基于对比学习增强的Lora微调超声影像分割 模型

张雨萌,李逸凡

上海理工大学光电信息与计算机工程学院,上海

收稿日期: 2025年3月23日; 录用日期: 2025年4月16日; 发布日期: 2025年4月24日

摘要

超声影像分析在现代医学中扮演着至关重要的角色,但精确分割是其面临的主要挑战之一。尽管现有的 深度学习模型如SAM在自然图像上表现出色,但在医学图像分割上仍存在性能差距。本研究提出了一种 基于对比学习增强的LoRA微调SAM-Med3D超声影像分割模型(USCL-Med3D),旨在提高3D超声影像分 割的精确度和效率。为此,设计了一种半监督伪标签数据集训练方法,通过自动化获取标注数据,降低 了标注难度并保证了标注效果。同时,引入对比学习架构VCL-head,增强了模型对3D超声影像上下文信 息的提取能力。此外,还对SAM-Med3D模型进行了LoRA微调,从而使模型具有更好的分割能力。实验 结果表明,所提方法在3D超声数据集和一些公开的3D医疗影像数据集上取得了优异的分割效果。

关键词

3D超声分割,数据集,对比学习,LoRA微调

A Contrastive Learning-Enhanced Lora Fine-Tuned SAM-Med3D Model for Ultrasound Image Segmentation

Yumeng Zhang, Yifan Li

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 23rd, 2025; accepted: Apr. 16th, 2025; published: Apr. 24th, 2025

Abstract

Ultrasound image analysis plays a critical role in modern medicine, but precise segmentation

remains one of its major challenges. Although existing deep learning models like SAM perform well on natural images, there is still a performance gap in medical image segmentation. This study proposes a contrastive learning-enhanced LoRA fine-tuned SAM-Med3D ultrasound image segmentation model (USCL-Med3D) to improve the accuracy and efficiency of 3D ultrasound image segmentation. We designed a semi-supervised pseudo-label dataset training method to automatically obtain annotated data, reducing annotation difficulty while ensuring annotation quality. Additionally, a contrastive learning architecture was introduced to enhance the model's ability to extract contextual information from 3D ultrasound images. Furthermore, we fine-tuned the SAM-Med3D model using LoRA, effectively incorporating the feature representation abilities of the 3D ultrasound dataset. Our method achieved excellent segmentation performance on the 3D ultrasound dataset and several publicly available 3D medical imaging datasets.

Keywords

3D Ultrasound Segmentation, Dataset, Contrastive Learning, LoRA Fine-Tuning

Copyright © 2025 by author(s) and Hans Publishers Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <u>http://creativecommons.org/licenses/by/4.0/</u>

CC Open Access

1. 引言

医学影像分析在现代医学中扮演着至关重要的角色,它是医生诊断和治疗疾病不可或缺的工具。通过各种影像技术,如X射线、计算机断层扫描(CT)、磁共振成像(MRI)和超声影像等,医生能够洞察人体内部结构和功能,从而为患者提供更准确的诊断和更有效的治疗方案。

超声影像分析作为医学影像分析的一个重要领域,最重要的挑战之一是对超声影像进行精确分割。 传统上,像[1]这样的深度学习模型及其变体[2][3]在特定任务和数据集上表现出色,但缺乏交互能力导致 它们在实际使用时作用有限。尽管一些工作考虑了交互式分割方法作为克服狭窄范围、特定任务模型限 制的一种手段[4]。但 Segment Anything Model (SAM)[5]的引入标志着基于提示的交互式分割范式的转变, 由于基础模型的固有泛化能力,SAM 在自然图像上取得了不错的表现。SAM 由一个预训练的视觉变换 器(ViT)编码器[6]、一个提示编码器和一个小解码器组成,通过在超过 10 亿个掩膜和 1100 万张图像上训 练,SAM 获得了能够适应新任务而无需重新训练的能力[5]。

虽然在自然图像上取得了成功,但将 SAM 直接应用于医学分割揭示了与特定任务[7]相比的性能差距。为了缓解这个问题,MedSAM[8]和 SAM-Med2D[9]通过模型微调为包括超声医学影像在内的 2D 医学数据量身定制。虽然这些方法在 2D 医学数据上获得了成功,但考虑到相当多的医学数据,比如超声影像,计算机断层扫描(CT)影像在时间或空间上具有连续性,导致 2D 数据定制的 2D 方法忽略了这些数据隐含的特征。因此,当应用于 3D 体积数据时,这些方法由于逐片处理[10]而表现不佳。为了缓解这个问题,SAM-Med3D[11]引入了 SAM 组件的 3D 对应物和端到端的 3D 训练。SAM-Med3D 作为一种具有完全可学习的 3D 架构的模型,并在完全处理的大规模体积医学数据集上得到了训练。但 SAM-Med3D 在 3D 超声影像数据上的表现不佳。经过分析认为,3D 体积数据,尤其是 3D 超声影像数据中含有一些上下文信息[12],SAM-Med3D 架构并没有学习体积数据所含上下文信息的能力,所以该方法在应用到 3D 超声影像时没有收获理想的效果。

在视觉和语言基础模型(LFM)如 GPT-4 [13]的启发下,近期研究涌现了一批视觉基础模型(VFM) [14]-[18]。这些 VFM 通过大规模预训练和定制化微调方法,展现出了卓越的零样本和少样本泛化能力,能够 适应各种任务或领域。CLIP [15]模型通过在大量语言 - 图像对上预训练,提供了一对视觉和语言模型, 它在分类、视频理解和图像操作等多种任务上表现出色。DALLE3 [18]基于大规模变换器架构训练,展现 了卓越的文本到图像生成能力。DINOv2 [16]在大量图像数据上进行预训练,以学习鲁棒的视觉特征,无 需额外的监督即可在多种图像识别和分割任务上展现出卓越的性能。SegGPT [14]将不同的分割数据转换 为标准化图像格式,为多样化的分割任务开创了统一方法。SEEM [17]提供了一种全能方法,使用不同的 提示一次性分割和识别图像或视频中的项目。SAM 模型通过在惊人的 10 亿个掩模上预训练,成为了一 个多用途的 VFM,用于可提示的图像分割,在各种视觉任务中展现了令人印象深刻的零样本性能。尽管 SAM 在通用性方面表现出色,它在医学成像、伪装和阴影[19]等领域仍有局限性,并且这些 VFM 并未专 门为医学图像分析量身定制。

