

基于双模态知识协同驱动的教师-学生模型 无造影剂CT肝脏肿瘤分割

刘捷^{1,2}, 徐晨初^{3*}, 孙怡宁^{1,2}, 熊焰¹

(1. 中国科学技术大学, 安徽合肥 230026; 2. 中国科学院合肥物质科学研究院, 安徽合肥 230031;
3. 安徽大学, 安徽合肥 230601)

摘要: 无造影剂分割的CT(Computed Tomography)肝脏肿瘤图像在推进结直肠癌伴肝转移瘤筛查方面显示出巨大潜力,能够直接从无造影剂的CT图像中提供可靠的肝脏肿瘤分割结果,避免了造影剂的毒性、CT扫描的辐射和高昂的造影剂费用。本文提出了一种创新的“双模态知识协同驱动的教师-学生模型(Bimodal Knowledge Collaborative Driven Teacher-Student Model, BKC-TS)”,用于精准分割无造影剂的肝脏CT图像中的肝脏肿瘤,显著提升诊疗的安全性、准确性和时效性。BKC-TS利用教师网络学习显性肝脏肿瘤知识,指导学生网络从无造影剂图像中识别几乎不可见的肿瘤。它还协同临床检查文本数据和医学影像数据构建肝脏肿瘤知识,文本作为先验信息指引CT图像中肿瘤的学习过程,提升知识的精准性及指导的准确性。首先,文本-影像协同学习的教师学生框架通过引入文本知识,改进CT图像分辨率低的问题,提高无造影剂图像的肿瘤分割准确性。其次,双模态知识融合传递模块通过知识提炼、知识融合和知识传递,深度整合影像和临床数据,有效支持学生网络在无造影环境中的肿瘤定位和识别。最后,高斯分布约束的学生自主学习策略采用学生网络异步迭代计算分割分布的方法,评估并筛选出有益知识,提高学生网络的自主学习能力、泛化能力和鲁棒性。所有实验均在一个泛化的数据集上进行,该数据集包含各1140张增强前后的CT肝脏图像。实验结果表明,BKC-TS在肝脏肿瘤分割任务上获得了最佳性能(IOUS提升至少2.17个百分点),证明了其在无造影剂肝脏肿瘤分割技术发展中的重要作用。

关键词: 肝脏分割;无造影剂;师生网络;协同驱动;知识传递

基金项目: 国家自然科学基金项目(No.62476001)

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112(2025)05-1596-11

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20240782

Bimodal Knowledge Collaborative Driven Teacher-Student Model for Non-Contrast Agent CT Liver Tumor Segmentation

LIU Jie^{1,2}, XU Chen-chu^{3*}, SUN Yi-ning^{1,2}, XIONG Yan¹

(1. University of Science and Technology of China, Hefei, Anhui 230026, China; 2. Hefei Institutes of Physical Science, Chinese Academy of Science, Hefei, Anhui 230031, China; 3. Anhui University, Hefei, Anhui 230601, China)

Abstract: Non-contrast CT (Computed Tomography) liver tumor imaging shows great potential in advancing the screening of colorectal cancer with liver metastasis. It provides reliable liver tumor segmentation from non-enhanced CT images, avoiding contrast agent toxicity, radiation, and costs. In this paper, we propose an innovative “teacher-student model driven by dual-modal knowledge collaboration (BKC-TS)” for accurately segmenting liver tumors in non-contrast CT images, significantly improving the safety, accuracy, and efficiency of liver tumor diagnosis and treatment. BKC-TS employs a teacher network to acquire explicit liver tumor knowledge and guide a student network in recognizing nearly invisible tumors from non-contrast images. It integrates clinical examination text data with medical imaging data. Text data, as prior information, guides tumor learning in CT images, enhancing precision and accuracy. The text-image collaborative learning teacher-student framework improves liver tumor segmentation accuracy in non-contrast images by integrating text knowledge and addressing CT image resolution issues. The dual-modal knowledge fusion and transmission module combines im-

aging and clinical data through knowledge extraction, fusion, and transmission, effectively supporting tumor localization and recognition in non-contrast images. The gaussian distribution-constrained student self-learning strategy boosts the student network’s independent learning, generalization, and robustness by iterating segmentation distribution and selecting beneficial knowledge. All experiments were conducted on a generalized dataset containing 1 140 CT liver images before and after enhancement. Experimental results show that BKC-TS achieved optimal liver tumor segmentation (at least a 2.17 percentage points IOU improvement), demonstrating its importance in non-contrast technology development.

Key words: liver segmentation; no contrast agent; teacher-student network; collaborative drive; knowledge transfer
Foundation Item(s): National Natural Science Foundation of China (No.62476001)

1 引言

精准的无造影剂肝脏肿瘤分割对临床医生在诊断和治疗肝脏肿瘤中至关重要^[1,2],它可以帮助医生进行肿瘤分期、简化手术规划和评估治疗效果^[3].更重要的是,无造影剂肝脏肿瘤分割技术可以带来多方面的好处(见图 1):

- (1)消除造影剂的毒性. 避免使用含碘的 CT(Computed Tomography)造影剂,减少患者可能发生的造影剂诱导肾病风险及各种造影剂注射的副作用,提升患者诊疗的安全性.
- (2)避免 CT 扫描带来的辐射. 减少增强 CT 扫描的次数,降低对患者身体造成的辐射伤害,避免潜在的健康问题,同时也有助于降低累积辐射暴露.
- (3)减少患者经济负担. 造影剂的成本在我国通常较高,降低造影剂开销可以大幅减轻患者的经济压力,增加患者的治疗可及性.

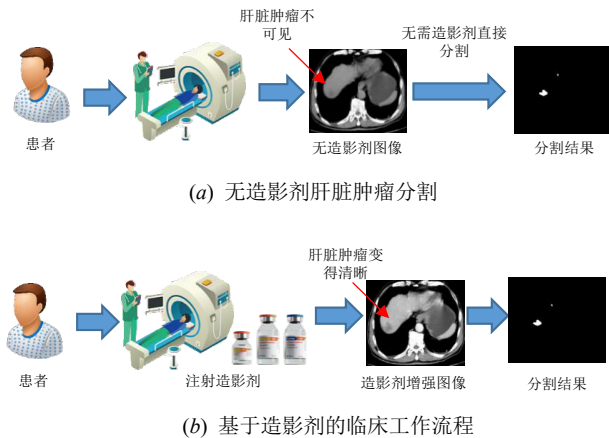


图 1 无造影剂 CT 肝脏肿瘤分割方法与基于造影剂的临床工作流程

目前,已有少数方法基于肝脏 MRI(Magnetic Resonance Imaging)图像实现无造影剂肝脏肿瘤分割^[4,5],并获得了临床的高度关注.然而,这些方法依赖于多模态 MRI 图像,无法应用于 CT 图像的无造影剂肝脏肿瘤分割(见图 2).无造影剂肝脏肿瘤分割的最大挑战在于肿瘤在无造影剂图像上几乎不可见,这极大地限制了从

无造影剂图像中学习和分割肿瘤知识的准确性.现有的无造影剂肝脏肿瘤分割方法依赖于多模态 MRI 成像序列,通过结合多模态图像中对肝脏和肿瘤的知识,实现无造影剂肝脏肿瘤分割.然而,CT 图像只有一种成像模态,这使得直接应用基于多模态 MRI 的无造影剂肝脏肿瘤分割方法在 CT 图像上变得不可能.因此,临床上迫切需要开发适用于 CT 图像的无造影剂肝脏肿瘤分割方法.

