

# 利用微调的大规模视觉模型进行乳腺超声图像分割

## 摘要

深度学习技术在医学图像分割方面具有巨大潜力，但通常需要为不同类型的数据集定制模型，这在临床应用中带来了不便。而像 SAM 这样的视觉大模型虽然在自然图像分割方面表现出强大的泛化能力，但在医学图像分割领域的表现却不尽如人意。SAMUS 是基于大规模图像分割模型 segmentation Anything model (SAM) 的医学图像分割研究新范式。SAMUS 继承了 SAM 的 ViT 图像编码器、提示编码器和掩码解码器，并对图像编码器进行了调整，加入了 cnn 分支图像编码器和跨分支关注模块以弥补 transformer 在提取局部信息上的不足，同时使用一个特征适配器和一个位置适配器来微调图像编码器。此外，我在此基础上融合图像的高频信息以提升分割性能。实验证明，SAMUS 在医学图像分割方面比原 SAM 模型有更好的性能。通过在乳腺超声图像数据集上的实验评估证明，SAMUS 在医学图像分割方面比原 SAM 模型有更好的分割性能。总而言之，各种基于 SAM 的微调模型给医学图像分割以新的启发，并有望在医学图像分割领域上发挥更大作用，成为医生和医疗研究人员的有力工具，提供更为精确和个性化的智能医疗辅助诊断，实现更为智能的医疗影像分析，为患者提供更优质的医疗服务。

**关键词：**医学图像分割；超声图像；大规模视觉模型；适配器

## 1 引言

分割是医学成像分析中的一项基本任务，它涉及识别和描绘各种医学图像 (如器官、病变和组织) 中的感兴趣区域 (ROI)，以便于计算机辅助诊断和智能临床手术。准确的分割可以帮助医生准确地识别和定位病理区域，对于许多临床应用至关重要，包括疾病诊断、治疗计划和疾病进展监测 [13]。由于计算能力和医疗数据资源的快速扩展，基于深度学习的医学图像分割在准确率和速度上都取得了巨大的进步，但仍存在一些可改进的地方。人工分割长期以来一直是描绘解剖结构和病理区域的黄金标准，但这一过程耗时，劳动密集，并且通常需要高度的专业知识。半自动或全自动分割方法可以显著减少所需的时间和人工，提高一致性，并能够分析大规模数据集。此外，由于医学图像的特点，如模式众多，组织和器官结构复杂，以及可用的注释数据很少，大多数基于深度学习的模型仅限于特定的模式、器官或病变 [4, 21]。这些模型通常是针对特定的分割任务设计和训练的，当应用于新任务或不同类型的成像数据时，它们的性能可能会显著下降。这种通用性的缺乏对这些模型在临床实践中的广泛应用构成了实质性的障碍。

如今，无论是在自然语言领域还是视觉领域，DALL-E [14]、GPT-4 [12]、SAM [7]、

SegGPT [18] 等大规模模型大量涌现, 使得研究者可以在一个统一的框架下解决多种问题。由于其显著的泛化能力, 这种大规模模型在工业用途上更有前景。在图像分割领域, 分割任意模型 (Segmentation Anything Model, SAM) [7] 作为一种强大而通用的视觉分割模型获得了广泛的关注, 它能够根据用户提示生成各种细粒度的分割掩码。利用其泛化性能, 无需部署单一的医学图像分割模型, 可以将医学图像分割的解决方案直接集成到此大型 CV 模型中。但最近的许多研究表明, 尽管它在自然图像上表现出色, 但它在医学图像分割上表现不佳。SAM 在医学图像上失败的主要原因是虽然 SAM 在训练中建立了一个复杂而高效的数据引擎, 但它收集的大多为自然图像, 缺乏专业的医学图像训练数据。由于自然图像和医学图像之间存在明显的领域差距, SAM 难以推广到多模态和多对象的医学数据集。医学图像是从某些协议和扫描仪收集的, 并且由于其特定的临床目的而以不同的模式 (电子、激光、x 射线、超声波、核物理和磁共振) 呈现。因此, 医学图像是基于一系列基于物理的属性和能量来源, 与自然图像在像素强度、颜色、纹理等分布特征上存在显著差异。因此, SAM 不能直接应用于医学领域的局限性是意料之中的。

