

基于双侧学习分支聚合上下文和局部特征的肝包虫病病灶分割

陆定^{1*}, 王舒², 尚婧晔¹, 夏勋³, 熊静⁴, 邓春梅⁴, 魏丽^{4*}

¹四川省疾病预防控制中心寄生虫病防治所, 四川 成都

²四川省计算机研究院网络及信息化研究所, 四川 成都

³成都医学院第一附属医院神经外科, 四川 成都

⁴甘孜藏族自治州人民医院放射科, 四川 甘孜

收稿日期: 2024年11月27日; 录用日期: 2024年12月21日; 发布日期: 2024年12月31日

摘要

利用人工智能算法对医学CT图像中的肝包虫病变区域进行检测和分析, 对医生的工作效率提高具有重要意义。然而, 由于数据集的大小有限、局部细节模糊、形状不规则等因素, CT图像中肝包虫病区域的分割仍然具有挑战性。因此, 本文提出了一种新的肝包虫病分割网络, 该网络结合了两个特征提取分支以实现速度和准确性的平衡。首先我们设计了一个上下文分支(CB), 通过类似Transformer的模块来保留尺度不变的全局上下文信息。然后, 基于深度聚合(DAP)模块的浅层细节分支(DB)来提供详细信息。我们在私有数据集上进行了大量实验。实验结果表明, 所提出的网络优于现有方法, 并在准确性和推理速度之间取得了良好的平衡。

关键词

肝包虫病, 上下文和局部特征, 双侧学习

Segmentation on Pathological Regions of Hepatic Echinococcosis Based on Bilateral Learning Branch Aggregation of Context and Local Features

Ding Lu^{1*}, Shu Wang², Jingye Shang¹, Xun Xia³, Jing Xiong⁴, Chunmei Deng⁴, Li Wei^{4*}

¹Parasitic Disease Prevention and Control Institute, Sichuan Provincial Center for Disease Control and Prevention, Chengdu Sichuan

*通讯作者。

文章引用: 陆定, 王舒, 尚婧晔, 夏勋, 熊静, 邓春梅, 魏丽. 基于双侧学习分支聚合上下文和局部特征的肝包虫病病灶分割[J]. 临床个性化医学, 2024, 3(4): 2627-2637. DOI: 10.12677/jcpm.2024.34374

²Institute of Network and Information Technology, Sichuan Institute of Computer Sciences, Chengdu Sichuan

³Neurosurgery Department of The First Affiliated Hospital of Chengdu Medical College, Chengdu Sichuan

⁴Radiology Department of Ganzi Tibetan Autonomous Prefecture People's Hospital, Ganzi Sichuan

Received: Nov. 27th, 2024; accepted: Dec. 21st, 2024; published: Dec. 31st, 2024

Abstract

Utilizing artificial intelligence algorithms for detecting and analyzing hepatic echinococcosis regions in medical CT images holds significant importance in improving the efficiency of medical practitioners. However, segmentation of echinococcosis regions in CT images remains challenging due to factors such as limited dataset size, blurred local details, and irregular shapes. Therefore, this paper proposes a novel echinococcosis lesion segmentation network that integrates two feature extraction branches to achieve a balance between speed and accuracy. Firstly, we design a contextual branch (CB) to preserve scale-invariant global contextual information using a Transformer-like module. Additionally, a shallow detail branch (DB) based on the Deep Aggregation Pyramid (DAP) module is introduced to provide detailed information. Extensive experiments are conducted on a private dataset. Experimental results demonstrate that the proposed network outperforms existing methods and achieves a good balance between accuracy and inference speed.

Keywords

Hepatic Echinococcosis, Contextual and Local Characteristics, Bilateral Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

包虫病是一种严重危害人类健康的人畜共患性疾病。引起人体包虫病的绦虫主要有细粒棘球绦虫幼虫感染所致的囊型包虫病(cystic echinococcosis, CE)和多房棘球绦虫幼虫感染所致的泡型包虫病(alveolar echinococcosis, AE)两种类型[1], 两类包虫病均主要侵犯肝脏[2]。四川省西部是包虫病流行最严重的地区之一, 截至 2021 年底, 全国棘球蚴病患者有 40.56%分布在四川, 居全国之首[3], 四川省 92.36%的患者来自甘孜州(2023 年监测数据)。经有效干预, 2017 年抽样调查发现, 包虫病流行最为严重的石渠县肝囊型包虫病患病率仍高达 3.45%, 这部分患者患病时间长, 病情复杂, 相当一部分病人错失最佳诊断和治疗时机, 严重影响患者的生活质量[4]。因此, 如何有效诊断肝包虫病是一个迫切需要解决的重大民生问题。

