

基于边缘增强的选择性特征融合肾癌三维CT图像分割

王涛, 白雪飞, 王文剑

引用本文

王涛, 白雪飞, 王文剑. [基于边缘增强的选择性特征融合肾癌三维CT图像分割](#)[J]. 计算机科学, 2025, 52(3): 41-49.

WANG Tao, BAI Xuefei, WANG Wenjian. [Selective Feature Fusion for 3D CT Image Segmentation of Renal Cancer Based on Edge Enhancement](#) [J]. Computer Science, 2025, 52(3): 41-49.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[一种基于TVM的自动调度搜索优化方法](#)

Automatic Scheduling Search Optimization Method Based on TVM

计算机科学, 2025, 52(3): 268-276. <https://doi.org/10.11896/jsjcx.240100126>

[基于元学习的半监督声音事件检测方法](#)

Semi-supervised Sound Event Detection Based on Meta Learning

计算机科学, 2025, 52(3): 222-230. <https://doi.org/10.11896/jsjcx.240100191>

[基于深度学习的气象预报模型研究综述](#)

Survey on Deep Learning-based Meteorological Forecasting Models

计算机科学, 2025, 52(3): 112-126. <https://doi.org/10.11896/jsjcx.240900095>

[基于注意力机制与对比损失的单视图草图三维重建](#)

3D Reconstruction of Single-view Sketches Based on Attention Mechanism and Contrastive Loss

计算机科学, 2025, 52(3): 77-85. <https://doi.org/10.11896/jsjcx.240200102>

[基于区域编码的可驱动头部虚拟化身重建算法](#)

Animatable Head Avatar Reconstruction Algorithm Based on Region Encoding

计算机科学, 2025, 52(3): 50-57. <https://doi.org/10.11896/jsjcx.240200060>

基于边缘增强的选择性特征融合肾癌三维 CT 图像分割

王 涛¹ 白雪飞¹ 王文剑^{2,3}

1 山西大学计算机与信息技术学院 太原 030006

2 计算智能与中文信息处理教育部重点实验室(山西大学) 太原 030006

3 山西警察学院网络安全保卫系 太原 030401

(wt2418385260@163.com)

摘 要 针对肾癌三维 CT 图像存在病变区域多尺度、边缘像素稀疏、对比度低以及肿瘤形状复杂且不规则等问题,提出一种基于边缘增强的选择性特征融合肾癌三维 CT 图像分割网络(EE-SFF U-Net)。EE-SFF U-Net 采用基于 U-Net 的对称编解码网络架构,编码路径中包含一个用于强化边缘信息的边缘增强模块,可有效挖掘、利用浅层特征信息以缓解边缘像素稀疏问题,同时避免小目标的漏检。此外,在网络的跳跃连接中,设计一个选择性特征融合模块,使得深浅层特征相互补充,实现不同信息的有效聚合。最后提出一个综合 Generalized Dice Loss 和 Focal Loss 的混合损失函数,利用动态权重调整策略,实现损失函数的优化训练,并降低病变区域多尺度和肿瘤形状大小不规则带来的影响。所提方法在保证病变区域整体定位准确的同时,强化对小目标特征信息的挖掘利用,从而提高分割的准确性和鲁棒性。在 KiTS19 公开数据集上的实验结果表明,与其他分割算法相比,该方法各项指标表现良好,分割性能有显著提升。

关键词: 肾癌三维 CT 分割;边缘增强;选择性特征融合;3D U-Net;深度学习

中图分类号 TP391

Selective Feature Fusion for 3D CT Image Segmentation of Renal Cancer Based on Edge Enhancement

WANG Tao¹, BAI Xuefei¹ and WANG Wenjian^{2,3}

1 School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China

2 Key Laboratory of Computational Intelligence and Chinese Information Processing (Shanxi University), Ministry of Education, Taiyuan 030006, China

3 Department of Network Security, Shanxi Police College, Taiyuan 030401, China

Abstract Aiming at the problems of multi-scale lesion areas, sparse edge pixels, low contrast, as well as complex and irregular tumor shape in 3D CT images of renal cancer, this paper proposes a selective feature fusion 3D CT image segmentation network based on edge enhancement (EE-SFF U-Net). EE-SFF U-Net adopts the symmetric encoder-decoder network architecture based on U-Net, and the encoding path contains an edge enhancement module for strengthening edge information, which can effectively mine and utilize shallow feature information to alleviate the sparsity problem of edge pixels and avoid missing detection of small targets. In addition, in the skip connections of the network, a selective feature fusion module is designed to make the deep and shallow features complement each other and realize the effective aggregation of different information. Finally, a hybrid loss function with Generalized Dice Loss and Focal Loss is proposed. The dynamic weight adjustment strategy is used to realize the optimal training of the loss function, and to improve the influence of multi-scale lesions and irregular tumor shape and size. The proposed method not only ensures the accuracy of the overall localization of the lesion area, but also strengthens the mining and utilization of small target feature information, so as to improve the accuracy and robustness of segmentation. The experimental results on KiTS19 public dataset show that the proposed method performs well in various indexes and significantly improves the segmentation performance compared with other segmentation algorithms.

Keywords 3D CT segmentation of renal cancer, Edge enhancement, Selective feature fusion, 3D U-Net, Deep learning

到稿日期:2024-03-13 返修日期:2024-08-13

基金项目:国家自然科学基金(U21A20513, 62076154)

This work was supported by the National Natural Science Foundation of China (U21A20513, 62076154).

通信作者:王文剑(wjwang@sxu.edu.cn)

1 引言

肾癌,又称肾细胞癌、肾腺癌,是起源于肾上皮组织的一种恶性肿瘤,占肾恶性肿瘤的 80%~90%。在过去二十年中,肾癌全球发病率每年增加 2%,肾脏肿瘤变得越来越常见^[1]。目前肾癌在我国泌尿生殖系统肿瘤中占第二位,仅次于膀胱癌,占成人恶性肿瘤的 2%~3%、儿童恶性肿瘤的 20%左右^[2]。因此,肾癌早期的及时诊断治疗显得尤为重要。对肾癌的诊断通常是由专业的医生利用磁共振成像(MRI)或者 CT 成像来进行分析,其中 CT 成像由于可以生成具有良好解剖特征的高分辨率图像,并且能够提供具有出色对比度和卓越空间分辨率的图像,成为诊断肾癌的重要工具^[3]。

使用计算机辅助诊断系统来帮助临床医生分析病灶区域的临床状态,可为临床医生在诊断和治疗过程中提供定量依据,快速地分析病情并制定治疗计划。然而,肾癌 CT 图像中的病变区域存在多尺度、边缘像素稀疏、对比度低以及肿瘤形状复杂、不规则等问题,这些问题都给肾癌 CT 图像的三维分割任务带来了较大的困难和挑战。