目前的研究主要集中在对特定医学分割数据集的 SAM 进行视觉调优。一般来说, 最近的视觉调优方法可以分为五大类, 包括微调、参数调优、重新映射调优、提示调优和自适应调优。具体来说, 微调方法既包括调整预训练模型的整个参数集, 也包括对预训练模型的特定部分进行选择性地微调 [10]。参数调优方法直接修改模型参数的权重或偏差 [9]。重新映射方法通过知识蒸馏、基于权重的重新映射或基于体系结构的重新映射, 将学习到的信息从预训练模型转移到下游模型 [23]。提示调优通过将一组可学习参数与输入结合或设计一个子网络来生成视觉提示来引入下游任务的知识 [17]。适配器调优是最广泛采用的策略, 它通过将额外的可学习参数与固定的预训练模型结合起来, 促进下游任务的学习 [19]。SAMUS 的整体架构继承自 SAM, 保留了提示编码器和掩码解码器的结构和参数, 未做任何调整。相比之下, 对图像编码器进行了精心修改, 以解决局部特征不足和计算内存消耗过多的挑战, 使其更适合临床友好的分割。主要修改包括减小输入大小、重叠补丁嵌入、向 ViT 分支引入适配器、增加 CNN 分支、引入跨分支关注 (crossbranch attention, CBA)。由于 SAMUS 在调优过程中只添加和更新一小部分参数, 因此在实际使用中, 它的部署成本和存储成本非常小。

## 2 相关工作

### 2.1 医学分割模型

早期的医学图像分割方法主要采用显式轮廓特征或马尔可夫随机场来获得较好的分割效果。随着深度学习的快速发展, U-Net [15] 作为开创性的工作, 开启了医学图像分割的新时代。继 U-Net 之后, 研究人员设计了多种变体来提高医学图像分割的性能, 包括 Res-UNet [11]、Dense-UNet [8]、U-Net++ [24]、3D-Unet [25] 等。近年来, Transformer [16] 在自然图像中的应用也引发了医学图像分割的研究, 尤其是 Vision Transformer [5] 的出现深

刻影响着视觉分割领域的发展。主流的网络设计策略是将 Transformer 模块集成到 U-Net 框架中, 包括 TransUnet [3]、SwinUnet [2]、Hi-Former [6]、DAE-Former [1] 等。与基于 cnn 的方法相比, Transformer 块的集成使得可以全局关注和聚合相似特征, 从而在医学图像分割方面取得了前所未有的性能。这些基于 U-Net 和 Transformer 的医学图像分割方法取得了显著的性能提升, 但仍然存在一些挑战和问题。首先, 这些方法通常需要大量的标注数据进行训练, 而医学图像的标注成本非常高昂。其次, 现有的方法往往忽略了医学图像中的上下文信息, 导致分割结果不够准确。此外, 这些方法在处理不同类型的医学图像时, 如 CT、MRI 和 X 光图像, 可能需要重新设计和调整网络结构, 这增加了模型的复杂性和计算成本。

## 2.2 CV 大模型

在网络规模的数据集上预训练的大型语言模型具有强大的泛化能力, 这些“基础模型”[8] 的使用可以推广到训练期间看到数据分布之外的任务。此功能通常通过提示工程实现, 其中使用手工编写的文本来提示语言模型为手头的任务生成有效的文本。受 ChatGPT 和 GPT4 等此类大型语言模型的启发, 研究人员开发了类似的大规模视觉模型, 这些模型能够通过预训练和微调范式快速适应和扩展到目标任务或领域。例如, 最近在 1B 掩模上预训练的 SAM 已经成为一种通用的图像分割大模型。SAM 展示了强大的泛化性能, 允许任何对象的交互式或自动分割。它首先定义一个提示的分割任务, 此任务需要一个支持灵活提示的模型, 并且可以在提示时实时输出分割掩码以允许交互使用。其次, 为了训练模型, 其构建了一个“数据引擎”, 在使用高效模型来协助数据收集和使用新收集的数据来改进模型之间进行迭代, 由此创建了含 1B 高质量的分割掩码的数据集。最终模型由三个组件构成, 包括一个图像编码器, 一个灵活的提示编码器和一个快速掩码解码器。SAM 可以根据不同的用户提示, 包括点、边界框和粗掩码, 分割相应的对象, 适应各种分割应用程序。虽然 SAM 具有惊人的泛化能力, 但由于缺乏专业数据, 它们无法理解和提供目标组织的精确解剖结构, 尚未针对医学图像分割领域进行明确优化。