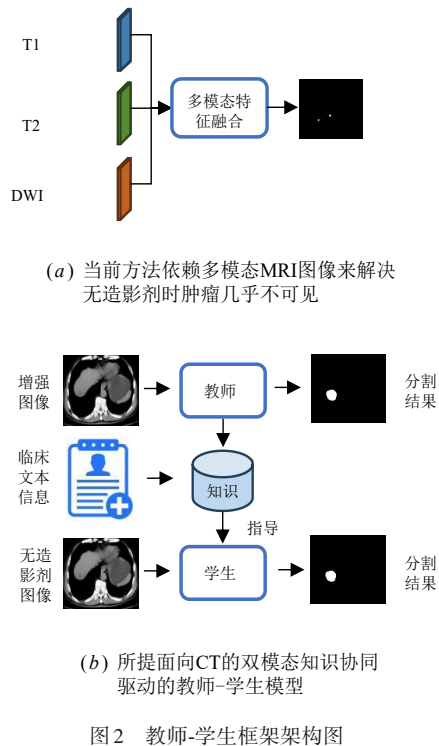


图 2 教师-学生框架架构图

此外,CT 图像的无造影剂肝脏肿瘤分割还面临着图像高噪声、肿瘤形态多样性及复杂组织干扰的三大挑战:(1)由于扫描过程中的物理限制和患者的生理运动,CT 图像具有固有的高噪声特性,这些噪声会导致肝脏和肿瘤的边界模糊不清,影响分割模型对肿瘤边界特征的学习,最终降低分割的准确性.(2)肝脏肿瘤在 CT 图像中的表现形式多种多样,不同患者的肿瘤形态、大小、位置各不相同,且肿瘤内部的组织密度也存

在显著差异. 这种多样性增加了肿瘤分割的复杂性, 降低了分割的泛化性. (3) CT 图像对肝脏内部复杂组织的对比度较低, 肝脏中的血管、脂肪和其他组织在 CT 图像上与低对比度的肿瘤具有高度相似的灰度值, 这种高度相似性导致特征学习间的相互干扰, 引发大量的假阳性分割结果.

本文提出了一种创新的“双模态知识协同驱动的教师-学生模型(Bimodal Knowledge Collaborative Driven Teacher-Student Model, BKC-TS)”, 用于精准分割未注射造影剂的肝脏 CT 图像中的肝脏肿瘤, 从而显著提升肝脏肿瘤诊疗的安全性、准确性和时效性. BKC-TS 模型的创新之处在于其独特的教师-学生架构设计, 专门针对无造影剂条件下肿瘤难以辨识的问题. 在该模型中, 教师网络负责提取并学习显性的肝脏肿瘤特征知识, 这些知识随后用于指导学生网络, 使其能够在未注射造影剂的 CT 图像中准确识别出几乎隐形的肿瘤, 进而完成肿瘤的精确分割. 更重要的是, BKC-TS 设计了双模态知识协同机制, 即协同患者的临床检查文本数据(包括个人信息、生化信息、病理信息等)和医学影像数据(如注射造影剂后的 CT 肝脏图像)这两种模态来构建显性肝脏肿瘤知识. 文本作为先验信息, 用于指引

和约束 CT 图像中肿瘤的学习过程, 避免 CT 高噪声、肿瘤形态多样性及复杂组织间的干扰问题, 从而提升知识的精准性及对学生网络的指导准确性.

所提出的 BKC-TS 模型(见图 3)构建了一个教师-学生框架, 旨在解决无造影剂肝脏肿瘤分割的难题. 该模型利用教师网络从增强图像中学习显性肝脏肿瘤知识, 并通过三个核心特性, 指导学生网络在非增强图像中进行精准分割. 双模态知识协同首次将文本信息和图像信息结合, 构建双模态知识, 并用于指导学生网络进行肿瘤分割. 文本信息作为先验知识, 有效解决了无造影剂图像分辨率低、肿瘤形态多样和复杂组织间干扰等问题, 提升了知识的精准性及指导的准确性. 知识融合与传递: 创新性地设计了知识提炼、知识融合和知识传递机制, 有效将教师网络从增强图像中学习到的肿瘤知识传递给学生网络. 知识提炼提取病变和形态知识, 知识融合整合影像和临床数据, 知识传递则通过知识蒸馏引导学生网络在非增强图像上进行精准分割. 学生自主学习: 引入高斯分布约束的学生自主学习策略, 通过计算知识传递前后的分布差异, 评估并筛选出有益知识, 使学生网络能够自主学习, 逐步摆脱对教师网络的依赖, 提高模型的泛化能力和鲁棒性.

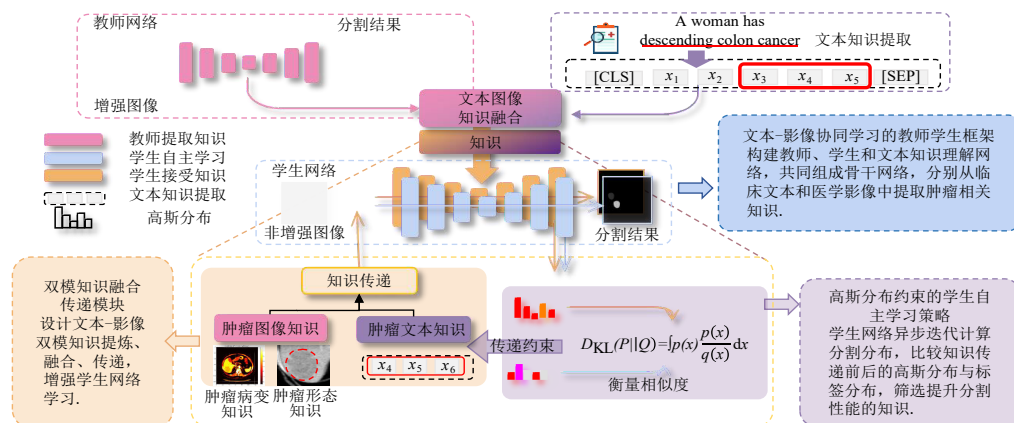


图 3 BKC-TS 方法架构

2 相关工作

2.1 现存的肝脏肿瘤图像分割方法

尽管已经发表了许多关于肝脏肿瘤分割的研究, 大部分研究往往依赖于增强图像^[6,7]. 增强图像因为具有更高的对比度, 能够清晰地显示肿瘤的形状和轮廓, 尤其是对恶性肿瘤的识别更为有效, 因此, 大多数现有方法专注于这类图像^[8,9]. 然而, 这些方法在应用于非增强图像时效果通常较差, 而且注射造影剂会对患者的身体和经济状况造成负担. 因此, 开发一种能够在未注射造影剂的情况下有效分割肝脏肿瘤的方法变得尤为重要^[10-14], 这种方法可以减轻患者的身体和经

济负担. 尽管这些方法取得了进展, 但他们主要集中于图像分析, 而忽视了医疗实践中非常关键的文本信息, 如诊断报告和病史记录. 这些文本数据能提供肿瘤的临床信息, 对于在图像质量不高时辅助识别和分割肿瘤至关重要. 因此, 未来的研究需要集中于图像与文本信息的有效融合, 这种多模态方法有望在无造影剂的情况下优化诊断和治疗过程, 从而提升患者的生存率.

