

基于局部上下文关系与自适应感知学习的细粒度肺部分割

全福, 张智高, 王郝日钦, 关强, 王庆虎

(内蒙古民族大学计算机科学与技术学院, 内蒙古通辽 028000)

摘要: 肺部CT图像分割是计算机辅助诊断系统识别肺癌等疾病的基础步骤, 但肺部组织的多样性易造成CT图像的肺部区域出现局部噪点, 且心脏等器官的介入也易导致肺部边缘模糊。为解决上述问题, 提出一种结合局部上下文关系建模与自适应感知学习的U型卷积网络模型。针对肺部区域的局部噪点, 模型通过周边信息对肺部特征构建多层次的上下文关系, 增强网络对多样肺细胞的识别能力。针对边界难以区分的问题, 设计了自适应感知学习模块: 模块设计了混合注意力从通道层面与空间层面引导模型赋予肺部区域更多的关注; 自适应感知模块之间由下至上的特征融合机制增强了模型识别肺部特征的鲁棒性。该方法在LUNA和SHCXR两个肺部分割数据集上取得了98.58%和97.68%分割精度, 较其他分割方法平均提升了0.34%和0.25%, 可为进一步的肺部疾病分析提供有力支持。

关键词: 图像分割; 肺部图像分割; U型网络; 上下文关系学习; 自适应感知学习

DOI: 10.11907/rjdk.241823

开放科学(资源服务)标识码(OSID):

中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2025)001-0115-08



Fine-Grained Lung Segmentation Based on Local Context Relation and Adaptive Perception Learning

QUAN Fu, ZHANG Zhigao, WANG Haoriquin, GUAN Qiang, WANG Qinghu

(College of Computer Science and Technology, Inner Mongolia Minzu University, Tongliao 028000, China)

Abstract: Automatic lung image segmentation is a key step for computer-aided diagnosis systems to detect diseases such as lung cancer. However, the diversity of lung cells easily introduces local noise in the lung region of CT images, and organ interventions like the heart often blur the lung boundaries. To address these issues, this paper proposes a novel U-Net framework that combines local context relation learning modules and adaptive perception modules. To reduce the influence of local noises, this paper constructs a multi-level context relationship for the lung feature extraction by exploiting the surrounding information of specific lung region, which enhances the recognition ability of U-Net to diverse lung cells. To address the issue of blurry lung boundaries, this paper proposes an adaptive perception learning module working at the skipping connections. This module includes a mixed attention mechanism to guide the model for focusing more on the lung area at both the channel and spatial dimensions. In addition, the designed bottom-up feature fusion path can further enhance the robustness of the learned lung features. The proposed method achieves 98.58% and 97.68% accuracy on LUNA and SHCXR dataset, which are 0.34% and 0.25% higher than other segmentation methods on average. The proposed approach can provide help for further analysis of lung diseases.

Key Words: image segmentation; lung image segmentation; U-Net; context relation learning; adaptive perception learning

收稿日期: 2024-09-12

扫描二维码阅读全文:



基金项目: 内蒙古自然科学基金项目(2022MS06029); 内蒙古自治区科技计划项目(2020GG0189); 内蒙古民族大学信息技术赋能智慧城乡建设产学研用创新团队项目(2021)

作者简介: 全福(1984-), 男, 内蒙古民族大学计算机科学与技术学院硕士研究生, 研究方向为医学图像处理; 张智高(1980-), 男, 硕士, 内蒙古民族大学计算机科学与技术学院副教授, 研究方向为推荐系统与大数据处理; 王郝日钦(1988-), 男, 博士, 内蒙古民族大学计算机科学与技术学院高级实验师, 研究方向为智慧农业与自然语言处理; 关强(1989-), 男, 博士, 内蒙古民族大学计算机科学与技术学院讲师, 研究方向为模式识别与图像处理; 王庆虎(1983-), 男, 硕士, 内蒙古民族大学计算机科学与技术学院副教授, 研究方向为计算机视觉与普适计算。本文通讯作者: 王庆虎。

0 引言

科技的发展使得医学影像的解读成为当前疾病诊断过程的关键工作,医疗从业人员可能由于经验欠缺等因素导致解读结果出现偏差,以至出现疾病漏诊和误诊等情况。计算机辅助诊断(Computer Aided Diagnosis, CAD)系统通过计算机视觉和人工智能等技术,有效提高了疾病诊断效率与准确率,为医生的诊治决策提供了有价值的参考^[1-3]。得益于分辨率高、计算角度多等优势,计算机断层扫描(Computed Tomography, CT)结果已广泛应用于CAD系统。CT技术是医生诊断筛查肺部疾病的常用影像技术,在诊断过程中,预先精准分割CT图像中的肺部区域对最终检测识别肺癌等疾病具有重要作用。

肺部在CT图像中具有尺寸和形状不规则、肺部边界模糊等特点,这使得肺部分割成为医学图像分割领域的挑战性任务之一。截至目前,已有研究主要基于传统机器学习方法和深度学习方法完成该任务。传统机器学习方法可分为基于区域的分割方法和基于轮廓的分割方法^[4]。基于区域的分割方法主要通过计算特征相似性引导肺部分割,该类算法的计算速度快,但在处理肺部边界模糊区域时准确度会明显降低^[5-6]。基于轮廓的分割方法需要应用人工知识确定初始轮廓位置,此类方法计算过程稳定,抗噪性能强,但分割效果与初始的轮廓位置关联性高^[7-8]。同时,在应对形态不规则的肺部区域时,基于轮廓的方法容易出现收敛难的问题。

