

中图法分类号: 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-15

论文引用格式: Zhong Keyi, Huang Xuhui, Wan Huan, Miao Xuanlei, Wei Xin. Annotation-free liver tumor segmentation framework guided by clinical imaging knowledge[J/OL]. Journal of Image and Graphics, XXXX: 1-15. DOI: 10.11834/jig.250632. (钟可毅, 黄旭慧, 万欢, 缪轩磊, 魏欣. 临床影像知识引导的无标注肝肿瘤分割框架[J/OL]. 中国图象图形学报, XXXX: 1-15. DOI: 10.11834/jig.250632.) [DOI: 10.11834/jig.250632]

临床影像知识引导的无标注肝肿瘤分割框架

钟可毅¹, 黄旭慧¹, 万欢², 缪轩磊^{3,4}, 魏欣¹

1. 南昌大学软件学院, 南昌 330047; 2. 江西师范大学人工智能学院, 南昌 330022; 3. 江西中医药大学, 南昌 330004; 4. 江西中医药大学附属医院, 南昌 330006

摘要: 目的 肝肿瘤的精确分割常受限于高质量标注数据的稀缺。尽管数据合成技术提供了替代方案, 但现有的一些合成方法容易生成缺乏解剖合理性的样本。此外, 合成图像中存在的非自然伪影, 可能导致模型陷入捷径学习, 进而影响其在真实临床数据上的泛化性能。为应对这些挑战, 设计一种无需真实肿瘤标注的肝肿瘤分割框架 Know2Seg。方法 首先, 结合相关临床影像知识, 在不同的合成环节设置相应约束, 实现临床知识引导的合成策略: 利用椭球先验与弹性变形等方法生成肿瘤掩膜, 通过全变分去噪与小波变换等处理构建多尺度的合成纹理, 进而生成符合临床解剖规律的CT(computed tomography)扫描及其分割标签。其次, 为缓解合成伪影对模型训练的干扰, 利用合成过程的中间态真值, 以及模型输出的肿瘤后验概率, 作为辅助监督信号。通过基于图像解耦思想的训练架构引导模型将输入图像显式解耦为健康解剖与病灶纹理两个成分, 引导模型将注意力从合成伪影转移至深层语义病理特征。结果 实验表明, Know2Seg在不使用真实肿瘤标注的前提下, 分割性能与同等实验条件下的全监督方法相当, 并且优于一些其他的无标注方法。在LiTS数据集上, Know2Seg的Dice相似系数(dice similarity coefficient, DSC)达到62.9%、归一化表面Dice(normalized surface dice, NSD)达到65.5%、表面距离(surface distance, SD)降至14.0 mm、豪斯多夫距离(95% hausdorff distance, HD95)降至40.4 mm。结论 本文通过将相关临床影像知识转化为可计算的合成约束, 并利用图像解耦思想缓解了捷径学习问题, 构建了一个兼顾可解释性与鲁棒性的无标注肿瘤分割框架。

关键词: 无标注学习; 3D肿瘤分割; 临床知识; 数据合成; 图像解耦

Annotation-free liver tumor segmentation framework guided by clinical imaging knowledge

Zhong Keyi¹, Huang Xuhui¹, Wan Huan², Miao Xuanlei^{3,4}, Wei Xin¹

1. School of Software, Nanchang University, Nanchang 330047, China; 2. School of Artificial Intelligence, Jiangxi Normal University, Nanchang 330022, China; 3. Jiangxi University of Chinese Medicine, Nanchang 330004, China; 4. The Affiliated Hospital of Jiangxi University of Chinese Medicine, Nanchang 330006, China

Abstract: **Objective** Liver cancer is one of the malignant tumors with the highest mortality worldwide. Accurate tumor segmentation based on computed tomography (CT) is of great significance for clinical diagnosis, surgical planning, and effi-

收稿日期: 2025-12-17; 修回日期: 2026-08-23

基金项目: 国家自然科学基金(62566040); 江西省自然科学基金(20232BAB212027, 20242BAB23013, 20252BAC200194); 江西省千人计划(jxsg2023101085); 江西省教育厅青年项目(GJJ2400824); 江西中医药大学重点学科建设经费资助课题(2023jzdxk034)

Supported by: The National Natural Science Foundation of China (62566040); The Natural Science Foundation of Jiangxi (20232BAB212027, 20242BAB23013, 20252BAC200194); The Thousand Talents Plan of Jiangxi Province (jxsg2023101085); The Youth Project of Jiangxi Provincial Department of Education (GJJ2400824); The Key Discipline Construction Fund of Jiangxi University of Chinese Medicine. (2023jzdxk034)

©中国图象图形学报版权所有

cacy evaluation. Although deep learning has been applied in the field of medical image segmentation, under the fully supervised training paradigm, precise segmentation of liver tumors is constrained by the scarcity of high-quality annotated data. Specifically, the complex shapes and fuzzy boundaries of liver tumors, the high cost of manual labeling, and inter-observer variability lead to the scarcity of available high-quality training data, which has become a key bottleneck restricting the clinical application of such models. Although data synthesis provides a way to alleviate data scarcity, some existing methods may generate samples inconsistent with clinical anatomical rules. Furthermore, unnatural artifacts in synthetic images may cause models to learn shortcut cues, limiting their applicability to real tumor CT cases. To address these challenges, an annotation-free liver tumor segmentation framework, named Know2Seg, is developed. The objective of this framework is to alleviate the dependence of segmentation model training on real tumor annotations while making the synthesis process consistent with clinical rules and guiding the model to learn meaningful tumor-related features from synthetic data. **Method** Know2Seg mainly includes clinical knowledge-guided liver tumor synthesis and a training architecture. In the synthesis stage, rather than relying on generative networks, relevant clinical imaging knowledge is incorporated into the synthesis process in the form of computable constraints. The synthesis process contains four steps: tumor location sampling, morphological simulation, texture synthesis, and image fusion. These steps are designed to make the generated tumors more consistent with clinical and pathological characteristics. In location sampling, ellipsoid-constrained sampling and local anatomical checking are used to reduce anatomically unreasonable tumor placement. In morphological simulation, synthetic tumors are divided into different size categories, including micro, small, medium, and large lesions. Elastic deformation are used to generate irregular and anisotropic tumor shapes that conform to pathological features. In texture synthesis, random noise is processed by total variation denoising, wavelet decomposition, inverse reconstruction, and gaussian filtering to simulate the heterogeneous density patterns of liver tumors. In image fusion, a density attenuation field modulated by tumor texture is applied to healthy liver CT images. This strategy preserves high-density anatomical structures, such as vessels, and reduces unrealistic truncation artifacts caused by simple copy-paste operations. In the training stage, a training architecture informed by the idea of image disentanglement is used to support the model's generalization performance on real tumor CT images. Instead of directly learning the mapping from synthetic images to segmentation labels, the training objectives encourage the network to learn representations of healthy anatomy and lesion texture. The network outputs reconstruction maps and semantic segmentation logits. The segmentation logits are further converted into a tumor posterior probability map, which can be used as the blending weight for the reconstruction maps. Through this design, the model learns segmentation together with the reconstruction of healthy anatomy and tumor texture, which distinguishes normal structures from abnormal lesions. Four losses are jointly optimized, including healthy liver reconstruction loss, tumor texture reconstruction loss, segmentation loss, and blending consistency loss. With the intermediate ground truths generated during synthesis, this multi-task design makes the model learn more distinguishable representations of normal anatomy and abnormal lesions. **Result** Experiments are conducted on public three-dimensional abdominal CT datasets. Know2Seg is compared with representative annotation-free tumor segmentation methods, including anomaly detection-based methods and synthesis-based methods. The results show that Know2Seg achieves promising segmentation performance. It obtains a dice similarity coefficient (DSC) of 62.9%, a normalized surface dice (NSD) of 65.5%, a surface distance (SD) of 14.0 mm, and a 95% hausdorff distance (HD95) of 40.4 mm. These results indicate that the framework performs favorably among the compared annotation-free methods. Since no real tumor annotations are used during training, the performance also suggests that the knowledge-guided tumor synthesis strategy can provide effective supervisory signals for downstream tumor segmentation, while the training architecture can improve the model's generalization ability on real data. The DSC and NSD results suggest good volumetric overlap and boundary consistency. The SD and HD95 results further indicate that the predicted tumor boundaries are relatively consistent with anatomical structures. Ablation experiments and hyperparameter sensitivity experiments further verify the effectiveness and robustness of the framework. These analyses indicate that both the synthesis strategy and the training architecture contribute to the overall performance. **Conclusion** By transforming relevant clinical imaging knowledge into computable synthesis constraints and using a training strategy informed by image disentanglement to reduce the risk of shortcut learning, an annotation-free learning framework characterized by both interpretability and robustness is constructed. It does not require the pretraining of complex generative networks and helps reduce the reli-

ance of model training on manual annotations.