国内外学者针对肾癌图像的分割任务开展了大量的研究,早期的研究成果包括基于阈值^[4]、基于水平集^[5]、基于活动轮廓^[6]、基于聚类^[7-8]、基于区域增长^[9]等分割算法。这些方法自我学习能力较弱,需要手动调整参数或者人工制定规则才能实现分割目标的提取,无法实现完全的自动分割。而且,图像的分割效果容易受到模糊、噪声等干扰信息的影响。

近年来,由于深度学习算法具有出色的表征学习能力、高度自动化和良好的泛化性能等优点,在肾癌图像分割领域中得到了广泛应用。例如 Çicek 等提出一个 3D U-Net 网络^[10],将传统 U-Net^[11]中的 2D 卷积全部替换为 3D 卷积,输入数据也相应地从二维数据变为三维数据,有效保留了相邻切片之间的空间层面特征信息。Hatamizadeh 等提出了 UNE-TR^[12],采用 Transformer 作为分割网络的主要编码器,可以更有效地捕获远距离依赖关系和上下文信息。这些方法忽略了多尺度目标在编码端下采样过程中随着分辨率逐渐降低而丢失特征信息的问题,无法有效提取边界等局部细节信息,同时模型较为复杂,需要较多运算资源。

如何更好地提取利用边缘信息,降低边缘像素稀疏带来的影响一直是肾癌分割网络设计的难点之一。Zhou 等提出了 U-Net++^[13],通过增加跳跃连接数量将浅层特征与深层特征进行连接,帮助网络更好地保留和利用边缘等细节特征信息。Lin 等设计了一种边缘引导解码器网络 CTO^[14],通过 Sobel 边缘检测算子获得边界信息,对特征图进行边缘增强,然后利用边缘引导模块融合前景和背景信息实现边缘引导^[15]。这些方法本质上都是通过特征传输过程中叠加特征图数量来保证有效特征的传输,但同时会带来冗余特征,导致计算量和训练难度的增加,分割性能提升有限。

为了更好地传输特征信息,减轻冗余特征带来的影响,Schlemper 等提出了 Attention U-Net^[16]网络。该网络在 U-Net 网络跳跃连接部分添加了一个注意力门机制(AG)用来减少冗余特征的传输,避免编码端传输非病灶区域的特征并引导解码器关注特定局部区域的显著特征。Myronenko 等通过在网络中引入残差单元^[17],将每个残差块的输出与输入

相加,使得浅层特征能够更快速地传递到深层,并且在深层中被合并,以生成更高级别的特征表示。这些网络仅仅对编码端传递的特征进行简单过滤,没有将浅层特征与深层特征合理高效地利用起来,并且在特征传输过程中由于没有对特征进行选择,很容易造成错误积累,在应对多尺度目标的分割任务时表现不佳。

针对现有算法中存在的问题,本文提出了一种基于边缘增强的选择性特征融合肾癌三维 CT 图像分割方法(Selective Feature Fusion Based on Edge Enhancement U-Net, EE-SFF U-Net)。该方法采用多层级的特征融合策略,将浅层和深层特征以及边缘特征信息有机地结合起来,充分考虑不同尺度病变区域的全局特征与局部细节,保证分割结果的准确性和稳定性,实现高效、精确的肾癌三维 CT 图像分割。

本文的主要贡献包括以下 3 个方面:

(1)针对肾癌病灶边界难以识别的问题,设计了边缘增强模块。相较于已有方法,本文方法通过融合浅层特征和互补特征以及利用 3 个不同方向的卷积,可有效强化边缘特征提取能力,为后续深层特征信息的挖掘利用奠定基础。

(2)为了减少深浅层特征相互融合产生的特征冗余,提出了选择性特征融合模块。采用通道注意力机制选择性地保留有效特征,以增强网络对多尺度病变区域的学习能力。

(3)考虑到多尺度的肾癌病变区域,设计了混合损失函数。采用动态调整混合损失函数权重的策略,使得网络在训练前期着重于肾脏和肿瘤整体区域的定位,在训练后期则重点关注较难识别的病变区域。在 KiTS19 公开数据集上的实验结果验证了本文方法的优越性。

2 本文工作

本文提出一种基于边缘增强的选择性特征融合肾癌三维 CT 图像分割方法(EE-SFF U-Net),网络整体结构如图 1 所示。

该网络以 3D U-Net 的编码-解码架构为基干,按照特征图分辨率由大到小顺序,编码端共分为 5 层,通过多个连续的卷积层捕捉图像中病灶区域的多尺度特征。随着卷积层的增加,特征图的分辨率逐渐降低而感受野逐渐增大。同时,特征通道数量从 16 逐渐增加至 256,有效提升了网络的特征提取能力。网络右侧 4 层为解码阶段,为了更加准确地实现特征图重建,使用反卷积操作(Up-Conv)来恢复特征图,将输入低分辨率的特征图映射为高分辨率的特征图,特征通道数量减半,逐步恢复特征信息。同时,为了增强网络的鲁棒性和泛化能力,EE-SFF U-Net 在每个卷积层之后都添加了批归一化(Batch Normalization, BN)层用以对特征图进行批标准化处理。

考虑到肾癌三维 CT 图像边缘像素稀疏,以及下采样带来的边缘信息衰减问题,EE-SFF U-Net 在编码端浅层增加了边缘增强模块(Edge Enhancement Module, EEM),通过强化病灶区域的边缘轮廓等细节信息来辅助深层特征的提取与多尺度特征的选择性融合。在跳跃连接的第一层到第四层中加入选择性特征融合模块(Selective Feature Fusion Module, SFFM),通过减少冗余特征的传输,实现多层次、多尺度语义信息的有效聚合。最后将融合后的更加丰富有效的特征信息传输到解码端,反卷积以生成最终分割结果。

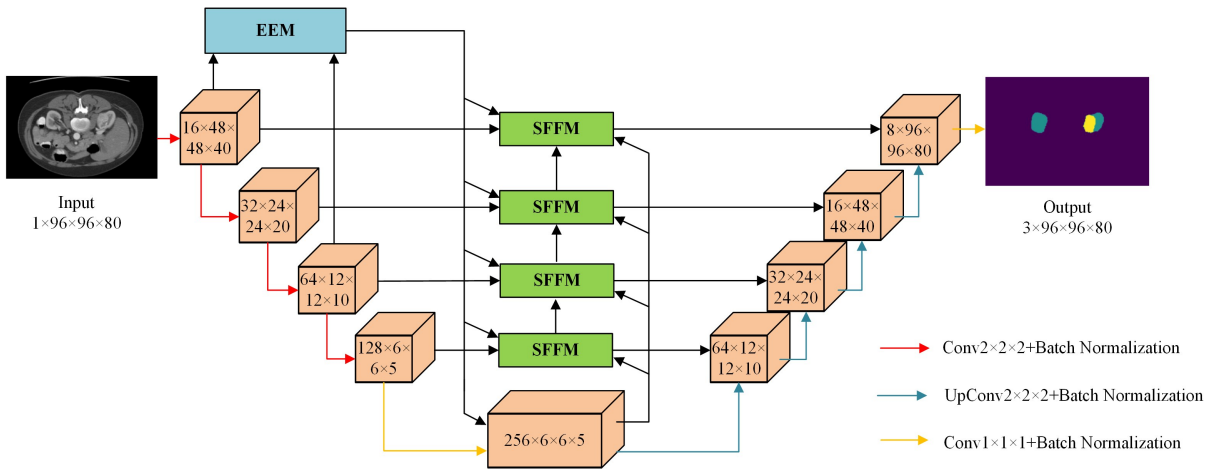


图 1 EE-SFF U-Net 网络结构

Fig. 1 EE-SFF U-Net architecture

2.1 边缘增强模块 (EEM)