基于传统机器学习算法的肺分割方法需要依赖手工设计特征,无法处理日渐累积的医学图像数据,分割性能受限。以卷积神经网络为代表的深度学习方法突破了这一限制,它能实现端到端的图像分割训练和推理过程,同时能从百万级的医学图像中自动提取代表性特征,分割效果提升明显^[9]。Long等^[10]提出的全卷积网络(Fully Convolutional Network, FCN)是经典的图像分割结构,能满足不同场景下的分割需求。以FCN为基础,Ronneberger等^[11]设计了U型网络(U-Net)架构。U-Net具有对称的编码器与解码器,同级编码特征与解码特征间添加了创新性的跳层连接结构,使其能在少样本的医学数据集上取得较优的分割性能。基于此,陆续有研究人员改进U-Net结构。Zhou等^[12]设计了更密集的跳跃连接,提出了U-Net++网络,以减轻编码器与解码器间的语义鸿沟;Chen等^[13-14]在U-Net结构中融入了擅长处理全局信息的Transformer模块,以捕捉不同肺部区域间的内在关联性。然而,在处理高分辨率的肺部CT图像时,两种网络相较于初始的U-Net计算负载提升明显。Oktay等^[15]提出了Attention U-Net,该网络在解码过程加入注意力机制提取任务相关的显著特征。然而,识别肺部区域往往需要依赖分散在周边的上下文特征,集中在显著区域无法应对肺部组织多样性带来

的噪点污染难题,如图1所示(彩图扫OSID码可见,下同)。其中,绿色框内的肺部区域在CT图像上的显示效果与非肺部区域相似,通过黄色框内提供的上下文信息能提升网络识别该区域为肺部的置信度)。Gu等^[16]在高层特征空间中设计了密集的空洞卷积平行计算模块,该模块能较好地利用上下文融合信息提升肺部区域像素级的分割精度。由于建模的上下文距离过长,该模块适用在高层特征空间提升网络学习抽象肺部特征的能力,忽略了对分割任务极其重要的细节所富有的低层特征,分割效果受限。Li等^[17]设计了跨尺度特征聚合模块在结果层融合不同尺度的肺部特征,然而由于缺乏中间的交互过程,这类过于直接的融合方式可能一定程度地减弱了融合特征的表达能力。上述改进模型难以解决肺部组织细胞多样和肺部区域边界模糊带来的难题,但均具有一定的启发意义和借鉴价值。

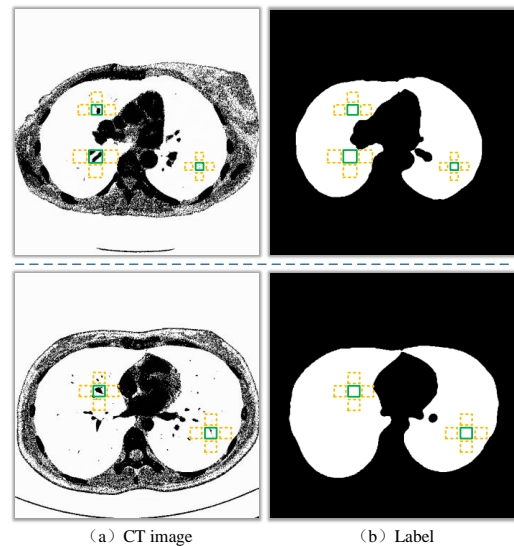


Fig. 1 Illustration of incorporating local context information

图1 结合局部上下文信息的示意图

为解决上述问题,本文基于U-Net提出了一种基于局部上下文关系建模与自适应感知学习的网络框架。首先,在低层到高层的特征空间中设计了分层的局部上下文关系学习模块,并在该模块中融入了中间特征与结果特征融合机制,引导网络在各级编码和解码空间中利用局部上下文信息增强对肺部区域的理解。其次,在U-Net跳层连接中,提出了一种集成混合注意力机制与跨层融合的自适应感知模块。在该模块内,模型从通道和空间视角学习识别肺部区域的显著特征,一定程度地缓解了肺部边缘模糊导致的分割失准;在模块间,跨层融合机制通过结合多层混合注意力特征提升了模型识别肺部区域的稳定性。综上所述,本文主要贡献点如下:①为解决肺部组织多样导致的肺部噪点问题,提出了一种分层的局部上下文关系学习模块,模块充分利用局部上下文信息建模肺部特征表达,在编码空间和解码空间引导模型学习提升CT图像肺部区域的像素级分割精度;②提出了一种自适应感知模块作用

在 U-Net 跳层连接过程, 模块设计了混合注意力和跨层融合机制, 通过给予肺部区域更高的学习权重和跨层特征融合带来的语义丰富度, 提升了网络对肺部边缘的识别精度和鲁棒性; ③在现有肺部分割数据集上的实验结果表明, 本文方法能有效提升 U-Net 网络的整体分割性能, 达到了较优水平。

1 本文方法

本文方法旨在解决肺部 CT 图像分割普遍面临的肺部组织多样和肺部边缘不易区分的难题。所提方法基于经典的 U-Net 框架, 并在该网络的特征空间施加精心设计的引导模块以提升分割精度。本文设计了分层的局部上下文关系学习 (Local Context Relation Learning, LCRL) 模块和自适应感知 (Adaptive Perception, AP) 模块嵌入在 U-Net 网络。前者作用在 U-Net 网络的编码空间和解码空间构建局部上下文关系以提升模型识别肺部区域的准确度; 后者衔接在同级的编码特征与解码特征的跳层连接过程, 旨在缓解由心脏等其他器官介入导致的肺部边缘模糊问题。

1.1 网络结构

本文提出的网络结构如图 2 所示。该网络基于纯粹的 U-Net 结构, 总共包含 5 个层级 (L_1 - L_5) 的编码和解码阶段。输入一张肺部 CT 图像, 依次经过 U-Net 的 L_1 阶段到 L_5 阶段进行编码和解码。编码过程的每个阶段均设有卷积模块 (包含 3×3 卷积和 ReLU 激活函数^[18])、对应层级的 LCRL 模块以及最大池化。编码过程每经过一个阶段, CT 图像特征的大小减半, 通道数变为上一阶段的两倍。与编码过程相反, 解码过程通过上采样操作逐步恢复特征大小, 最后经过 1×1 卷积层和 Sigmoid 激活函数预测每个像素属于肺部的概率。本文设计的 LCRL 模块和 AP 模块属于即插即用模块, 输入特征和输出特征的形状大小一致。LCRL 模块 (图 2 中的蓝色箭头) 作用在编码和解码网络中的 3×3 卷积之后; AP 模块 (图 2 中的绿色模块) 放置在 U-Net 网络各层编码和解码的连接过程中。