Key words: annotation-free learning; 3D tumor segmentation; clinical knowledge; data synthesis; image disentanglement

论文引用格式: Zhong Keyi, Huang Xuhui, Wan Huan, Miao Xuanlei, Wei Xin. Annotation-free liver tumor segmentation framework guided by clinical imaging knowledge — SCID[J/OL]. Journal of Image and Graphics. DOI: 10.11834/jig.250632. (钟可毅, 黄旭慧, 万欢, 缪轩磊, 魏欣. 临床影像知识引导的无标注肝肿瘤分割框架—SCID[J/OL]. 中国图象图形学报. DOI: 10.11834/jig.250632.)

0 引言

肝癌是全球范围内高致死率的恶性肿瘤之一(Hu等, 2024), 尽早诊断与治疗是提升患者生存率的关键。计算机断层扫描(computed tomography, CT)等影像技术能提供精细的解剖结构信息, 已经成为肝癌临床评估的重要工具。在医学影像上精确地进行肿瘤分割, 不仅是实现病灶量化分析、疗效评估与手术规划的前提, 更是当前医学图像分析领域亟待突破的核心挑战(Antonelli等, 2022)。然而, 肝肿瘤在门静脉期CT中可表现为低密度区域, 与肝实质之间的密度差异不均匀(Lee等, 2004; Liang等, 2021), 并且其边界模糊, 形态不规则。这些特性使得传统的人工标注肿瘤十分耗时、繁琐, 且结果依赖医生的个人经验, 导致不同医生之间的手工标注结果可能存在差异, 难以标准化(Gotra等, 2017; Schmidt-Mengin等, 2024)。

面对人工筛查与标注肿瘤存在的局限性, 以传统机器学习与深度学习为代表的人工智能技术被引入医学图像处理领域, 并已经取得了一定应用成果(Yang等, 2020; Mokrane等, 2020; Castaldo等, 2021; Wei等, 2023)。但在全监督训练范式下, 深度学习模型的性能水平通常依赖于大规模、高质量的像素级或体素级标注数据, 而针对特定器官的公开高质量数据集的规模较小、病例多样性相对匮乏, 这构成了当前人工智能在医学图像分割任务中的关键瓶颈之一。在这种标注受限的条件下, 模型的训练与泛化受到了挑战(Isensee等, 2021; Tang等, 2022)。为了缓解对大规模人工标注的依赖, 数据合成技术, 特

别是基于生成对抗网络(generative adversarial networks, GANs)与扩散模型(diffusion models)的方法, 已经成为一种有效的替代方案(Li等, 2025)。虽然这类以数据驱动为主的方法在合成常规的二维图像时的视觉逼真度较高, 但合成医学图像时面临一定挑战, 主要问题在于其缺乏显式的解剖约束, 难以精准地控制肿瘤的边界浸润与纹理异质性等临床特征, 易生成违反医学常识的样本, 限制了合成数据在医疗场景中的应用价值。因此如何合成符合临床解剖规律的数据, 是医学图像合成领域亟待解决的核心问题之一。

此外, 由于图像合成的部分环节常涉及图像的线性叠加与局部融合, 合成过程可能会引入与真实数据分布不同的统计特征。深度神经网络(尤其是卷积架构)在优化过程中存在纹理偏差(texture bias)现象, 即倾向于依赖图像中的局部高频伪影(Geirhos等, 2019), 例如边缘的拼接痕迹或异常的梯度突变, 而非学习真正的病理语义特征(Zhang等, 2023b), 导致模型在真实临床数据上的泛化性能下降, 这一现象被称为捷径学习(shortcut learning)(Geirhos等, 2020)。

针对上述局限, 本文设计了一种无需真实肿瘤标注的三维肝肿瘤分割框架Know2Seg。该框架的合成阶段不依赖大规模数据训练生成器, 而将医生诊断肿瘤的临床知识算法化, 能生成大量合成肿瘤, 并将其融合至健康肝脏图像中。该框架的训练阶段采用了基于图像解耦思想的训练架构, 引导网络执行合成阶段的近似逆过程, 将输入图像显式解耦为健康背景与病变纹理两个不同成分, 有助于网络理解解剖结构的完整性与病灶的独立性, 相比单纯端到端学习体素级的分类, 模型具有更强的泛化能力。在本文实验中, 该框架表现出优秀的分割性能。

本文的主要工作可概括如下。

1) 实现了临床知识引导的肿瘤合成策略: 结合相关临床影像知识, 在各合成环节设置相关约束, 能够生成形态比较符合临床规律的病灶。

2) 设计了基于图像解耦思想的训练架构: 利用合成过程的中间态真值, 以及模型输出的肿瘤后验概率, 引导网络同时学习重建与分割的多任务, 可以

缓解模型在训练过程针对合成伪影的捷径学习问题,赋予模型一定区分正常解剖背景与异常病变区域的可解释能力。

3) 构建了一个相对高效的无标注分割框架: Know2Seg 基于临床知识合成数据,合成阶段无需采集真实病变图像,也无需预训练生成模型,可节省训练前的计算开销与时间成本。

4) 验证了算法化与解耦化临床特征的有效性: 实验结果表明,相较于从有限数据中以端到端方式间接学习隐式特征,通过学习显式的解耦化临床特征,模型能够更有效地捕捉肿瘤复杂的分布规律。

1 相关工作

1.1 无标注肿瘤检测与分割

在缺乏真实标注数据的情况下,无标注病灶检测与分割的可利用基于重建的方法,其原理是利用健康数据训练自编码器或生成对抗网络,通过计算输入图像与修复后图像的残差来定位异常区域。例如, Schlegl 等(2019)提出的 f-AnoGAN 利用生成对抗网络,在视网膜 OCT(optical coherence tomography) 图像的异常检测与定位中取得了突破; Baur 等(2021)系统性地验证了自编码器在脑部 MRI(magnetic resonance imaging) 无监督异常分割中的有效性。

虽然上述异常检测方法在脑部或眼底等背景结构相对规律的医学影像中取得了一定成果,但将其直接应用于腹部器官时却面临挑战。在腹部 CT 影像中,以肝脏为例,肝脏及其周边的解剖结构相对复杂,门静脉与肝静脉等血管网络的密度范围与部分病灶存在重叠,容易导致基于重建误差的方法产生假阳性结果。此外,肝脏肿瘤在尺寸、边界浸润及内部纹理上呈现一定异质性,使得缺乏明确医学先验的通用异常检测算法可能难以准确解构肝肿瘤的特病理特征。

近期还出现了一些基于合成数据的无标注学习研究,通过在健康器官图像中植入合成肿瘤来构建训练对。Hu 等(2023)的研究通过植入具有特定纹理和形状的椭球体来模拟肿瘤,其视觉图灵测试证实了其合成方法能够达到临床级的视觉逼真度。为了模拟肿瘤在发育过程中的复杂的形态, Lai 等(2024)引入了元胞自动机理论,通过设定细胞生长

与侵袭规则,成功生成了具有不规则边缘的病灶形态。

然而,上述合成方法主要侧重于提升样本的视觉逼真度或形态复杂性,但对于模型在利用合成数据训练时的捷径学习风险,缺乏专门的讨论。针对这个问题, Know2Seg 在合成阶段融合了临床知识,也在模型的训练阶段利用了基于图像解耦思想的架构,通过显式的解耦机制引导模型缓解合成伪影的干扰,学习更深层的病理语义特征,从而实现更加鲁棒的无标注肝肿瘤分割。

1.2 医学图像合成

为应对高质量标注医学数据稀缺、采集成本高昂、患者隐私保护等核心挑战,医学图像合成技术近年来受到广泛关注(Wang 等, 2021; Li 等, 2025)。生成式模型,例如生成对抗网络和扩散模型,已在从脑部 MRI 到皮肤镜图像等多种模态的数据生成中证明了其应用潜力(Han 等, 2018; Zhang 等, 2023a; Sindhura 等, 2024)。

尽管二维合成技术已相对成熟,但三维医学影像因其能提供完整的空间解剖上下文信息,在肿瘤检测等任务中有重要价值。近期研究在三维腹部肿瘤及脑肿瘤(Zhang 等, 2023c; Hu 等, 2023; Chen 等, 2024; Lai 等, 2024)的合成方面取得了进展。鉴于三维体素计算带来的高昂显存与算力开销,一些轻量化的生成策略也应运而生(Peng 等, 2022; Han 等, 2023; Zhu 等, 2023)。最近, Shao 等(2025)提出了一种基于扩散生成对抗网络的 MRI 到 CT 跨模态合成方法,有效地保留了医学图像的结构细节; Zhao 等(2026)开发的 MAISI-v2 框架,支持高分辨率 3D 多模态数据的快速生成。

然而,一些合成方法比较侧重视觉上的逼真程度,而对合成数据的解剖合理性关注度较低。例如,一些生成对抗网络虽能生成纹理相对自然的图像,但在控制肿瘤的异质性分布、边界浸润程度等对诊断关键的细微特征方面可能存在一定进步空间。而 Know2Seg 将相关临床影像知识转化为合成约束,使合成肿瘤在临床特征上合理与多样化,从而提升合成数据在训练分割模型时的可靠性。

