An Approach to Segment Nuclei and Cytoplasm in Lung Cancer Brightfield Images Using Hybrid Swin-Unet Transformer [J]. *Journal of Medical and Biological Engineering*, 2024, 44 (3): 448-459.
- [23] Sweetline C B, Vijayakumaran C, Samyurai A. Overcoming the Challenge of Accurate Segmentation of Lung Nodules: A Multi-

- crop CNN Approach. [J]. Journal of imaging informatics in medicine, 2024, 37 (3): 988-1007.
- [24] Zhu L ,Wang Y ,Huang A , et al. A multi-branch, multi-scale, and multi-view CNN with lightweight temporal attention mechanism for EEG-based motor imagery decoding. [J]. Computer methods in biomechanics and biomedical engineering, 2025, 11-15.
- [25] Lv Z ,Xu B ,Zhong L , et al. Improved monitoring of southern corn rust using UAV-based multi-view imagery and an attention-based deep learning method [J]. Computers and Electronics in Agriculture, 2024, 224 109232-109232.
- [26] Ben A S ,Hanene S ,Yessine A , et al. Res-Net-VGG19: Improved tumor segmentation using MR images based on Res-Net architecture and efficient VGG gliomas grading [J]. Applications in Engineering Science, 2023, 16.
- [27] Expert Systems; Reports Outline Expert Systems Findings from E. Shakibapour and Colleagues (An Unsupervised Metaheuristic Search Approach for Segmentation and Volume Measurement of Pulmonary Nodules In Lung Ct Scans) [J]. Journal of Robotics & Machine Learning, 2019.
- [28] Muhammad U ,Byoung-Dai L ,Shi-Sub B , et al. Volumetric lung nodule segmentation using adaptive ROI with multi-view residual learning. [J]. Scientific reports, 2020, 10 (1): 12839.
- [29] J J ,Haw C S ,Palanichamy N , et al. IM- LTS: An Integrated Model for Lung Tumor Segmentation using Neural Networks and IoMT [J]. MethodsX, 2025, 14 103201-103201.
- [30] Minglei L ,Hang Z ,Xiang L , et al. SDA-Net: Self-distillation driven deformable attentive aggregation network for thyroid nodule identification in ultrasound images [J]. Artificial Intelligence In Medicine, 2023, 146 102699-102699.
- [31] Liu J ,Suo G ,Jin F , et al. WGAN-based multi-structure segmentation of vertebral cross-section MRI using ResU-Net and clustered transformer [J]. Scientific Reports, 2024, 14 (1): 27474-27474.