1.2 局部上下文关系学习模块

由于肺部细胞的多样性, 局部肺区域在 CT 图像上的视觉显示可能与非肺区域一致, 这为像素级的肺分割带来了极大挑战。在 CT 图像中, 尽管部分肺区域从视觉上难以辨认, 但四周通常也是其他肺部组织, 如图 1 所示。因此, 有效利用上下文信息是解决该问题的突破口。本文提出了局部上下文关系学习 (LCRL) 模块, 该模块在特征空间中可以针对每个待建模的局部区域, 融入周围不同上下文距离的特征以提升分割效果。

具体而言, 本文通过堆叠不同膨胀因子的空洞卷积模块建模上下文关系^[19]。空洞卷积是普通卷积的拓展形式, 可以设定不同的膨胀因子数值改变卷积核作用间距。当膨胀因子值为 1 时, 其与普通卷积一致, 作用面积和卷积

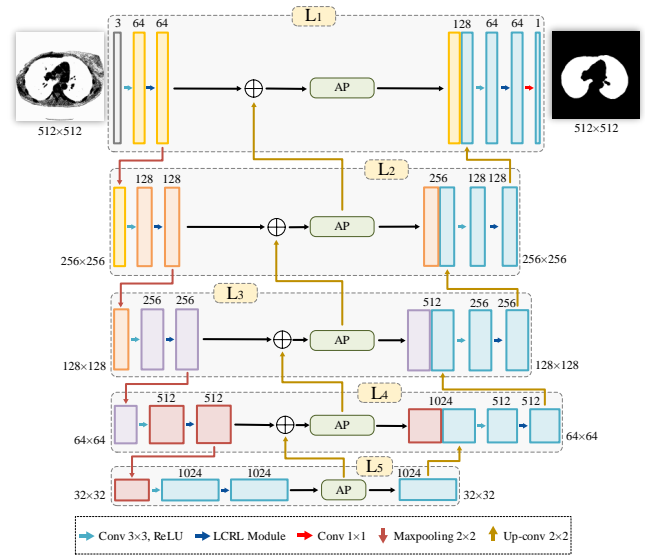


Fig. 2 Structure of the proposed network

图 2 本文提出的网络架构

核大小相同。随着膨胀因子值增大, 卷积核覆盖的上下文距离也越大。本文在 LCRL 模块中设定了 5 种上下文距离建模路径, 并将其归类为 5 个层级 (L_1 - L_5) 分别作用在 U-Net 网络的 5 个编码解码层中, 如图 3 所示。从左至右每条路径使用的空洞卷积单元依次增加, 建模的上下文距离也依次增大。

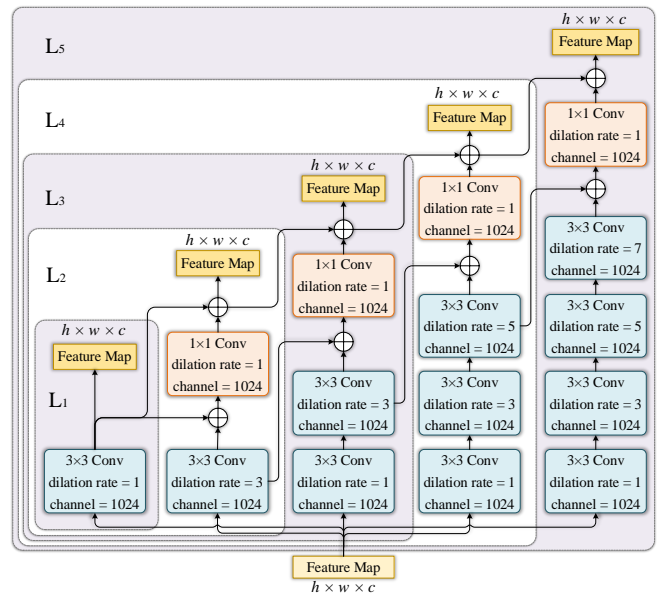


Fig. 3 The proposed local context relation learning module

图 3 本文提出的局部上下文关系模块

通过在 3×3 卷积中设定与膨胀因子值相同的填充距离, LCRL 模块中每个卷积单元的输入特征与输出特征大小保持一致。借助此特性, 本文在 LCRL 模块中嵌入了中间特征融合和结果特征融合机制以提升局部特征的上下文表达能力。对于中间特征融合, 本文设计了由左往右的特征融合策略, 并将该融合结果进一步输入至 1×1 卷积中得到对应路径的最终特征 (除了 L_1 层级)。结果特征融合

机制则将每条路径的输出特征相加融合得到整个LCRL模块的输出特征。相较于初始特征,该特征充分融入了各局部特征的上下文信息,具有更强大的肺部区域识别能力,实验结果也验证了LCRL模块的有效性。

1.3 自适应感知模块

肺部CT扫描结果通常也包含邻近的器官,如心脏和动脉大血管等,因此在成像结果中易出现肺部边缘难以区分等不足。为缓解上述问题,本文从特征学习的视角,在U-Net架构的跳层连接中设计了自适应感知(AP)模块,该模块内部设计了混合注意力学习机制提升网络对肺部区域的识别能力。

AP模块的具体结构如图4所示,由空间注意力(Spatial Attention, SA)、通道注意力(Channel Attention, CA)^[20]和自适应门(Adaptive Gate, AG)组成。本文将SA和CA组合成4种注意力学习模式,包括单一的空间注意力、单一的通道注意力、空间注意力+通道注意力和通道注意力+空间注意力(图4中的紫色线条)。自适应门依据输入肺部CT图像特征自动推导的4种特征权重值(图4中的红色线条),与4种注意力产生的结果融合得到表征肺部能力更强的混合注意力特征。