2.2 教师-学生框架中的知识传递策略

教师-学生框架在迁移学习特别是知识蒸馏方面得到了广泛应用^[15,16], 例如在图像分类^[17]、目标检测^[18]

和语音识别^[19]等任务中.这种框架能够有效传递教师模型的丰富知识,提升学生模型的性能.而大部分现有研究都采用了“直接映射”策略.这种策略通过教师网络直接指导学生网络复制其输出,使学生网络模仿教师的决策逻辑,从而简化了知识转移过程,提升了学习效率和泛化能力^[20,21].直接映射的优势在于高效的知识传递和学生网络的快速收敛^[22],但其有效性依赖于教师和学生网络处理的数据属于同一分布^[22,23],因为在相同分布下,教师网络的输出特征与学生网络所需处理的输入特征具有一致性,这种一致性确保了知识的直接迁移能够保持高效和准确.不同分布时,教师网络学到的特征和知识在学生网络中可能不再适用,导致知识传递效果显著降低.然而,在无造影剂 CT 肝脏肿瘤分割任务中,这种分布的一致性并不存在,传统的知识传递方法可能效果有限.因此,开发能够在不同数据分布之间有效进行知识传递的新策略变得尤为关键.

2.3 图像分割中的不确定性

不确定性在图像分割中得到了广泛应用,例如在小样本分割^[24]、半监督分割^[25]以及强化学习分割^[26]等任务中.它反映了模型对预测结果的置信度及预测的可靠性,有助于识别潜在的错误区域并指导模型优化.不确定性通过采样或概率分布捕获低置信度区域,然后利用增加数据、优化模型结构或采用主动学习策略等方式提升模型在这些区域的预测准确性和可靠性,从而提高整体性能^[27,28].在医学图像分析中,不确定性有着非常重要的作用,通过不确定性分析,可以帮助医生更准确地评估医学图像数据的可信度和精确性,特别是在面对复杂病例或图像质量较低的情况下^[29,30].然而当前不确定性的方法主要应用在输入输出基于线性映射的单网络模型中,针对本文使用存在不同输入输出的教师学生双网络模型仍缺乏探索.因此,需要探索一种新的方法,将不确定性估计有效地引入到教师学生双网络模型中,从而避免盲目接受知识,只保留对结果有利的部分知识,以提高其在图像分割任务中的鲁棒性和稳定性.

3 方法

所提的“BKC-TS”(见图3)构建了一个教师-学生框架,使得教师网络和学生网络能够分别探索增强和非增强的肝脏肿瘤 CT 图像.特别是,“BKC-TS”利用教师网络从增强图像中提取肿瘤知识,并指导学生网络,使其在测试阶段能够独立地从非增强图像中检测和分割肝肿瘤.同时,从文本信息中提取的文本特征进一步加强了对肝肿瘤的检测和分割.如图3所示,“BKC-TS”包含三个模块:(1)文本-影像协同学习的教师-学生框架,利用教师网络从增强肝脏肿瘤图像中提取肿瘤图像知

识,并结合临床数据中提取的肿瘤相关文本知识,指导学生网络对非增强肝脏肿瘤图像进行准确分割;(2)双模知识融合传递模块,利用教师网络的空间注意力权重和形状特征提取肿瘤影像的病变与形态知识,再与文本知识融合后传递给学生,指导学生网络在无造影剂图像中准确定位肿瘤区域,从而显著提升了肝脏肿瘤分割的准确性和效率;(3)高斯分布约束的学生自主学习策略,通过计算知识传递前后的分布,衡量分割效果的变化,确保学生能独立于教师进行自主学习,从而提升学生网络的泛化性和鲁棒性.

3.1 文本-影像协同学习的教师学生框架

文本-影像协同学习的教师学生框架创造性地构建了一种教师网络和本知识理解网络协同驱动、共同指导学生网络的双模架构,从而实现学生网络能够独立完成 CT 肝脏肿瘤图像分割的任务.教师网络和学生网络分别用来对增强图像和非增强图像进行肝肿瘤的分割,文本知识理解网络用来提取临床数据中和肿瘤相关的文本知识,从而实现基于临床数据的肿瘤文本知识和增强图像双模态驱动的非增强肝脏肿瘤图像分割.如图4所示,教师网络和学生网络都由肿瘤图像知识编码器、肿瘤图像知识解码器组成,肿瘤图像知识编码器通过卷积加上最大池化的四层下采样组成,肿瘤图像知识解码器通过反卷积和跳跃链接组成的四层下采样组成,这两个网络采用相同的模型来建立他们之间强一致性,促进了图像知识在模型的不同位置下从教师转移到学生网络.文本知识提取网络,通过分词、投影和对齐实现对临床数据中肿瘤文本知识的准确提取,并将其与图像知识融合,进而帮助非增强图像实现肝脏肿瘤的准确分割.具体来说,对于教师的输入增强图像以及学生的输入非增强图像 $[N, C, H, W]$,本文对其进行四次下采样分别得 $[N, C_1, H/2, W/2]$ 、 $[N, C_2, H/4, W/4]$ 、 $[N, C_3, H/8, W/8]$ 和 $[N, C_4, H/16, W/16]$, N 是输入的 Batch Size, C 是特征维度, H 和 W 是原始图像的高度和宽度,在添加位置编码后,本文使用自注意来增强其中所获得的视觉信息,以获得进化的视觉特征.该过程如式(1)所示:

$$F_i = I + \text{MHSA}(I) \quad (1)$$

其中, $\text{MHSA}()$ 为多头自注意层, $L()$ 为层归一化, 最后得到具有残差的进化视觉特征 $F_i \in \mathbb{R}^{H \times W \times C}$.

而对于输入的文本 T , 本文先分词然后使用 BioBERT 来提取文本知识 $F_t [N, C, L]$. 其中 N 是输入的 Batch Size, C 是特征的维度, L 是文本信息的长度. 之后本文采用 Project Embedding 和 Position Embedding 将文本的特征维度与图像的特征维度对齐, 这个过程如式(2)所示:

$$F_i = \sigma(\text{Conv}(TW_T)) \quad (2)$$

其中, W_T 为可调整参数的矩阵, $\text{Conv}(\cdot)$ 表示 1×1 的卷积层, $\sigma(\cdot)$ 表示 ReLU 激活函数. 给定一个输入特征 $[C, L]$, 输出投影特征为 $[C_i, M]$, 其中 M 为投影后的标记数, C_i 为投影特征的维度, 与图像标记的维度一致.

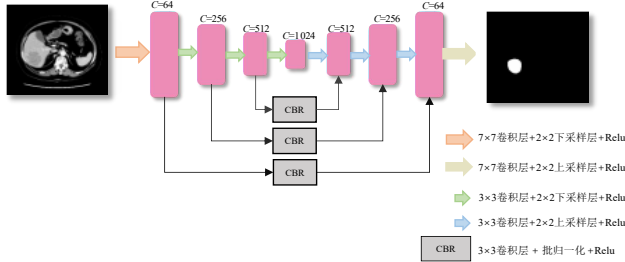


图4 教师-学生网络编解码结构图

3.2 双模知识融合传递模块

双模知识融合传递模块(如图5)包括知识提炼、融合和传递三个部分,用以更好地构建与整合知识. 知识提炼通过在增强图像上提取肝脏肿瘤特征,并从中提炼出有用的影像知识. 知识融合将影像和临床数据知识作为输入进行深度整合. 知识传递通过知识蒸馏将融合内容传递给学生网络,指导其在非增强图像上进行准确的肝脏肿瘤分割. 特别地,知识提炼引入病变与形态知识的提取. 病变知识通过注意力机制计算空间位置关系并传递给学生网络,以确保病变定位在肝脏内. 形态知识在下采样阶段提取并融合肿瘤的边缘、纹理和轮廓特征. 知识融合通过多头交叉注意力机制加权计算影像与临床数据特征,生成融合特征,提升模型识别复杂病理特征的能力. 知识传递采用均方误差损失函数,使学生网络特征分布逼近教师网络特征分布,提高模型精度和稳定性. 该模块充分利用临床数据与影像知识,极大地优化了学生网络在非增强图像下的学习表现.