在 3D 医学超声图像分析领域,已经有尝试使用医学数据对 SAM (Segment Anything Model) [5]架构 进行微调。然而,这些方法通常涉及在 2D 框架内处理和微调 3D 数据,这可能无法充分利用 3D 超声图 像中存在的深度信息。例如,一些研究已经使用适配器和大量掩模在特定的医学分割数据集上对 SAM 进行微调,但这些方法尚未针对 3D 超声成像的独特特性进行优化[8] [20]。相反,也有一些方法使用 3D 架构对 3D 数据进行预训练,旨在更有效地捕获空间信息。然而,这些方法在 3D 超声数据集上并未展现出 令人满意的性能,可能由于缺乏针对超声的特定优化或数据集的多样性和规模不足,限制了模型的泛化 能力[21] [22]。同样值得注意的是,一些 3D 方法在超声图像分割上取得了值得称赞的结果,但通常需要 更多的提示点,这给实际应用带来了较大的困难。相比之下,SAM-Med3D [11]因其交互能力而脱颖而出,能够在显著减少提示点的情况下提供出色的分割结果。尽管如此,到目前为止,还没有研究者成功地将 SAM 架构的强大功能与 3D 超声图像分析的具体需求相结合,特别是在高效的交互式分割且提示点最少 的背景下。总结来说,虽然在将 SAM 应用于 3D 超声图像分析方面已经取得了进展,但仍需要能够以最 少的交互有效利用 3D 数据的全部潜力的方法。

自监督学习(Self-Supervised Learning, SSL)的出现彻底改变了三维医学图像分析领域[23]-[26],特别 是超声成像,由于其实时采集的特性和固有噪声,超声成像面临独特的挑战。SSL 方法能够无需大量手 动注释就能从三维超声数据的复杂结构中学习[27][28],这些方法通常依赖于利用图像内部存在的几何和 上下文关系的预设任务[29][30]。在三维医学超声成像的背景下,SSL 已被用于学习可以泛化到各种下游 任务的鲁棒特征。例如,像[31]和[24]这样的方法通过仿射增强利用多尺度恢复和几何相似性学习来从医 学图像中学习不变特征。然而,这些方法可能没有完全针对超声数据的独特特性进行优化,例如其各向 异性和斑点噪声。针对医学图像分析的位置相关 SSL 方法也已被探索,尽管对超声成像的特定关注较少。 这些方法旨在预测图像块在更大框架中的上下文位置或位置。挑战在于有效地捕获超声数据中存在的复 杂几何关系,这些关系可能不像其他成像方式那样一致[32]。尽管 SSL 在医学图像分析中具有潜力,但 文献中关于将这些方法应用于三维超声图像的应用仍然存在差距。现有作品尚未完全探索超声数据的细 微差别,例如其实时采集过程以及需要高级语义理解来准确分割和分析图像的需求[33]-[35]。因此,需要 针对三维超声图像的特定特性量身定制的 SSL 方法。这些方法虽然在不同的数据集上取得了一定的效果, 但多数方法缺乏可交互能力。SAM 及其衍生方法虽然具备可交互能力,但由于没有表示数据中隐含的上 下文信息的能力,导致这些方法在 3D 超声数据集上效果不佳(表 1)。

为了缓解上述问题,本研究设计了 USCL-Med3D 模型,这项工作的贡献有三个方面。1) 提出了一种 使用自动化获取的半监督伪标签数据集训练的方法,有效降低了 3D 超声影像数据的标注难度,同时保 证了一定的标注效果。2) 通过为模型加入对比学习架构,从而使其具备提取 3D 超声影像上下文信息的 能力,有效增强了模型在 3D 超声数据集和一些 3D 医疗数据集上的表现。3) 对具有一定通用 3D 医学影 像分割能力的 SAM-Med3D 模型设计了微调方法,有效利用 SAM-Med3D 从大量医疗数据上训练所获得 的特征表示能力并引入了表示 3D 超声数据集所隐含的特征的能力。USCL-Med3D 在 3D 超声数据集和 一些公开的 3D 医疗影像数据集上获得了出色的效果。

2. 方法

2.1. 半监督伪标签数据集

2.1.1. 数据获取

在构建专注于超声影像分析的数据集过程中,积累并精心挑选了具有高度针对性的数据集。本数据 集专注于体积超声图像,总计囊括了来自 39 名患者的 500 个体积超声扫描,这些患者均表现出不同程度 的积液病症。为了确保数据的全面性和代表性,对样本来源进行了细致划分: 25 名患者的数据集中于胸 腔积液的观测,通过医师沿胸腔冠状面,握持超声探头平稳扫描获取;而另外 14 名患者则提供包含腹腔 积液及腹腔器官的超声影像,这些影像不仅记录了积液的状况,还额外覆盖了肝、双肾及脾脏这三种关 键解剖结构,从而为研究增添了丰富的病理与解剖学信息维度。图 1 展示了医师采集和标注的这几种数 据。这些特定病理状态下的超声影像组成数据集,有助于训练模型为诊断和治疗提供更准确的信息。



Figure 1. Six distinct sets of ultrasonographic images obtained through standardized probe manipulation by board-certified physicians, accompanied by their respective diagnostic annotations 图 1.6 种由专业医师握持超声探头扫查得到的画面及其对应标注

2.1.2. 利用 SegGPT 生成伪标签数据

SegGPT 可以作为生成伪标签数据集的过程充分利用了临床医生的标注,这些标注作为 SegGPT 模型的提示(prompt)。这些代表性的帧作为参考, SegGPT 据此推断并分割未标注数据的其余部分,从而创建出丰富的伪标签集。首先收集一系列医学影像视频,每个视频在选定的关键帧上都有标注,这些关键帧 捕获了分割目标包含的信息。这些帧被选为输入 SegGPT, SegGPT 利用其先进的上下文学习能力生成初始的分割掩码(mask)。然后, SegGPT 将其理解应用于整个视频序列的分割,如图 2,确保分割尊重医学影像的空间和时间连续性。通过将分割掩码与相应的视频帧组织成三维数据对,如图 3 所示,在保持原