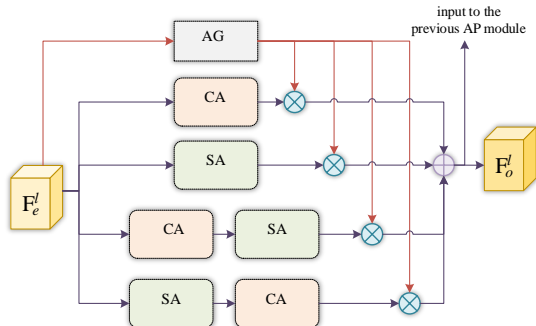


Fig. 4 Structure of adaptive perception module

图4 自适应感知模块结构

在训练过程中,空间注意力单元和通道注意力单元能逐步学习并赋予关键肺部位置和重要通道高权重值,增强模型对图像肺部区域的理解。空间注意力的详细结构如图5所示,计算过程可用如下公式表示:

$$F_s^l = \text{Sigmoid}(\text{conv}(\text{concat}(\text{Avgpool}(F_e^l), \text{Maxpool}(F_e^l)))) \times F_e^l \quad (1)$$

其中, F_e^l 表示输入特征,是U-Net在层级 l 的编码特征,其高、宽和通道大小分别为 h 、 w 和 c 。Avgpool(\cdot)和Maxpool(\cdot)分别表示平均池化和最大池化操作,两者分别计算单个空间位置的平均值和最大值,且都得到大小为 $h \times w \times 1$ 的特征。将其从通道维度合并(式(1)中的concat操作),此时特征大小为 $h \times w \times 2$ 。应用 1×1 卷积(式(1)中的conv(\cdot))将该串联特征转换为大小为 $h \times w \times 1$ 的空间特征,再通过Sigmoid激活函数得到输入特征在每个空间位置的权重大小。将该权重值与输入特征在空间维度进行点对点相乘,得到关键位置增强的肺部特征。

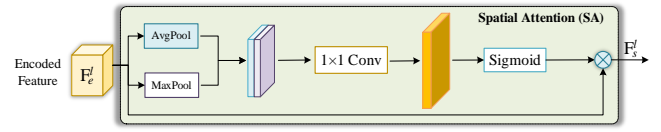


Fig. 5 Structure of spatial attention module

图5 空间注意力结构

与空间注意力的功能类似,通道注意力机制生成肺部特征在每个通道位置的权重大小,其详细结构如图6所示,整个处理过程如下:

$$F_c^l = \text{Sigmoid}(\text{conv}(\text{Avgpool}(F_e^l)) + \text{conv}(\text{Maxpool}(F_e^l))) \times F_e^l \quad (2)$$

通道注意力的输入特征与空间注意力相同。由于输出的权重值对应每个特征通道,因此式(2)中的平均池化和最大池化均作用在通道维度,中间过程产生的特征大小均为 $1 \times 1 \times c$,最后通过Sigmoid函数输出权重,并与输入特征在通道维度相乘得到通道维度的重要信息。

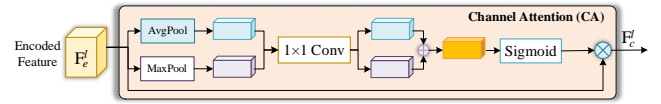


Fig. 6 Structure of channel attention module

图6 通道注意力结构

自适应门依据输入特征产生4种注意力特征的融合权重值,其结构如图7所示,产生权重值的过程可用如下公式表示:

$$w^l = \text{softmax}(\text{FC}(\text{Maxpool}(F_e^l) + \text{Avgpool}(F_e^l))) \quad (3)$$

自适应门中最大池化和平均池化的处理过程与通道注意力相同,均从通道维度提取大小为 $1 \times 1 \times c$ 的特征。FC表示全连接层(Fully Connected Layer),负责将高维向量特征映射至4维空间,最后通过Softmax函数得到4种注意力特征的权重值 w_i^l ($1 \leq i \leq 4$)。 w_i^l 与对应注意力特征相乘相加,得到自适应感知模块的输出特征 F_o^l :

$$F_o^l = F_c^l * w_1^l + F_s^l * w_2^l + F_{c+s}^l * w_3^l + F_{s+c}^l * w_4^l \quad (4)$$

其中, F_c^l 、 F_s^l 、 F_{c+s}^l 和 F_{s+c}^l 分别表示通道注意力的生成特征、空间注意力的生成特征、通道注意力串联空间注意力的生成特征,以及空间注意力串联通道注意力的生成特征。与单一的注意力机制相比,混合注意力特征 F_o^l 能学习更多元的肺部特征表示,在促进模型识别肺部区域的重要信息方面更占优势。

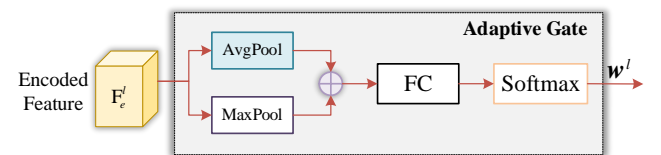


Fig. 7 Structure of adaptive gate

图7 自适应门结构

此外,为了进一步增强U-Net模型学习识别肺部区域的鲁棒性,本文在5个AP模块之间设计了由下至上的特征融合过程。与解码过程相同,融合过程首先将高层特征上

采样为与上层特征大小相同的特征, 然后进行特征相加输入至 AP 模块并作下一步处理。整个融合过程注重将高层的抽象特征逐步与更低层的编码特征融合, 在一定程度上增加了各阶段肺部特征的语义丰富度, 促进了模型分割性能的提升。

1.4 损失函数

肺部 CT 图像分割任务属于像素级的二分类任务, 即对每个像素进行肺部和非肺部划分。本文使用 Dice 系数损失函数引导模型训练过程, 该损失的计算表达如下^[21]:

$$L_{dice} = 1 - \frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \quad (5)$$

