医学影像分割中的U-Net改进算法研究——以肝脏CT图像为例

陈诗颖

广州医科大学 第一临床学院，广州市 中国511400

摘要针对传统U-Net在肝脏CT图像分割中存在的边界模糊、小病灶漏检及三维结构建模不足等问题，本研究提出了一种融合残差密集模块、双路径注意力机制与三维混合编码的改进模型（ResAtt-3DUNet）。首先，通过在编码器中引入残差密集模块（Res-Dense Block），增强血管纹理等局部特征的跨层传递能力；其次，提出通道-空间双路径注意力机制，利用SE模块动态校准特征通道权重，并结合Sobel算子驱动的边缘响应函数抑制肝周脂肪组织的噪声干扰；最后，设计三维混合编码器，融合V-Net的3D卷积与TransUNet的Transformer模块，有效捕获连续断层影像的时空关联性。实验表明：在LiTS和3D-IRCADb数据集上，本模型肝脏分割Dice系数达96.3%（较传统U-Net提升8.2%），肿瘤边界平均表面距离（ASD）降至1.2mm，三维重建空洞率由12%压缩至2.5%，对<5mm微小结节的敏感度达92.7%。消融实验验证了多尺度密集残差模块（Dice提升4.7%）与双路径注意力机制（假阳性减少62%）的有效性，临床验证显示其与金标准标注的Kappa一致性达0.92。本研究为肝脏肿瘤精准分割提供了高鲁棒性解决方案，可显著降低手术规划中的误诊风险。

‌关键词‌：肝脏CT分割；注意力机制；三维卷积；残差密集网络；Transformer

**Study on improved U-Net algorithm in medical image segmentation —— Taking liver CT image as an example**

Chen Shiying

First Clinical College, Guangzhou Medical University, 511400, Guangzhou, China

**Abstract** In response to issues such as boundary blurring, missed detection of small lesions, and insufficient 3D structure modeling in traditional U-Net for liver CT image segmentation, this study proposes an improved model (ResAtt-3DUNet) that integrates residual dense modules, dual-path attention mechanisms, and 3D hybrid coding. First, by introducing a residual dense module (Res-Dense Block) into the encoder, it enhances the cross-layer transfer capability of local features like vascular textures. Second, a channel-space dual-path attention mechanism is proposed, using SE modules to dynamically calibrate feature channel weights and combining Sobel operator-driven edge response functions to suppress noise interference from perisinusoidal fat tissue. Finally, a 3D hybrid encoder is designed, integrating V-Net's 3D convolution with TransUNet's Transformer module to effectively capture the spatiotemporal correlation of continuous slice images. Experiments show that on LiTS and 3D-IRCADb datasets, the liver segmentation Dice coefficient reaches 96.3% (an improvement of 8.2% over traditional U-Net), the average surface distance (ASD) of tumor boundaries decreases to 1.2 mm, and the 3D reconstruction void rate is reduced from 12% to 2.5%. The sensitivity to micro-nodules <5mm is 92.7%. Ablation experiments verify the effectiveness of multi-scale dense residual modules (a Dice improvement of 4.7%) and dual-path attention mechanisms (a false positive reduction of 62%). Clinical validation demonstrates a Kappa consistency of 0.92 with the gold standard annotation. This study provides a highly robust solution for accurate segmentation of liver tumors, which can significantly reduce the risk of misdiagnosis in surgical planning.