现有基于 U-Net 的肾癌分割模型利用连续的卷积和下采样操作扩大感受野并获得更多的全局语义信息。然而,编码端进行下采样后,特征图分辨率逐渐降低,部分细节信息会

丢失,在边缘部分的预测往往不太准确。因此,本文提出一种边缘增强模块 (EEM) 来准确定位病灶的边缘像素区域,在编码端下采样之后增强特征图边缘像素权重,模块结构如图 2 所示。

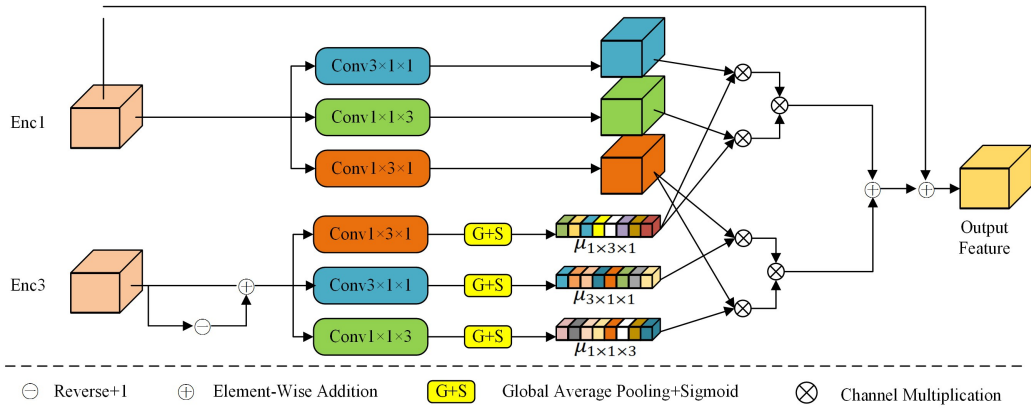


图 2 边缘增强模块

Fig. 2 Edge enhancement module

在网络训练的过程中,浅层特征图包含较多的轮廓、线条等边缘特征信息,深层特征图则由于卷积操作导致边缘信息衰减,获取到的边缘特征信息可能会不一致。因此,EEM 选择编码端第一层和第三层的输出特征图作为输入,分别记为 Enc1 分支和 Enc3 分支。对二者的特征信息进行处理并融合,保证模型获得丰富又准确的边缘特征信息。

为了避免小尺寸的病灶目标被漏检并准确获得特征图的边缘位置信息,EEM 首先对 Enc3 分支的特征图进行 Sigmoid 激活函数处理,然后进行取反加 1 操作,获得原始特征图的互补特征图,并将其与原始特征图求和。通过取反操作,可以有效地利用特征图中的互补信息,防止小尺寸目标漏检,从而获取更为全面的特征表示,进一步提高对病灶区域的分割精度。

Mou 等^[18]的研究表明,使用 $1 \times 1 \times 3$ 卷积、 $1 \times 3 \times 1$ 卷积和 $3 \times 1 \times 1$ 卷积,可以从不同的方向提取曲线结构的边缘信息。受此启发,对 Enc3 分支中互补求和得到的特征图分别经过 $1 \times 3 \times 1$ 卷积、 $3 \times 1 \times 1$ 卷积和 $1 \times 1 \times 3$ 卷积,然后利用全局平均池化和 Sigmoid 函数得到权重图 $\mu_{1 \times 3 \times 1}$ 、 $\mu_{3 \times 1 \times 1}$ 、 $\mu_{1 \times 1 \times 3}$,这些权重图代表了每个通道的重要性。将 Enc1 分支

中的特征图也进行相似的卷积,然后将卷积结果分别乘以上述权值,得到 4 组各个方向边缘特征的集成结果。将集成结果再次相乘并逐像素相加,得到最终的边缘特征集成结果,以此来获取边缘信息丰富的通道,有效捕捉图像中的边缘特征信息。最后将其与富含边缘特征的 Enc1 分支原始特征图相加,作为边缘增强后的特征图输出。EEM 的实现过程如下:

$$F_{EEM1} = [Conv_{3 \times 1 \times 1}(F_{Enc1}) \otimes \mu_{1 \times 3 \times 1}] \otimes [Conv_{1 \times 1 \times 3}(F_{Enc1}) \otimes \mu_{1 \times 3 \times 1}] \quad (1)$$

$$F_{EEM3} = [Conv_{1 \times 3 \times 1}(F_{Enc1}) \otimes \mu_{3 \times 1 \times 1}] \otimes [Conv_{1 \times 3 \times 1}(F_{Enc1}) \otimes \mu_{1 \times 1 \times 3}] \quad (2)$$

$$F_{EEM} = F_{EEM1} + F_{EEM3} + F_{Enc1} \quad (3)$$

其中, $Conv_{3 \times 1 \times 1}(\cdot)$, $Conv_{1 \times 1 \times 3}(\cdot)$ 和 $Conv_{1 \times 3 \times 1}(\cdot)$ 分别表示 $3 \times 1 \times 1$, $1 \times 1 \times 3$, $1 \times 3 \times 1$ 卷积; $\otimes$ 表示通道乘法操作, F_{Enc1} 表示 Enc1 分支的输入特征图; $\mu_{1 \times 3 \times 1}$, $\mu_{3 \times 1 \times 1}$ 和 $\mu_{1 \times 1 \times 3}$ 表示 Enc3 分支获得的权重图; F_{EEM1} 和 F_{EEM3} 分别表示 Enc1 分支和 Enc3 分支的边缘特征集成结果。 F_{EEM} 表示边缘增强后的特征图,即 EEM 的最终输出结果。

2.2 选择性特征融合模块 (SFFM)