其中, N 为整张肺部 CT 图像的像素数量。 p_i 的取值范围为 $[0, 1]$, 表示像素 i 处为肺部区域的预测概率; g_i 的取值范围为 $\{0, 1\}$, 是像素 i 处的真实标签。Dice 系数损失函数通过提升样本预测值与真实值的覆盖率, 能较好地应对图像分割任务中正负样本不平衡现象。

2 实验结果与分析

2.1 数据集与评价指标

2.1.1 数据集

本文使用 Lung Nodule Analysis (LUNA) 竞赛中提供的分割数据集 (<https://www.kaggle.com/kmader/finding-lungs-in-ct-data/data/>) 和 Shenzhen Hospital Chest X-ray (SHCXR) 数据集 (<https://www.kaggle.com/datasets/yoctoman/shcxlung-mask>) 验证所提方法的有效性。LUNA 数据集包含了 2 维和 3 维的肺部 CT 图像和对应的分割标签, 本文仅使用该数据集提供的 2 维图像进行实验, 总共包含 267 张 CT 图像和 267 张对应的标签图像; SHCXR 包含 566 张胸部 X 光图像和对应的标签。本文将两个数据集的 80% 作为训练集, 其余 20% 作为测试集。

2.1.2 评价指标

本文使用 Dice 相似性系数 (Dice Similarity Coefficient, DSC)、交并比 (Intersection over Union, IoU)、灵敏度 (Sensitivity, Sen) 和准确度 (Accuracy, Acc) 评估不同方法的性能。4 种指标的计算公式如下:

$$\begin{aligned} DSC &= \frac{2|G \cap P|}{|G| + |P|} \\ IoU &= \frac{|G \cap P|}{|G \cup P|} \\ Sen &= \frac{TP}{TP + FN} \\ Acc &= \frac{TP + TN}{TP + TN + FP + FN} \end{aligned} \quad (6)$$

其中, G 和 P 分别表示输入样本的真实标签和预测结果; $|G|$ 和 $|P|$ 表示真实标签和预测结果中的元素总和; TP 、 TN 、 FP 和 FN 常用于衡量二分类模型的性能, 它们在本文

中分别表示真值是肺部区域预测结果也是肺部区域的像素数量 (TP)、真值是肺部区域预测结果是非肺部区域的像素数量 (TN)、真值是非肺部区域预测结果是肺部区域的像素数量 (FP) 和真值是非肺部区域预测结果也是非肺部区域的像素数量 (FN)。

在 4 项指标中, DSC 和 IoU 是评估图像分割模型的重要指标^[22]。本文使用 Sen 和 Acc 是为了进一步评估模型的分割鲁棒性。

2.2 实现细节

本文使用 PyTorch 框架实现所提方法, 所有实验均在一块显存为 24 GB 的 Nvidia GeForce RTX 4090 显卡上运行。在训练过程中, 批大小设置为 6, 输入图像的宽和高均为 512。本文使用随机梯度下降算法优化模型, 总共训练 150 轮。在整个训练过程中, 初始学习率 (learning rate, lr) 设置为 0.004, lr 按照以下策略逐步更新:

$$lr = lr * \left(1 - \frac{iter}{max_iter}\right)^{power} \quad (7)$$

其中, $iter$ 表示当前所在的训练轮次, max_iter 表示总共的训练轮数, $power$ 为衰减指数, 在本文中设置为 0.9。本文使用随机水平、垂直和对角翻转策略扩充训练图像数量。同时, 进一步通过图像缩放、随机平移和颜色抖动 3 种随机变换策略预处理训练图像。在测试过程中, 本文未使用任何数据增强策略。

2.3 与其他分割方法比较

为了验证所提方法的有效性, 本文与其他流行的分割方法进行对比, 包括 U-Net^[11]、U-Net++^[12]、R2U-Net^[23]、Attention U-Net^[15] 和 CE-Net^[16]。为保证公平性, 所有对比方法使用的数据集、训练策略和测试策略均与本文一致, 实验结果如表 1 所示。在 LUNA 和 SHCXR 数据集上, 本文方法在 DSC、IoU、Sen 和 Acc 4 项指标上的结果均更优。

Table 1 Result comparison of the proposed method and other methods

方法	DSC		IoU		Sen		Acc	
	LUNA	SHCXR	LUNA	SHCXR	LUNA	SHCXR	LUNA	SHCXR
U-Net ^[11]	0.9594	0.9487	0.9348	0.9047	0.9714	0.9325	0.9814	0.9708
U-Net++ ^[12]	0.9607	0.9494	0.9394	0.9065	0.9701	0.9346	0.9823	0.9750
R2U-Net ^[23]	0.9715	0.9512	0.9524	0.9089	0.9709	0.9353	0.9855	0.9756
Attention U-Net ^[15]	0.9592	0.9501	0.9352	0.9070	0.9642	0.9340	0.9815	0.9751
CE-Net ^[16]	0.9602	0.9504	0.9342	0.9074	0.9718	0.9334	0.9814	0.9750
本文方法	0.9731	0.9515	0.9555	0.9094	0.9735	0.9359	0.9858	0.9768

在 LUNA 数据集上, 本文方法的 DSC 和 IoU 分别达到了 0.973 1 和 0.955 5, 比次优方法 R2U-Net 分别高了 0.16% 和 0.31%; 在 SHCXR 数据集上, 所提方法的 Sen 值达到了 0.935 9, Acc 值为 0.976 8, 与同样具有注意力模块的 Attention U-Net 分别高出了 0.19% 和 0.17%, 与仅注重局部关系建模的 CE-Net 相比两项指标分别高出了 0.25% 和 0.18%。实验结果充分体现了本文方法对不同性质肺部图像的鲁棒性和分割有效性。