具体来说,在病变知识提取中,输入特征来自第四层下采样后的特征 $F \in \mathbf{R}^{H \times W \times C}$, 其中 H 、 W 、 C 分别表示特征图的高度、宽度和通道数量. 首先,调整特征 F 的维度,得到查询矩阵 $Q \in \mathbf{R}^{N \times C}$ 、键矩阵 $K \in \mathbf{R}^{N \times C}$ 和值矩阵 $V \in \mathbf{R}^{N \times C}$, 其中 N 表示特征图所有像素点的数量. 然后,计算 Q 与 K 转置的矩阵乘法,并应用 softmax 层生成通道注意力图 $\theta \in \mathbf{R}^{C \times C}$, 如式(3)所示:

$$\theta_{ij} = \frac{\exp(Q_i \cdot K_j)}{\sum_{j=1}^C \exp(Q_i \cdot K_j)} \quad (3)$$

式(3)表示的是通道 i 和通道 j 之间的注意力权重 θ_{ij} , 其中 Q_i 是查询向量 Q 的第 i 个通道, j 是键向量 K 的第 j 个通道, 通过计算 Q 和 K 的内积并取指数, 然后对

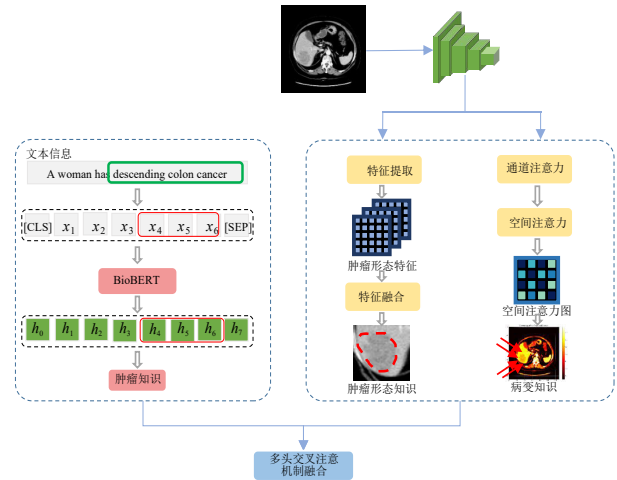


图5 影像-文本融合图

所有通道的指数和进行归一化,得到的 θ_{ij} 是归一化后的注意力权重. 最后,将矩阵 θ 与 V 进行矩阵乘法,并调整为 $\mathbf{R}^{H \times W \times C}$, 执行恒等映射操作, 获得最终通道注意力块输出 $F' \in \mathbf{R}^{H \times W \times C}$, 如式(4)所示:

$$F'_i = \sum_{j=1}^C \theta_{ij} V_j + F_i \quad (4)$$

式(3)和式(4)中通道与空间注意力机制的计算基于已有的自注意力机制理论. 然后,将通道注意力块的输出特征 F' 作为空间注意力块的输入. 对输入特征进行三个 1×1 卷积操作, 生成特征图 Q' 、 K' 和 V' , 其中 Q' 、 $K' \in \mathbf{R}^{N \times C'}$, $C' = C/8$, 以及 $V' \in \mathbf{R}^{N \times C}$. 同理,使用式(3)和式(4)可以获得最终输出 $F''_{\text{spatial}} \in \mathbf{R}^{H \times W \times C}$. 基于 F' , F''_{spatial} 进一步获得所有位置之间的语义关联, 从而增强特征的语义表示.

在肿瘤形态知识提取中,教师网络在不同下采样阶段提取肿瘤的边缘、纹理和轮廓特征,即第一层、第二层和第三层的下采样特征中获取这些信息. 设第一层、第二层和第三层的特征图分别为 F_1 、 F_2 和 F_3 . 将这些特征图上采样和卷积处理,使其具有相同的尺寸和通道数. 设目标尺寸为 $\mathbf{R}^{H \times W \times C}$, 该过程如式(5)所示:

$$F'_j = \gamma \left(\delta(F_j, \alpha = (H, W), \beta = C) \right) \quad (5)$$

其中, F_j 表示第 j 层的原始特征图, α 和 δ 分别表示尺寸调整参数和上采样操作, 用于将特征图调整到相同的空间尺寸 (H, W) , β 和 γ 分别表示通道数调整参数和卷积操作, 根据参数 β 将上采样后的特征图调整到通道数 C , 得到调整后的特征图 F'_j . 然后,通过级联操作整合这些特征图, 并利用卷积操作 γ 形成综合的特征表示 F''_{shape} :

$$F''_{\text{shape}} = \gamma([F'_1, F'_2, F'_3]) \quad (6)$$

式(5)和式(6)创新性地调整多层特征图并进行级

联,突出形态知识提取.接着,将作为输入的临床数据与影像知识采用多头交叉注意力进行融合,并对输出进行归一化和残差连接,如式(7)所示:

$$\mathbf{F}_c = \mathbf{F}_{\text{shape}}'' + \mathbf{F}_{\text{spatial}}'' + \mu \left(L \left(\phi(\mathbf{F}_t, \mathbf{F}_{\text{shape}}'', \mathbf{F}_{\text{spatial}}'') \right) \right) \quad (7)$$

式(7)创新性地结合多头交叉注意力机制与残差连接,强化双模态数据的融合与互补特征表达.其中, $\phi(\cdot)$ 表示多头交叉注意机制,用于对临床数据和影像数据进行特征交互; L 表示归一化操作, μ 是一个可学习的参数,用于控制残差连接的权重,平衡原始特征与融合特征的影响.知识传递部分采用了均方误差损失函数,将教师网络提取的特征 $\mathbf{F}_{\text{teacher}}$ 和学生网络提取的特征 $\mathbf{F}_{\text{student}}$ 比较,计算损失:

$$L_{\text{distill}} = \frac{1}{N} \sum_{i=1}^N (\mathbf{F}_{\text{teacher}}^i - \mathbf{F}_{\text{student}}^i)^2 \quad (8)$$

其中, N 表示特征的总数量, $\mathbf{F}_{\text{teacher}}^i$ 和 $\mathbf{F}_{\text{student}}^i$ 分别表示教师网络和学生网络提取的第 i 个特征.通过最小化这个损失函数,学生网络逐渐学习并逼近教师网络的特征分布.

3.3 高斯分布约束的学生自主学习策略

高斯分布约束的学生自主学习策略创新性地构建了学生网络异步迭代计算分割分布的方法,不仅有效筛选并保留了有益知识,而且显著提升了学生网络的分割性能.该策略通过计算知识传递前后分割结果的高斯分布,并将这些分布与真实分布之间的距离作为评估指标,从而筛选出有助于提升非增强图像分割效果的知识.具体而言,该策略以分割结果为均值,以分割不确定性作为方差构建高斯分布,并利用KL(Kullback-Leibler)散度与真实分布进行相似性度量,以评估该知识对分割结果的有益程度.如果知识传递后的分布与标签分布更为接近,则保留该知识,否则丢弃.高斯分布约束的学生自主学习策略使学生网络能够独立于教师网络进行自主学习,逐步摆脱对教师网络的依赖,并专注于学习对分割结果有帮助的知识,从而提升学生网络的泛化能力和鲁棒性.