最近, 不同的卷积神经网络被应用于分割医学图像。U-Net [5]是众多分割网络中最具代表性的网络模型之一。U-Net 使用由多个卷积和池化运算符组成的编码器从输入中提取抽象特征, 然后由多个转置卷积或上采样运算符组成的解码器尝试从特征中识别目标。此外, 编码器和解码器之间设计了快捷连接以促进信息流动。在 U-Net 的基础上, V-Net [6]将运算符的维度从二维扩展到三维, 并添加了残差连接, 有效缓解了网络深度增加导致的梯度消失问题。为了减少不同抽象特征级别之间的语义差异, Fan 等人 [7]提出了一种名为 Inf-Net 的网络, 用于分割新型冠状病毒性肺炎的 CT 图像。该网络使用一组隐式反向注意力模块和显式边缘注意力模块来建立区域和边界之间的关系。此外, Gu 等人[8]基于注意机制和条件随

机场设计了一个名为 CA-Net 的网络，用于调节信息流以进行脑胶质瘤的分割。

尽管上述提到的方法在医学图像中分割病变方面表现出令人印象深刻的性能，但其中很少有考虑到由于卷积运算符固有的局部性而导致的全局和局部信息之间的差异[9]。目前，Transformer 中的自注意力机制[10]用于建模全局信息，已广泛应用于自然语言处理任务。Transformer 在图像分类、目标检测和分割等任务中超越了基于卷积的模型的性能[11]。然而，Transformer 的成功应用依赖于大规模数据集，这对医学图像分割任务提出了挑战。因此，我们提出了一种新型的双侧学习机制，以平衡全局和局部信息。设计了一个基于扩张卷积的上下文分支(CB)来保留尺度不变的全局信息，同时提出了一个利用深度聚合金字塔模块(DAP)学习局部信息的细节分支(DB)。最后，通过多个实验验证了所提方法的有效性。

2. 基于双侧学习分支的网络架构

本文提出的网络是一个像素级别的分割网络，主要由两个组成部分组成：① 上下文分支(Context Branch, CB): 由轻量级的 RepVGG 和特征融合模块(MLPF)组成。最后三个特征映射(即 1/8, 1/16, 1/32 分辨率)被融合表示全局信息。② 细节分支(Detailed Branch, DB): DB 建立在 DAP 模块上，用于在相同分辨率下提取特征图，专门设计用于捕获局部信息。从 CB 和 DB 分支提取的特征图然后与输入图像具有相同的分辨率融合。最后，使用一个 1×1 卷积层，后跟 sigmoid 函数来生成肝包虫病变区域的概率图。此外，在训练过程中使用标签来优化 CB 和网络的输出。整体网络结构如图 1 所示，其中橙色矩形框对应多通道特征图，框的高和宽表示特征图的空间大小和通道数。MLPF 代表基于 MLP 融合两个特征图，DAP 代表深度金字塔聚合模块，Label 代表在训练期间进行的深度监督。

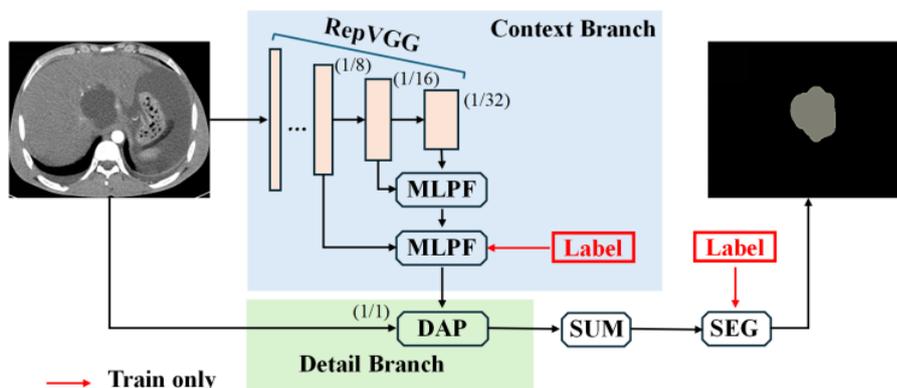


Figure 1. Overall network architecture diagram
图 1. 整体网络架构图

2.1. 改进的 RepVGG

为了在肝包虫病变区域识别中实现速度和准确性的良好平衡，本文采用扩张卷积来改进 RepVGG [12]的模块以提取多尺度的特征图，并将其作为 CB 的骨干网络。改进后的 RepVGG 模块与常规卷积进行了比较，如图 2 所示。