1.3 合成数据在下游任务中的应用

高质量的合成数据已成为增强医学影像模型性能的关键驱动力之一(Li 等, 2022)。合成数据可应用于重建、配准等环节(Pan 等, 2025),并且在图像

分类、语义分割等多项下游任务中,引入合成数据可带来一定性能提升(He等,2022;Zhao等,2023)。

合成数据的核心优势在于其域适应能力与针对长尾分布(long-tailed distribution)的补全能力(Sankaranarayanan等,2018;Wu等,2023;Guo等,2025)。合成数据能够按需生成不同形态的训练样本,弥补真实数据集中可能存在的分布失衡问题。

近期研究表明,使用合成图像扩充小型医学数据集,能有效提升下游分割任务的性能(Vu等,2025);Xie等(2026)提出了一种扩散模型微调框架MedDiff-FT,通过合成图像与标注,提高了其在多种医学数据集上的性能。

上述研究主要验证了合成数据作为增强手段的有效性,这也启发了本文进一步探索其作为核心训练数据的潜力,利用合成图像为下游模型提供丰富

的训练数据,缓解其对稀缺的真实病变标注数据的依赖。

1.4 模型训练中的图像特征解耦

合成图像为模型提供了新的训练条件,但模型未必总是能学习到其中真正与病灶有关的影像特征。Geirhos等(2020)指出,深度学习模型在训练过程中存在捷径学习现象,模型可能优先学习与预测目标存在统计学相关性,但缺乏实际语义的特征,因此在训练数据上能够获得较好的验证结果,但在分布不同的数据上评估的性能下降。Lin等(2024)的研究也表明,在医学图像分割任务中模型同样存在捷径学习现象,医学图像的预处理痕迹以及数据裁剪方式等因素,都可能在训练过程中被模型当作判别依据,这种与病灶无关的特征会削弱模型在真实临床数据上的泛化能力。

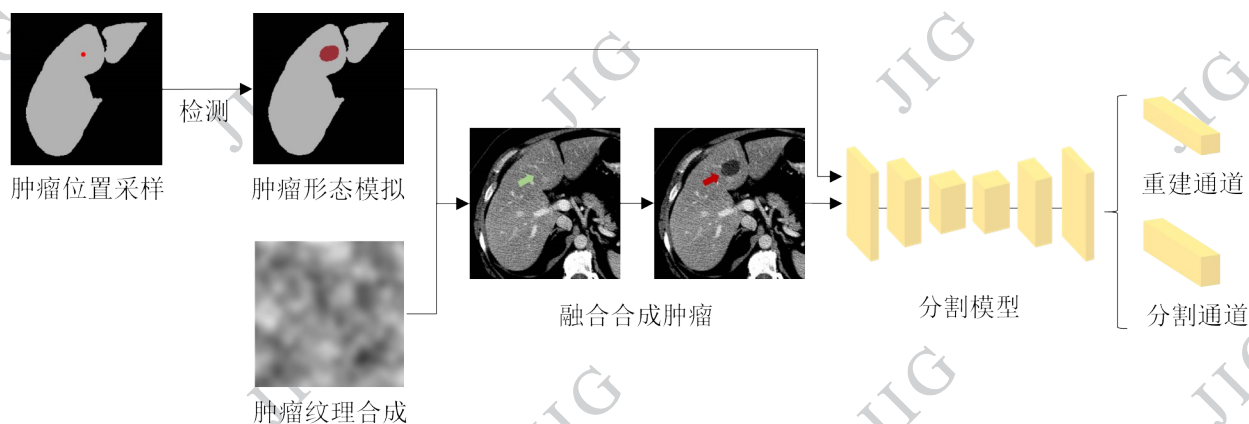


图1 Know2Seg的示意图

Fig. 1 Schematic diagram of Know2Seg

为了减少图像中无关因素对分割结果的干扰,有一些研究尝试将图像的特征解耦。Chartsias等(2019)在心脏图像分析中,将表征解剖结构的空间信息与表征成像外观的模态信息分开建模,减少了二者在特征表示中的混杂。这类研究说明,在训练过程中适当地区分图像中的不同成分,有助于模型学习并在潜空间中保留与解剖结构相关的特征。Vorontsov等(2022)在半监督肿瘤分割研究中,将病变信息从背景组织信息分离开来,并利用训练健康与患病图像之间的双向转换任务辅助模型训练,这说明对图像中的背景信息和病灶信息进行分开建模,有助于改善标注有限条件下模型的分割性能。此外,Zhang等(2023b)在自监督肿瘤分割研究中也

将图像中不同信息解耦,采用预训练和再适配的两阶段训练方式增强模型的泛化性能,强调预训练任务的目标也应尽量靠近下游分割任务。上述研究表明,在肿瘤分割任务中,若能在训练阶段适当解耦图像中的不同信息,则有利于模型学习到稳定的病灶特征。

基于上述认识,本文认为,无标注肿瘤分割不仅依赖于合成样本的解剖合理性,还依赖于训练阶段对特征学习的合理引导。所以,本文将输入的合成图像中的解剖背景与病灶纹理分别建模,并利用合成过程中可获得的中间信息作为监督信号,从而提高模型在真实临床数据上的泛化能力。

两项核心内容:

2 方法

无标注肝脏肿瘤分割框架 Know2Seg 主要包含

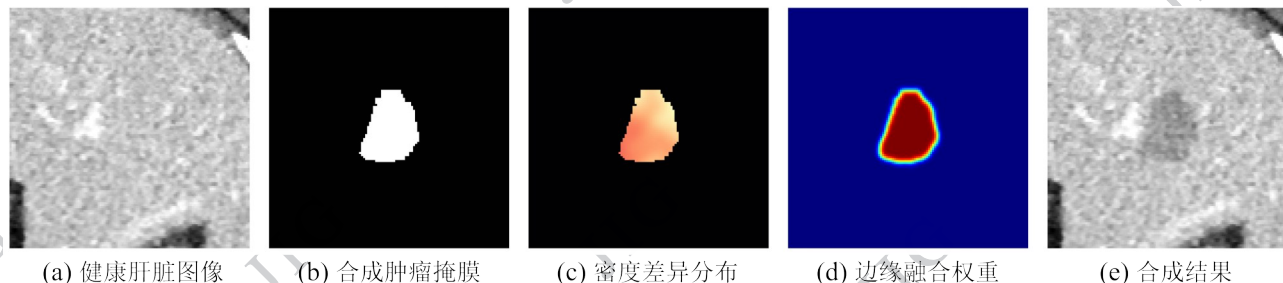


图2 临床知识引导的肿瘤合成过程可视化示例

Fig. 2 Visualization of the clinical knowledge-guided tumor synthesis process

2.1 节介绍了临床知识引导的合成策略,该策略将相关临床影像知识转化为可计算的合成约束,以生成合成肿瘤,作为后续模型的训练数据(见图1)。

2.2 节阐述了基于图像解耦思想的训练架构,利用合成阶段的中间态真值作为监督信号,以及模型输出的后验概率,引导模型将输入图像显式解耦为健康背景与病灶纹理。这种机制可以缓解模型利用合成数据训练时存在的捷径学习问题。

2.1 临床知识引导的肝肿瘤图像合成

本文实现了一种基于临床知识的合成策略,能够在正常肝脏CT扫描图像中生成大量符合临床解剖规律的肿瘤。如图1所示,Know2Seg的合成过程由四个核心环节构成:肿瘤位置采样、形态模拟、纹理合成以及融合。图2展示了从健康背景到最终合成样本的部分中间过程。

在图2中,图(a)为原始CT图像中的健康肝脏区域,进行了适当裁剪与放大。图(b)为定义的肿瘤掩膜。图(c)为热力图,颜色越浅,则代表该处相对于肝实质的密度差异程度越大;颜色越深,则代表密度差异程度越小,该处的密度更接近原始健康组织。图(d)也为热力图,颜色越偏红,则代表该处越靠近肿瘤中心;颜色越偏蓝,代表该处越靠近肿瘤边界,体现了肿瘤边缘的模糊与平滑过渡效果。

2.1.1 肿瘤位置采样

肝肿瘤的位置分布具有一定随机性,例如肝细胞癌(hepatocellular carcinoma, HCC)因其右叶体积优势等因素而更常见于右叶;转移性肿瘤受到血流