始二维数据的完整性的同时,还增加了深度信息,增强了数据的体积连贯性。通过将这些分割与临床医 生的标注进行严格对比,确保了其准确性和可靠性。这一过程的最终结果是生成了一个高保真度的伪标 签数据集,这不仅有利于模型训练,还为正在进行的研究和临床实施增加了资源。通过这种方法,有效 地将未标注的医学影像数据转化为宝贵的资产,为模型训练扩充了数据集,并为未来的研究和临床应用 提供了丰富的信息源。



Figure 2. The methodological framework for generating pseudo-label data using SegGPT 图 2. 使用 SegGPT 生成伪标签数据的方法框架

2.2. USCL-Med3D: 基于对比学习增强的 LoRA 微调超声影像分割模型

2.2.1. SAM-Med3D

SAM-Med3D 模型是对 Segment Anything Model (SAM)的创新性改进,专为 3D 医学图像分割任务设计。该模型通过引入完全可学习的 3D 架构,克服了 SAM 原始 2D 架构在处理 3D 医学图像时的局限性,这对于捕捉 3D 图像中的复杂空间关系至关重要。SAM-Med3D 的架构可分为三个核心组件,每个组件都 针对 3D 数据处理进行了调整。

3D 图像编码器: SAM-Med3D 使用 3D 卷积层,该层设计用于有效嵌入体素补丁。结合可学习的 3D 绝对位置编码(PE),将 2D PE 概念扩展到三维,使模型能够编码图像的深度空间信息。编码后的嵌入随 后通过 3D 注意力块进行处理,这些注意力块使用 3D 相对 PE 来捕捉详细的空间特征。

3D 提示编码器: 该组件旨在解释 3D 医学图像中的稀疏(点、框)和密集(掩码)提示。稀疏提示使用 3D 位置编码表示,反映了数据的三维特性,而密集提示则通过 3D 卷积处理,生成对 3D 空间细节敏感的嵌入。

3D 掩码解码器: SAM-Med3D 中的 3D 掩码解码器结构从 SAM 的掩码解码器修改而来,使用 3D 转置卷积进行特征图上采样。该解码器中的变换层执行考虑 3D 空间信息的操作,确保生成的分割掩码与输入图像的 3D 上下文一致。解码器的输出随后通过逐点多层感知器(MLP)进行细化,产生最终的分割掩码。值得注意的是,SAM-Med3D 的训练数据中并不包含超声数据。同时,预训练中缺乏高级语义信息严重阻碍了 SAM-Med3D 应用于下游任务时的表现。

2.2.2. VCL-Head: 将上下文位置信息加入 SAM-Med3D

虽然 SAM-Med3D 在 3D 医学图像分割任务中表现出色。但是由于缺乏从样本中提取有效上下文信息的能力, SAM-Med3D 在 3D 超声影像数据上表现不佳。为了缓解 SAM-Med3D 缺乏提取上下文信息的能力和在 3D 超声影像数据集上泛化能力差的问题,本文提出 USCL-Med3D 模型, USCL-Med3D 的架构如图 4 所示。USCL-Med3D 模型由两个部分组成: 1) 使用 VCL-head 将 3D 超声影像所含的上下文信息



Figure 3. The video frames and segmentation masks are organized into three-dimensional data 图 3. 将视频帧和分割掩码组织成三维数据

嵌入 Image Encoder 与 Mask Decoder 的 LoRA 层。2) 微调 Image Encoder 和 Mask Decoder 的 LoRA 层, 对齐超声影像数据在 Image Encoder 和 Mask Decoder 中表达的特征。经过这两个部分的处理, USCL-Med3D 模型可以较好地提取超声影像的上下文信息,并在 3D 超声影像数据上获得更好的表现。

3D 超声影像中包含相对一致的上下文位置信息,即不同器官之间一致的几何关系。虽然 SAM-Med3D 获得了建模 3D 空间信息的能力,但 SAM-Med3D 的 3D 空间信息中仍然缺乏上下文位置信息,导致其在 3D 超声影像数据上的性能次优。为了学习三维超声数据中的上下文信息,实验设计中引入了 Voco 中的 对比学习方法,并针对此方法进行适配 SAM-Med3D 的修改,命名为 VCL-head。该方法通过利用三维医 学图像中不同器官之间相对一致的上下文位置信息,实现了在预训练阶段学习一致的语义表示。图 5 展 示了使用 VCL-head 提取三维超声影像中上下文信息的过程。具体而言,首先从输入体积的不同位置裁剪 出 n 个不重叠的基础体积(Base Volume),并增强这些基础体积之间的特征差异性,以此作为不同区域的 类别分配。随后,随机裁剪子体积(Crop Volume),并通过对不同基础体积的相似性进行对比,预测这些 子体积位于哪个区域,即预测不同子体积的上下文位置。这一过程通过一个预文本任务(Pretext Task)隐式 地将上下文位置先验编码到模型表示中,有效提升了任务性能。

VCL-head 包含预测分支和正则化分支。预测分支用于预测不同裁剪体积之间的上下文位置关系,而 正则化分支则用于增强不同基础体积之间的特征差异性,以学习更具辨别性的特征表示。最后,这些学 习的特征表示都被编码到 ImageEncoder 的 LoRA 层中。