**Key words**  liver CT segmentation; attention mechanism; 3D convolution; residual dense network; Transformer

1.前言

肝癌严重威胁人类健康，早诊断、早治疗是提高肝癌治愈率的关键。CT 检查凭借其高分辨率，在肝癌诊断中发挥着重要作用，能够有效检出微小癌灶，对肝癌的诊断符合率可达 90% 以上[1] 。肝脏 CT 图像分割作为肝癌诊断与手术规划的核心环节，精准分割结果对于病变部位的精确定位和诊断误差的缩小至关重要。传统的人工勾画肝脏 CT 图像的方式，不仅耗费医生大量时间和精力，而且主观性强，易受医生经验水平的影响。为解决这一问题，U-Net 模型被引入医学影像分割领域。U-Net 凭借其独特的编码器 - 解码器结构和跳跃连接机制，在一定程度上能够捕捉多尺度上下文信息，实现端到端的像素级预测，在医学影像分割中取得了一定成果[3]。然而，在实际应用于肝脏 CT 图像分割时，U-Net 暴露出诸多问题。从数据准备角度看，多模态数据泛化性不足，CT 与其他模态成像物理特性差异大，混合模型难以捕捉单一模态深层特征，导致分割结果不稳定，易出现误判和漏检；标注数据稀缺且类别分布不均衡[17]，使得模型对病灶边界和多样性特征学习不充分，小病灶漏检问题突出[14]。模型结构方面，跳跃连接存在语义鸿沟，不同层级特征直接拼接产生融合噪声，影响病灶精准识别；池化操作造成信息丢失，降低了结构边界定位精度，微小病灶容易被漏检[6]。训练优化过程中，类别不平衡的损失函数偏差使得模型偏向保守预测，忽视小目标信号，导致肝脏分割精度下降。[14]评估环节，常用的 Dice 系数等指标存在局限性，无法全面准确地评估分割结果的边界局部精度。鉴于此，本研究聚焦于改进 U-Net 算法，提出融合残差密集模块、双路径注意力机制与三维混合编码的 ResAtt-3DUNet 模型，旨在突破传统 U-Net 的局限，提升肝脏及肝肿瘤影像分割精度，为肝癌的精准诊断和手术规划提供有力支持，降低临床误诊风险。[11][12][16]

2.相关工作‌

2.1‌U-Net核心结构‌

‌编码器（收缩路径）‌：通过卷积层、最大池化逐级提取图像特征，逐步降低空间分辨率，捕获全局语义信息（如肝脏整体形态）‌。

‌解码器（扩展路径）‌：通过反卷积或上采样操作恢复图像分辨率，结合跳跃连接（skip connections）融合编码器的高分辨率细节特征，提升局部边界精度‌。

‌跳跃连接机制‌：将编码器的低级特征（如边缘、纹理）与解码器的高级语义特征（如器官轮廓）逐层拼接，弥补下采样过程中的空间信息丢失，增强小病灶的定位能力‌。

2.2U-Net在肝脏CT图像分割中的应用原理

‌多尺度上下文捕捉‌：编码器通过多级下采样来提取肝脏区域的总体特征（如体积、形状等），解码器与跳跃连接结合以恢复局部细节（如血管分布），这种结构可适应CT图像中肝脏与周围组织的复杂对比关系。

‌端到端像素级预测‌：采用全卷积解码器与条件随机场（CRF）后处理模块[9]，联合优化像素级分类与形态学连续性，直接对每个像素分类（肝脏/非肝脏），适用于CT图像的灰度特征分析和器官边界判定

2.2U-Net在肝脏CT图像分割中遇到的问题

U-Net在肝脏CT图像分割中遇到的问题可以分为数据准备、模型结构、训练优化、后处理与评估四个方面。

2.3.1数据准备方面

①多模态数据泛化性不足

核心原因：CT与MRI、T1与T2加权成像的物理特性差异显著（如T1强调解剖结构、T2对病变敏感）。使用混合模型难以捕捉单一模态的深层特征，将导致过拟合或分割效果不稳定，分割结果可能出现严重偏差。

问题表现：将正常组织误判为病灶，或漏检小肿瘤区域；三维重建结果出现断层或空洞，增加手术风险等。

②标注数据稀缺、数据类别分布不均[8]

核心原因：数据数量不足；肝脏与病灶像素占比差异大。模型对病灶边界（如肝硬化再生结节与正常肝组织交界处）学习不充分、难以学习病灶多样性特征（如不同强化模式的肝癌）。