动力学等因素的作用,可分布在肝脏的外周。由于肝脏右叶体积稍大,因此,在体积等概率采样的条件下,肝脏右叶自然会获得更大的采样权重。考虑到椭球体的体积随半径呈立方增长,将随机采样向量 u 的模长校正为 $r = u^{1/3} (u \sim \mathcal{U}[0, 1])$,避免采样点过度集中在球体中心区域。随后,通过仿射变换将采样点映射至目标解剖空间,生成候选的采样点坐标 P_{cand} :

$$P_{\text{cand}} = C + R \cdot A \cdot u \quad (1)$$

式中 C 为椭球中心向量, R 为由主轴方向决定的旋转变换矩阵, A 为控制半轴长度的缩放系数。

本文还设计了一种采样点的检测机制。定义肝脏掩膜为 $M_{\text{liver}}(x)$, 且 $\Omega(P, s)$ 表示以 P 为中心、邻域边长为 s 的局部三维邻域,采样点需要满足以下条件,以避免采样点落入肝脏边缘、非肝实质等不符合临床解剖规律的区域,并能一定程度上避开大血管密集区域:

$$P_{\text{cand}} \in \left\{ P \left| \frac{1}{|\Omega|} \sum_{v \in \Omega(P_{\text{cand}}, s)} M_{\text{liver}}(v) > \tau_{\text{vol}} \wedge \|\nabla I(P_{\text{cand}})\| < \tau_{\text{grad}} \right. \right\} \quad (2)$$

式中 τ_{vol} 为肝实质体素阈值; $\|\nabla I\|$ 为梯度的模, τ_{grad} 为梯度阈值。

2.1.2 形态模拟

肝肿瘤在生长时会受周围组织的物理挤压,且可呈现非规则的各向异性形态。临床上,这些肿瘤可表现为不同尺度的结节状聚集(Lee等,2004),其

形态多近似于球体或椭球体。因此, 本文将合成肿瘤按照尺寸划分为不同类别, 每个类别对应特定的半径和生成数量, 如表1所示。此外, 在合成肿瘤时, 还可以在同一肝脏中同时生成不同尺寸的病灶, 以模拟多发性肝癌与卫星灶场景。

表1 肝肿瘤尺寸分级

Table 1 Classification for liver tumor dimensions

类别	基础半径(体素)	等效半径(mm)	生成数量
微型	4	3.0	$n \sim \mathcal{U}\{3, 10\}$
小型	8	6.0	$n \sim \mathcal{U}\{3, 10\}$
中型	16	12.0	$n \sim \mathcal{U}\{2, 5\}$
大型	32	24.0	$n \sim \mathcal{U}\{1, 3\}$

注: 基于0.75 mm体素分辨率进行换算。

本文将肝肿瘤的形态近似为椭球体。对于给定的基础半径 r , 分别在三个半轴方向 $i \in \{x, y, z\}$ 上随机扰动, 得到实际半轴长度 r_i 。但规则的椭球体缺乏真实感, 所以需要引入一定形态形变来增强合成肿瘤的多样性与真实性。对于利用 r_i 生成的初始掩膜 $M_{\text{geo}}(\mathbf{x})$, 引入弹性形变(Simard等, 2003), 且肿瘤越大形变程度越强:

$$M_{\text{tumor}}(\mathbf{x}) = M_{\text{geo}}(\mathbf{x} + \Phi(\mathbf{x})) \quad (3)$$

式中 $\Phi(\mathbf{x})$ 是位移场。这种形变使得合成肿瘤边缘呈现出适当的凹凸不平形态, 可以有效提升模型对肿瘤形态异质性的泛化能力。

2.1.3 纹理合成

肝肿瘤在CT图像中表现为相对复杂的异质性纹理。为模拟真实的肿瘤纹理特征, 首先生成服从均匀分布的随机噪声场 $N_{\text{uni}} \sim \mathcal{U}[0, 1]$, 为去除离散的噪点并构建主要的纹理结构特征, 利用全变分去噪处理一次初始噪声场, 将无序的噪声转化为若干灰度值相对均匀的区块, 在视觉上模拟肿瘤内部因坏死、纤维分隔或血供不均等病理改变导致的宏观密度不均现象。随后, 利用离散小波变换将图像的成分进行分解, 削弱纹理合成过程中产生的高频成分, 同时保留中低频的组织结构信息。再经过逆小波重构与高斯滤波处理, 得到初步的纹理 $T_{\text{base}}(\mathbf{x})$:

$$T_{\text{base}}(\mathbf{x}) = \mathcal{G}_{\sigma_r} \left(\mathcal{W}^{-1} \left(\Gamma \cdot \mathcal{W}(\mathcal{D}_{\text{TV}}(N_{\text{uni}})) \right) \right) \quad (4)$$

式中 $\mathcal{D}_{\text{TV}}$ 为全变分去噪算子, $\mathcal{W}$ 与 $\mathcal{W}^{-1}$ 分别为小波分解与重构, Γ 为衰减系数, 用于滤除图像高频成分; $\mathcal{G}_{\sigma_r}$ 为高斯滤波, σ_r 可以控制纹理的整体平滑程

度。 $T_{\text{base}}(\mathbf{x})$ 后续会再被归一化, 得到纹理 $T_{\text{syn}}(\mathbf{x})$ 。

2.1.4 融合

肝脏固有的解剖结构, 例如门静脉等血管, 在门静脉期因造影剂通常表现为相对肿瘤更高的密度。这些结构特征使得本身密度较低的肿瘤在图像检测和分割任务中更具挑战性。常规的Copy-Paste方案直接将肿瘤图案叠加到肝脏图像上, 可能破坏原有解剖结构, 产生不真实的截断伪影。为了避免这一问题, 本小节的核心思想是在原始肝脏CT值的基础上, 减去一个受纹理调制的衰减场。

首先, 为了模拟CT影像中真实肿瘤的边缘模糊现象, 对合成肿瘤的掩膜进行高斯滤波, 生成描述体素属于肿瘤概率的Alpha掩膜, 它在肿瘤中心趋近于1, 而在边缘处逐渐衰减为0, 可用于后续的融合:

$$\alpha(\mathbf{x}) = \mathcal{G}_{\sigma_a}(M_{\text{tumor}}(\mathbf{x})) \quad (5)$$

随后, 在肝脏区域生成肿瘤纹理图 $I_{\text{lesion}}(\mathbf{x})$, 以表征病灶在无背景干扰下的灰度特征。同时构建衰减场 $D(\mathbf{x})$, 式中 δ 是肿瘤的密度衰减幅度, $I_{\text{orig}}(\mathbf{x})$ 为原始CT图像:

$$I_{\text{lesion}}(\mathbf{x}) = I_{\text{orig}}(\mathbf{x}) - \delta \cdot T_{\text{syn}}(\mathbf{x}) \quad (6)$$

$$D(\mathbf{x}) = \delta \cdot T_{\text{syn}}(\mathbf{x}) \cdot \alpha(\mathbf{x}) \quad (7)$$

最后利用肝脏掩膜 M_{liver} 将衰减场融合进原始CT图像 $I_{\text{orig}}(\mathbf{x})$ 中, 得到最终合成图像 $I_{\text{syn}}(\mathbf{x})$:

$$I_{\text{syn}}(\mathbf{x}) = I_{\text{orig}}(\mathbf{x}) - D(\mathbf{x}) \cdot M_{\text{liver}}(\mathbf{x}) \quad (8)$$

由公式(8)可知, 密度衰减仅在肝脏区域内的肿瘤位置(即 $\alpha(\mathbf{x}) > 0$ 的区域)执行, 而肝脏内的血管等较高密度结构由于在 $I_{\text{orig}}(\mathbf{x})$ 中其CT初始值较高, 在减去 $D(\mathbf{x})$ 后依然能保持相对高密度, 从而自然地保留了解剖细节。

2.2 基于图像解耦思想的训练架构

传统的合成数据训练范式通常直接在合成图像与分割标签间建立端到端的映射关系。然而, 这种单一的监督模式存在一定的局限性。合成图像可能包含一些不自然的生成痕迹, 而模型在优化过程中存在纹理偏差现象, 即倾向于关注合成过程中产生的特定噪声模式或边缘伪影, 而非具有临床泛化意义的病理特征, 陷入捷径学习。仅依赖单一的分割损失函数可能难以规范模型的特征提取行为, 导致模型难以稳定区分背景与前景, 进而限制了模型在真实临床数据上的泛化性能。