具体来说,首先,该策略计算三个高斯分布,分别是经过教师网络知识传递后的学生网络做一次前向传播的分割结果分布:

$$\epsilon_{T_s} = \mathcal{N}(\mu_{T_s}, \sigma_{T_s}^2) \quad (9)$$

其中, μ_{T_s} 是教师网络传递后学生网络的预测值均值, σ_{T_s} 是分割不确定性.未经过教师网络知识传递的学生网络做一次前向传播的分割结果分布:

$$\epsilon_S = \mathcal{N}(\mu_S, \sigma_S^2) \quad (10)$$

其中, μ_S 是学生网络的预测值均值, σ_S 是分割不确定性.真实图像的高斯分布(Ground Truth, GT):

$$\epsilon_{GT} = \mathcal{N}(\mu_{GT}, \sigma_{GT}^2) \quad (11)$$

其中, μ_{GT} 和 σ_{GT} 分别是真实图像标签的均值和方差.其次,该策略分别计算经过教师知识传递后的高斯分布与GT高斯分布的KL散度,以及未经过教师传递后的高斯分布与GT高斯分布的KL散度,以度量教师知识传递后的分布与GT分布的相似性.KL散度的计算公式为

$$D_K(P||Q) = \int_{-\infty}^{+\infty} p(x) \log \frac{p(x)}{q(x)} dx \quad (12)$$

其中, P 和 Q 分别代表两种不同的概率分布.在这里,本文需要计算:

$$D_{KL}(\epsilon_{T_s} || \epsilon_{GT}) \quad (13)$$

和

$$D_{KL}(\epsilon_S || \epsilon_{GT}) \quad (14)$$

最后,该策略判断如果经过教师知识传递后的学生网络分割结果的高斯分布与GT的高斯分布结果更为接近,即

$$D_{KL}(\epsilon_{T_s} || \epsilon_{GT}) < D_{KL}(\epsilon_S || \epsilon_{GT}) \quad (15)$$

则接受教师传递的知识;否则丢弃教师传递的知识,只依赖GT进行学习.通过这种方式,可以筛选出那些增强图像特征有助于提升非增强图像分割效果的个体并进行特征传递.该策略确保了学生网络在学习过程中能够有效过滤不利于分割结果的教师知识,从而提升最终的分割性能.

4 实验

4.1 数据集

BKC-TS在一个泛化的数据集(来自三个中心)进行训练和测试.泛化数据集包含1140个增强前和增强后CT肝脏图像.所有扫描均采用GE Revolution 256排CT.患者均采用仰卧位,扫描序列包括腹盆腔CT平扫及多期增强扫描,其中多期增强扫描包括动脉期、静脉期、延迟期扫描.腹盆腔CT平扫参数:管电压120KV,管电流270mA,扫描范围自膈顶部至腹股沟区,层厚5mm.增强扫描采用的是双筒高压注射器,使用留置针以流速5mL/s于肘静脉注射对比剂碘佛醇.注射造影剂后于25、70、180s自动触发扫描,收集动脉期、静脉期、延迟期图像.层厚为5mm,横轴位观察,窗位35HU,窗宽240HU.采用静脉期图像为增强对比图像.

4.2 实验细节

BKC-TS随机选择约4/5患者的数据进行训练,剩下1/5患者的数据进行独立测试.所有代码都基于PyTorch,并对所有参数进行调优以获得最佳结果.采用SGD优化器,初始学习速率为 1×10^{-3} ,学习率衰减为0.9,从 $[1 \times 10^{-3}, 1 \times 10^{-4}, 1 \times 10^{-5}]$ 挑选最优学习率.该模

型被训练了 200 个周期,这是通过在 $[100, 150, 200, 250, 300]$ 范围内评估不同周期后决定的,在前 100 个周期中,初始学习率设置为 1×10^{-3} ,然后在剩余的周期内线性衰减至 0,在 4x NVIDIA GeForce RTX 3090 GPUs 上,处理一个尺寸为 416×416 的测试图像平均需要 0.021 s(大约 48 FPS).

4.3 实验设置和评价指标

为了全面证明 BKC-TS 的优越性,本文首先进行了对比实验,以证明它在无造影剂肝脏肿瘤分割中取得了新的最先进的性能.因此,实验将本文的 BKC-TS 与最近 8 种最先进的方法进行了比较.这些方法都使用注意力机制优化模型,分割肝脏肿瘤.其次,本文进行了消融实验,以评估 BKC-TS 的每个主要组件的性能,包括文本-影像协同学习的教师学生框架、双模知识融合传递模块和高斯分布约束的学生自主学习策略.这使本文能够证明每个组件的个体优越性.在本文的实验中,本文采用了 5 个常见的评估指标,它们分别是交并比(Intersection over Union, IOU)、Dice 系数、精确度(Precision)、召回率(Recall)以及 F1 分数(F1 Score).

4.4 实验结果

本文的 BKC-TS 模型仅使用 1 140 张非增强肝脏肿瘤 CT 图像作为输入,就实现了对肝脏肿瘤的精准分割. BKC-TS 在分割任务中的 IOU 达到了 75.34%, Dice 系数为 82.15%,精确度为 92.73%,召回率为 92.68%, F1 分数为 92.70%. 这些结果表明, BKC-TS 展现出巨大的潜力,能够作为一种准确的肝肿瘤诊断的临床替代方法,并有可能完全消除对造影剂的需求.

4.4.1 无造影剂的肝肿瘤分割的准确性

从图 6 和表 1 中得到的结果表明,本文的 BKC-TS 能够直接从非增强肝脏肿瘤 CT 图像中准确分割肿瘤.分割结果(即肿瘤边界)在视觉上与放射科医生基于造影剂增强 CT 图像标注的标签高度一致.图 6 中,红色标注表示肿瘤的真实位置,右下角的二值图为肿瘤的分割结果图, BKC-TS 在两者之间实现了高度重叠,进一步突出了其分割性能的优越性.此外,表 1 展示了本文的 BKC-TS 在各种分割指标中的高性能,本文的 BKC-TS 实现了 IOU 为 75.34%, Dice 系数 82.15%,精确度为 92.73%,召回率为 92.68%, F1 分数为 92.70%,其中标注的 * 和 ** 表示与其他方法相比在统计学上具有显著差异($p < 0.05$ 和 $p < 0.01$).此外,图 7 显示,本文的 BKC-TS 在无增强肝脏肿瘤分割中取得了良好的 ROC 曲线, AUC 值达到 91.21%. 如图 7 所示,红色曲线位置更靠近左上角和右上角,分别代表更优的 ROC 和 PR 性能,进一步证明了 BKC-TS 的分割效果优于其他方法.