为了更直观地展示不同分割方法的效果,本文在测试集中随机挑选了4张CT图像和X光图像的分割结果,如图8和图9所示。

(1)CT图像。如图8所示,绿色框表示该区域的分割效果较优,红色框表示分割一般或较差。在第1行结果中,除本文方法和R2U-Net网络外,其余方法均出现大面积的肺部像素归类错误。在第2行结果中,所有方法在分类肺部像素时,均不同程度地将非肺部区域识别为肺部区域,本文方法的误差最低。在第3行和第4行结果中,本文方法在识别肺部边缘方面优于多数方法。如第3行中右肺中的小孔区域,本文方法对该小孔的分割较为精准。在第4行结果中,标签图像中双肺内侧均出现了不同程度的缺口,本文方法的分割结果对该缺口大小及形状的预测优

于多数对比方法。

(2)X光图像。如图9所示,绿色框表示该区域的分割效果较优,红色框表示分割一般或较差。总体上,本文方法对肺部区域的边缘像素分类与对比方法相比更为准确。在第1行结果中,所提方法更精准地识别了右肺叶左下角的尖角区域。第2行结果中,在右肺叶底部区域,对比方法均不同程度地出现了分割模糊,本文方法对该区域的边缘分割更为平滑。同理,在第3行结果中,本文方法的左肺叶右边缘的分割结果和标签的平滑程度更相似,但在U-Net和Attention U-Net等方法的分割结果中,右边缘均出现了不同程度的缺口。在第4行结果中的左肺叶上半部分区域,本文方法与对比方法相比,分割结果的完整性优势明显。

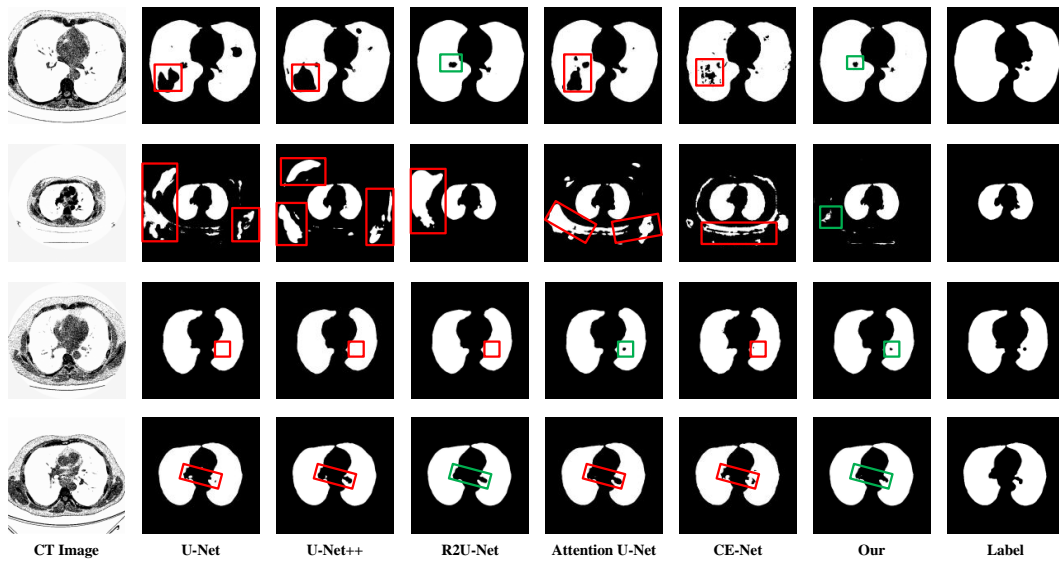


Fig. 8 Segmentation results of different methods on lung CT images

图8 不同方法下的CT肺部图像分割结果

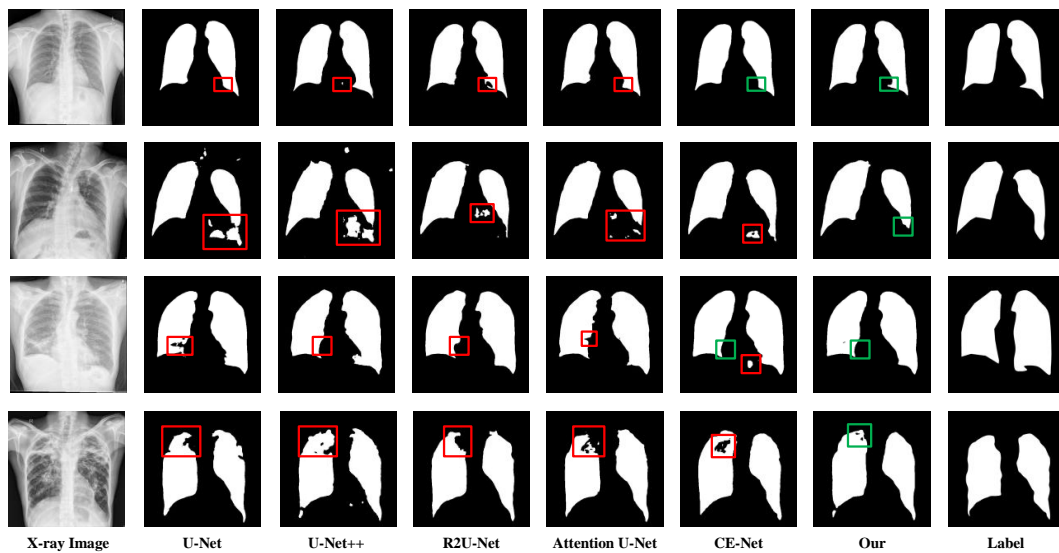


Fig. 9 Segmentation results of different methods on lung X-ray images

图9 不同方法下的X光肺部图像分割结果

在两个数据集上的实验结果充分表明了本文方法在识别肺部区域的准确率更高。主要原因在于:①所提出的

局部上下文关系学习模块能充分结合局部肺组织的周边信息提升模型对肺部区域的识别能力;②所设计的自适应

感知模块能从通道视角和空间视角重视肺部特征区域, 跨层融合机制提升了肺部特征的语义丰富度, 在一定程度上提升了模型对肺部边缘的分割准确率。

2.4 消融实验

为了验证本文方法的有效性, 对不同模块和局部改进的组件进行消融实验, 具体包括 LCRL 模块和 AP 模块、不同层级的 LCRL 模块 (L_1 - L_5)、LCRL 模块的中间特征融合机制、AP 模块中的自适应门和由下至上的跨层特征融合机制。所有实验均在 LUNA 数据集上进行, 实验环境和参数设定与 2.1 节和 2.2 节描述一致。