对于一个输入 $x_i \in \mathbb{R}^{128\times 128\times 128}$,首先将其裁剪成 4 个不重叠的基础体积,如图 5 所示。将基础体积输入骨干网络,这里是 SAM-Med3D 的 ImageEncoder,输出编码特征 base^(x_i)。然后,按照之前的 SSL 工作 [36] [37],使用具有线性层的映射模块将 base^(x_i)投影到潜在特征 $b^{(x_i)}$ 中。然后,随机裁剪一个体积并将其 转换为高维特征空间作为 $c^{(x_i)}$ 。骨干网络和映射模块也用于从随机裁剪的体积中投影特征。然后,计算 $b^{(x_i)} n c^{(x_i)}$ 之间的相似度 logits, $l^{(x_i)}$,如下所示:

$$I_{i}^{(x_{i})} = \frac{\operatorname{CosSim}\left(f_{stu}\left(c^{(x_{i})}\right), f_{tea}\left(b_{i}^{(x_{i})}\right)\right) + \operatorname{CosSim}\left(f_{tea}\left(c^{(x_{i})}\right), f_{stu}\left(b_{i}^{(x_{i})}\right)\right)}{2}, \ i \in n.$$

$$(1)$$

上式中, $f_{stu}(x)$ 和 $f_{tea}(x)$ 表示将 x 输入骨干网络的学生和教师模型。特别的,在进行 $f_{tea}(x)$ 之前,会利 用学生模型先更新教师模型的参数,如下:

$$\rho_{\text{teacher}} = m \times \rho_{\text{teacher}} + (1 - m) \times \rho_{\text{student}}, \qquad (2)$$

超参数 *m* 取 0.9。得到 logits 之后, 计算预测分支损失 L_p 。首先计算 logits 和标签 *y* 之间的距离: $d_i^{(x_i)} = \left| y_i^{(x_i)} - l_i^{(x_i)} \right|, i \in n,$ (3)

|.|指的是计算绝对值,然后计算L_n:

$$L_{p}^{(x_{i})} = -\frac{1}{n} \sum_{i=n}^{n} \log\left(1 - d_{i}^{(x_{i})}\right).$$
(4)

从直观上看,不同的基础体积 *b*^(x_i)之间往往包含不同的器官,病理区域(语义差异)。因此,优化目标 是扩大不同基底之间的高维特征差异。为此,设计正则化损失 *L*,来放大不同基础体积 *b*^(x_i)之间的特征差 异,首先计算不同基础体积之间的相似度:

$$s_{ij} = \operatorname{CosSim}(b_i, b_j) = \frac{b_i \cdot b_j}{\|b_i\| \|b_j\|}, \ i, j \in n, i \neq j,$$
(5)

计算得到 s_{ii} 后, L_{reg} 以将 s_{ii} 优化至0为目标

$$L_{reg} = \frac{2}{n(n-1)} \sum_{i,j=n,i\neq j}^{n} \left| s_{ij} \right|.$$
(6)

通过这两种方法,VCL-head 不仅提高了模型对超声数据上下文信息的捕捉能力,而且显著增强了模型在后续任务中的语义理解能力。



Figure 4. The overall architecture of USCL-Med3D 图 4. USCL-Med3D 的整体架构

2.2.3. 使用 LoRA 微调 SAM-Med3D

ImageEncoder 和 MaskDecoder 作为 SAM-Med3D 中参数量最大的部分,在微调过程中对图像编码器 进行全局更新需要大量的计算开销。为了将医学领域的知识以较低的成本整合到图像编码器中,引入 LoRA 微调技术。具体来说,先冻结 ImageEncoder 的全部参数,作为 VCL-head 的 backbone,使用 LoRA 微调。这个步骤可以使网络学习在微调期间冻结原始图像编码器中的所有参数。考虑到 ImageEncoder 和 Mask Decoder 需要对 3D 超声数据进行编码和解码,需要为 ImageEncoder 和 MaskDecoder 中每个 Transformer 块的注意力层增加微调 LoRA 层,以此对齐编码器和解码器对特征的表达,如图 4 所示。

LoRA 使用编码器 - 解码器结构对权重更新施加低秩约束。它冻结了预训练的模型权重,并将小的可训练秩分解矩阵注入到 Transformer 体系结构的每一层。具体来说,给定一个预训练的权矩阵 $W_0 \in \mathbb{R}^{bxa}$, LoRA 增加了一对线性编码器 W_e , W_d ,即可训练的秩分解矩阵。 W_e 和 W_d 满足低阶约束, $W_e \in \mathbb{R}^{rxa}$, $W_d \in \mathbb{R}^{bxr}$, $r \ll \min(a,b)$ 。在 SAM-Med3D 中,注意力层的正向传递由 $h = W_0 x$ 变为:

$$h = W_0 x + W_d W_e x \tag{7}$$

通过 LoRA 微调,在较好地保留 SAM-Med3D 原有的图像编码,提示编码和掩膜解码能力的同时引入表达 3D 超声影像数据特征的能力,实验设计方法在 3D 超声数据集上获得了优秀的效果。



Figure 5. The VCL-head is utilized to extract contextual information from three-dimensional ultrasound data 图 5. 使用 VCL-head 提取三维超声数据的上下文信息

3. 实验

3.1. 对比实验

在构建专注于超声影像分析的数据集过程中,积累并精心挑选了具有高度针对性的数据集。本数据 集专注于体积超声图像,总计囊括了来自 39 名患者的 500 个体积超声扫描,这些患者均表现出不同程度 的积液病症。为了确保数据的全面性和代表性,对样本来源进行了细致划分: 25 名患者的数据集中于胸 腔积液的观测,通过医师沿胸腔冠状面,握持超声探头平稳扫描获取;而另外 14 名患者则提供包含腹腔 积液及腹腔器官的超声影像,这些影像不仅记录了积液的状况,还额外覆盖了肝、双肾及脾脏这三种关 键解剖结构,从而为研究增添了丰富的病理与解剖学信息维度。图 1 展示了医师采集和标注的这几种数 据。这些特定病理状态下的超声影像组成数据集,有助于训练模型为诊断和治疗提供更准确的信息。

表1给出了 FastSAM3D、SAM-Med3D 和 USCL-Med3D 在三维超声影像数据验证集上的性能指标。

实验表明,利用三维超声影像数据对 SAM-Med3D 进行微调的 USCL-Med3D 在超声数据上的表现明显优于 SAM-Med3D。随着提示点数量的增加,USCL-Med3D 在各类别数据上的表现越来越优于 SAM-Med3D, 当提示点数量为 10 时,USCL-Med3D 在各类别数据上平均领先 SAM-Med3D 方法 23.14% Dice 分数。至 关重要的是,USCL-Med3D 在不同的提示点数上始终超过 FastSAM3D 和 SAM-Med3D。这在一定程度上证明,USCL-Med3D 学习了三维超声影像数据中由医生扫查动作形成的时间序列特征和患者身体内部的空间特征,使 USCL-Med3D 在三维超声数据集上获得了优秀的表现。