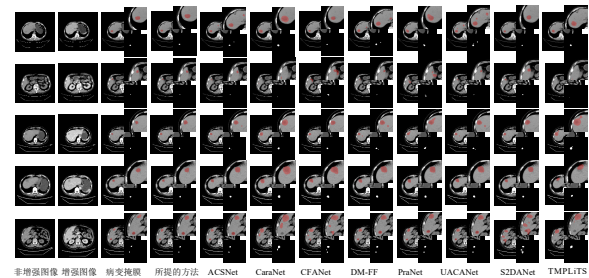


图 6 BKC-TS 与其他方法对比

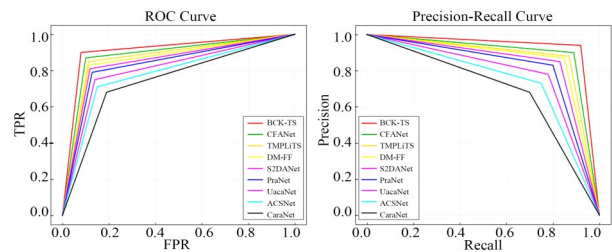


图 7 BKC-TS 与其他方法在无造影剂肝脏肿瘤分割中的 ROC 曲线与 PR 曲线对比

4.4.2 BKC-TS 性能

图 6 和表 1 的可视化结果及分割指标评估表明,本文提出的 BKC-TS 在所有 8 种比较方法中表现优越,具有最高的分割性能.在表 1 中,与其他 8 种比较方法相比, BKC-TS 在 IOU 指标上提高了 2.17~7.87 个百分点,在 Dice 系数上提升了 1.86~8.53 个百分点,在精确度上提高了 13.32~21.22 个百分点,在召回率上提升了 2.66~8.26 个百分点,在 F1 分数上提升了 9.46~15.59 个百分点.这些结果证明了本文提出的 3 个模块在准确表示教师网络知识并有效转移至学生网络方面的有效性.此外,在图 7 中, BKC-TS 展示了所有 8 种比较方法中最佳的 ROC 曲线(曲线越接近左上角,性能越佳)和最大的 AUC, AUC 值提高了 6.0%~11.5%.最后,在表 2 中,我们将教师学生网络中编解码框架替换成 Deeplab 和 Backbone 进行了消融,在 5 种评价指标上均取得了最佳性能表现,其中 IOU 提升了约 7.04~12.03 个百分点, Dice 提升了约 5.46~17.02 个百分点, Precision 提升了约 6.28~15.52 个百分点, Recall 提升了约 3.75~10.43 个百分点, F1 分数提升了约 6.56~8.60 个百分点,其中标注的 * 和 ** 表示与其他方法相比在统计学上具有显著差异($p < 0.05$ 和 $p < 0.01$).

4.4.3 文本-图像提示的双模态分割模块的优越性

图 8 和表 3 中的定量结果显示,与 Text-Free 和 No MHSA 相比, BKC-TS 在无造影剂 CT 肝脏肿瘤分割中展现了显著的性能优势.使用文本信息能够更全面地反映文本-图像提示的双模态分割模块的卓越表现.引入文本信息可以有效改善 CT 图像的低分辨率问题,从而提高肝肿瘤分割的精度. Text-Free 指的是

表 1 BKC-TS 在无造影剂肝肿瘤分割领域取得了最先进的性能

单位:%

模型	评价指标				
	IOU ↑	Dice ↑	Precision ↑	Recall ↑	F1 Score ↑
CaraNet ^[31]	67.47** ± 2.89	73.62** ± 2.22	70.96** ± 3.27	84.42** ± 2.99	77.11** ± 2.30
CFANet ^[32]	73.17* ± 3.95	80.29* ± 3.25	79.41** ± 4.13	87.46** ± 1.80	83.24** ± 2.41
PraNet ^[33]	71.26** ± 4.33	78.65** ± 3.70	73.80** ± 4.22	90.02* ± 0.98	81.11** ± 2.58
UACANet ^[34]	70.89** ± 3.86	77.99** ± 3.23	77.43** ± 3.65	85.49** ± 1.18	81.26** ± 2.08
DM-FF ^[35]	72.03** ± 3.76	76.82** ± 3.14	78.84** ± 3.74	79.78** ± 1.63	79.31** ± 2.42
ACSNet ^[36]	70.13** ± 3.61	74.59** ± 3.43	72.68** ± 2.64	86.05** ± 1.10	78.81** ± 3.04
TMPLITS ^[37]	72.58** ± 4.49	79.44** ± 5.05	71.39** ± 7.01	88.23** ± 4.71	80.62** ± 8.21
S2DANet ^[38]	71.33** ± 4.83	78.29** ± 4.92	75.89** ± 5.78	88.56** ± 4.22	78.53** ± 9.34
BKC-TS	75.34 ± 4.74	82.15 ± 4.14	92.73 ± 1.27	92.68 ± 0.63	92.70 ± 0.71

表 2 所提的 BKC-TS 编解码结构的优越性

单位:%

Backbone	评价指标				
	IOU ↑	Dice ↑	Precision ↑	Recall ↑	F1 Score ↑
SegNet	63.31** ± 12.72	65.13** ± 14.25	77.21** ± 22.71	82.25** ± 18.43	84.10** ± 3.75
DeepLabV3+	68.30** ± 19.61	76.69** ± 3.12	86.45** ± 13.35	88.93** ± 3.22	86.14** ± 3.34
Ours	75.34 ± 4.74	82.15 ± 4.14	92.73 ± 1.27	92.68 ± 0.63	92.70 ± 0.71

从原始 BKC-TS 模型中去除了文本信息,而 No MHSA 则指去除了多头注意力机制,其他组件保持不变. 结果显示, BKC-TS 在五种评价指标中均表现最佳, IOU 提升了约 2.17~6.61 个百分点, Dice 系数提高了约 2.85~6.94 个百分点, 精确度提升了约 3.71~5.49 个百分点, 召回率提升了约 4.33~4.75 个百分点, F1 分数提升了约 3.85~5.12 个百分点, 其中标注的 * 和 ** 表示与其他方法相比在统计学上具有显著差异 ($p < 0.05$ 和 $p < 0.01$). 这些改进可以归因于文本-图像提示的双模态分割模块的集成, 该模块创新性地利用肿瘤的文本知识来引导图像分割, 从而实现了更准确分割.

4.4.4 空间-形状知识传递模块的优越性

图 9 和表 3 中的 BKC-TS 和 Simple Fusion 及 No SA 结果表明, 使用肿瘤形态知识和病变知识可以更好地全面反映空间-形状知识传递模块在无造影剂 CT 肝脏肿瘤分割中的优异性能. Simple Fusion 表示在融合临床数据与影像知识没有使用多头交叉注意力机制, 直接使用拼接的方式. No SA 意味着没有使用空间注意力

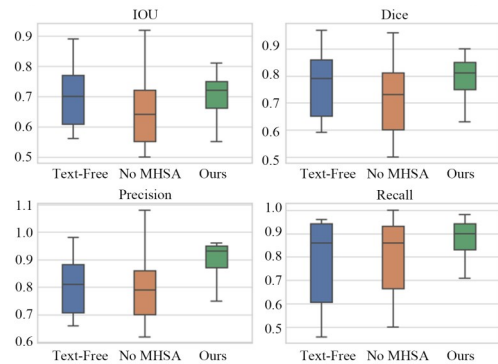


图 8 BKC-TS 通过文本信息引导图像分割, 在无造影剂肝脏肿瘤分割中优于各消融版本

机制, 而 BKC-TS 的其他组件保持不变. BKC-TS 在 5 种评价指标中取得了最好的性能. IOU 提升了约 4.38~7.52 个百分点, Dice 提升了约 4.11~8.31 个百分点, Precision 提升了约 4.30~5.84 个百分点, Recall 提升了约 3.76~5.45 个百分点, F1 分数提升了约 4.87~5.31 个百分点. 这些改进可以归因于空间-形状知识引导模块的优越性的集成, 它创造性地将教师网络学习到的肿瘤形态