## 2.3 适配器

针对 SAM 在医学图像分割上表现差的问题, 已经提出了几种方法来使用有限的下游医学数据集来调整 SAM 并使其有效地适应医学图像领域。MedSAM 通过冻结图像编码器和提示编码器, 在大量医学图像上训练 SAM, 重点是调整 SAM 的掩模解码器 [10]。MSA 在 ViT 图像编码器的每个变压器层上采用两个上下转换器来引入特定于任务的信息 [20]。SAMed 在图像编码器上应用了低秩 (low-rank-based, LoRA) 策略, 以更低的计算成本对 SAM 进行调优, 使其更适用于医学图像分割 [22]。SAMUS 为了将训练好的 SAM 图像编码器 (即 ViT 分支) 推广到更小的输入尺寸和医学图像域, 引入了一个位置适配器和五个特征适配器。位置适配器负责调整位置嵌入以匹配嵌入序列的分辨率。特征适配器使 ViT 分

支能够更好地适应医学图像域的特征分布。这些适配器可以有效地调优 ViT 分支，同时只需要更少的参数，以便方便且准确地部署于临床场景中。

## 2.4 图像高频信息

# 3 本文方法

## 3.1 模型概述

SAMUS 的整体架构继承自 SAM，保留了提示编码器和掩码解码器的结构和参数，未做任何调整。相比之下，对图像编码器进行了精心修改，以解决局部特征不足和计算内存消耗过多的挑战，使其更适合临床友好的分割，如图1所示。