Method	Points	Pleural effusion	Seroperitoneum	Liver	Spleen	Kidney
	1pt	12.19	11.54	13.73	12.51	14.42
	3pt	14.22	13.17	15.71	16.04	16.53
FastSAM5D	5pt	17.38	17.52	18.15	17.28	20.42
	10pt	23.31	22.62	24.15	23.04	25.72
	1pt	13.85	14.54	16.5	16.85	17.41
CAM M- 12D	3pt	19.3	21.54	22.04	22.66	25.72
SAM-Med3D	5pt	32.74	33.28	39.64	42.51	46.31
	10pt	38.24	37.41	54.73	47.58	54.68
	1pt	21.03	24.97	30.85	25.46	26.31
Ours	3pt	23.17	26.43	32.94	27.89	29.53
	5pt	66.22	63.58	71.02	63.38	68.29
	10pt	70.28	66.08	74.2	66.39	71.41

Table 1. A comparative analysis of the proposed methodology with other approaches on the 3D ultrasound dataset 表 1. 设计方法与其他方法在 3D 超声数据集上的对比

为了验证 USCL-Med3D 在跨域数据集上的泛化性,使用 AMOSdataset [13]等公开数据集进行测试, 这些公开数据集的测试集中不包含超声影像数据,实验结果如表 2 所示。在提示点较少时,USCL-Med3D 优于其他 2D 方法和 FastSAM3D,但落后于 SAM-Med3D,但随着提示点数量增加,USCL-Med3D 逐渐 缩小了与 SAM-Med3D 的差距,在提示点数量为 10 时,USCL-Med3D 在 AMOSdataset 上的表现超过了 SAM-Med3D,这可能是由于提示点数量增加后,提示点有更大的概率较为均匀地分布在三维超声影像数 据在训练时划分的上下文区域中,使 ImageEncoder 能更好地表达三维超声影像数据的上下文关系,从而 获得了较好的表现。

Table 2. Comparison of the proposed method with other methods on the 3D u dataset (CT, MRI) 表 2. 所提方法与其他方法在三维影像数据集(CT, MRI)上的对比

Method	AMOS				TotalSegmentator			BraTS				
	1 pt	3 pt	5 pt	10 pt	1 pt	3 pt	5 pt	10 pt	1 pt	3 pt	5 pt	10 pt
SAM	4.9	9.3	11.4	14.5	20.2	27.9	31.1	34.8	10.8	19.2	21.7	23.7
MobileSAM	4.1	5.6	6.3	7	14.9	17	18.2	21.2	7.9	13.2	15.6	18.6
TinySAM	4.9	7.7	8.9	10.1	17.1	22.5	24.3	26.2	10.3	16.5	18.7	21.2
SAM-Med2D	9.7	12.7	12.9	13.2	8.3	8.1	10	12.8	10.3	7.6	8.2	8.4

张雨萌,	李逸凡
------	-----

续表												
SAM-Med3D	31.9	69.4	71.1	74.2	29.3	70.8	73.6	75.9	35.8	70.3	72.8	75.5
FastSAM3D	27.3	36.8	40.2	43.7	25	37.8	44.5	51.9	33.3	40.1	42.1	44.5
Ours	29	67.7	70.3	71.6	27.2	69.1	72.4	73.8	33.1	68.6	70.3	72.7

3.2. 可视化

USCL-Med3D 和 SAM-Med3D 在三维超声影像数据上对各分类目标进行分割的可视化结果如图 6 所示,进一步展示了 USCL-Med3D 在三维超声影像上的优秀表现。可以看出,SAM-Med3D 只对脾脏有一个粗略的分割,但难以提供较好的分割结果。相比之下,USCL-Med3D 能够较为准确的覆盖目标区域。



Figure 6. A visual comparison between the proposed method and SAM-Med3D on the three-dimensional ultrasound imaging dataset 图 6. 提出方法和 SAM-Med3D 在三维超声影像数据集上的可视化对比

3.3. 消融实验

为了验证 USCL-Med3D 模型中不同模块和不同参数设置的表现效果,在三维超声影像数据集上进行 消融研究。

训练方法。USCL-Med3D 模型包含两个最重要的训练策略: VCL-head 上下文信息提取和 SAM-Med3D 架构微调。表 5 展示了使用 VCL-head 微调 3D Image Encoder 并嵌入三维超声数据上下文信息后,将 3D Image Encoder 并入 SAM-Med3D 架构,和 3D Mask Decoder 一起微调,并在三维超声影像数据集上进行

测试的结果和仅在 SAM-Med3D 架构中对 3D Image Encoder 和 3D Mask Decoder 进行微调的结果。实验 结果表明,使用 VCL-head 架构对 3D Image Encoder 进行为微调相比仅使用 SAM-Med3D 架构进行微调, 在三维超声影像数据集上的测试效果提升了 3.16%。这在一定程度上证明 VCL-head 架构帮助 3D Image Encoder 获取了三维超声影像中的上下文信息,提升了模型的表现。

 Table 3. The impact of fine-tuning the ImageEncoder alone versus fine-tuning both the ImageEncoder and Mask Decoder on the proposed method

Fine tuning Module	Using VCL-head	Dice
ImagaEncodar	\checkmark	64.13
IntageEncoder	Х	62.47
	\checkmark	60.72
Mask Decoder	Х	61.29
	\checkmark	69.67
ImageEncoder + Mask Decoder	Х	66.51