图 2(A)表示 ResNet [13]的原始残差块，它由两个 3×3 卷积层的残差分支和一个快捷分支构成。然后将这两个分支求和作为块的输出。残差结构用于缓解深度网络中的梯度消失问题，并使网络加快收敛速度。图 2(B)表示训练时使用的 RepVGG 的卷积块，它由一个 3×3 卷积层的残差分支、一个 1×1 卷积层的快捷分支和一个快捷连接分支组成，最终输出由三个分支的输出线性相加。这种结构由多条分支传播梯度信息，可以更好地模拟梯度消失。同时，它还可以在网络不同阶段监督训练过程，有助于提升网络鲁棒性。图 2(C)

表示在推理中使用的 RepVGG 的卷积块，它由两个 3×3 卷积层重新构建而成，用于替代训练过程中的三个分支。这种方法丢弃了叠加过程和多分支并行计算的开销，并可以更好地提高内存利用率，从而提高模型的推理速度。其中图中红色矩形框代表批归一化层， \oplus 代表通道维度特征的线性相加。

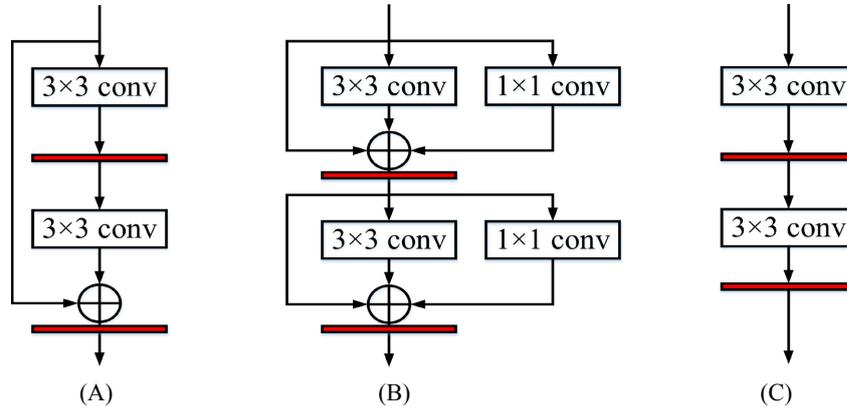


Figure 2. Comparison of convolution blocks
图 2. 卷积块比较

由于在分割具有大尺寸变化和不规则形状的肝包虫病灶任务中，获得较大的感受野和高维上下文语义信息非常重要。为了进一步增强特征并提取主干网络的全局感受野，本文使用步长为 2 的 5×5 扩张卷积层来替代步长为 1 的标准 3×3 卷积层，最终可使输出与输入的大小保持一致。扩张卷积基于普通卷积，感受野的扩大。其中普通卷积计算原理如公式(1)所示：

$$z_{v,u}^{l+1} = \sum_{c=0}^{C-1} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} a_{v+p,u+q,c}^l \cdot w_{p,q,c}^l \tag{1}$$

其中， P 、 Q 和 C 分别表示卷积核的维度， p 、 q 和 c 表示内核 w^l 的索引。 $z_{v,u}^{l+1}$ 表示通过卷积操作 w^l 计算得到的 $l+1$ 层中位置 (v, u) 的输入，以及来自上一层 a^l 的输出。扩张卷积基于普通卷积引入了扩张率 r 来控制卷积核中的间隔。扩张卷积可以被看作在卷积核的元素之间插入 $r-1$ 个零。扩张卷积的计算原理如公式(2)所示：

$$z_{v,u}^{l+1} = \sum_{c=0}^{C-1} \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} a_{v+r \cdot p, u+r \cdot q, c}^l \cdot w_{p,q,c}^l \tag{2}$$

通过比较公式(1)和公式(2)，可以看出当 $r=1$ 时，扩张卷积退化为标准卷积。

2.2. 基于 MLP 的上下文分支

本文改进的 RepVGG 提取的最后三个特征图由 MLP 和卷积(MLPF)构成的特征融合模块进行融合。具体位置如图 1 所示，其中两个 MLPF 用于融合三个特征图(即 $1/8$, $1/16$, $1/32$)，上下文分支的输出是输入图像的 $1/8$ 分辨率。