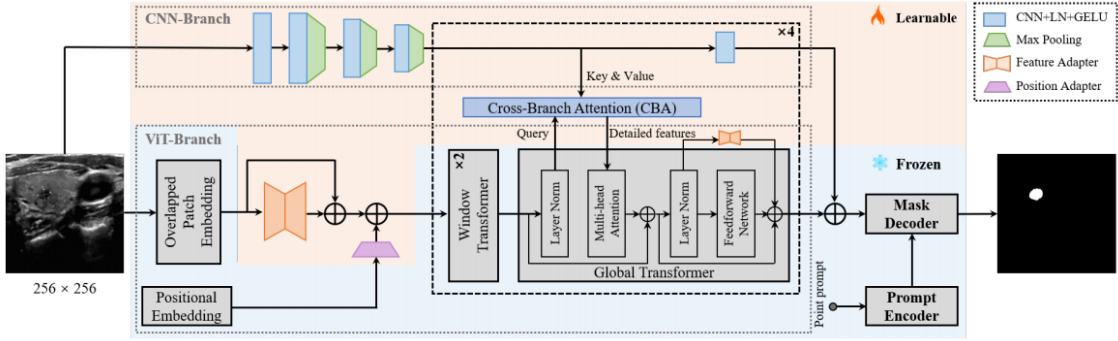


图 1. 模型框架示意图

其主要修改包括减小输入大小、重叠补丁嵌入、向 ViT 分支引入适配器、增加 CNN 分支、引入跨分支关注 (crossbranch attention, CBA)。具体来说，输入空间分辨率从  $1024 \times 1024$  像素缩小到  $256 \times 256$  像素，由于变压器的输入序列较短，从而大大降低了 GPU 内存成本。重叠的块嵌入与 SAM 中的块嵌入使用相同的参数，而其步幅是原始步幅的一半，很好地保留了块边界的信息。此外，SAMU 还在 ViT 分支中引入一个位置适配器和五个功能适配器。位置适配器是为了适应更短序列的全局位置嵌入，因为输入尺寸更小。第一特征适配器遵循重叠的补丁嵌入，将输入特征与预训练的 ViT 图像编码器的所需特征分布对齐。剩余的特征适配器连接到全局变压器中前馈网络的剩余连接，以微调预训练的图像编码器。CNN 分支与 ViT 分支并行，通过 CBA 模块向后者提供互补的局部信息，CBA 只集成到每个全局转换器中，以 ViT 分支的特征为查询，与 CNN 分支的特征建立全局依赖关系。最后，将两个分支的输出合并作为 SAMUS 的最终图像特征嵌入。

### 3.2 ViT 分支中的适配器

为了将训练好的 SAM 图像编码器推广到更小的输入尺寸和医学图像域，引入了一个位置适配器和五个特征适配器。具体来说，位置适配器负责调整位置嵌入以匹配嵌入序列的分辨率。它首先通过最大池化对位置嵌入进行下采样，步长和核大小为 2，从而获得与嵌入序列相同的分辨率。随后，应用核大小为  $3 \times 3$  的卷积运算来调整位置嵌入，进一步帮助 ViT 分支更好地处理较小的输入。所有功能适配器都具有相同的结构，该结构由三个组件组成：向下线性投影、激活函数和向上线性投影。每个特征适配器的程序可表述为：

$$A(x) = \mathcal{G}(xE_d)E_u, \quad (1)$$

式中， $\mathcal{G}$  表示 GELU 激活函数， $E_d \in \mathbb{R}^{d \times \frac{d}{4}}$ ， $E_u \in \mathbb{R}^{\frac{d}{4} \times d}$  为投影矩阵， $d$  为特征嵌入维数。通过这些简单的操作，特征适配器使 ViT 分支能够更好地适应医学图像域的特征分布。这些适配器可以有效地调优 ViT 分支，同时只需要更少的参数，部署成本和存储成本大大减小。

### 3.3 CNN 分支和跨分支关注模块

CNN 分支由顺序连接的卷积池块组成。具体来说，输入最初通过单个卷积块，然后通过三个卷积池块进行处理。然后，CNN 分支中的特征图与 ViT 分支的特征图具有相同的空间分辨率。在 CNN 分支的其余部分，这样的单个卷积块按顺序重复 4 次。

跨分支关注模块 (CBA) 则在 CNN 分支和 ViT 分支之间搭建了一座桥梁，进一步补充了 ViT 分支缺少的局部特征提取功能。对于来自 ViT 分支  $F_v$  和 CNN 分支的一对特征映射，单个头部的跨分支关注可以表示为：

$$\mathcal{F}(F_v, F_c) = (S(\frac{F_v E_q (F_c E_k)^T}{\sqrt{d_m}}) + R)(F_c E_v), \quad (2)$$

其中  $S$  表示 Softmax 函数。 $E_q \in \mathbb{R}^{d \times d_m}$ ， $E_k \in \mathbb{R}^{d \times d_m}$ ， $E_v \in \mathbb{R}^{d \times d_m}$  是可学习的权重矩阵，用于将  $F_c$  和  $F_v$  投影到不同的特征子空间。 $R \in \mathbb{R}^{hw \times hw}$  为相对位置嵌入， $d_m$  为 CBA 维数。CBA 的最终输出是  $g$  这种单头注意力的线性组合。

### 3.4 高频信息提取模块

二维图像本质上是二维离散非平稳信号，包含不同的频率范围和空间位置信息。小波变换可以在分解信息的同时有效地保留这些信息。具体以二维图像为例。我利用小波变换将原始图像分解为低频、水平高频、垂直高频和对角高频分量 (LL、HL、LH 和 HH)。分别保存原始图像的低频和不同高频信息。我用高频分量在不同方向上的和表示高频图像  $H$ 。 $H$  定义为：

$$H = HL + LH + HH. \quad (3)$$



为什么要用小波变换？与其他方法（如傅里叶变换）相比，小波变换是生成高频信息的有效方法。虽然高频信息噪声更多，但物体边界更清晰，这有助于模型更关注细节。如图2所示。