表 3 所提的 BKC-TS 每个组件的个体优越性

单位:%

	IOU	Dice	Precision	Recall	F1 Score
Text-Free	73.17* ± 4.53	79.30** ± 3.92	89.02*** ± 3.54	88.35*** ± 1.39	88.85*** ± 4.11
No MHSA	68.73** ± 5.57	75.21*** ± 6.74	87.24*** ± 6.82	87.93*** ± 4.03	87.58*** ± 7.94
No SA	67.82** ± 5.13	78.04*** ± 3.53	86.89*** ± 3.00	88.92*** ± 1.36	87.39*** ± 3.59
Simple Fusion	70.96** ± 4.20	73.84*** ± 8.17	88.43*** ± 4.58	87.23*** ± 4.06	87.83*** ± 6.12
Dependency Learning	72.99*** ± 4.48	80.19* ± 3.99	87.82*** ± 3.52	89.86*** ± 0.86	88.83*** ± 1.77
MSE	71.53** ± 4.63	76.93*** ± 6.66	89.05* ± 4.23	87.74*** ± 5.42	88.32*** ± 3.95
BKC-TS	75.34 ± 4.74	82.15 ± 4.14	92.73 ± 1.27	92.68 ± 0.63	92.70 ± 0.71

知识和病变知识传递给学生,从而导引学生能够独立完成对未增强肝肿瘤图像的准确分割.

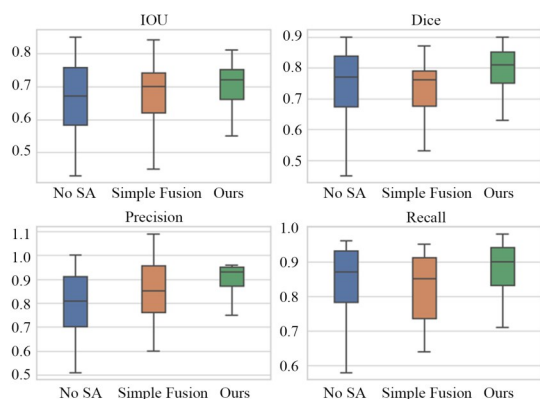


图9 BKC-TS结合空间与形状知识,在无造影剂肝脏肿瘤分割中优于各消融版本

4.4.5 高斯分布约束的学生自主学习策略的优越性

图10和表3中的BKC-TS和Dependency Learning及MSE的定量结果表明,使用不确定性可以更好地全面反映不确定性引导的学生自主学习模块在无造影剂CT肝脏肿瘤分割中的优异性能,其通过对接受到的肿瘤知识进行筛选,进而解决学生网络对不确定性信息的盲目依赖问题.完全依赖教师学习意味着学生网络将接收教师的全部信息,不会再进行筛选.MSE表示使用均方误差代替KL散度进行度量,BKC-TS其他组件保持不变.BKC-TS在5种评价指标中取得了最好的性能.IOU提升了约2.35个百分点,Dice提升了约1.96~5.22个百分点,Precision提升了约3.68~4.91个百分点,Recall提升了约2.82~4.94个百分点,F1分数提高了约3.87~4.84个百分点.这些改进可以归因于高斯分布约束的学生自主学习策略的集成,该策略创造性地利用在学生网络中引入的不确定性对教师网络传递的知识进行筛选,进而帮助学生网络对未增强图像进行准确分割.

5 讨论

虽然本文的BKC-TS框架在非增强CT肝脏肿瘤图像的分割方面取得了显著进展,但本文还存在一定的局限性和未来值得探索的领域.首先,模型对高质量且多样化数据集的依赖性较强.在实际应用中,不同医院和设备生成的CT图像质量和特征可能存在差异,这可能会影响模型的泛化能力.其次,文本信息的质量和准确性对模型性能有重要影响.如果诊断报告和病史记录中存在错误或不完整的信息,可能会导致分割结果的准确性下降.此外,在处理无造影剂图像上非常微小的肿瘤时,本文的方法可能需要进一步改进,以提高对这

些小病灶的识别和分割能力,确保在复杂和细微的临床场景中依然表现出色.未来的研究可以探索更广泛的数据来源和更多样化的临床场景,同时改进多模态信息融合的方法,以进一步提升模型的实际应用价值.此外,可以针对无造影剂图像上非常微小肿瘤的识别和分割,开发更精细的检测和分割技术,并优化不确定性引导的学习模块,以更好地处理复杂和细微的临床场景.

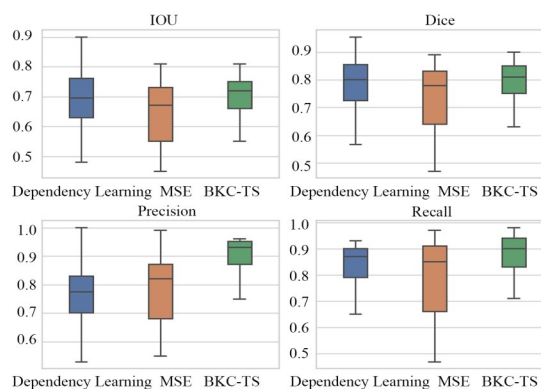


图10 BKC-TS引入不确定性分析优化学生网络,在无造影剂肝脏肿瘤分割中优于各消融版本

6 结论

在本文中,本文提出了一种新的双模态知识协同驱动的教师-学生模型(BKC-TS),用于直接从非增强的肝脏肿瘤CT图像中分割肝脏肿瘤.BKC-TS利用教师网络学习增强图像肿瘤相关的知识来指导学生网络实现非增强图像的肝脏肿瘤准确分割,并且协同临床文本信息共同指导学生网络在无造影剂CT图像中肿瘤的学习过程.该方法在1140张CT图像上取得了最优的肝脏肿瘤分割性能,结果表明,BKC-TS的IOU为75.34%,Dice系数82.15%,精度为92.73%,召回率为92.68%.这些结果表明,BKC-TS具有巨大的潜力,可以作为一种准确的肝肿瘤诊断的临床替代方法,有可能消除对造影剂的需要.

参考文献

- [1] HUANG Q, SUN J F, DING H, et al. Robust liver vessel extraction using 3D U-Net with variant dice loss function[J]. Computers in Biology and Medicine, 2018, 101: 153-162.
- [2] VORONTSOV E, TANG A, PAL C, et al. Liver lesion segmentation informed by joint liver segmentation[C]// 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Piscataway: IEEE, 2018: 1332-1335.
- [3] ALMOTAIRI S, KAREEM G, AOUF M, et al. Liver tumor segmentation in CT scans using modified SegNet[J].