问题表现：微小结节、早期肿瘤等小病灶漏检，漏检率甚至在30%以上；病灶体积测量偏差超过15%。

‌2.3.2 模型结构的局限性‌

①跳跃连接的语义鸿沟

概念：跳跃连接（Skip Connection）是U-Net对称架构中的关键设计，通过将编码器（下采样路径）的浅层特征图与解码器（上采样路径）的对应层级特征图直接相加连接，实现‌多尺度特征融合。‌

核心原因：低级特征（边缘、纹理）与高级语义（器官/病变）直接拼接，它们之间的层级差异会引发融合噪声。

问题表现：肿瘤或血管等结构的边缘分割不连续，甚至出现假阳性或假阴性区域，被噪声干扰无法精准识别病灶。

②池化操作的信息丢失[13]

概念：池化是卷积神经网络（CNN）中的一种降维操作，减小数据尺寸以降低计算复杂度，主要用于压缩特征图尺寸并保留关键信息。

核心原因：池化操作（如最大池化或平均池化）通过局部区域采样压缩特征图，导致像素级空间坐标信息被破坏；大尺寸池化核（如3×3或更大）会直接湮灭小目标的有效表征；池化后的高级特征偏向全局语义（如肝脏整体轮廓），但缺乏对局部细节的精细刻画。

问题表现：结构边界的定位精度降低；微小病灶漏检；形态复杂区域分割出错。

2.3.‌3训练优化难题‌

类别不平衡的损失函数偏差

概念：类别不平衡损失函数偏差‌指因肝脏区域（目标类）与背景区域（非目标类）的像素数量差异过大，导致交叉熵等损失函数的梯度更新被背景类主导，

核心原因：背景像素占比过高导致模型偏向保守预测（假阴性），从而抑制模型对肝脏区域的关注能力，模型通过简单预测背景即可获得较低损失值；未对不同类别分配差异化的权重。

问题表现：出现“伪优化”现象；忽略小目标（如肝硬化结节）的信号；使肝脏分割Dice系数下降15%-30%

2.3.‌4评估瓶颈‌

评估指标的局限性

问题描述：Dice系数对整体重叠敏感，但忽略边界局部精度[1]。

典型误区：高Dice值可能掩盖边缘分割错误（如平均表面距离过大）。

2.3模型‌改进方向‌

‌多尺度特征融合‌：通过残差连接、密集模块增强上下文信息提取‌；

‌注意力机制‌：在跳跃连接中嵌入通道-空间双路径注意力机制，通过Sobel算子驱动的边缘响应函数抑制肝周脂肪组织的噪声干扰。

‌三维建模优化‌：结合V-Net、TransUNet处理连续断层影像的时空关联性‌。

3. 主要成果论述

本研究提出的改进U-Net算法（命名为ResAtt-3DUNet）在肝脏及肝肿瘤CT影像分割中取得精度提升、边界优化、建模改进等显著成果。

3.1‌精度提升‌

在LiTS公开数据集上，肝脏分割Dice系数达96.3%（较传统U-Net提升8.2%），肿瘤分割Dice系数达95.1%，漏检率降至5.3%。[2]

3.2‌边界优化‌

通过边缘注意力机制，肿瘤边界平均表面距离（ASD）从3.8mm降至1.2mm，假阳性区域减少62%。

3.3‌三维建模改进‌

结合跨层时空特征融合，三维重建的肝血管连续性提升，分割结果空洞率由12%降至2.5%；对<5mm微小结节的敏感度达92.7%（原模型为68.4%），体积测量误差<8%。

3.4创新点

‌多尺度-注意力混合架构‌：在编码器中引入残差密集模块（Res-Dense Block），增强血管纹理等局部特征提取；

‌双路径注意力机制‌：在跳跃连接中融合通道注意力（SE模块）与空间边缘注意力，抑制肝脏周围器官的干扰信号；

‌三维-Transformer混合编码器‌：利用V-Net的3D卷积捕获层间解剖结构，结合TransUNet的Transformer模块建模全局上下文关联。

4. 关键实现技术

4.1 残差密集模块

在U-Net编码器的每个下采样阶段前加入残差密集连接，通过跨层梯度传播解决特征丢失问题。具体结构：

输入经3×3卷积→批量归一化→ReLU后，与后续4个密集连接子模块拼接，输出特征图公式：

=

其中为初始输入，n=4。[7]