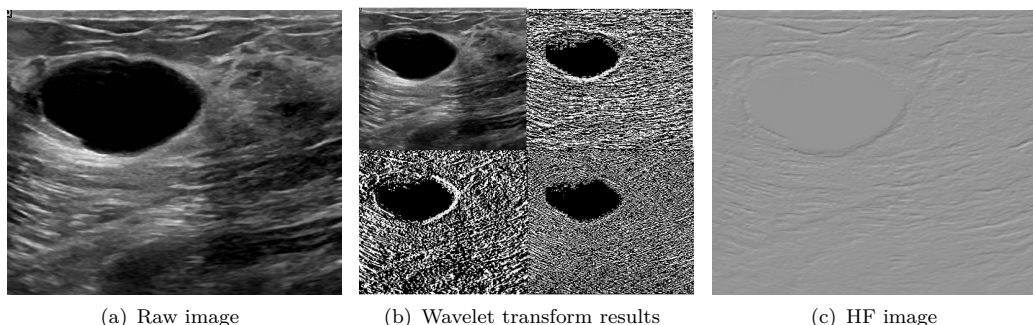


图 2. 以 benign(26) 为例，可视化 LF 和 HF 结果。(a) 原始图像。(b) 小波变换结果。(c) 高频图像。

## 4 复现细节

### 4.1 模型框架

该复现过程使用了 SAMUS 源论文中的模型的框架代码作为基础框架，在此基础上进行模块创新。

### 4.2 信息提取

主要的创新工作在于从高频组件中学习更局部细节信息，更关注病灶边界与纹理。同时，提出了多尺度注意门，并将其与 SAMUS 中的 CNN 分支相结合，它用于抑制不重要的特征，增强有价值的特征。高频信息具体提取过程见如下代码。

```
1 def wavelet(x):
2     LL, (LH, HL, HH) = pywt.dwt2(x, 'haar')
3
4     LH = (LH - LH.min()) / (LH.max() - LH.min()) * 255
5     HL = (HL - HL.min()) / (HL.max() - HL.min()) * 255
6     HH = (HH - HH.min()) / (HH.max() - HH.min()) * 255
7
8     merge = LH + HL + HH
9     merge = (merge - merge.min())
10            / (merge.max() - merge.min()) * 255
11     merge = cv2.resize(merge, [x.shape[1], x.shape[0]])
12
```

### 4.3 高频信息融合

涉及论文 idea 实现，暂不开源，请谅解。

## 5 实验结果分析

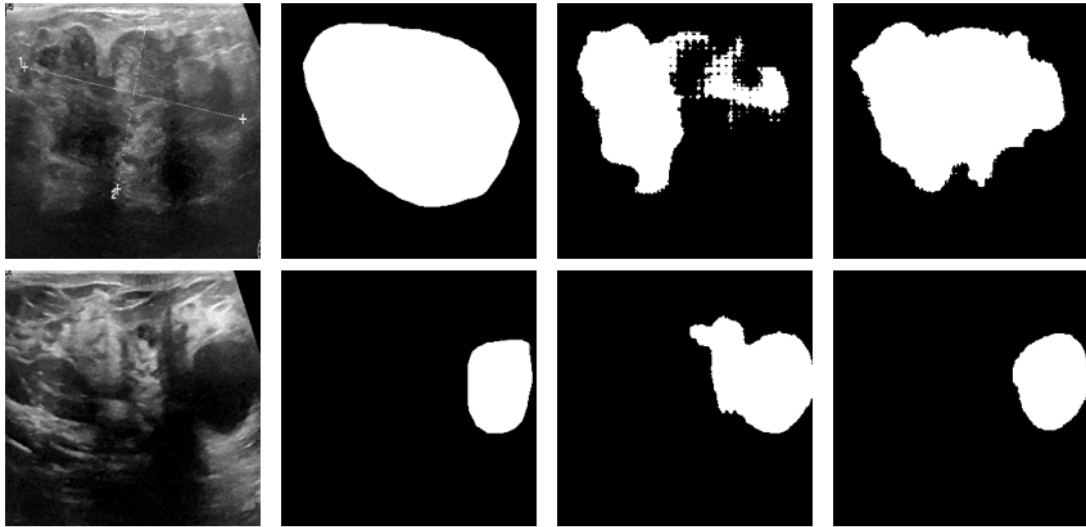
在本节中，我们将分析提出的 High Frequency Component 在医学超声图像分割中的有效性。