- Sensors, 2020, 20(5): 1516.
- [4] KAVUR A E, GEZER N S, BARIS M, et al. CHAOS challenge-combined (CT-MR) healthy abdominal organ segmentation[J]. *Medical Image Analysis*, 2021, 69: 101950.
- [5] KÜSTNER T, HEPP T, FISCHER M, et al. Fully automated and standardized segmentation of adipose tissue compartments via deep learning in 3D whole-body MRI of epidemiologic cohort studies[J]. *Artificial Intelligence*, 2020, 2(6): e200010.
- [6] LI B N, CHUI C K, CHANG S, et al. A new unified level set method for semi-automatic liver tumor segmentation on contrast-enhanced CT images[J]. *Expert Systems with Applications*, 2012, 39(10): 9661-9668.
- [7] SUN C J, GUO S X, ZHANG H M, et al. Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs[J]. *Artificial Intelligence in Medicine*, 2017, 83: 58-66.
- [8] NASEEM R, KHAN Z A, SATPUTE N, et al. Cross-modality guided contrast enhancement for improved liver tumor image segmentation[J]. *IEEE Access*, 2021, 9: 118154-118167.
- [9] ZHOU J Y, WONG D W K, DING F, et al. Liver tumour segmentation using contrast-enhanced multi-detector CT data: Performance benchmarking of three semiautomated methods[J]. *European Radiology*, 2010, 20(7): 1738-1748.
- [10] ZHANG D, CHEN B, CHONG J, et al. Weakly-Supervised teacher-Student network for liver tumor segmentation from non-enhanced images[J]. *Medical Image Analysis*, 2021, 70: 102005.
- [11] LYU P J, WANG C, FANG W, et al. Registration-propagated liver tumor segmentation for non-enhanced CT-based interventions[M]//*Image and Graphics*. Cham: Springer Nature Switzerland, 2023: 124-136.
- [12] ZHANG J N, LUO S C, QIANG Y, et al. Edge constraint and location mapping for liver tumor segmentation from nonenhanced images[J]. *Computational and Mathematical Methods in Medicine*, 2022, 2022: 1248311.
- [13] XIAO X J, HU Q V, WANG G H. Edge-Aware Multi-Task Network for Integrating Quantification Segmentation and Uncertainty Prediction of Liver Tumor on Multi-Modality Non-Contrast MRI[M]//*Medical Image Computing and Computer Assisted Intervention-MICCAI 2023*. Cham: Springer Nature Switzerland, 2023: 652-661.
- [14] WANG W Z, SONG Q Y, ZHOU J R, et al. Dual-level selective transfer learning for intrahepatic cholangiocarcinoma segmentation in non-enhanced abdominal CT[M]//*Medical Image Computing and Computer Assisted Intervention-MICCAI 2020*. Cham: Springer International Publishing, 2020: 64-73.
- [15] GJESTANG H L, HICKS S A, THAMBAWITA V, et al. A self-learning teacher-student framework for gastrointestinal image classification[C]//*2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*. Piscataway: IEEE, 2021: 539-544.
- [16] WANG L, YOON K J. Knowledge distillation and student-teacher learning for visual intelligence: A review and new outlooks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(6): 3048-3068.
- [17] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//*2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Piscataway: IEEE, 2021: 9992-10002.
- [18] WANG T, YUAN L, ZHANG X P, et al. Distilling object detectors with fine-grained feature imitation[C]//*2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE, 2019: 4933-4942.
- [19] YANG X Y, LI Q J, ZHANG C, et al. Knowledge distillation from multiple foundation models for end-to-end speech recognition[EB/OL]. (2023-03-20) [2025-3-11]. <https://arxiv.org/abs/2303.10917v1>.
- [20] ANAND D, GUPTA V, PARUCHURI P, et al. An enhanced advising model in teacher-student framework using state categorization[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(8): 6653-6660.
- [21] CHEN H T, WANG Y H, XU C, et al. Learning student networks via feature embedding[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(1): 25-35.
- [22] GOU J P, YU B S, MAYBANK S J, et al. Knowledge distillation: A survey[J]. *International Journal of Computer Vision*, 2021, 129(6): 1789-1819.
- [23] RADOSAVOVIC I, KOSARAJU R P, GIRSHICK R, et al. Designing network design spaces[C]//*2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE, 2020: 10428-10436.
- [24] LANG C B, CHENG G, TU B F, et al. Learning what not to segment: A new perspective on few-shot segmentation[C]//*2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE, 2022: 8047-8057.
- [25] YU L Q, WANG S J, LI X M, et al. Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium

- segmentation[M]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Cham: Springer International Publishing, 2019: 605-613.
- [26] AHMAD USMANI U, WATADA J, JAAFAR J, et al. A reinforced active learning algorithm for semantic segmentation in complex imaging[J]. IEEE Access, 2021, 9: 168415-168432.
- [27] RAKESH V, JAIN S. Efficacy of Bayesian neural networks in active learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2021: 2601-2609.
- [28] FENG D H, CHEN X, WANG X Y, et al. Penalized Entropy: A novel loss function for uncertainty estimation and optimization in medical image classification[C]//2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS). Piscataway: IEEE, 2022: 306-310.
- [29] LIU X B, SONG L P, LIU S, et al. A review of deep-learning-based medical image segmentation methods[J]. Sustainability, 2021, 13(3): 1224.
- [30] ZHANG G S, ZHAO K H, HONG Y F, et al. SHA-MTL: Soft and hard attention multi-task learning for automated breast cancer ultrasound image segmentation and classification[J]. International Journal of Computer Assisted Radiology and Surgery, 2021, 16(10): 1719-1725.
- [31] LOU A G, GUAN S Y, KO H, et al. CaraNet: Context axial reverse attention network for segmentation of small medical objects[C]//Medical Imaging 2022: Image Processing. California: SPIE, 2022: 81-92.
- [32] ZHOU T, ZHOU Y, HE K L, et al. Cross-level feature aggregation network for polyp segmentation[J]. Pattern Recognition, 2023, 140: 109555.
- [33] FAN D P, JI G P, ZHOU T, et al. PraNet: Parallel reverse attention network for polyp segmentation[M]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2020. Cham: Springer International Publishing, 2020: 263-273.
- [34] Kim T, LEE H, KIM D. Uacanet: Uncertainty augmented context attention for polyp segmentation[EB/OL]. (2021-07-22)[2025-03-11]. <https://arxiv.org/abs/2107.02368v3>.
- [35] XU C C, WU X, WANG B Y, et al. Accurate segmentation of liver tumor from multi-modality non-contrast images using a dual-stream multi-level fusion framework[J]. Computerized Medical Imaging and Graphics, 2024, 116: 102414.
- [36] ZHANG R F, LI G B, LI Z, et al. Adaptive context selection for polyp segmentation[M]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2020. Cham: Springer International Publishing, 2020, 1: 253-262.
- [37] HU C F, XIA T Y, CUI Y, et al. Trustworthy multi-phase liver tumor segmentation via evidence-based uncertainty[J]. Engineering Applications of Artificial Intelligence, 2024, 133: 108289.
- [38] LIU H X, YANG J, JIANG C, et al. S2DA-Net: Spatial and spectral-learning double-branch aggregation network for liver tumor segmentation in CT images[J]. Computers in Biology and Medicine, 2024, 174: 108400.

作者简介



刘 婕 女, 1994年1月生, 安徽亳州人。博士, 安徽省立医院南区影像中心主治医师。主要研究方向为医学人工智能。
E-mail: 528579064@qq.com



徐晨初 男, 1987年6月生, 安徽合肥人。安徽大学计算机科学与技术学院副教授、硕士生导师。主要研究方向为医疗人工智能、机器学习、医学图像分析。
E-mail: xcc@ahu.edu.cn



孙怡宁 男, 1962年12月生, 安徽太湖人。中科院合肥物质科学研究院研究员、博士生导师, 合肥智能所首席科学家, 运动与健康信息技术学科带头人。主要研究方向为光机电一体化技术、机器人技术、传感技术。
E-mail: ynsun@iim.ac.cn



熊 焰 男, 1960年8月生, 安徽合肥人。中国科学技术大学计算机科学与技术学院教授、博士生导师。主要研究方向为计算机网络与信息安全、移动计算与移动网络、分布式处理等。
E-mail: yxiong@ustc.edu.cn