表 3. 微调 ImageEncoder/和微调 ImageEncoder + Mask Decoder 对设计方法的影响

微调模型组件。SAM-Med3D 模型的两个重要组成部分是 3D Image Encoder 和 3D Mask Decoder。本 节探讨这两个模块的贡献。表 3 中展示了在训练后半段,即特征对齐阶段,仅微调 3D Image Encoder,仅 微调 3D Mask Decoder 和同时微调两种模块的实验结果。由此可以观察到,仅微调 3D Image Encoder,可 能使 3D Mask Decoder 无法对 3D Image Encoder 编码的特征进行正确解码,原因可能是 3D Image Encoder 在 VCL-head 训练阶段重新微调并注入了三维超声影像数据的上下文关系特征,需要重新微调 3D Mask Decoder 使该模块能对编码特征进行正确解码。而仅微调 3D Mask Decoder,模型性能较差,推测是由于 训练目标不一致,导致由 VCL-head 训练得到的 3D Image Encoder 缺失了对三维超声影像进行精确分割 的能力。综上所述,需要同时微调 3D Image Encoder 和 3D Mask Decoder 来对齐三维超声影像数据特征

 Table 4. The impact of different fine-tuning methods on the model's performance across four datasets

 表 4. 不同微调方法对模型在四个数据集上表现的影响。

Tuning Method	3D US	AMOS	TotalSegmentator	BraTS
Baseline	46.53	74.2	75.9	75.5
LoRA (ImageEncoder + Mask Decoder) on 3DUS	69.67	71.6	73.8	72.7
Full-tuning on 3DUS	88.63	56.1	58.2	57.8
AdaptFormer [38]	61.94	47.9	50.3	52.4

微调方法。表 4 展示了选择不同微调方法对模型在三维超声数据集,AMOS,TotalSegmentator和 BraTS 数据集上表现的影响。对基线模型在三维超声数据集上进行全微调,获得了最好的表现,Dice分数达到了 88.63%。但全微调严重影响了模型的泛化性能,导致模型在公开三维医学影像(MRI,CT)数据集上的表现退化严重。这是由于仅使用三维超声数据全微调基线模型,导致基线模型原有的从其他领域数据学习的特征信息和通用知识混乱,导致模型虽然提升了在特定领域数据的表现,但削弱了模型在其他领域的泛化性能。使用适配器 AdaptFormer 进行微调,使基线模型在三维超声数据集上的表现相比原模 型提高了 15.41%,但在公开数据集上的表现相比原模型退化严重。AdaptFormer 简单的结构虽然在一般 任务(图像分类,文本分类)上表现良好,但对于 3D 医学影像分割这种复杂任务,AdaptFormer 简单的结 构无法显式的建模三维超声数据中的特征信息和上下文关系,导致模型在三维超声数据上的表现不够理 想,并影响了微调后模型在其他公开医学影像数据集上的泛化性能。使用 LoRA 模块微调基线模型的图 像编码器与掩膜解码器,引入表达 3D 超声影像数据特征和上下文关系信息的能力,极大地提升了模型 在三维超声数据集领域的表现,在三维超声影像数据集上相较基线模型 Dice 分数提升 23.14%,并较好 的保留了基线模型中原有的图像编码能力,使模型在其他公开的三维医学影像数据集上也表现出良好的 泛化性能。

LoRA rank	LoRA tuning Module	Dice
	ImageEncoder	54.85
8	Mask Decoder	53.39
	ImageEncoder + Mask Decoder	65.04
	ImageEncoder	64.47
16	Mask Decoder	60.14
	ImageEncoder + Mask Decoder	67.81
	ImageEncoder	64.13
32	Mask Decoder	60.72
	ImageEncoder + Mask Decoder	69.67
	ImageEncoder	56.38
64	Mask Decoder	54.72
	ImageEncoder + Mask Decoder	69.13

Table 5. The performance o	f the proposed method on the 3D ultrasound dataset under different LoRA ranks
表 5. 不同 LoRA rank 下,	设计方法在 3D 超声数据集上的效果

超参数 LoRA rank 的选择对模型性能有显著影响。较低的 rank 可能导致模型无法充分捕捉 3D 超声影像中的复杂特征,而较高的 rank则可能引入过多的噪声,导致模型过拟合。为了确定最佳的 LoRA rank,我们进行了多次实验,测试了在微调不同模块(ImageEncoder, Mask Decoder, ImageEncoder + Mask Decoder)的情况下,使用不同的 rank 值(8, 16, 32, 64)的微调效果。实验结果表明,当微调模块为 ImageEncoder + Mask Decoder,且 rank 为 32 时,模型在 3D 超声数据集上的表现最佳(见表 5)。具体来说,当 rank = 32 时,对含有丰富 3D 超声影像特征信息的和上下文关系的编码器与解码器进行微调,Dice 分数达到了 69.67%。这说明设定 LoRA rank = 32 可以较好地帮助模型捕捉 3D 超声影像数据中多且复杂的特征关系。在 LoRA rank = 32 的参数条件下纵向对比,同等条件下仅对 ImageEncoder 或 Mask Decoder 微调,由于可训练权重的减少,LoRA 模块向模型提取的少量特征关系中引入了冗余噪声,导致模型性能下降。与其他 LoRA rank 值横向对比,在仅微调 ImageEncoder 或 Mask Decoder 时,rank = 16 达到了甚至超越了 rank = 32 的实验效果; rank = 8 由于取值过低,导致 LoRA 模块难以有效微调基础模型中的三个目标模块,但仍然获得了接近 rank = 64 的微调效果;而 rank = 64 由于过大的取值,导致 LoRA 模块向模型中引入了更多的噪声,使得模型性能退化到了 rank = 8 的水平。在微调 ImageEncoder + Mask Decoder 模块时,可以观察到模型表现随着 LoRA rank 值的增加而提高,直到 rank = 64 时,模型表现退化。在微调模型组件

的选择中,讨论了微调 ImageEncoder 和 Mask Decoder 对实验效果的影响和原因。上述实验充分证明了 rank = 32 在特征编码和上下文信息提取之间取得了良好的平衡。

训练效率。使用 LoRA 微调方法,极大地提升了模型的训练效率和 GPU 显存利用率。LoRA 微调在 提升基线模型在三维超声影像数据集上表现的同时,较好地保留了模型原有的泛化能力。设置 epoch = 100 的情况下,平均训练时长 2.7 个小时。全微调参数量不变,平均需要训练 38.2 小时。AdaptFormer 训 练参数量相比全微调参数量减少,但由于需要在每个 Transformer 模块中插入包含两个线性层的 Adapter 适配器,其训练参数量仍远大于 LoRA 微调方法。同时,在 Batch = 2 的情况下,LoRA 微调训练仅占用 9.6G 显存,远小于全微调 68.1G 显存占用。LoRA 微调方法带来的高效训练和资源节约大幅减少了训练 参数量和计算开销,降低了内存占用和训练时间,特别适合资源有限的环境。LoRA 微调在保持基线模型 泛化性能的同时,能够快速适应新任务,提升了模型的灵活性和可扩展性。