MLPF 的计算原理如图 3 所示。在融合过程中，较小的特征图通过双线性插值进行两次上采样(UPS)以匹配较大特征图的大小。然后，使用两个 3×3 卷积层顺序地将两个特征图的通道数减少到较小的特征图。两个相同大小的特征图在通道维度上连接起来。最后，分别在令牌维度和通道维度上使用两个 MLP 来融合相同大小的特征图。MLP 层基于 Transformer，它在建模特征之间的长距离关系方面具有优势。Transformer 的计算原理如公式(3)所示：

$$A' = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (3)$$

假设输入特征为 A ， Q 、 K 和 V 表示特征 A 的转换。其中， d_k 表示特征 A 的维度，除以 $\sqrt{d_k}$ 的目的是增强数值稳定性。以特征 Q 为例，其计算原理可以总结为 $Q = W^Q A$ ，其中 W^Q 表示可学习参数，通过矩阵乘法将特征 A 的转换为特征 Q 。对于特征 K 和 V ，也是类似的原理。 softmax 激活函数旨在引入非线性，并完成归一化。

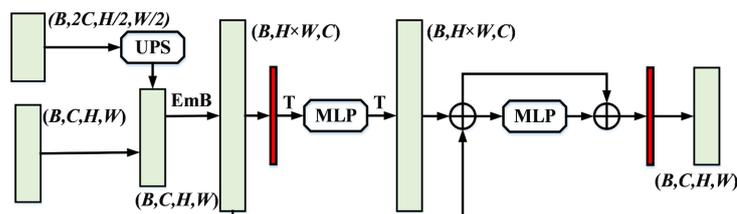


Figure 3. Feature fusion module
图 3. 特征融合模块

2.3. 基于 DAP 的细节分支

局部细节信息对于实现精确分割至关重要。CB 提取的特征具有低分辨率，并且缺乏详细信息。因此，我们提出了基于深度聚合金字塔(DAP)模块的 DB 分支[14]，以提取具有高分辨率的特征，弥补了 CB 分支的局限性。DB 分支的架构如图 4 所示。DB 分支的结构包括五个路径。首先，使用三个平均池化操作符将特征图减小到不同的尺寸。其次，使用卷积层来提取多尺度的局部细节。第三，将这些特征图上采样到相对于输入相同的分辨率。最后，这五个路径通过 1×1 卷积进一步处理，并沿着通道维度串联起来。其中，每个卷积单元由卷积层、批量归一化和 ReLU 激活函数组成。

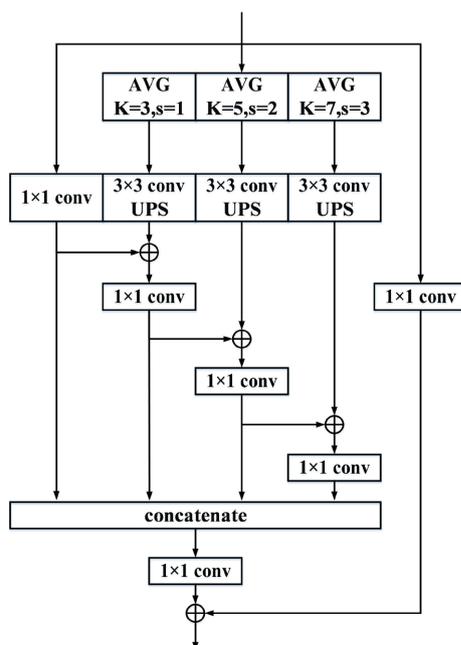


Figure 4. Deep pyramid aggregation module
图 4. 深金字塔聚合模块

2.4. 语义分支和深度监督损失

为了提高该网络不同特征提取分支和不同卷积阶段的特征表达能力, 本文设计了一个深度监督模块来指导训练过程中使用的不同特征提取阶段。深度监督模块的架构如图 5 所示。