针对上述问题, 现有的一些研究理论提供了新

的视角。Chen 等人(2024)的研究表明,不同器官间的早期肿瘤具有相似的外观特征。这提示病灶纹理与所处器官背景具备可解耦性,医学影像可近似视

为健康解剖背景与病灶纹理的独立线性叠加。Zhang 等(2023b)从解耦学习的视角指出,通过显式图像解耦可有效引导模型学习这种独立性表征。

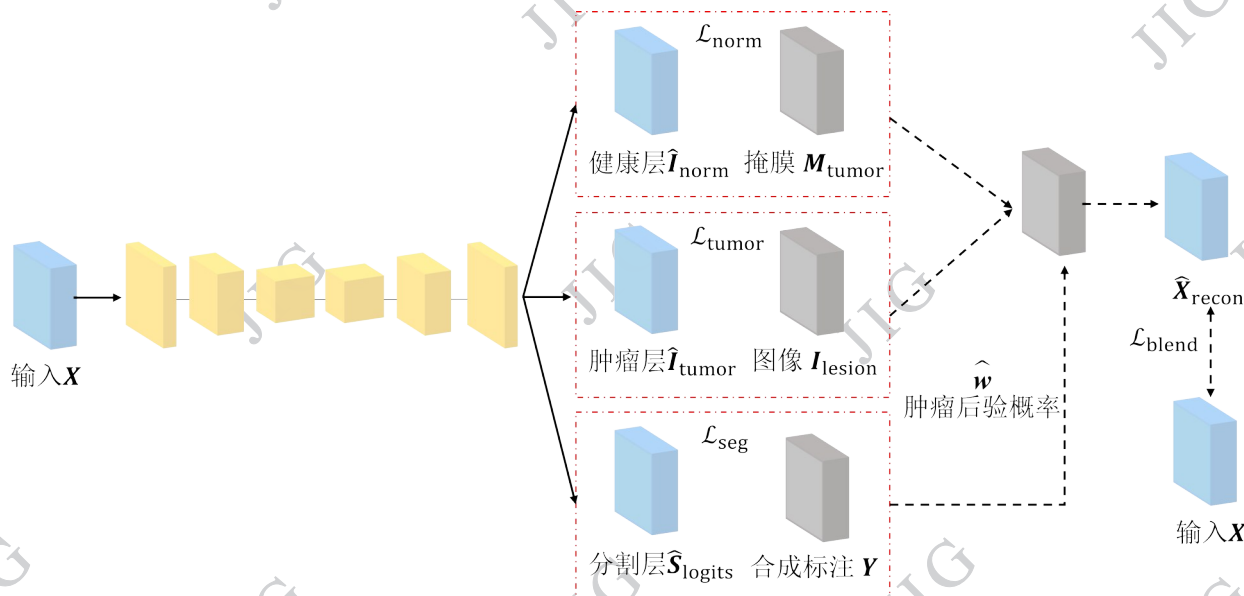


图3 训练架构示意图

Fig. 3 Training architecture diagram

这些理论揭示了数据合成可视为叠加过程,那么模型对图像的语义理解可视为特征解耦过程。

2.1 节所述的肿瘤合成策略是白盒的,除了获得合成图像及其标注,还能同步获取过程中的全部中间态真值,用来约束模型的特征学习。因此,本文通过一种基于图像解耦思想的训练架构,引导模型将输入图像显式解耦为健康解剖与病灶纹理。通过这种显式的解耦约束,模型不再只依赖黑箱式的判别机制,而是被引导分别掌握健康器官与肿瘤的独立表征,有助于模型学习到更具鲁棒性与泛化能力的深层语义特征。

2.2.1 架构细节

定义输入的三维CT图像为 $X \in \mathbf{R}^{1 \times H \times W \times D}$ (即2.1节中的 $I_{syn}(x)$),为了实现图像的解耦,网络的输出被设计为一个五通道的三维张量 $O \in \mathbf{R}^{5 \times H \times W \times D}$ 。通道0与通道1分别对应健康肝脏重建图 $\hat{I}_{norm}$ 和肿瘤纹理重建图 $\hat{I}_{tumor}$ 。通道2至4依次为分别对应背景、肝脏、肿瘤三个类别的语义分割对数几率 $\hat{S}_{logits}$ 。

协同优化上述多重学习目标,本文设计了一个由四项损失函数构成的复合损失 $\mathcal{L}$ 。定义输入图像为 X 、其合成标注为 Y 。根据2.1节定义的合成中间态真值,肿瘤纹理图为 I_{lesion} 、肿瘤掩膜为 M_{tumor} 。

1) 健康肝脏重建损失 $\mathcal{L}_{norm}$:约束重建的健康肝脏图 $\hat{I}_{norm}$ 与非肿瘤区域保持一致,引导模型学习健康肝脏的分布:

$$\mathcal{L}_{norm} = \|(\hat{I}_{norm} - X) \odot (1 - M_{tumor})\|_1 \quad (9)$$

2) 肿瘤纹理重建损失 $\mathcal{L}_{tumor}$:要求肿瘤重建图 $\hat{I}_{tumor}$ 还原合成过程中肿瘤纹理图 I_{lesion} ,引导模型关注肿瘤内部的异质性特征:

$$\mathcal{L}_{tumor} = \|\hat{I}_{tumor} - I_{lesion}\|_1 \quad (10)$$

3) 分割损失 $\mathcal{L}_{seg}$:标准化的Dice损失与交叉熵损失的组合(DiceCELoss),直接对分割通道的输出进行监督,以优化类别不平衡与分割边界:

$$\mathcal{L}_{seg} = \mathcal{L}_{DiceCE}(\hat{S}_{logits}, Y) \quad (11)$$

4) 混合一致性损失 $\mathcal{L}_{blend}$:作为一种自洽性约束,将分割通道输出的肿瘤后验概率作为混合重建图的权重 $\hat{w}$:

$$\hat{w} = \frac{\exp(\hat{S}_{logits,2})}{\sum_{k=0}^2 \exp(\hat{S}_{logits,k})} \quad (12)$$

式中 $\hat{S}_{logits,k}$ 为第 k 类分割通道的输出 ($k = 0, 1, 2$ 分别对应背景、肝脏与肿瘤三种不同类别)。随后,利用 $\hat{w}$ 将解耦出的两个重建图重新加权融合,并要求融合结果逼近原始输入 X :

$$\mathcal{L}_{\text{blend}} = \|(1 - \hat{w}) \odot \hat{I}_{\text{norm}} + \hat{w} \odot \hat{I}_{\text{tumor}} - X\|_1 \quad (13)$$

该项能防止信息在解耦过程中丢失,形成监督闭环。

最终的总损失函数为上述四项损失的加权和:

$$\mathcal{L} = \lambda_0 \mathcal{L}_{\text{norm}} + \lambda_1 \mathcal{L}_{\text{tumor}} + \lambda_2 \mathcal{L}_{\text{seg}} + \lambda_3 \mathcal{L}_{\text{blend}} \quad (14)$$

式中 $\lambda_0, \lambda_1, \lambda_2, \lambda_3$ 为各项损失的权重。

2.2.2 训练与推理策略

为减少冗余运算,提升计算效率,训练和推理均限定在肝脏的感兴趣区域(region-of-interest, ROI)内进行:

$$R_{\text{liver}} = \{(x, y, z) | x \in [s_1, e_1], y \in [s_2, e_2], z \in [s_3, e_3]\} \quad (15)$$

式中 (s_i, e_i) 分别为各维度上肝脏掩膜的最小与最大坐标。

在训练阶段,优化器采用 AdamW,并配合余弦退火学习率调度策略,以实现稳定收敛。

在推理阶段采用重叠滑窗策略,对重叠区域的体素预测采用高斯加权平均进行融合,生成分割概率图。针对初步预测结果,再执行后处理:首先保留肝脏分割结果中最大的连通区域,以消除背景噪声;再利用膨胀后的肝脏掩码作为空间约束,剔除所有位于肝脏区域之外的假阳性肿瘤预测。

3 实验与结果

3.1 实验设置

3.1.1 数据集

本文实验依托于多个公开可获取的医学影像数据集,包括 LiTS (Bilic 等, 2023)、CHAOS (Kavur 等, 2021)、Pancreas-CT (Roth 等, 2016) 及 BTCV (Landman 等, 2015),具体信息如表 2 所示。实验共采用上述数据集中的 234 例三维腹部 CT 扫描数据,所有病例图像都经过统一的预处理,先重定向至标准的 RAS 坐标系,再重采样为 $1.0 \times 1.0 \times 1.0$ 毫米的体素间距。

实验数据划分为两个子集:肝脏肿瘤数据集与健康肝脏数据集。其中,肝脏肿瘤数据集选自 LiTS 数据集中 118 例带有肿瘤标注的病例,用于在五折交叉验证实验中评估分割模型的性能。健康肝脏数据集则整合了全部数据集中的 116 例健康 CT 扫描,用于合成含有肝肿瘤的图像以支持分割模型在无标注条件下训练。

特别地,实验中会使用真实肿瘤训练模型作为对比组。具体地,上述肝脏肿瘤数据集中的 118 例病变样本分配成 5 种不同的组合,用于训练全监督模型。每个组合中的 118 例病变样本按约 8:2 的比例交替充当训练集和评估集,这样每个样本都能在恰好一次的划分中用作评估,其余四次均参与训练。

3.1.2 评估指标

为对模型的分割效果的全面分析,本文使用了医学图像分析领域普遍认可的评价指标:

1) Dice 相似系数 (dice similarity coefficient, DSC) 用于计算预测区域与真实标注之间的体积重叠度,反映肿瘤分割一致性,其值越高表明二者体积重合度越好。

2) 归一化表面 Dice (normalized surface dice, NSD) 在设定容差的条件下评估表面重叠情况,可以衡量模型分割的肿瘤边界的对齐精度。

3) 表面距离 (surface distance, SD) 量化预测边界与真实边界间的距离,值越低说明定位越准确、边界分割越精细。

4) Hausdorff 距离 (95% hausdorff distance, HD95) 计算 95 分位数距离可以降低极端异常值的影响,能更稳定地衡量分割边界的最大偏差及模型的鲁棒性。