针对肾癌三维 CT 图像中存在的多尺度目标问题,为了

避免冗余特征的传输,更有效地结合多尺度、多层次的语义

信息,本文设计了选择性特征融合模块(SFFM),如图3所示。

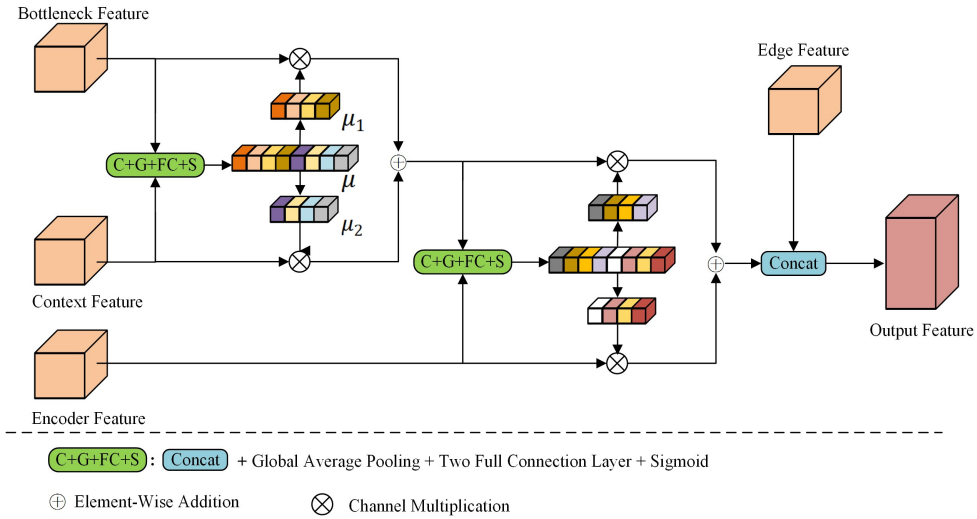


图3 选择性特征融合模块

Fig. 3 Selective feature fusion module

SFFM 的输入主要包含 4 部分特征,分别为:瓶颈模块特征(Bottleneck Feature)、跳跃连接上下文特征(Context Feature)、编码端传递特征(Encoder Feature)和边缘增强模块特征(Edge Feature)。SFFM 放置在第一层到第四层的跳跃连接处,对输入的 4 部分特征进行选择性聚合,充分挖掘和利用多尺度特征信息,在降低冗余的同时保证特征信息的有效传播。为了更加有效地提取深层信息,本文在瓶颈部分引入一个双注意力模块,该模块通过学习与任务有关的特征,准确定位并突出目标区域,以此作为瓶颈模块特征。需要注意的是,由于第四层跳跃连接上下文特征与瓶颈模块特征相同,为了避免冗余,该层 SFFM 的输入为 3 部分特征:瓶颈模块特征、编码端传递特征、边缘增强模块特征,其余操作与其他 SFFM 一致。

SFFM 首先对瓶颈模块特征和跳跃连接上下文特征进行选择性融合。瓶颈模块的主要作用是帮助解码端在上采样过程中尽快恢复特征信息,强制模型学习恢复原始图像所需的有用信息。而跳跃连接中也包含了帮助解码端恢复原始图像的许多重要特征,所以需要对其进行处理,去除冗余特征。SFFM 首先对两部分特征进行通道维度的拼接操作,采用通道注意力机制^[19]经过全局平均池化、两个全连接层、Sigmoid 函数以生成通道权重,对不同特征通道之间的相互依赖关系进行编码。将生成的权重在通道维度进行拆分,得到两部分输入特征各自对应的权重。获得的权重分别与对应的输入特征相乘并逐像素相加,实现选择性融合。由于权重是根据特征通道计算的,因此可以自适应地调整不同输入的特征图,对最终分割任务有帮助的特征通道权重增大,其余通道权重减小,以此实现对特征的选择。

将融合后的特征与编码端传递的特征再次进行上述选择性融合操作。通过传递编码端特征,丰富解码端的特征表示,提高模型对噪声、变形等干扰因素的鲁棒性,使得模型更加稳定,收敛速度更快。随后与边缘增强模块传输过来的特征进行直接拼接,通过增加特征图的通道数来提高模型的表达能力,避免选择性融合丢失增强后的边缘特征信息,有效改善

通道表征利用率不足的问题,提升模型泛化能力,增强对复杂场景以及多尺度目标的适应性。SFFM 的实现过程如下:

$$\mu = \sigma(FC_2(\max(0, FC_1(GAP(F_1 \oplus F_2)))))) \quad (4)$$

$$F_{SFF} = (\mu_1 \oplus F_1) + (\mu_2 \oplus F_2) \quad (5)$$

$$F_{SFFM} = SFF(SFF(F_{Bottleneck}, F_{Context}), F_{Encoder}) \oplus F_{Edge} \quad (6)$$

其中, σ 表示 Sigmoid 函数, FC_1 和 FC_2 表示两个全连接层, $GAP(\cdot)$ 表示全局平均池化, $\oplus$ 表示通道乘法操作, $SFF(\cdot, \cdot)$ 表示进行选择性特征融合操作, $\oplus$ 表示通道拼接操作, F_1 和 F_2 表示 SFFM 输入的两部分特征, μ 为两部分特征按通道拼接的通道权重图, μ_1 和 μ_2 表示两部分输入特征各自的通道权重图, F_{SFF} 表示进行选择性特征融合操作得到的特征图, $F_{Bottleneck}$, $F_{Context}$, $F_{Encoder}$, F_{Edge} 分别表示瓶颈模块特征、跳跃连接上下文特征、编码端传递特征、边缘增强模块特征, F_{SFFM} 表示 4 部分特征经过选择性特征融合获得的特征图,即 SFFM 的最终输出结果。

2.3 混合损失函数

在肾癌三维 CT 图像分割任务中,损失函数的设计和选择也是一项重要的任务。目前, Dice Loss^[20] 是用于评估医学图像分割任务的流行性能指标。Dice Loss 通过目标值和预测值的交叉项来计算损失,可以有效地处理样本不平衡的情况。但是当分割目标较小时,训练过程中的微小变化可能导致 Dice Loss 的剧烈波动并且难以收敛,分割性能较差。Sudre 等提出了 Generalized Dice Loss^[21],该函数考虑到了多分类任务中每个类别在整个图像中所占比例,从而可以处理不同类别之间的权重不平衡问题。相较于 Dice Loss, Generalized Dice Loss 更适合多类分割任务。Focal Loss^[22] 是一种针对类别不平衡的二分类损失函数,与传统的交叉熵损失函数相比,其引入了一个自适应的调节因子,从而能够更好地处理类别不平衡的问题,并且能够有效地抑制易分类样本对损失函数的影响,从而提高难分类样本的学习效果。

考虑到肾癌三维 CT 图像中存在较多的多尺度病变区域,本文设计了一种混合损失函数,在训练过程中动态更新两部分损失函数的权重,保证网络在提取肾脏和肿瘤特征

的同时学习小目标样本的特征,能够实现针对性的监督,从而实现对多尺度病变区域特征的挖掘和完善。混合损失函数定义为:

$$Loss = \alpha Loss_{Generalized\ Dice} + (1 - \alpha) Loss_{Focal} \quad (7)$$

其中, α 是 Generalized Dice Loss 和 Focal Loss 之间的权重平衡参数,初始值为 1。此时,网络着重于病变区域肾脏和肿瘤的特征提取。 α 在训练过程中自动更新,逐渐减小。Generalized Dice Loss 权重逐渐减小为 0, Focal Loss 权重逐渐增加到 1。在准确定位识别肾脏和肿瘤的基础上,网络更加关注难以分割的复杂目标特征的提取,从而进一步提升多尺度病变区域的分割精度。