4. 总结

本研究提出了 USCL-Med3D 模型,这是一种结合对比学习和 SAM-Med3D 模型的 LoRA 微调的新方法,用于 3D 超声图像分割。通过增强现有模型捕捉 3D 超声图像中固有的上下文信息的能力,并通过半监督伪标签技术减少手动标注的负担,从而缓解了现有模型的局限性。

实验结果表明, USCL-Med3D 在各种 3D 超声数据集上表现优越, 达到了更高的分割准确性和效率。 对比学习的引入显著增强了模型的特征提取能力, 而 LoRA 微调则有效利用了 SAM-Med3D 的广泛预训 练知识, 较好地适应了 3D 超声成像的特定挑战。

USCL-head 的优秀效果展现了将先进的机器学习技术与医学成像相结合的潜力。未来的探索可以集中在改进半监督学习框架和扩展模型在其他类型医学成像数据中的适用性。

总结,USCL-Med3D 在医学图像分割领域,特别是针对 3D 超声图像方面,表现优异。通过缓解关键挑战并利用最先进的技术,为提高医学实践中的诊断准确性和效率做出了贡献。

致 谢

在本论文的完成过程中,我们得到了许多人的帮助和支持,在此向他们表示最诚挚的感谢。

我们要感谢实验室的各位同学和同事。他们在我们遇到困难时给予了热情的帮助,并与我们分享了 他们的经验和见解。特别感谢李逸凡同学在实验设计和数据分析方面提供的帮助,以及在论文修改过程 中提出的宝贵意见。

此外,我们要感谢我的家人和朋友。他们在我们攻读学位期间给予了我无条件的支持和鼓励,使我 们能够专注于研究工作。他们的理解和关爱是我们不断前进的动力。

谨以此致谢,表达对所有帮助过我们的人的感激之情。

参考文献

- 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, 5-9 October 2015, 234-241. <u>https://doi.org/10.1007/978-3-319-24574-4_28</u>
- [2] Ke, J., Lu, Y., Shen, Y., Zhu, J., Zhou, Y., Huang, J., et al. (2023) ClusterSeg: A Crowd Cluster Pinpointed Nucleus Segmentation Framework with Cross-Modality Datasets. *Medical Image Analysis*, 85, Article 102758. https://doi.org/10.1016/j.media.2023.102758
- [3] Gao, H., Li, Y., Long, K., *et al.* (2024) A Survey for Foundation Models in Autonomous Driving. arXiv: 2402.01105. https://doi.org/10.48550/arXiv.2402.01105
- [4] Amrehn, M., Gaube, S., Unberath, M., et al. (2017) UI-Net: Interactive Artificial Neural Networks for Iterative Image Segmentation Based on a User Model. arXiv: 1709.03450. <u>https://doi.org/10.48550/arXiv.1709.03450</u>

- [5] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., et al. (2023) Segment Anything. 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, 1-6 October 2023, 3992-4003. <u>https://doi.org/10.1109/iccv51070.2023.00371</u>
- [6] Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2020) An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv: 2010.11929. <u>https://doi.org/10.48550/arXiv.2010.11929</u>
- [7] Zhang, Y., Shen, Z. and Jiao, R. (2024) Segment Anything Model for Medical Image Segmentation: Current Applications and Future Directions. *Computers in Biology and Medicine*, **171**, Article 108238. https://doi.org/10.1016/j.compbiomed.2024.108238
- [8] Ma, J., He, Y., Li, F., Han, L., You, C. and Wang, B. (2024) Segment Anything in Medical Images. *Nature Communications*, 15, Article No. 654. <u>https://doi.org/10.1038/s41467-024-44824-z</u>
- [9] Cheng, J., Ye, J., Deng, Z., et al. (2023) SAM-Med2d. arXiv: 2308.16184. <u>https://doi.org/10.48550/arXiv.2308.16184</u>
- [10] Mazurowski, M.A., Dong, H., Gu, H., Yang, J., Konz, N. and Zhang, Y. (2023) Segment Anything Model for Medical Image Analysis: An Experimental Study. *Medical Image Analysis*, 89, Article 102918. https://doi.org/10.1016/j.media.2023.102918
- [11] Wang, H., Guo, S., Ye, J., et al. (2023) SAM-Med3D. arXiv: 2310.15161. https://doi.org/10.48550/arXiv.2310.15161
- [12] Wu, L., Zhuang, J. and Chen, H. (2024) VoCo: A Simple-Yet-Effective Volume Contrastive Learning Framework for 3D Medical Image Analysis. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 22873-22882. <u>https://doi.org/10.1109/cvpr52733.2024.02158</u>
- [13] Achiam, J., Adler, S., Agarwal, S., et al. (2023) GPT-4 Technical Report. arXiv: 2303.08774. <u>https://doi.org/10.48550/arXiv.2303.08774</u>
- [14] Wang, X., Zhang, X., Cao, Y., Wang, W., Shen, C. and Huang, T. (2023) SegGPT: Towards Segmenting Everything in Context. 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, 1-6 October 2023, 1130-1140. <u>https://doi.org/10.1109/iccv51070.2023.00110</u>
- [15] Radford, A., Kim, J.W., Hallacy, C., et al. (2021) Learning Transferable Visual Models from Natural Language Supervision. Proceedings of the 38th International Conference on Machine Learning, Virtual, 18-24 July 2021, 8748-8763.
- [16] Oquab, M., Darcet, T., Moutakanni, T., et al. (2023) DiNOv2: Learning Robust Visual Features without Supervision. arXiv: 2304.07193. <u>https://doi.org/10.48550/arXiv.2304.07193</u>
- [17] Zou, X., Yang, J., Zhang, H., et al. (2024) Segment Everything Everywhere All at Once. Proceedings of the 37th International Conference on Neural Information Processing System, New Orleans, 10-16 December 2023, 19769-19782.
- [18] Betker, J., Goh, G., Jing, L., et al. (2023) Improving Image Generation with Better Captions. Computer Science. https://cdn.openai.com/papers/dall-e-3.pdf
- [19] Chen, T., Zhu, L., Ding, C., Cao, R., Wang, Y., Zhang, S., et al. (2023) SAM-Adapter: Adapting Segment Anything in Underperformed Scenes. 2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Paris, 2-6 October 2023, 3359-3367. <u>https://doi.org/10.1109/iccvw60793.2023.00361</u>
- [20] Wu, J., Ji, W., Liu, Y., et al. (2023) Medical SAM Adapter: Adapting Segment Anything Model for Medical Image Segmentation. arXiv: 2304.12620. <u>https://doi.org/10.48550/arXiv.2304.12620</u>
- [21] Gong, S., Zhong, Y., Ma, W., Li, J., Wang, Z., Zhang, J., et al. (2024) 3DSAM-Adapter: Holistic Adaptation of SAM from 2D to 3D for Promptable Tumor Segmentation. *Medical Image Analysis*, 98, Article 103324. https://doi.org/10.1016/j.media.2024.103324
- [22] Chen, C., Miao, J., Wu, D., Zhong, A., Yan, Z., Kim, S., et al. (2024) MA-SAM: Modality-Agnostic SAM Adaptation for 3D Medical Image Segmentation. *Medical Image Analysis*, 98, Article 103310. https://doi.org/10.1016/j.media.2024.103310
- [23] He, K., Fan, H., Wu, Y., Xie, S. and Girshick, R. (2020) Momentum Contrast for Unsupervised Visual Representation Learning. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, 13-19 June 2020, 9726-9735. <u>https://doi.org/10.1109/cvpr42600.2020.00975</u>
- [24] He, Y., Yang, G., Ge, R., Chen, Y., Coatrieux, J., Wang, B., et al. (2023) Geometric Visual Similarity Learning in 3D Medical Image Self-Supervised Pre-Training. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, 17-24 June 2023, 9538-9547. <u>https://doi.org/10.1109/cvpr52729.2023.00920</u>
- [25] 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. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, 18-24 June 2022, 20698-20708. <u>https://doi.org/10.1109/cvpr52688.2022.02007</u>
- [26] Du, H., Dong, Q., Xu, Y. and Liao, J. (2023) Weakly-Supervised 3D Medical Image Segmentation Using Geometric Prior and Contrastive Similarity. *IEEE Transactions on Medical Imaging*, 42, 2936-2947. https://doi.org/10.1109/tmi.2023.3269523