3.1.3 实现细节

为建立标准的评估基准,本文方法以 3D-UNet (Ronneberger 等, 2015) 为骨干网络。实验在 NVIDIA A40 48GB GPU 上完成。相关超参数的设置见表 3。

表 2 数据集

Table 2 Details of datasets

数据集	是否含肿瘤标签	使用/总 CT 数
LiTS	✓	129/131
CHAOS	×	20/20
Pancreas-CT	×	38/80
BTCV	×	47/50

3.2 实验结果

3.2.1 整体性能表现

如表 4 所示,本文方法 (Know2Seg) 与当前一些无标注肿瘤分割方法进行了对比。对比方法涵盖了基于异常检测的范式,例如 PatchCore (Roth 等,

表3 超参数
Table 3 Hyper-parameters

参数	值
C	(190, 140, 90)
R	$\begin{bmatrix} -0.8 & 0.6 & 0.2 \\ 0.6 & 0.8 & -0.2 \\ -0.3 & 0.0 & -1.0 \end{bmatrix}$
Λ	[230, 130, 80]
τ_{vol}	10
τ_{grad}	400
Γ	0.3
σ_{α}	$\mathcal{U}(1,2)$
δ	$\mathcal{U}(65,145)$
λ_0	0.3
λ_1	0.5
λ_2	2.0
λ_3	0.8

2022)和Schlegl等(2019)提出的f-AnoGAN,以及基于合成数据的无标注学习范式,例如SynTumor(Hu等,2023),Pixel2Cancer(Lai等,2024)。

实验结果表明,在无标注方法中,本文方法的多项核心指标上均展现出比较优异的性能。并且,在

同等实验条件下,其性能水平优于用真实肿瘤训练的全监督方法,这与有限的真实数据中肿瘤特征分布不均匀有关。本文的合成策略能够生成特征分布相对均衡的训练样本,结合训练架构的约束来缓解捷径学习问题,从而提升模型在真实数据上的泛化性能。

在衡量形状一致性的距离指标上,本文方法表现同样优秀,表明模型不仅关注体素分类是否正确,还能相对稳定地学习肿瘤的形态特征,分割的肿瘤边界更加符合真实情况与解剖学规律。

3.2.2 消融实验

使用传统图像指标来量化合成肿瘤质量的方式面临挑战,如果依赖医生对合成样本逐例验证,耗时耗力,并且过程繁琐(Wu等,2024)。因此,本文借鉴了OASIS(Sushko等,2022)的评估思想,利用分割模型的性能水平来评估合成质量。

本文的消融实验评估了不同模块对模型性能的影响趋势:椭球约束采样、边界避免、边缘平滑、弹性形变,以及训练架构中的各约束。表5和表6提供了消融实验的定量结果。

结果表明,每个模块均有助于提升模型性能。椭球约束采样实现了正确的肿瘤定位,是保障合成样本解剖合理的基础;边界避免进一步减小合成肿瘤位于过于靠近肝脏边缘、大血管等不符合临床规

表4 总体性能
Table 4 Overall performances

方法	无肿瘤标注	有/无肿瘤标注CT数	DSC (%) ↑	NSD (%) ↑	SD (mm) ↓	HD95 (mm) ↓
PatchCore(Roth等,2022)	✓	0/116	16.0	16.4	—	—
f-AnoGAN(Schlegl等,2019)	✓	0/116	19.0	16.9	—	—
VAE(Kingma等,2013)	✓	0/116	24.6	23.6	—	—
Yao等(2021)	✓	0/116	32.8	31.3	—	—
Zhang等(2023c)	✓	0/40	53.0	—	—	—
SynTumor(Hu等,2023)	✓	0/116	59.8	61.3	—	—
SynTumor†(Hu等,2023)	✓	0/116	60.1	61.3	16.9	46.5
DiffTumor†(Chen等,2024)	✓	0/116	59.3	60.1	—	—
Pixel2Cancer(Lai等,2024)	✓	0/116	58.9	63.7	17.9	52.4
使用真实肿瘤训练†	×	118/0	59.4	60.0	17.1	44.1
Know2Seg(本文)	✓	0/116	62.9	65.5	14.0	40.4

注:†表示参照了Chen等(2024)的实验设置(五折交叉验证,每折验证病例数分别为23、23、23、23和26)得到的结果,其他结果来源于相关参考文献。最佳结果以粗体标出。

表 5 合成肿瘤策略的消融实验

Table 5 Ablation studies of synthetic tumor strategy

椭球约束采样	边界避免	边缘平滑	弹性变形	DSC (%)↑	NSD (%)↑	SD (mm)↓	HD95 (mm)↓
✓	-	-	-	35.9	40.3	25.9	60.3
✓	-	-	✓	37.2	37.5	32.7	71.1
✓	✓	-	-	44.3	44.2	26.0	63.2
-	-	✓	-	57.3	59.4	19.2	51.6
-	-	✓	✓	59.1	60.4	17.2	48.0
✓	-	✓	✓	61.0	63.1	16.3	47.7
✓	✓	✓	✓	61.5	63.7	15.2	44.4

注:损失函数使用DiceCELoss。每一行表示某种模块是否被应用(✓)或被禁用(-)。若禁用椭球约束采样,实验时会在肝脏的中心区域随机选择位置放置肿瘤。

律区域的概率;利用 Alpha 掩膜来平滑肿瘤的边缘,则可模拟真实 CT 影像中由于浸润性生长呈现的边缘模糊现象,增强合成肿瘤的真实度;弹性形变引入了更丰富的肿瘤形态变化,提升了合成肿瘤的形态多样性。

表 6 训练架构中损失函数的消融实验

Table 6 Ablation studies on loss functions in training architecture

$\mathcal{L}_0$	$\mathcal{L}_1$	$\mathcal{L}_2$	$\mathcal{L}_3$	DSC (%)↑	NSD (%)↑	SD (mm)↓	HD95 (mm)↓
-	✓	✓	✓	61.6	64.3	15.0	41.2
✓	-	✓	✓	62.2	65.5	14.2	42.3
✓	✓	✓	-	61.7	64.5	14.9	45.4
✓	✓	✓	✓	62.9	65.5	14.0	40.4

注:应用了所有合成肿瘤模块。

从表 6 中可以看出,不启用训练架构中的全部约束时,模型的分割结果有所下降,这是因为端到端训练的模型可能过拟合于合成数据中不自然的伪影。其中,健康肝脏重建提供了健康组织的参照,有助于模型区分肝实质与肿瘤区域,增强抗干扰能力;肿瘤重建能引导模型关注肿瘤内部的异质性纹理特征,增强模型对微小或边界模糊病灶的捕捉能力;混合一致性约束是训练架构中实现监督闭环的核心,进一步增强了模型的分割性能。

3. 2. 3 超参数敏感性分析

为了评估本文方法的鲁棒性,本节对控制肿瘤合成的一些关键超参数进行了敏感性分析,在默认参数设置的基础上进行调整,以观察这些变化对下游分割性能的影响趋势。

表 7 关键超参数的敏感性分析

Table 7 Sensitivity analysis of key hyperparameters

参数	参数设置	DSC (%)↑	NSD (%)↑
τ_{vol}	5	61.5	63.7
	10	62.5	64.7
	15	61.2	63.7
τ_{grad}	350	61.0	62.9
	400	62.5	64.7
	450	62.2	64.9
σ_α	$\mathcal{U}[0.5,1.5]$	62.3	64.3
	$\mathcal{U}[1.0,2.0]$	62.5	64.7
	$\mathcal{U}[1.5,2.5]$	61.8	63.5
δ	$\mathcal{U}[55,95]$	61.1	62.7
	$\mathcal{U}[65,145]$	62.5	64.7
	$\mathcal{U}[90,160]$	60.3	62.4

如表 7 所示,本框架在不同的参数设置下均保持了相对稳定的性能,具体分析如下。

1) 肿瘤位置采样(τ_{vol}, τ_{grad}):

放宽 τ_{vol} 或 τ_{grad} 允许肿瘤出现在更靠近肝脏边界或血管的区域,增加了其位置的多样性;但同时也可能引入相对不符合解剖规律的位置,导致模型性能轻微下降。过于严格的阈值会限制训练样本的多样性,导致模型泛化能力降低。实验表明,默认的阈值设置在解剖合理性与样本多样性之间取得了平衡。

2) 肿瘤融合(σ_α):

利用高斯滤波平滑合成肿瘤的边缘对于模拟真实肿瘤的浸润性生长有关键作用。较小的滤波范围会降低合成肿瘤边缘的模糊程度与多样性,进而降低模型对肿瘤边界的判定能力。较大的模糊范围则

可能生成模糊、失真的伪影,同样会干扰特征学习。

3) 密度衰减(δ):

该参数影响合成肿瘤的密度衰减程度,其数值越大,则肿瘤与肝脏背景之间的对比度越大,反之则与肝脏背景的对比度越小。实验结果显示,适当调整其取值范围,模型的分割性能依然相对稳定。