4.2 双路径注意力机制

‌通道注意力路径‌：采用SE模块动态校准特征通道权重，公式：

=σ(δ(GAP(F)))

（GAP为全局平均池化，为全连接层）[5]

‌空间边缘注意力路径‌：利用Sobel算子提取特征图边缘响应E[10]，生成空间掩膜，

抑制非目标区域激活。

4.3 三维混合编码器设计

‌3D卷积层‌：使用5×5×5卷积核处理连续CT断层，提取肝血管的时空特征；

‌Transformer模块‌：在瓶颈层嵌入多头自注意力机制（MSA），公式：

其中dk为维度，Q/K/V由特征图展平后线性映射生成。[4]

5. 验证

5.1 实验设置

‌数据集‌：LiTS（131例肝脏CT）、3D-IRCADb（20例肝肿瘤CT），均包含专家标注的肝脏/肿瘤掩膜；

‌评估指标‌：Dice系数、ASD、敏感度（SEN）、特异度（SPE）；

‌对比模型‌：U-Net、V-Net、AttUNet、TransUNet。

5.2 结果分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | 肝脏Dice (%) | 肿瘤Dice (%) | ASD (mm) | 空洞率 (%) |
| U-Net (基线) | 88.1 | 86.9 | 3.8 | 12.0 |
| V-Net | 91.5 | 89.2 | 2.6 | 8.4 |
| ResAtt-3DUNet | ‌96.3‌ | ‌95.1‌ | ‌1.2‌ | ‌2.5‌ |

‌消融实验‌：单独使用残差密集模块使Dice提升4.7%，增加双路径注意力后进一步提升3.5%；

‌临床验证‌：在合作医院50例患者数据中，分割结果与金标准（3D Slicer手动标注）的Kappa一致性达0.92。[15]

6.结论与展望

6.1技术总结

首先，本研究提出的 ResAtt-3DUNet 模型通过引入残差密集模块，增强了血管纹理等局部特征的跨层传递能力，有效解决了特征丢失问题。其次，双路径注意力机制的应用，利用 SE 模块动态校准特征通道权重，结合 Sobel 算子驱动的边缘响应函数抑制噪声干扰，显著优化了分割边界。最后，三维混合编码器融合了 V-Net 的 3D 卷积与 TransUNet 的 Transformer 模块，有效捕获了连续断层影像的时空关联性，提升了三维建模能力。

6.2应用展望

在 1 年内，该模型有望在部分医疗机构进行小规模应用，辅助医生进行肝脏 CT 图像的初步分析，提高诊断效率。3 - 5 年内，随着技术的进一步优化和完善，可推广至更多医院，成为肝脏疾病诊断和手术规划的常规辅助工具，同时不断拓展应用场景，如肝脏疾病的早期筛查和预后评估等。

6.3伦理思考

随着 AI 技术在医学领域的广泛应用，AI 治理至关重要。在医学影像分割中，要确保模型的训练数据来源合法合规，保护患者的隐私信息。同时，对于 AI 辅助诊断的结果，医生应保持审慎态度，避免过度依赖，确保医疗决策的科学性和合理性，保障患者的安全和权益。

参考文献

[1] Heimann, T., B. van Ginneken, M. A. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, et al. "Comparison and Evaluation of Methods for Liver Segmentation from Ct Datasets." [In English]. Ieee Transactions on Medical Imaging 28, no. 8 (Aug 2009): 1251-65. https://doi.org/10.1109/Tmi.2009.2013851. <Go to ISI>://WOS:000268525500012.

[2] Bilic, P., P. Christ, H. B. Li, E. Vorontsov, A. Ben-Cohen, G. Kaissis, A. Szeskin, et al. "The Liver Tumor Segmentation Benchmark (Lits)." [In English]. Medical Image Analysis 84 (Feb 2023)

.https://doi.org/ARTN10268010.1016/j.media.2022.102680. <Go to ISI>://WOS:000928244200004.