- [27] Cui, J., Zhong, Z., Tian, Z., et al. (2023) Generalized Parametric Contrastive Learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46, 7463-7474. <u>https://doi.org/10.1109/TPAMI.2023.3278694</u>
- [28] Caron, M., Touvron, H., Misra, I., Jegou, H., Mairal, J., Bojanowski, P., et al. (2021) Emerging Properties in Self-Supervised Vision Transformers. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, 10-17 October 2021, 9630-9640. <u>https://doi.org/10.1109/iccv48922.2021.00951</u>
- [29] Taleb, A., Loetzsch, W., Danz, N., *et al.* (2020) 3D Self-Supervised Methods for Medical Imaging. *Advances in Neural Information Processing Systems*, **33**, 18158-18172.
- [30] Zhou, H., Lu, C., Chen, C., Yang, S. and Yu, Y. (2023) A Unified Visual Information Preservation Framework for Self-Supervised Pre-Training in Medical Image Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45, 8020-8035. <u>https://doi.org/10.1109/tpami.2023.3234002</u>
- [31] Zhou, X., Gao, H., Xu, X., et al. (2022) PCRL: Priority Convention Reinforcement Learning for Microscopically Sequencable Multi-Agent Problems. 36th Conference on Neural Information Processing Systems (NeurIPS 2022), New Orleans, 28 November-9 December 2022.
- [32] Zhang, Z. and Gong, X. (2023) Positional Label for Self-Supervised Vision Transformer. Proceedings of the AAAI Conference on Artificial Intelligence, 37, 3516-3524. <u>https://doi.org/10.1609/aaai.v37i3.25461</u>
- [33] Tao, X., Li, Y., Zhou, W., Ma, K. and Zheng, Y. (2020) Revisiting Rubik's Cube: Self-Supervised Learning with Volume-Wise Transformation for 3D Medical Image Segmentation. *Medical Image Computing and Computer Assisted Intervention—MICCAI* 2020, Lima, 4-8 October 2020, 238-248. <u>https://doi.org/10.1007/978-3-030-59719-1_24</u>
- [34] He, K., Chen, X., Xie, S., Li, Y., Dollar, P. and Girshick, R. (2022) Masked Autoencoders Are Scalable Vision Learners. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, 18-24 June 2022, 15979-15988. <u>https://doi.org/10.1109/cvpr52688.2022.01553</u>
- [35] He, Z., Unberath, M., Ke, J. and Shen, Y. (2023) TransNuSeg: A Lightweight Multi-Task Transformer for Nuclei Segmentation. *Medical Image Computing and Computer Assisted Intervention—MICCAI* 2023, Vancouver, 8-12 October 2023, 206-215. <u>https://doi.org/10.1007/978-3-031-43901-8_20</u>
- [36] Chen, T., Kornblith, S., Norouzi, M., et al. (2020) A Simple Framework for Contrastive Learning of Visual Representations. International Conference on Machine Learning. PmLR, 1597-1607.
- [37] Chen, X. and He, K. (2021) Exploring Simple Siamese Representation Learning. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, 20-25 June 2021, 15745-15753. https://doi.org/10.1109/cvpr46437.2021.01549
- [38] Yang, N., Zhang, Y., Wang, Y., Tang, D., Li, Y. and Yuan, D. (2024) Adaptformer: An Adaptive Multimodal Deep Decomposition Approach for Power Consumption Forecasting. *Advanced Data Mining and Applications*, Sydney, 3-5 December 2024, 48-62. <u>https://doi.org/10.1007/978-981-96-0847-8_4</u>