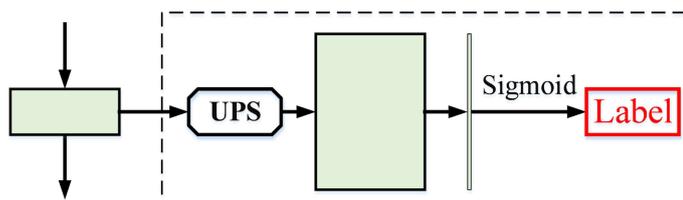


Figure 5. Deep supervision module
图 5. 深度监督模块

在深度监督模块中, 监督的特征图被上采样到 1 分辨率并通过 1×1 卷积层处理, 在训练中通过 sigmoid 函数根据真值计算损失。根据整体网络架构, 本文使用两个深度监督模块来监督从 CB 中获取的全局上下文特征图和最终的概率图。在推理过程中, 将第一个深度监督模块舍弃, 最终结果被直接视为预测结果, 不需要任何额外的模型参数和计算。

在每个深度监督模块中, 特征图经过一个 1×1 卷积层, 然后是 sigmoid 函数处理。然后, 相对于地面真实值计算损失。在训练阶段, 使用两个深度监督模块来监督来自 CB 的预测和最终的概率图。在推理阶段, 省略第一个辅助监督模块, 并将最终结果视为预测结果。语义分割任务通常使用交叉熵损失[15]地面真实值之间的不相似性。然而, 在病变分割任务中出现了类别不平衡问题, 其中背景构成 CT 图像的大部分, 从而减慢了网络的学习过程。此外, 背景中存在大量具有挑战性的体素, 如纹理和结节, 这增加了分割任务的复杂性。(肝包虫病在 CT 图像中呈现多种影像学特征, 这些特征与别的多种肝脏病变相似, 包括单纯性肝囊肿、细菌性肝脓肿、肝囊腺瘤以及肝细胞癌, 这增加了分割任务的复杂性。)因此, 在训练阶段使用 Dice 损失, 因为与交叉熵相比, 它可以更好地解决类不平衡问题。Dice 系数(DSC)是评估两个集合相似性的指标[16]。两个集合之间的 DSC 如公式(3)所示:

$$DSC = 2 \frac{|P \cap Y|}{|P| + |Y|} \quad (3)$$

其中, P 和 Y 分别表示模型的分割结果和真实值的体积。DSC 被限制在 $[0, 1]$ 范围内, 其中数值为 1 表示预测与真实值完全匹配, 反之亦然。Dice 损失可以描述如公式(4)所示:

$$L_{DSC} = 1 - 2 \frac{|P \cap Y|}{|P| + |Y|} \quad (4)$$

根据提出的深度监督模块, 损失函数由两部分组成, 即用于分割性能的主要损失和旨在稳定训练的辅助损失。这两个损失被定义为 DSC 损失, 公式如(5)所示:

$$Loss = L_f + \alpha L_m \quad (5)$$

其中 L_m 代表辅助损失, L_f 代表最终结果和真值之间的主要损失, α 表示辅助监督损失的权重, 本文设置为 0.4。

3. 实验与分析

本节首先介绍数据集和采用的数据增强方法。接下来详细讨论实验环境以及评估指标。最后展示和

讨论与其他模型比较结果和消融实验。

3.1. 数据集和评估指标

3.1.1. 数据集和评估指标

本文利用甘孜藏族自治州人民医院提供的自标注的肝包虫病 CT 影像数据集对该方法进行了验证。数据集由专业医生详细标注。转换后的数据集包含 4320 张灰度图像，每张图像的分辨率为 512×512 。其中，随机选取 1440 张作为验证数据集，另外 2880 张作为训练数据集。

3.1.2. 数据增强

模型的泛化能力可以通过使用数据增强来提高[17]。根据研究和分析，肝包虫病变图像中的噪声分布在显著差异。针对这种情况，本文在训练上采用了水平翻转、几何变形等方法来扩大训练数据集。其中几何变形是根据肝包虫病变区域的不规则性，应用三角函数改变每个像素点的位置。该方法公式如(4)所示：

$$\begin{aligned}\hat{h} &= h - 10 \sin\left(\frac{2\pi w}{152} + 10\right) \\ \hat{\omega} &= \omega - 10 \sin\left(\frac{2\pi h}{152} + 10\right)\end{aligned}\quad (4)$$