综上所述,Know2Seg框架对关键超参数的波动表现出较低的敏感性,证实了其具有一定的鲁棒性。

3.2.5 局限性

在实验过程中,本文也认识到所述方法仍然存在一定的局限性,具体如下。

1) 目前实验主要以肝脏肿瘤为研究对象,暂未将方法论完整地迁移至其他器官。Chen等(2024)的研究表明,肝脏、肾脏及胰腺的早期肿瘤均表现出相似的形态学特征,并且即使是经验丰富的医生若仅凭局部的肿瘤纹理图像块,也很难准确地判断其来自于哪个上述器官。而本文方法是程序化的,理论上具备向其他器官迁移的可能。

2) 在实验过程中,本文发现合成策略总是会具有一定的随机性,并不是所有的合成样本都能达到理想的质量。有的样本可能缺乏真实的病理特征,或是可能引入了不符合解剖规律的特征。如果将此低质量样本纳入训练,模型会学到错误的病理特征,不仅不能提升模型的性能,反而会因为干扰模型正常的特征学习,造成模型在真实数据上的分割精度下降。虽然在理论上,可以通过筛选并去除这些低质量样本,但面对大规模的合成数据,这种方式的时间和精力成本很高,因此建立自动化的质量控制机制是必要的。目前,已经有研究验证了基于对抗性判别的方案在过滤合成样本方面的有效性(Wu等,2024),因此也可以考虑一些相关的方案来过滤低质量的合成样本。

4 结 论

本文设计了一种基于临床知识引导的无标注肝肿瘤分割框架Know2Seg。实验结果表明,本文方法在性能上与基于真实肿瘤数据训练的模型相当,并优于一些主流的无标注方法,展现了算法化临床知识的有效性,体现了利用合成肿瘤在相对容易获取的健康数据集上训练分割模型的潜力。

未来工作将聚焦于以下几个方向:将Know2Seg

框架迁移并适配至其他器官;研究自动质量控制机制来筛选合成样本;以及将本方法与蒸馏学习或提示学习相结合,从而更有效地融合临床知识,提升肿瘤分割的性能。

参考文献(References)

- Antonelli M, Reinke A, Bakas S, Farahani K, Kopp-Schneider A, Landman B A, et al. 2022. The medical segmentation decathlon. *Nature Communications*, 13 (1): 4128 [DOI: 10.1038/s41467-022-30695-9]
- Baur C, Denner S, Wiestler B, Navab N and Albarqouni S. 2021. Auto-encoders for unsupervised anomaly segmentation in brain MR images: A comparative study. *Medical Image Analysis*, 69: 101952 [DOI: 10.1016/j.media.2020.101952]
- Bilic P, Christ P, Li H B, Vorontsov E, Ben-Cohen A and Kaissis G. 2023. The liver tumor segmentation benchmark (LiTS). *Medical Image Analysis*, 84: 102680 [DOI: 10.1016/j.media.2022.102680]
- Castaldo A, De Lucia D R, Pontillo G, Gatti M, Cocozza S, Ugga L, et al. 2021. State of the art in artificial intelligence and radiomics in hepatocellular carcinoma. *Diagnostics*, 11 (7): 1194 [DOI: 10.3390/diagnostics11071194]
- Chartsias A, Joyce T, Papanastasiou G, Semple S, Williams M, Newby D E, et al. 2019. Disentangled representation learning in cardiac image analysis. *Medical Image Analysis*, 58: 101535 [DOI: 10.1016/j.media.2019.101535]
- Chen Q, Chen X, Song H, Xiong Z, Yuille A, Wei C, et al. 2024. Towards generalizable tumor synthesis//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, WA, USA: IEEE: 11147-11158 [DOI: 10.1109/CVPR52733.2024.01060]
- Geirhos R, Jacobsen J H, Michaelis C, Zemel R, Brendel W, Bethge M, et al. 2020. Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2 (11): 665-673 [DOI: 10.1038/s42256-020-00257-z]
- Geirhos R, Rubisch P, Michaelis C, Bethge M, Wichmann F A and Brendel W. 2019. ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness//*Proceedings of the 7th International Conference on Learning Representations (ICLR 2019, OrleansNew, USA: OpenReview.net*
- Gotra A, Sivakumaran L, Chartrand G, Vu K N, Vandenbroucke-Menu F, Kauffmann C, et al. 2017. Liver segmentation: indications, techniques and future directions. *Insights into Imaging*, 8: 377-392 [DOI: 10.1007/s13244-017-0558-1]
- Guo P F, Zhao C, Yang D, Xu Z Y, Nath V, Tang Y C, et al. 2025. MAISI: Medical AI for Synthetic Imaging//2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Tucson,

- AZ, USA: IEEE: 4430-4441 [DOI: 10.1109/WACV61041.2025.00435]
- Han C, Hayashi H, Rundo L, Araki R, Shimoda W, Muramatsu S, et al. 2018. GAN-based synthetic brain MR image generation//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Washington, DC, USA: IEEE: 734-738 [DOI: 10.1109/ISBI.2018.8363678]
- Han K, Xiong Y, You C, Khosravi P, Sun S, Yan X, et al. 2023. Med-gen3D: A deep generative framework for paired 3D image and mask generation//International Conference on Medical Image Computing and Computer-Assisted Intervention. Vancouver, BC, Canada: Springer: 759-769 [DOI: 10.1007/978-3-031-43907-0_72]
- He R, Sun S, Yu X, Xue C, Zhang W, Torr P, et al. 2022. Is synthetic data from generative models ready for image recognition? [EB/OL]. [2026-06-08].
<https://arxiv.org/abs/2210.07574>
- Hu Q, Chen Y, Xiao J, Sun S, Chen J, Yuille A L, et al. 2023. Label-free liver tumor segmentation//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, BC, Canada: IEEE: 7422-7432 [DOI: 10.1109/CVPR52729.2023.00717]
- Hu R, Song Y, Liu Y, Zhu Y, Feng N, Qiu C, et al. 2024. Imbalance multiclass problem: a robust feature enhancement-based framework for liver lesion classification. *Multimedia Systems*, 30(2): 104 [DOI: 10.1007/s00530-024-01291-2]
- Isensee F, Jaeger P F, Kohl S A A, Petersen J and Maier-Hein K H. 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2): 203-211 [DOI: 10.1038/s41592-020-01008-z]
- Kavur A E, Gezer N S, Bariş M, Aslan S, Conze P H, Groza V, et al. 2021. CHAOS Challenge - combined (CT-MR) healthy abdominal organ segmentation. *Medical Image Analysis*, 69: 101950 [DOI: 10.1016/j.media.2020.101950]
- Kingma D P and Welling M. 2013. Auto-encoding variational bayes. [EB/OL]. [2026-06-08].
<https://arxiv.org/abs/1312.6114>
- Lai Y, Chen X, Wang A, Yuille A and Zhou Z. 2024. From pixel to cancer: Cellular automata in computed tomography//Medical Image Computing and Computer Assisted Intervention -- MICCAI 2024. Marrakesh, Morocco: Springer: 36-46 [DOI: 10.1007/978-3-031-72378-0_4]
- Landman B, Xu Z, Iglesias J E, Styner M, Langerak T R and Klein A. 2015. MICCAI multi-atlas labeling beyond the cranial vault—workshop and challenge[DB/OL]. [2026-06-10]. <https://www.synapse.org/Synapse:syn3193805>
- Lee K, O'Malley M, Haider M and Hanbidge A. 2004. Triple-phase MDCT of hepatocellular carcinoma. *American Journal of Roentgenology*, 182(3): 643-649 [DOI: 10.2214/ajr.182.3.1820643]
- Li D, Ling H, Kim S W, Kreis K, Fidler S, Torralba A, et al. 2022. BigDatasetGAN: Synthesizing ImageNet with pixel-wise annotations//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, LA, USA: IEEE: 21330-21340 [DOI: 10.1109/CVPR52688.2022.02064]
- Li T H, He Z Q, He L, Zhang W and Huang Z F. 2025. Fusion of image enhancement and self-attention mechanism for the generation of active pulmonary tuberculosis CT image data. *Journal of Image and Graphics*, 30(7):2558-2569 (李天虎, 何志琴, 何力, 张蔚, 黄忠锋. 2025. 融合图像增强和自注意力机制的活动性肺结核CT影像数据生成. *中国图象图形学报*, 30(7):2558-2569) [DOI: 10.11834/jig.240518]
- Liang Y, Xu F, Guo Y, Lai L, Jiang X, Wei X, et al. 2021. Diagnostic performance of LI-RADS for MRI and CT detection of HCC: a systematic review and diagnostic meta-analysis. *European Journal of Radiology*, 134: 109404 [DOI: 10.1016/j.ejrad.2020.109404]
- Lin M, Weng N, Mikolaj K, Bashir Z, Svendsen M B S, Tolsgaard M G, et al. 2024. Shortcut learning in medical image segmentation//Medical Image Computing and Computer Assisted Intervention - MICCAI 2024. Marrakesh, Morocco: Springer: 623-633 [DOI: 10.1007/978-3-031-72111-3_59]
- Mokrane F Z, Lu L, Vavasour A, Otal P, Peron J M, Luk L, et al. 2020. Radiomics machine-learning signature for diagnosis of hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules. *European Radiology*, 30(1): 558-570 [DOI: 10.1007/s00330-019-06347-w]
- Pan Y S, Ma H J, Xia Y and Zhang Y N. 2025. Application of content generation in medical images. *Journal of Image and Graphics*, 30(6):1985-2000 (潘永生, 马豪杰, 夏勇, 张艳宁. 2025. 医学影像中的生成技术. *中国图象图形学报*, 30(6):1985-2000) [DOI: 10.11834/jig.240459]
- Peng W, Adeli E, Zhao Q and Pohl K M. 2022. Generating realistic 3D brain MRIs using a conditional diffusion probabilistic model. [EB/OL]. [2026-06-08].
<https://arxiv.org/abs/2212.08034v1>
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: Convolutional networks for biomedical image segmentation//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Roth H, Farag A, Turkbey E B, Lu L, Liu J and Summers R M. 2016. Data from pancreas-CT [DB/OL]. The Cancer Imaging Archive. [2026-06-08]. <https://doi.org/10.7937/k9/tcia.2016.tnb1kqbu>
- Roth K, Pemula L, Zepeda J, Schölkopf B, Brox T and Gehler P. 2022. Towards total recall in industrial anomaly detection//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, LA, USA: IEEE: 14318-14328 [DOI: 10.1109/CVPR52688.2022.01392]
- Sankaranarayanan S, Balaji Y, Jain A, Lim S N and Chellappa R.