2.4.1 LCRL 模块与 AP 模块的作用

为了验证 LCRL 模块和 AP 模块的有效性, 本文以 U-Net 网络的实验结果为基准, 分别验证两个模块单独作用和联合作用时的效果, 实验结果如表 2 所示。可以看出, 单独使用两模块后模型的 DSC、IoU、Sen、Acc 分别达到了 0.962 2、0.948 3、0.972 6、0.984 1 和 0.966 2、0.944 5、0.972 7、0.983 7。联合使用两个模块的分割效果提升更为明显, DSC、IoU、Sen、Acc 相较于基准模型 U-Net 分别提升了 1.37%、2.07%、0.21% 和 0.44%, 这充分验证了 LCRL 模块和 AP 模块对提升肺部分割质量具有正向作用。本文认为这是由于: ① LCRL 模块通过局部特征的上下文信息增强了模型对多样肺部细胞的理解; ② AP 模块的混合注意力和跨层融合机制进一步提升了模型在全局视角对 CT 图像肺部区域的识别能力。

Table 2 The influence of LCRL module and AP module

表 2 LCRL 模块与 AP 模块的影响

方法	DSC	IoU	Sen	Acc
U-Net	0.959 4	0.934 8	0.971 4	0.981 4
U-Net + LCRL	0.962 2	0.948 3	0.972 6	0.984 1
U-Net + AP	0.966 2	0.944 5	0.972 7	0.983 7
U-Net + LCRL + AP	0.973 1	0.955 5	0.973 5	0.985 8

2.4.2 不同层级 LCRL 模块的作用

为了验证在不同层级加入 LCRL 模块的作用, 本文将其替换成 3×3 卷积模块 (3×3 卷积和 ReLU 函数) 作为对比, 实验结果如表 3 所示。其中, L_i & L_{i+1} 表示模型在 L_i 层和 L_{i+1} 层均使用了 LCRL 模块。从实验结果可以发现, 在 U-Net 各层使用 LCRL 模块时, 分割性能随着层数的增加逐渐上升, 这表明 LCRL 模块在各层对提升肺部分割效果均有促进作用。本文认为, 这是由于 LCRL 模块可以通过建模局部上下文信息引导模型学习更鲁棒的肺部特征, 从而提升了整体分割性能。

2.4.3 LCRL 模块中间特征融合机制的作用

如表 4 所示, 在 LCRL 模块中使用中间融合机制的 DSC、IoU、Sen 和 Acc 比不使用时分别高出了 0.04%、0.16%、0.02% 和 0.21%, 这表明了在 LCRL 模块中添加中间融合机制的有效性。本文认为这是由于中间融合机制能够进一步增强所学肺部特征的语义丰富度, 从而提升了分割效果。

Table 3 The influence of LCRL module used in different levels

表 3 不同层级 LCRL 模块的影响

方法	DSC	IoU	Sen	Acc
U-Net	0.959 4	0.934 8	0.971 4	0.981 4
U-Net + LCRL (L_1)	0.960 2	0.938 1	0.971 7	0.982 2
U-Net + LCRL (L_1 & L_2)	0.960 7	0.940 6	0.971 9	0.982 7
U-Net + LCRL (L_1 & L_2 & L_3)	0.961 0	0.942 2	0.972 0	0.983 1
U-Net + LCRL (L_1 & L_2 & L_3 & L_4)	0.961 5	0.945 0	0.972 1	0.983 1
U-Net + LCRL (L_1 & L_2 & L_3 & L_4 & L_5)	0.962 2	0.948 3	0.972 6	0.984 1

Table 4 The influence of intermediate feature fusion in LCRL module

表 4 LCRL 模块中间融合机制的影响

方法	DSC	IoU	Sen	Acc
U-Net + LCRL 模块 (无中间融合)	0.961 8	0.946 7	0.972 4	0.982 0
U-Net + LCRL 模块 (有中间融合)	0.962 2	0.948 3	0.972 6	0.984 1

2.4.4 自适应门的作用

为了验证自适应门的有效性, 本文采用均值融合策略作为对照组, 即给 4 种注意力特征设置相同的融合权重值 (本文为 0.25), 实验结果如表 5 所示。结果显示, 通过自适应门融合得到的分割效果比使用均值融合有明显提升, DSC、IoU、Sen 和 Acc 分别高出 0.12%、0.24%、0.05% 和 0.06%, 这充分验证了自适应门的作用。本文认为, 这主要是由于自适应门通过训练后能根据不同的肺部特征分配更精准的融合权重值, 从而提升了肺部分割准确度。

Table 5 The influence of adaptive gate in AP module

表 5 AP 模块中自适应门的影响

方法	DSC	IoU	Sen	Acc
U-Net + AP 模块 (均值融合)	0.965 0	0.942 1	0.972 2	0.983 1
U-Net + AP 模块 (自适应融合)	0.966 2	0.944 5	0.972 7	0.983 7

2.4.5 由下至上特征融合的作用

由表 6 结果可知, 在 AP 模块间加入由下至上的融合机制后, DSC、IoU、Sen 和 Acc 4 项评估指标值分别提高了 0.09%、0.17%、0.05% 和 0.07%, 这表明了在 AP 模块间添加融合机制的有效性。本文认为, 这是由于不同层级特征的肺部信息抽象层级不同, 融合不同层级的肺部特征能够为分割网络提供更综合且更具判别力的肺部特征表示, 从而提升了分割精度。

Table 6 The influence of bottom-to-up feature fusion mechanism between AP modules