其中， \hat{h} 、和 $\hat{\omega}$ 分别为各像素点在变形前后的坐标位置。

3.1.3. 评价指标

本文使用四种常用的语义分割指标来评估该网络，指标如公式(5)~(8)所示：

$$Pa = \frac{TP}{TP + FP} \quad (5)$$

$$Rc = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2 \times Pa \times Rc}{Pa + Rc} \quad (7)$$

$$IoU = \frac{TP}{TP + FN + FP} \quad (8)$$

其中， TP 表示被正确分类的病变像素的数量， FN 表示被错误分类为背景的病变像素的数量， FP 表示被错误分类为病变的背景像素的数量。召回率(Rc)用于评价病灶像素被正确分类的性能，精度(Pa)用于评价分类为病灶的像素为真实病灶的性能。而 IoU (intersection over union)和 $F1$ (F1-score)作为重要的综合指标，可以看作是 Rc 和 Pa 的调和平均。四个指标的值分布在 0 和 1 之间，当它们的值接近于 1 时，模型对病变区域的分割效果更好。此外，为了评估网络的性能，使用参数数量(Par)和每秒浮点运算次数(Gfl)评估模型在推理过程中的计算资源成本。帧率(Fra)也用于表示模型完成一次推理的时间。

3.1.4. 实验细节

本文所有网络使用 Pytorch 深度学习框架实现和训练。所有基于 CNN 的模型都在 12G NVIDIA GTX TITAN XP 和 i7-7770 CPU@4.20GHz 的计算机上进行训练和验证。我们使用 Adam，学习率为 0.0001，将批量大小 minibatch 设为 4，优化所有分割网络。在 100 个 epoch 的训练过程中，选择验证数据集上平均损失值最小的模型作为最终模型。在推理阶段，输出值大于 0.5 的像素被识别为肝包虫病变，其他像素被识别为背景。

3.1.5. 模型比较

本文在私有数据集上使用我们的网络和其它五个经典分割网络进行了比较：SegNet [18]、Deeplab v3 [19]、U-Net [5]、U-Net++ [20]和 TransUNet [21]。其中 SegNet 中的解码器使用在编码器的最大池化步骤中计算的池化索引来执行非线性上采样，从而消除了学习上采样的需要并改善了内存与准确性的权衡；Deeplab v3 使用空卷积空间金字塔池化模块检测多尺度上的卷积特征图，捕获多尺度上下文信息；U-Net 的架构由一个捕获上下文的收缩路径和一个能够实现精确定位的对称扩展路径组成，具有高精度图像分割的能力；U-Net++改善了不同层级特征图之间的语义差距，在多场景下具有更好的分割精度。TransUNet 在 U-Net 的基础上使用 transformer 结构改进了特征融合模块。网络输入大小为 512×512 图像进行实验，详细结果如表 1 所示。

Table 1. Comparison of different networks

表 1. 不同网络的对比

Network	Par (M)	Gfl (GMac)	Fra (ms)	IoU (%)	F1 (%)
SegNet	29.4	16.1	23.5	57.5	71.6
Deeplab v3	15.3	15.6	26.2	68.3	83.2
UNet	17.3	16.3	24.9	78.1	87.9
UNet++	26.9	15.3	22.9	80.8	88.5
TransUnet	66.8	130.4	31.9	80.9	89.7
ours	7.3	12.7	17.9	81.9	90.3

从表中可以看出，本文提出的网络实现了每张图像 17.9 毫秒的最佳精度和实时速度。参数量和计算量达到了所有方法的最小值。并且，与其他分割网络相比，所提出的网络的 IoU 和 F1 中取得了最佳性能，证明了我们提出的网络在肝包虫病变区域特征提取方面的有效性和准确性。图 6 展示了图像在不同网络的分割结果。具体来说，基于 U-Net 的网络识别出微小的裂缝，并且比其他网络更准确地依赖于具有丰富细节信息的高分辨率特征图。而本文提出的网络对肝包虫病变区域的预测结果比 U-Net 具有更好的连续性和完整性，这主要是因为 CB 具有更大的感受野信息以及 DB 有丰富的细节信息，使得图像分割更完善，边界更清晰和流畅。