我们在 BUSI 数据集上进行消融实验。实验使用 Adam 优化器对网络进行优化。初始学习率设置为 0.0005，批量大小设置为 4，epoch 数设置为 400。使用 dice loss 和 mse loss 作为损失函数进行训练，在 400 轮训练过程中保存 loss 最小的模型用于测试，使用两个常见的分割评估指标，Dice 分数和 Hausdorff 距离进行测试评估。与 SAMUS 模型本身进行分割性能对比，结果如表1所示。

Components	BUSI	
	Dice	HD
×	85.56	27.00
√	86.41	24.25

表 1. 消融实验结果

从表中可以看出，我们的方法比 SAMUS 本身在 Dice 分数上高出约 0.85%，在霍斯多夫距离上有所下降。我们将成功归因于以下原因：高频信息的融合使模型更加关注超声图像形状边界等局部信息，针对超声图像对比度低和边界模糊的特点弥补了 Transformer 只能关注全局信息的不足。图3为数据集中良性与恶性乳腺肿瘤分割可视化结果。



(a) benign



(b) malignant

图 3. 可视化分割结果。(a) 良性。(b) 恶性。

## 6 总结与展望

深度学习方法在医学图像分割中具有巨大的潜力，然而，这些模型通常是不同类型的医学图像量身定制的，遇到不同数据集时一般需要调整模型结构以适应数据集特点，这给临床使用带来很大不便。而 SAM 这一类视觉大模型在自然图像分割上展现出了惊人的泛化能力，但由于模型缺乏医学专业知识，在医学图像分割领域表现不佳。SAMUS 以大规模图像



分割模型 segmentation Anything model (SAM) 为基础, 探索针对医学图像分割的定制化大规模模型的研究新范式。其在 SAM 的基础上加入了 cnn 分支图像编码器和跨分支关注模块以弥补 Transformer 在提取局部信息上的不足, 同时使用特征适配器和位置适配器来微调图像编码器以适应下游任务。SAMUS 在实验中展现出比原 SAM 模型更优越的性能, 证明了其在医学图像分割方面的有效性。另外, 我在 SAMUS 的基础上尝试融合超声图像的高频信息, 使其更关注关注图像的高频 (HF) 信息来提高性能。

未来, 随着技术的进步和数据集的不断丰富, 基于 SAM 的各种微调模型将会大量涌现, 基于大模型的分割模型有望在医学图像分割领域发挥更大的作用, 成为医生和医疗研究人员强有力的工具, 为患者提供更优质的医疗服务。例如, 在手术规划、肿瘤检测、器官定位等临床应用中, 可以提供更加精确和个性化的辅助, 帮助医生进行更准确的诊断和治疗。此外, 可能进一步与临床决策支持系统相结合, 实现更为智能的医疗影像分析, 从而推动个性化医疗和精准医疗的发展。

## 参考文献

- [1] Reza Azad, Ren'e Arimond, Ehsan Khodapanah Aghdam, Amirhosein Kazerouni, and Dorit Merhof. Dae-former: Dual attention-guided efficient transformer for medical image segmentation. ArXiv, abs/2212.13504, 2022.
- [2] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. Swin-unet: Unet-like pure transformer for medical image segmentation. In ECCV Workshops, 2021.
- [3] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan Loddon Yuille, and Yuyin Zhou. Transunet: Transformers make strong encoders for medical image segmentation. ArXiv, abs/2102.04306, 2021.
- [4] Junlong Cheng, Chengrui Gao, Fengjie Wang, and Min Zhu. Segnetr: Rethinking the local-global interactions and skip connections in u-shaped networks. In International Conference on Medical Image Computing and Computer-Assisted Intervention, 2023.
- [5] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. ArXiv, abs/2010.11929, 2020.
- [6] Moein Heidari, Amirhossein Kazerouni, Milad Soltany Kadarvish, Reza Azad, Ehsan Khodapanah Aghdam, Julien Cohen-Adad, and Dorit Merhof. Hiformer: Hierarchical multi-scale representations using transformers for medical image segmentation.