[3]‌Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." Chap. Chapter 28 In *Medical Image Computing and Computer-Assisted Intervention – Miccai 2015*. Lecture Notes in Computer Science, 234-41, 2015.

[4] Chen, Jieneng, Yongyi Lu , Qihang Yu , Xiangde Luo , Ehsan Adeli, Yan Wang , Le Lu *, et al.* "Transunet: Transformers Make Strong Encoders for Medical Image Segmentation." *arXiv:2102.04306*  (Feb 2021). https://doi.org/10.48550/arXiv.2102.04306.

[5]Jie Hu, Li Shen, Samuel Albanie, Gang Sun, Enhua Wu. "Squeeze-and-Excitation Networks." CVPR 2018 (2018). https://doi.org/10.48550/arXiv.1709.01507.

[6]‌Zhou Z, et al.‌ UNet++: Redesigning skip connections to exploit multiscale features[J]. IEEE Trans Med Imaging, 2020, 39(6):1856-1867.

[7]‌Isensee F, et al.‌ nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature Methods, 2021, 18:203-211.

https://www.nature.com/articles/s41592-020-01008-z

[8]‌Mokrane FZ, et al.‌ Radiomics machine-learning signature for diagnosis of hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules[J]. Eur Radiol, 2020, 30(1):558-570.

https://link.springer.com/article/10.1007/s00330-019-06347-w

[9]Krähenbühl, Philipp, and Vladlen Koltun. "Efficient Inference in Fully Connected Crfs with Gaussian Edge Potentials." Proceedings of the 25th International Conference on Neural Information Processing Systems, Granada, Spain, Curran Associates Inc., 2011.

[10]Xie, Saining, and Zhuowen Tu. "Holistically-Nested Edge Detection." Paper presented at the 2015 IEEE International Conference on Computer Vision (ICCV), 2016.

[11]Lambin, Philippe, Ralph T. H. Leijenaar, Timo M. Deist, Jurgen Peerlings, Evelyn E. C. de Jong, Janita van Timmeren, Sebastian Sanduleanu, et al. "Radiomics: The Bridge between Medical Imaging and Personalized Medicine." Nature Reviews Clinical Oncology 14, no. 12 (2017/12/01 2017): 749-62. https://doi.org/10.1038/nrclinonc.2017.141.

[12]Gillies, Robert J, Paul E Kinahan, and Hedvig Hricak. "Radiomics: Images Are More Than Pictures, They Are Data." Radiology 278, no. 2 (2016): 563-77.

[13]Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. "A Survey on Deep Learning in Medical Image Analysis." Medical Image Analysis 42 (2017/12/01/ 2017): 60-88. https://doi.org/https://doi.org/10.1016/j.media.2017.07.005. https://www.sciencedirect.com/science/article/pii/S1361841517301135.

[14]Tajbakhsh, Nima, Laura Jeyaseelan, Qian Li, Jeffrey Chiang, Zhihao Wu, and Xiaowei Ding. "Embracing Imperfect Datasets: A Review of Deep Learning Solutions for Medical Image Segmentation." (2019).

[15]Soler, Luc, Stephane Nicolau, Patrick Pessaux, Didier Mutter, and Jacques Marescaux. "Real-Time 3d Image Reconstruction Guidance in Liver Resection Surgery." [In eng]. Hepatobiliary Surgery and Nutrition 3, no. 2 (2014): 73-81. https://doi.org/10.3978/j.issn.2304-3881.2014.02.03. https://pubmed.ncbi.nlm.nih.gov/24812598.

[16]Lang, and Hauke. "Impact of Virtual Tumor Resection and Computer-Assisted Risk Analysis on Operation Planning and Intraoperative Strategy in Major Hepatic Resection." Arch Surg 140, no. 7 (2005): 629-38.

[17]Shorten, Connor, and Taghi M. Khoshgoftaar. "A Survey on Image Data Augmentation for Deep Learning." Journal of Big Data 6, no. 1 (2019). https://doi.org/10.1186/s40537-019-0197-0.