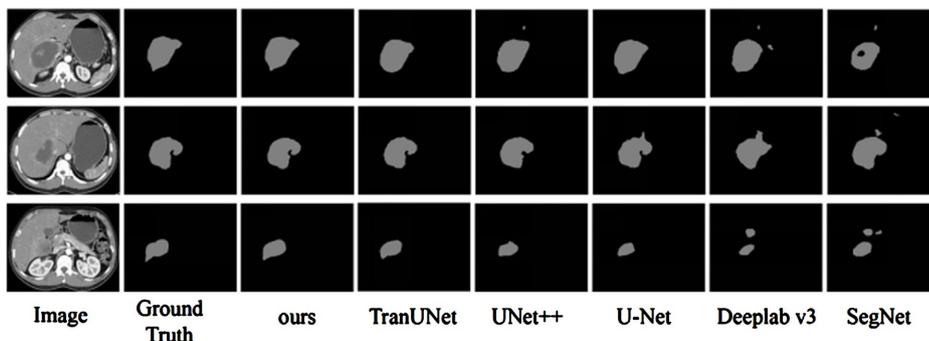


Figure 6. Image segmentation results of different networks

图 6. 不同网络的图像分割结果

3.1.6. 消融研究

本小节验证两个模块的有效性，在私有数据集上进行五个消融实验。实验分别验证了将网络结构改为 RepVGG-32、RepVGG-8、CB 以及没有深度监督模块的网络。实验结果如表 2 所示，分别展示每个模

块及其组合在 IoU 和 F1 指标的贡献。

Table 2. Comparison of segmentation networks with different improvement modules
表 2. 不同改进模块的分割网络比较

Network	IoU (%)	F1 (%)
RepVGG-32	66.8	80.1
RepVGG-8	78.8	88.3
Context Branch	79.1	89.1
Without Deep Supervision	81.3	89.8
ours	81.9	90.3

从表中可以看出,在不添加任何模块的 RepVGG-32 的基础上,我们的网络总共带来了 15.1%的 IoU 和 10.2%的 F1 提升。而仅使用 MLPF 的 CB 在相同 RepVGG-8 下提高了 0.3%的 IoU 和 0.8%的 F1。而在 RepVGG-32 的基础上,MLPF 提升了 12.3%的 IoU 和 9.0%的 F1,在此基础上,DB 提升了 2.2%的 IoU 和 0.7%的 F1。最后,使用深度监督模块提升了 0.6%的 IoU 和 0.5%的 F1。以上结论证明了改进的 RepVGG 以及提出的两个分支能提取丰富细节信息的高分辨率特征图,对于肝包虫病灶分割提升有显著提升。

4. 结论

目前,人工智能技术已用于肿瘤性疾病的辅助诊断,其中深度学习是一种基于数据的表示学习,可以学习高级抽象特征表达。国内外迅速将其应用到医学图像处理领域,并取得了较好的成果。基于深度学习的多模态医学影像关键技术快速发展,和传统的人工判定比较,利用计算机技术辅助过程极大地提高了效率,减少了人力、物力及财力的巨大消耗。目前,有学者[22]提出使用计算机辅助诊断(computer aided diagnosis, CAD)的方法来辅助医师对肝囊型包虫病做出诊断,以此弥补部分地区医疗资源不足的问题。随着深度学习网络的快速发展,图像处理中的目标检测任务也有了跨越性的进展[23]随着深度学习的快速发展,基于卷积神经网络(CNNs)等的目标检测算法取得了显著的成果,这不仅提高了检测的准确性,还加快了检测速度,使得实时应用成为可能。

本文针对肝包虫病变区域的全局背景特征、不同尺寸变化、小尺度局部细节和不规则形状等特点,提出了一种由上下文特征提取分支和详细特征提取分支组成的新型深度网络。在所提出的方法中,基于 RepVGG 的改进轻量级网络用于提取具有多尺度感受野的全局背景特征图。特征融合模块用于高效地融合特征图。同时,使用由深度聚合金字塔模块组成的详细特征提取分支来提供局部详细信息。所提出的网络在私有数据集上进行了验证。实验结果表明,所提出的网络达到了最佳的准确性。在未来的研究中,我们将在额外的医学图像分割任务中验证所提出的方法。