3 实验内容及结果分析

3.1 实验设置

3.1.1 实验数据与流程

本文使用 KiTS19^[23] 公开数据集对所提算法的分割性能

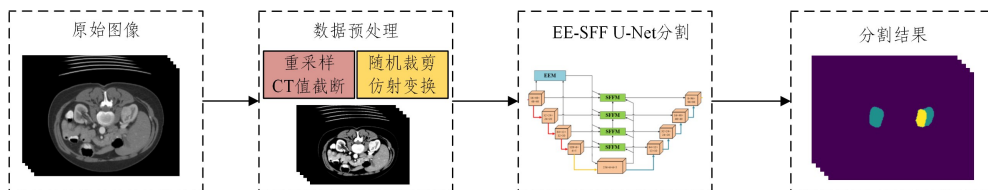


图 4 实验流程图

Fig. 4 Experimental flowchart

3.1.2 实验环境

本文实验基于 Pytorch 进行模型构建,硬件设置显卡型号为 NVIDIA A100 Tensor Core GPU,显存 40GB。深度学习框架为 Pytorch2.1.2 + cu121,图像预处理库为 MONAI^[26]。采用 Adam 优化算法^[27]对网络进行优化^[28],初始学习率为 1×10^{-3} ,权重衰减为 1×10^{-5} 。Max Epoch 为 30 000, Batch Size 为 8。混合损失函数权重平衡参数 α 每过 250 轮更新一次,直至为 0。以上参数设定均来自多次实验结果。对比实验和消融实验在相同环境下进行,参数设置保持一致,以此来验证本文算法的有效性。

3.1.3 评价指标

本文采用 Dice 相似性系数 (Dice Similarity Coefficient, Dice)^[29]、交并比 (Intersection Over Union, IOU)^[30] 和 95% 豪斯多夫距离 (95% Hausdorff Distance, HD95)^[31] 作为算法的性能评价指标。其中, Dice 系数和 IOU 主要用于衡量分割结果与实际结果的重合度,指标值越大,代表分割性能越好。HD95 则主要用于衡量预测分割区域边界与真实区域边界之间的距离,当 HD95 越小时,表示预测边界分割误差越小、分割性能越好。3 种评价指标的定义如下:

$$Dice = \frac{2 \cdot TP}{2 \cdot TP + FN + FP} \quad (8)$$

$$IOU = \frac{TP}{TP + FN + FP} \quad (9)$$

$$HD95 = \max(h(A, B), h(B, A)) \\ = \max\{\max_{a \in A} \{\min_{b \in B} \|a - b\|\}, \max_{b \in B} \{\min_{a \in A} \|b - a\|\}\} \quad (10)$$

进行实验,验证算法的有效性。KiTS19 数据集中包括 168 例训练数据,42 例验证数据,90 例测试数据。数据集中的每个样例都由成熟的专业医生手动标注肾脏器官和病变区域的肿瘤。

实验流程如图 4 所示,首先对原始数据进行预处理,包括重采样和 CT 值截断操作。实验中分别采用双线性插值和最近邻插值的方式进行图像和标签的像素大小调整,将图像中 x, y, z 方向的层间距调整为 (1.2 mm, 1.2 mm, 1.5 mm)^[24],以提高图像的清晰度。为了突出肾脏和病变区域的特征,减轻周围组织对分割任务的干扰,将每个样本序列图像的 CT 值限制在 $[-200, 400]$ 的区间内^[25]。考虑到训练数据尺寸大小不一,将输入图像和标签进行随机裁剪操作,随机裁剪尺寸为 (96, 96, 80)。除此之外,本文使用基于 PyTorch 的仿射变换方法来扩充数据集,包括缩放、旋转、平移等操作。处理完成之后,将数据集传入 EE-SFF U-Net 主网络中进行训练,最终得到分割结果。

其中, TP, FN 和 FP 分别代表真阳性、假阴性和假阳性的像素数量; A 和 B 分别代表预测边界上的点集和真实边界上的点集; $a \in A$ 和 $b \in B$ 分别表示各自点集上的一个点; $\|\cdot\|$ 表示两个点集之间的距离范式。

3.2 对比实验

为了验证本文方法的可行性和有效性,在 KiTS19 数据集上对其进行评估,选择了 5 种算法进行比较,包括经典算法 3D U-Net 以及肾癌分割较为先进的算法 UNETR, U-Net++, Attention U-Net 和 SegResNet。

图 5 给出了本文算法和 5 种对比算法在 KiTS19 数据集上的实验结果。其中,绿色区域表示肾脏的分割结果,黄色区域表示肿瘤的分割结果。根据肾癌 CT 图像中存在的问题,我们将实验数据分为 3 种类型。其中,前两组数据为常规类型,即病变区域大小适中,对比度较大,分割难度相对较低;中间三组数据为多尺度类型,病变区域更加复杂或者尺寸相对较小,模型难以识别;最后两组数据为边缘像素稀疏类型,病变区域与周围对比度较低。从图 5 可以看出,3D U-Net 由于忽略了下采样过程中特征信息衰减的问题,整体分割精度较低,在执行复杂任务时表现较差。UNETR 虽然将 Transformer 作为主要编码器,但是其局部信息的捕获能力较弱,忽略了边缘等细节信息,因此对于小目标以及边缘像素稀疏的病变区域检测效果不佳,对肿瘤整体定位不准确。U-Net++ 考虑到增加跳跃连接数量来细化目标特征信息,但是其传递了过多的冗余信息,所以产生了小目标和边缘像素稀疏的目标漏检以及检测出多个目标区域的问题。Attention U-Net 通过增加注