2018. Learning from synthetic data: Addressing domain shift for semantic segmentation//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE: 3752-3761 [DOI: 10.1109/CVPR.2018.00395]
- Schlegl T, Seeböck P, Waldstein S M, Langs G and Schmidt-Erfurth U. 2019. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Analysis*, 54: 30-44 [DOI: 10.1016/j.media.2019.01.010]
- Schmidt-Mengin M, Benichoux A, Belachew S, Komodakis N and Paragios N. 2024. ToNNO: Tomographic Reconstruction of a Neural Network's Output for Weakly Supervised Segmentation of 3D Medical Images//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE: 11428-11438 [DOI: 10.1109/CVPR52733.2024.01086]
- Shao H, Jing Y and Cui W. 2025. Cross modal translation of magnetic resonance imaging and computed tomography images based on diffusion generative adversarial networks. *Journal of Biomedical Engineering*, 42(3): 575-584 (邵虹, 荆一焯, 崔文成. 2025. 基于扩散生成对抗网络的核磁共振图像与计算机断层扫描图像跨模态转换. *生物医学工程学杂志*, 42(3): 575-584) [DOI: 10.7507/1001-5515.202404056]
- Simard P Y, Steinkraus D and Platt J C. 2003. Best practices for convolutional neural networks applied to visual document analysis//Seventh International Conference on Document Analysis and Recognition. Edinburgh, UK: IEEE: 958-963 [DOI: 10.1109/ICDAR.2003.1227801]
- Sindhura D, Pai R M, Bhat S N and Pai M M M. 2024. A review of deep learning and Generative Adversarial Networks applications in medical image analysis. *Multimedia Systems*, 30(3): 161 [DOI: 10.1007/s00530-024-01349-1]
- Sushko V, Schönfeld E, Zhang D, Gall J, Schiele B and Khoreva A. 2022. OASIS: only adversarial supervision for semantic image synthesis. *International Journal of Computer Vision*, 130(12): 2903-2923 [DOI: 10.1007/s11263-022-01673-x]
- Tang Y, Yang D, Li W, Roth H R, Landman B, Xu D, et al. 2022. Self-supervised pre-training of Swin transformers for 3D medical image analysis//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, LA, USA: IEEE: 20730-20740 [DOI: 10.1109/CVPR52688.2022.02007]
- Vorontsov E, Molchanov P, Gazda M, Beckham C, Kautz J and Kadoury S. 2022. Towards annotation-efficient segmentation via image-to-image translation. *Medical Image Analysis*, 82: 102624 [DOI: 10.1016/j.media.2022.102624]
- Vu M H, Tronchin L, Nyholm T and Löfstedt T. 2025. Using Synthetic Images to Augment Small Medical Image Datasets[EB/OL]. [2026-06-08].
<https://arxiv.org/pdf/2503.00962.pdf>
- Wang T, Lei Y, Fu Y, Wynne J F, Curran W J, Liu T, et al. 2021. A review on medical imaging synthesis using deep learning and its clinical applications. *Journal of Applied Clinical Medical Physics*, 22(1): 11-36 [DOI: 10.1002/acm2.13121]
- Wei J, Ji Q, Gao Y, Yang X, Guo D, Gu D, et al. 2023. A multi-scale, multi-region and attention mechanism-based deep learning framework for prediction of grading in hepatocellular carcinoma. *Medical Physics*, 50(4): 2290-2302 [DOI: 10.1002/mp.16127]
- Wu L, Zhuang J, Ni X and Chen H. 2024. FreeTumor: Advance Tumor Segmentation via Large-Scale Tumor Synthesis. [EB/OL]. [2026-06-08].
<https://arxiv.org/abs/2406.01264>
- Wu W, Zhao Y, Chen H, Gu Y, Zhao R, He Y, et al. 2023. DatasetDM: Synthesizing data with perception annotations using diffusion models//Advances in Neural Information Processing Systems 36 (NeurIPS 2023). OrleansNew, LA, USA: Curran Associates, Inc.: 54683-54695 [DOI: 10.52202/075280-2383]
- Xie J H, Zhang Z, Weng Z Y, Zhu Y S and Luo G B. 2026. MedDiff-FT: Data-Efficient Diffusion Model Fine-Tuning with Structural Guidance for Controllable Medical Image Synthesis//Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Daejeon, Republic of Korea: Springer: 306-316 [DOI: 10.1007/978-3-032-04965-0_29]
- Yang Q, Wei J, Hao X, Kong D, Yu X, Jiang T, et al. 2020. Improving B-mode ultrasound diagnostic performance for focal liver lesions using deep learning: A multicentre study. *EBioMedicine*, 56: 102777 [DOI: 10.1016/j.ebiom.2020.102777]
- Yao Q, Xiao L, Liu P and Zhou S K. 2021. Label-free segmentation of COVID-19 lesions in lung CT. *IEEE Transactions on Medical Imaging*, 40(10): 2808-2819 [DOI: 10.1109/TMI.2021.3066161]
- Zhang S, Liu J, Hu B and Mao Z. 2023a. GH-DDM: the generalized hybrid denoising diffusion model for medical image generation. *Multimedia Systems*, 29(3): 1335-1345 [DOI: 10.1007/s00530-023-01059-0]
- Zhang X, Xie W, Huang C, Zhang Y, Chen X, Tian Q, et al. 2023b. Self-supervised tumor segmentation with sim2real adaptation. *IEEE Journal of Biomedical and Health Informatics*, 27(9): 4373-4384 [DOI: 10.1109/JBHI.2023.3240844]
- Zhang Z, Deng H and Li X. 2023c. Unsupervised liver tumor segmentation with pseudo anomaly synthesis//Simulation and Synthesis in Medical Imaging (SASHIMI 2023). Cham, Switzerland: Springer: 86-96 [DOI: 10.1007/978-3-031-44689-4_9]
- Zhao C, Guo P, Yang D, He Y, Tang Y, Simon B, et al. 2026. MAISI-v2: Accelerated 3D high-resolution medical image synthesis with rectified flow and region-specific contrastive loss//Proceedings of the AAAI Conference on Artificial Intelligence. Singapore: AAAI Press: 13088-13098 [DOI: 10.1609/aaai.v40i15.38309]
- Zhao H, Sheng D, Bao J, Chen D, Chen D, Wen F, et al. 2023. X-Paste: Revisiting Scalable Copy-Paste for Instance Segmentation using CLIP and StableDiffusion//International Conference on Machine Learning, ICML 2023. Honolulu, HI, USA: PMLR:

42098-42109

Zhu L, Xue Z, Jin Z, Liu X, He J, Liu Z, et al. 2023. Make-a-volume: Leveraging latent diffusion models for cross-modality 3D brain MRI synthesis//International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2023. Cham, Switzerland: Springer: 592-601 [DOI: 10.1007/978-3-031-43999-5_56]

作者简介

钟可毅,男,硕士研究生,研究方向为机器学习、医学图像分

析。E-mail:zky@email.ncu.edu.cn

黄旭慧,女,副教授,主要研究方向为机器学习。E-mail:hxx@ncu.edu.cn

万欢,女,博士,主要研究方向为机器学习、计算机视觉。E-mail:HuanWan@jxnu.edu.cn

缪轩磊,男,中医师,主要研究方向为中西医结合临床医学。E-mail:miaoxuanlei@jxutcm.edu.cn

魏欣,通信作者,男,博士,副教授,主要研究方向为机器学习、计算机视觉、医学图像分析。E-mail:xinwei@ncu.edu.cn