利益冲突声明

全部作者声明无利益冲突。

作者贡献声明

陆定负责研究设计、论文撰写,尚婧晔参与图表绘制,夏勋、熊静、邓春梅参与数据收集,魏丽,王舒负责论文指导。

基金项目

四川省科技计划项目(2023YFG0281)四川省重点研发项目(2021YFG0316)。

参考文献

- [1] McManus, D.P., Gray, D.J., Zhang, W. and Yang, Y. (2012) Diagnosis, Treatment, and Management of Echinococcosis. *BMJ*, **344**, e3866-e3866. <https://doi.org/10.1136/bmj.e3866>
- [2] Holcman, B. and Heath, D.D. (1997) The Early Stages of *Echinococcus granulosus* Development. *Acta Tropica*, **64**, 5-17. [https://doi.org/10.1016/s0001-706x\(96\)00636-5](https://doi.org/10.1016/s0001-706x(96)00636-5)
- [3] 龚嫣, 薛垂召, 王旭, 等. 2021 年全国棘球蚴病防治进展[J]. 中国寄生虫学与寄生虫病杂志, 2023, 41(2): 142-148.
- [4] 喻文杰, 王谦, 廖沙, 等. 2017 年四川省石渠县包虫病患病情况现状调查[J]. 预防医学情报杂志, 2018, 34(5): 545-549.
- [5] 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: 18th International Conference*, Munich, 5-9 October 2015, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [6] Milletari, F., Navab, N. and Ahmadi, S. (2016) V-net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *2016 4th International Conference on 3D Vision (3DV)*, Stanford, 25-28 October 2016, 565-571. <https://doi.org/10.1109/3dv.2016.79>
- [7] Fan, D., Zhou, T., Ji, G., Zhou, Y., Chen, G., Fu, H., et al. (2020) Inf-net: Automatic COVID-19 Lung Infection Segmentation from CT Images. *IEEE Transactions on Medical Imaging*, **39**, 2626-2637. <https://doi.org/10.1109/tmi.2020.2996645>
- [8] Gu, R., Wang, G., Song, T., Huang, R., Aertsen, M., Deprest, J., et al. (2021) Ca-net: Comprehensive Attention Convolutional Neural Networks for Explainable Medical Image Segmentation. *IEEE Transactions on Medical Imaging*, **40**, 699-711. <https://doi.org/10.1109/tmi.2020.3035253>
- [9] Sang, H., Zhou, Q. and Zhao, Y. (2020) Pcanet: Pyramid Convolutional Attention Network for Semantic Segmentation. *Image and Vision Computing*, **103**, Article ID: 103997. <https://doi.org/10.1016/j.imavis.2020.103997>
- [10] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [11] Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2020) An Image Is Worth 16 × 16 Words: Transformers for Image Recognition at Scale.
- [12] Ding, X., Zhang, X., Ma, N., Han, J., Ding, G. and Sun, J. (2021) RepVGG: Making VGG-Style Convnets Great Again. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13733-13742. <https://doi.org/10.1109/cvpr46437.2021.01352>
- [13] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [14] Hong, Y., Pan, H., Sun, W., et al. (2021) Deep Dual-Resolution Networks for Real-Time and Accurate Semantic Segmentation of Road Scenes.
- [15] Lin, T., Goyal, P., Girshick, R., He, K. and Dollar, P. (2017) Focal Loss for Dense Object Detection. *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/iccv.2017.324>
- [16] Li, X., Sun, X., Meng, Y., Liang, J., Wu, F. and Li, J. (2020) Dice Loss for Data-Imbalanced NLP Tasks. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, July 2020, 465-476. <https://doi.org/10.18653/v1/2020.acl-main.45>
- [17] Wu, Q., Dai, P., Chen, P. and Huang, Y. (2019) Deep Adversarial Data Augmentation with Attribute Guided for Person Re-Identification. *Signal, Image and Video Processing*, **15**, 655-662. <https://doi.org/10.1007/s11760-019-01523-3>
- [18] Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 2481-2495. <https://doi.org/10.1109/tpami.2016.2644615>
- [19] Chen, L., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H. (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Computer Vision—ECCV 2018*, Munich, 8-14 September 2018, 833-851. https://doi.org/10.1007/978-3-030-01234-2_49
- [20] Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N. and Liang, J. (2020) Unet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation. *IEEE Transactions on Medical Imaging*, **39**, 1856-1867. <https://doi.org/10.1109/tmi.2019.2959609>
- [21] Chen, J., Lu, Y., Yu, Q., et al. (2021) Transunet: Transformers Make Strong Encoders for Medical Image Segmentation.
- [22] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) Imagenet Classification with Deep Convolutional Neural Networks.

Communications of the ACM, **60**, 84-90. <https://doi.org/10.1145/3065386>

- [23] Deng, J., Dong, W., Socher, R., Li, L., *et al.* (2009) ImageNet: A Large-Scale Hierarchical Image Database. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, 20-25 June 2009, 248-255. <https://doi.org/10.1109/cvpr.2009.5206848>