注意力门控机制来降低特征冗余,对病变区域的整体定位较为准确,但是仍然存在缺乏边缘细节的问题。SegResNet

利用残差连接传输特征信息,但是很容易造成错误积累,出现检测不到目标区域或检测出多个目标区域的问题。

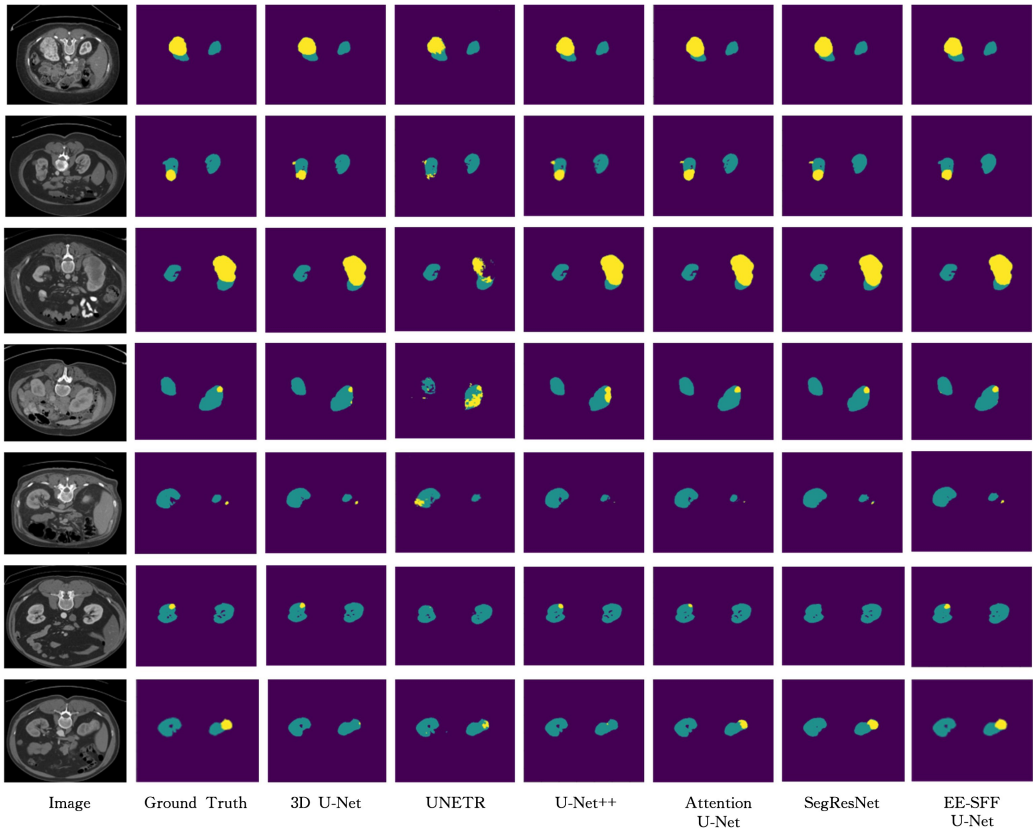


图 5 不同算法分割结果可视化(电子版为彩图)

Fig. 5 Visualization of segmentation results of different algorithms

与上述方法相比,EE-SFF U-Net 在聚合多层次特征信息的同时,强化了边缘特征信息,不仅能够准确识别多尺度的肿瘤,在边缘像素稀疏的区域也能够准确定位边界,在该数据集上的大部分图像中都能够获得更好的分割结果,与真实标签结果较为接近。

表 1 和表 2 列出了在 KiTS19 数据集上不同算法对肾脏和肿瘤测试的各项性能评估结果,加粗字体代表最优性能。从表 1 的结果可以看出,与其他算法相比,本文算法的整体性能最优,在 KiTS19 数据集上的 *Dice*, *IOU*, *HD95* 分别为 94.45%,88.55%和 8.1827 mm。相较于其他算法,EE-SFF U-Net 在 *Dice* 指标上最少提升了 0.53%,而 *IOU* 则属于次优值,与最优值相差不多,*HD95* 最少降低了 2.3090 mm;相较于 3D U-Net,*Dice* 和 *IOU* 分别提升了 2.27%和 2.71%,*HD95* 降低了 3.5930mm。

表 1 KiTS19 数据集上不同算法的肾脏分割性能
Table 1 Kidney segmentation performance of different algorithms on

KiTS19 dataset			
Methods	<i>Dice</i>	<i>IOU</i>	<i>HD95</i> /mm
3D U-Net	0.9218	0.8584	11.7757
UNETR	0.9212	0.8572	18.9129
U-Net++	0.9352	0.8802	21.4158
Attention U-Net	0.9392	0.8873	27.7096
SegResNet	0.9372	0.8834	10.4917
EE-SFF U-Net	0.9445	0.8855	8.1827

表 2 KiTS19 数据集上不同算法的肿瘤分割性能

Table 2 Tumor segmentation performance of different algorithms on KiTS19 dataset

Methods	<i>Dice</i>	<i>IOU</i>	<i>HD95</i> /mm
3D U-Net	0.5514	0.4331	117.4663
UNETR	0.4773	0.3702	129.7753
U-Net++	0.6236	0.5065	69.9455
Attention U-Net	0.6519	0.5370	66.8486
SegResNet	0.6024	0.4954	109.1184
EE-SFF U-Net	0.6733	0.5401	64.5423

观察表 2 中各算法对肿瘤的分割结果可知,在 KiTS19 数据集上,EE-SFF U-Net 的 *Dice*, *IOU*, *HD95* 分别为 67.33%,54.01%,64.5423mm。相较于其他算法,*Dice* 至少提升了 2.14%,*IOU* 至少提升 0.31%,*HD95* 至少降低了 2.3063 mm;相较于 3D U-Net,*Dice* 和 *IOU* 分别提升了 12.19%,10.70%,*HD95* 降低了 52.9240 mm,也证明了本文算法对肾脏肿瘤分割的有效性。

3.3 消融实验

为了验证本文算法各个模块的有效性,在 KiTS19 数据集上对其进行消融实验分析。如表 3 所列,加粗字体代表最优性能。其中,在 3D U-Net+EEM 的消融实验中,由于没有 SFFM 对特征进行处理,因此直接将 EEM 输出的特征图按通道维度拼接在原有跳跃连接传输的编码端特征图上。消融实验中未提到使用混合损失函数时,均采用 DiceCE Loss^[32]作为损失函数。

表 3 消融实验评价指标结果

Table 3 Evaluation index results of ablation experiment			
Methods	Kidney Dice	Tumor Dice	Composite Dice
3D U-Net	0.9294	0.5577	0.7436
3D U-Net+EEM	0.9429	0.6482	0.7956
3D U-Net+EEM+SFFM	0.9445	0.6634	0.8040
3D U-Net+EEM+SFFM+Loss	0.9416	0.6733	0.8075

由表 3 中的结果可知,在 3D U-Net 的基础上添加 EEM、SFFM 以及混合损失函数之后,对于肾脏和肿瘤,Dice 指标均有大幅提升。其中,添加了 EEM 的网络相较于 3D U-Net 在肾脏 Dice、肿瘤 Dice 以及平均 Dice 指标上分别有 1.35%,9.05%和 5.20%的提升。这说明本文提出的 EEM 具有较好的挖掘和利用边界细节特征的能力,能够有效增强边缘特征,尤其是针对复杂不规则的多尺度病变区域,具有更好的识别能力。添加了 EEM 和 SFFM 的网络相较于 3D U-Net 在肾脏 Dice、肿瘤 Dice 以及平均 Dice 指标上分别有 1.51%,10.57%和 6.04%的提升;在添加了 EEM 网络的基础上,则分别提升了 0.16%,1.52%,0.84%。这说明 SFFM 不仅能够将网络深层和浅层特征以及边缘特征有效利用起来,而且能够在其基础上减少冗余,避免无效特征的传输。相较于 3D U-Net,将 EEM、SFFM 以及混合损失函数全部添加到网络中在肾脏 Dice、肿瘤 Dice 以及平均 Dice 指标上分别有 1.22%,