- 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pages 6191–6201, 2022.
- [7] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross B. Girshick. Segment anything. ArXiv, abs/2304.02643, 2023.
  - [8] X. Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, and Pheng-Ann Heng. H-denseunet: Hybrid densely connected unet for liver and tumor segmentation from ct volumes. IEEE Transactions on Medical Imaging, 37:2663–2674, 2018.
  - [9] Dongze Lian, Daquan Zhou, Jiashi Feng, and Xinchao Wang. Scaling & shifting your features: A new baseline for efficient model tuning. ArXiv, abs/2210.08823, 2022.
  - [10] Jun Ma and Bo Wang. Segment anything in medical images. ArXiv, abs/2304.12306, 2023.
  - [11] Xiao nan Xiao, Sheng Lian, Zhiming Luo, and Shaozi Li. Weighted res-unet for high-quality retina vessel segmentation. 2018 9th International Conference on Information Technology in Medicine and Education (ITME), pages 327–331, 2018.
  - [12] OpenAI. Gpt-4 technical report. ArXiv, abs/2303.08774, 2023.
  - [13] David Ouyang, Bryan He, Amirata Ghorbani, Neal Yuan, Joseph E. Ebinger, C. Langlotz, Paul A. Heidenreich, Robert A. Harrington, David H. Liang, Euan A. Ashley, and James Y. Zou. Video-based ai for beat-to-beat assessment of cardiac function. Nature, 580:252 – 256, 2020.
  - [14] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. Zero-shot text-to-image generation. ArXiv, abs/2102.12092, 2021.
  - [15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. ArXiv, abs/1505.04597, 2015.
  - [16] Ashish Vaswani, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Neural Information Processing Systems, 2017.
  - [17] Haixin Wang, Jianlong Chang, Xiao Luo, Jinan Sun, Zhouchen Lin, and Qi Tian. Lion: Implicit vision prompt tuning. ArXiv, abs/2303.09992, 2023.

- [18] Xinlong Wang, Xiaosong Zhang, Yue Cao, Wen Wang, Chunhua Shen, and Tiejun Huang. Seggpt: Segmenting everything in context. ArXiv, abs/2304.03284, 2023.
- [19] Huisi Wu, Shih-Wen Chen, Guilian Chen, Wei Wang, Baiying Lei, and Zhenkun Wen. Fat-net: Feature adaptive transformers for automated skin lesion segmentation. Medical image analysis, 76:102327, 2021.
- [20] Junde Wu, Rao Fu, Huihui Fang, Yuanpei Liu, Zhao-Yang Wang, Yanwu Xu, Yueming Jin, and Tal Arbel. Medical sam adapter: Adapting segment anything model for medical image segmentation. ArXiv, abs/2304.12620, 2023.
- [21] Xiwang Xie, Xipeng Pan, Feng Shao, Weidong Zhang, and Jubai An. Mci-net: Multi-scale context integrated network for liver ct image segmentation. Comput. Electr. Eng., 101:108085, 2022.
- [22] Kaiwen Zhang and Dong Liu. Customized segment anything model for medical image segmentation. ArXiv, abs/2304.13785, 2023.
- [23] Borui Zhao, Quan Cui, Renjie Song, Yiyu Qiu, and Jiajun Liang. Decoupled knowledge distillation. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 11943–11952, 2022.
- [24] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE Transactions on Medical Imaging, 39:1856–1867, 2019.
- [25] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation. In International Conference on Medical Image Computing and Computer-Assisted Intervention, 2016.