表 6 AP 模块间由下至上特征融合机制的影响

方法	DSC	IoU	Sen	Acc
U-Net + AP 模块 (无特征融合)	0.965 3	0.942 8	0.972 2	0.983 0
U-Net + AP 模块 (有特征融合)	0.966 2	0.944 5	0.972 7	0.983 7

3 结语

本文旨在提高肺部 CT 图像的分割性能, 提出了局部上下文关系学习模块和自适应感知模块以缓解普遍遇到的肺部组织多样和肺部边缘模糊难题。局部上下文关系学习模块利用局部特征的周边信息提升网络对肺部区域的像素级识别能力。自适应感知模块包括混合注意力学

习机制和由下至上的特征融合过程,不仅能通过通道维度和空间维度增加网络对肺部区域的识别能力,还能融合不同抽象层级的特征增强肺部特征的表达能力,从而促进网络更好地辨别肺部区域与非肺部区域的差别。实验结果表明,相较于流行的分割方法,本文方法能够明显增强U-Net网络对肺部区域的像素级辨别能力,有效提升了模型分割性能。然而本文方法仅支持处理2D图像,高精度的3D肺部CT图像分割是下一步研究的重点。

参考文献:

- [1] ZHANG X J, CUI Y X, XU Y, et al. Assisted diagnosis of lumbar disc herniation based on deep learning [J]. *Software Guide*, 2024, 23 (7): 34-39.
张新莉, 崔亚轩, 胥义, 等. 基于深度学习的腰椎间盘突出症辅助诊断[J]. *软件导刊*, 2024, 23(7): 34-39.
- [2] GUO B B, GU X L, HU X F, et al. Research on intelligent assistant model for pancreatic cancer diagnosis based on CT image [J]. *Intelligent Computer and Applications*, 2023, 13(10): 9-14.
郭冰冰, 谷雪莲, 胡秀枋, 等. 基于CT图像的胰腺癌智能辅助诊断模型研究[J]. *智能计算机与应用*, 2023, 13(10): 9-14.
- [3] LIU J, ZHANG H X, FU L Y, et al. A preliminary study on the diagnosis of cracked teeth using machine learning [J]. *Journal of China Medical University*, 2023, 52(4): 348-352.
刘洁, 张洪晓, 付朗远, 等. 机器学习方法辅助诊断牙隐裂的研究[J]. *中国医科大学学报*, 2023, 52(4): 348-352.
- [4] FENG L F, CHEN Y, ZHOU T H, et al. Review of human lung and lung lesion regions segmentation methods based on CT images [J]. *Journal of Image and Graphics*, 2022, 27(3): 722-749.
冯龙锋, 陈英, 周滔辉, 等. CT图像肺及肺病变区域分割方法综述[J]. *中国图象图形学报*, 2022, 27(3): 722-749.
- [5] PRASAD J M N, CHAKRAVARTHY S, KRISHNA M V. WITHDRAWN: a novel approach to CAD for the detection of small cell and non-small cell lung cancers [J]. *Materials Today: Proceedings*, 2021(2): 233973290.
- [6] WANG X W, CHENG L Y, HUANG D Y, et al. Segmentation of pulmonary vessels based on MSFM method [C]// *IEEE International Conference on E-health Networking, Application & Services*, 2021: 1-4.
- [7] AFZALI A, BABAPOUR M F, POULADIAN M. 2D statistical lung shape analysis using chest radiographs: modelling and segmentation [J]. *Journal of Digital Imaging*, 2021, 34(3): 523-540.
- [8] MANICKAVASAGAM R, SELVAN S. GACM based segmentation method for lung nodule detection and classification of stages using CT images [C]// *International Conference on Innovations in Information and Communication Technology (ICIICT)*, 2019: 1-5.
- [9] LI Z W, LIU F, YANG W J, et al. A survey of convolutional neural networks: analysis, applications, and prospects [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 33(12): 6999-7019.
- [10] LONG J, SHELFHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 3431-3440.
- [11] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [C]// *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015: 234-241.
- [12] ZHOU Z W, RAHMAN S M M, TAJBAKSH N, et al. UNET++: a nested U-NET architecture for medical image segmentation [C]// *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018: 3-11.
- [13] CHEN J N, LU Y Y, YU Q H, et al. Transunet: transformers make strong encoders for medical image segmentation [DB/OL]. <https://arxiv.org/pdf/2102.04306>.
- [14] CHEN J N, MEI J R, LI X H, et al. TransUNet: rethinking the U-Net architecture design for medical image segmentation through the lens of transformers [J]. *Medical Image Analysis*, 2024, 97: 103280.
- [15] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: learning where to look for the pancreas [EB/OL]. <https://arxiv.org/pdf/1804.03999>.
- [16] GU Z W, CHENG J, FU H Z, et al. CE-NET: context encoder network for 2d medical image segmentation [J]. *IEEE Transactions on Medical Imaging*, 2019, 38(10): 2281-2292.
- [17] LI J H, LI H Y, ZHANG Y, et al. MCNet: a multi-level context-aware network for the segmentation of adrenal gland in CT images [J]. *Neural Networks*, 2024, 170: 136-148.
- [18] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines [C]// *Proceedings of the 27th International Conference on Machine Learning*, 2010: 807-814.
- [19] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [DB/OL]. <https://arxiv.org/pdf/1511.07122>.
- [20] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module [C]// *Proceedings of the European Conference on Computer Vision*, 2018: 3-19.
- [21] MILLETARI F, NAVAB N, AHMADI S A. V-Net: fully convolutional neural networks for volumetric medical image segmentation [C]// *2016 Fourth International Conference on 3D Vision (3DV)*, 2016: 565-571.
- [22] MÜLLER D, SOTO-REY I, KRAMER F. Towards a guideline for evaluation metrics in medical image segmentation [J]. *BMC Research Notes*, 2022, 15(1): 210.
- [23] ALOM M Z, HASAN M, YAKOPCIC C, et al. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation [DB/OL]. <https://arxiv.org/pdf/1802.06955>.

(责任编辑:孙娟)