11.56%和 6.39%的提升,与只添加 EEM 和 SFFM 且使用 DiceCE Loss 的网络相比,在肾脏部分有 0.29%的降低,而肿瘤 Dice 和平均 Dice 指标分别提升了 0.99%和 0.35%。这说明本文提出的混合损失函数在训练初期能够保证肾脏和肿瘤区域权重平衡,避免丢失小目标病变区域。随着训练的进行,模型也能够更加专注于难以分割的小目标,有效提升小尺寸病灶的分割性能。

消融实验结果可视化如图 6 所示,其中 EEM U-Net 表示在 3D U-Net 的基础上只添加 EEM,EEM+SFFM U-Net 表示在 3D U-Net 的基础上添加 EEM 和 SFFM 两个模块。

从图 6 可以看出,EEM U-Net 通过强化边缘特征的提取与利用,有效地补充了病变区域的边界信息,提高了对边缘模糊区域的识别能力。EEM+SFFM U-Net 则通过融合多层次、多尺度的特征来丰富特征信息的表示。从第三组可视化结果也可以看到,SFFM 在降低冗余信息的过程中存在将部分肿瘤区域误识别为肾脏的问题,但是从量化结果总体来看,SFFM 可以有效地增强对肾脏和肿瘤区域的识别能力。EE-SFF U-Net 通过在 EEM+SFFM U-Net 的基础上增加混合损失函数,保证在病变区域整体定位准确的同时,强化对难以分割区域的特征提取,有效提升分割精度。

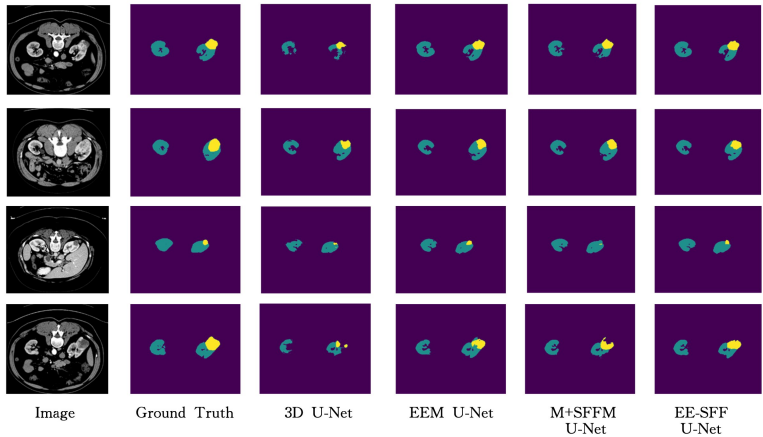


图 6 消融实验结果可视化
Fig. 6 Visualization of ablation experiment results

消融实验的结果验证了本文提出的两种模块以及混合损失函数的有效性,证明其可以很好地适应肾癌三维 CT 图像分割任务,有效提升模型分割精度。

3.4 模型复杂度分析

为了更好地评估模型复杂度,分别对每个模型的浮点运算数以及参数量进行分析,如表 4 所列。

表 4 不同算法的复杂度分析
Table 4 Complexity analysis of different algorithms

Methods	FLOPs	Parameters
3D U-Net	176.39×10^9	35.21×10^6
UNETR	6.88×10^9	92.62×10^6
U-Net++	133.61×10^9	6.98×10^6
Attention U-Net	14.65×10^9	5.91×10^6
SegResNet	3.66×10^9	1.17×10^6
EE-SFF U-Net	197.47×10^9	36.17×10^6

从表 4 可以看出,在所有对比模型中,UNETR 由于在

编码端加入了 Transformer 结构,因此参数量最大,并且综合表 1 和表 2 的分割结果来看,其性能也是最差的;U-Net++ 则因跳跃连接数量众多而运算速度较慢;其余模型复杂度和参数量虽然均低于 3D U-Net,但是在应对复杂分割任务时表现不佳。

EE-SFF U-Net 则由于在 SFFM 中挖掘利用多层次多尺度的特征信息,因此模型运算速度和参数量均有所提升,但其分割性能表现最好。这也侧面说明了 SFFM 的有效性以及本文模型整体的优越性。

结束语 本文针对肾癌三维 CT 图像中存在的病变区域多尺度、边缘像素稀疏以及对对比度低等问题,提出了一种基于边缘增强的选择性特征融合肾癌三维 CT 图像分割网络(EE-SFF U-Net)。该网络构建于 3D U-Net 的基础上,添加了 EEM 和 SFFM 两个模块,能够充分挖掘利用边缘信息并有效地融合深层特征和浅层特征。此外,针对小尺寸病灶难以

分割的问题,本文设计了一种混合损失函数,综合利用 Generalized Dice Loss 和 Focal Loss,在训练过程中动态调整权重,有效提升了小目标病变区域的分割性能。

肾癌 CT 成像由于其图像具有卓越的空间分辨率,成为肾癌诊断的重要工具。然而在实际的临床过程中,获取有真实标签的数据集相对而言较为困难,由专业的医生进行勾画标签比较耗时耗力。因此,未来研究工作中将考虑利用半监督学习或迁移学习等方式进一步优化分割模型,以适应无标签样本或者拥有部分标签样本的数据集,并尝试将分割模型应用于肾癌亚型的分类任务中。

参 考 文 献

[1] ZHANG Y, HUA X L, SHI H Q, et al. Systematic analyses of the role of prognostic and immunological EIF3A, a reader protein, in clear cell renal cell carcinoma[J]. *Cancer Cell International*, 2021, 21(1): 680.

[2] Chinese Anti-Cancer Association Urological Tumor Committee. The 2018 Annual Report of China Cancer[M]. Beijing: People's Medical Publishing House, 2019.

[3] KAUR R, JUNEJA M. A Survey of Kidney Segmentation Techniques in CT Images[J]. *Current Medical Imaging*, 2016, 14(2): 238-250.

[4] LEE H S, HONG H, KIM J. Detection and segmentation of small renal masses in contrast-enhanced CT images using texture and context feature classification[C]// 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). New York: IEEE, 2017: 583-586.

[5] BADURA P, WIECLAWEK W, PYCINSKI B. Automatic 3D Segmentation of Renal Cysts in CT[J]. *Information Technologies in Medicine*, 2016, 47(1): 149-163.

[6] LINGURARU M G, WANG S J, SHAH F, et al. Automated noninvasive classification of renal cancer on multiphase CT[J]. *Medical Physics*, 2011, 38(10): 5738-5746.

[7] BAE K, PARK B, SUN H L, et al. Segmentation of individual renal cysts from MR images in patients with autosomal dominant polycystic kidney disease[J]. *Clinical Journal of the American Society of Nephrology*, 2013, 8(7): 1089-1097.

[8] KAUR R, JUNEJA M, MANDAL A K. A hybrid edge-based technique for segmentation of renal lesions in CT images[J]. *Multimedia Tools and Applications*, 2018, 77(21): 1-21.

[9] IHLESIAS J E, SABUNCU M R. Multi-atlas segmentation of biomedical images: a survey[J]. *Medical Image Analysis*, 2015, 24(1): 205-219.

[10] ÇICEK O, ABDULKADIR A, LIENKAMP S S, et al. 3D U-Net: Learning dense volumetric segmentation from sparse annotation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2016: 424-432.

[11] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer Assisted Intervention (MIC-CAD). Cham: Springer, 2015: 234-241.

[12] HATAMIZADEH A, TANG Y C, NATH V, et al. UNETR:

Transformers for 3D medical image segmentation[C]// 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). New York: IEEE, 2022: 1748-1758.

[13] ZHOU Z W, RAHMAN S 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 (DLIA ML-CDS 2018). Cham: Springer, 2018: 3-11.

[14] LIN Y, ZHANG D, FANG X, et al. Rethinking boundary detection deep learning models medical image segmentation[C]// Information Processing in Medical Imaging (IPMI). Cham: Springer, 2023: 730-742.

[15] LI J, HU J P, QIAO M, et al. An Improved Method for Intensity Inhomogeneity of Brain Image Segmentation[J]. *Journal of Chongqing Technology and Business University (Natural Science Edition)*, 2023, 40(1): 34-39.

[16] SCHLEMPER J, OKTAY O, SCHAAP M, et al. Attention gated networks: Learning to leverage salient regions in medical images[J]. *Medical Image Analysis*, 2019, 53: 197-207.

[17] MYRONENKO A, SIDDIQUEE M, YANG D et al. Automated head and neck tumor segmentation from 3D PET/CT HECKTOR 2022 challenge report[C]// Head and Neck Tumor Segmentation and Outcome Prediction. Cham: Springer, 2022: 31-37.

[18] MOU L, ZHAO Y, FU H. CS2-Net: Deep learning segmentation of curvilinear structures in medical imaging[J]. *Medical Image Analysis*, 2021, 67: 101874.

[19] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 7132-7141.

[20] MILLETARI F, NAVAB N, AHMADI S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation[C]// 3D Vision (3DV). New York: IEEE, 2016: 565-571.

[21] SUDRE C H, LI W, VERCAUTEREN T et al. Generalized Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations[C] // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (DLIA ML-CDS 2017). Cham: Springer, 2017: 240-248.

[22] LIN T, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection[C]// 2017 IEEE International Conference on Computer Vision (ICCV). New York: IEEE, 2017, 324: 2999-3007.

[23] NICHOLAS H, FABIAN I, KLAUS H. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge[J]. *Medical Image Analysis*, 2021, 67: 101821.

[24] ISENSEE F, MAIER-HEIN K H. An attempt at beating the 3D U-Net[J]. *arXiv*: 1908. 02182, 2019.

[25] HSIAO C, SUN T, LIN P, et al. A deep learning-based precision volume calculation approach for kidney and tumor segmentation on computed tomography images[J]. *Computer Methods and Programs in Biomedicine*, 2022, 221: 106861.

[26] CARDOSO M J, LI W, BROWN R, et al. Monai: An open-source framework for deep learning in healthcare [J]. *arXiv*: 2211.

02701,2022.

[27] REYAD M,SARHAN A M,ARAFA M. A modified Adam algorithm for deep neural network optimization[J]. Neural Computing and Applications,2023,35(23):17095-17112.

[28] ARWA M,AHMED A,FADWA A. The Impact of Multi-Optimizers and Data Augmentation on TensorFlow Convolutional Neural Network Performance[C]// 2018 IEEE Conference on Multimedia Information Processing and Retrieval(MIPR). IEEE,2018:140-145.

[29] ZHAO R R,ZHOU Z J,YAO N,et al. Some novel Dice similarity measures for picture fuzzy sets and their applications[J]. Engineering Applications of Artificial Intelligence, 2024, 138: 109385.

[30] YU J,JIANG Y,WANG Z,et al. Unitbox: An advanced object detection network[C]//Proceedings of the 24th ACM International Conference on Multimedia. 2016:516-520.

[31] ZHANG W L,LI R J,DENG H T,et al. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation[J]. NeuroImage,2015,108:214-224.

[32] MINAEE S,BOYKOV Y,PORIKLI F,et al. Image Segmentation Using Deep Learning:A Survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2022,44(7):3523-3542.



WANG Tao, born in 2000, master candidate, is a member of CCF (No. P3207G). His main research interests include machine learning and image processing.



WANG Wenjian, born in 1968, Ph.D, professor, is an outstanding member of CCF(No. 16143D). Her main research interests include machine learning, computing intelligence and image processing.

(责任编辑:何杨)

“普通高等学校人工智能通识”系列教材征集启事

为加快推进自主知识体系、学科专业体系、教材教学体系建设,全面加强教材建设和管理,系统构建中国特色、世界水平的高等教育本科规划教材体系,CCF 联合机械工业出版社组织成立“普通高等学校人工智能通识”系列教材编委会,打造人工智能通识课程体系,赋能理工农医文等各类人才培养。

现面向全国高校教师征集“普通高等学校人工智能通识”系列教材编委会编委与选题,具体要求如下。

一、教材定位与内容要求

系列教材面向普通高等学校(如理科、工科、商科、文科、医学类、农学类、艺术类等)低年级学生,旨在培养大学生具备了解人工智能、使用人工智能和创新人工智能的综合能力。教材内容以“普及知识、启发思维、引导责任”为核心,注重概念普及与跨学科融合,结合应用案例、实践操作、行业动态与技术趋势,融入科技伦理、国家安全、文化遗产等课程思政元素,形成理论与实践结合、技术与人文并重、层次与特色鲜明的“人工智能通识”课程体系。

二、教材形态

充分利用现代化技术手段,提倡多样化的教材建设形态,如新形态教材、纸质教材、数字教材。为系列教材建设人才培养方案、教学大纲、知识图谱、教学视频、电子课件、习题试题、教学案例、实验项目、实训项目、数据集等教学资源,形成优质共享的人工智能通识教学资源库,降低高校开课的难度,方便广大教师教学和学生学

三、评审与出版

申报人需扫码填写“选题申报表”,将有工作人员与您联系提交相关的评审资料,通过评审的选题列入“普通高等学校人工智能通识”系列教材,符合条件的作者入选“普通高等学校人工智能通识”系列教材编委会。系列教材作为中国计算机学会与机械工业出版社联合建设的重点项目,将由双方联合出版发行。

四、截止时间

本次申报第一批截止时间为 2025 年 3 月 20 日。

五、联系方式

策划编辑:机械工业出版社 李永泉
联系方式:010-88379619 15801292968

策划编辑:中国计算机学会 杨少勇
联系方式:010-62570586 13261571271

扫码申请

