

# Differential-UMamba: 在数据有限的情况下重新思考 肿瘤分割

Dhruv Jain<sup>a,\*</sup>, Romain Modzelewski<sup>c</sup>, Romain Hérault<sup>b</sup>, Clement  
Chatelain<sup>a</sup>, Eva Torfeh<sup>d</sup>, Sebastien Thureau<sup>d</sup>

<sup>a</sup>*Normandie Univ. INSA Rouen LITIS 76801 Saint Etienne du Rouvray France*

<sup>b</sup>*Normandie Univ. UNICAEN ENSICAEN CNRS GREYC 14000 Caen France*

<sup>c</sup>*Nuclear Medicine Department Henri Becquerel Cancer Center and AIMS-Quantif  
Laboratory Rouen France*

<sup>d</sup>*Radiotherapy Department Henri Becquerel Cancer Center and AIMS-Quantif  
Laboratory Rouen France*

---

## Abstract

在数据稀缺的场景中，深度学习模型常常对噪声和无关的模式过拟合，这限制了其对未见样本的泛化能力。为了解决医疗图像分割中的这些挑战，我们引入了 Diff-UMamba，一种新的架构，将 UNet 框架与 mamba 机制结合以建模长距离依赖关系。Diff-UMamba 的核心是一个噪声抑制模块 (NRM)，该模块采用信号差分策略来抑制编码器内的噪声或无关激活。这鼓励模型过滤掉伪特征并增强与任务相关的表示，从而提高其对临床有意义区域的关注。因此，该架构在低数据环境中显著提高了分割精度和鲁棒性。Diff-UMamba 在多个公共数据集上进行了评估，包括 MSD (肺和胰腺) 和 AIIB23，在各种分割任务中相对于基线方法实现了 1-3 % 的一致性能提升。为了进一步评估在有限数据条件下的性能，还在 BraTS-21 数据集上通过改变可用训练样本比例进行了额外实验。该方法还在一个小型内部非小细胞肺癌 (NSCLC) 数据集上验证，用于在锥形束 CT (CBCT) 中进行肿瘤总体积 (GTV) 分割，在基线之上实现了 4-5 % 的提升。

**Keywords:** Tumor Segmentation, Mamba, Limited Data Scenario, Noise Reduction, nnUNet

---

---

\*Corresponding author

Email address: [dhruv.jain@insa-rouen.fr](mailto:dhruv.jain@insa-rouen.fr) (Dhruv Jain)

## 1. 引言

深度学习已经在医学图像分割中带来了革命性的变化，其中卷积神经网络 (CNNs) 在广泛的临床应用中提供了最先进的性能 [1]。大多数用于三维医学图像分割的架构都是建立在基础的 UNet 设计之上。在这些架构中，nnUNet [2] 作为一个强大的自配置框架脱颖而出，能够在多个基准测试中持续取得优异结果而无需手动调整，确立了其作为标准基线的地位。SegResNet [3] 通过结合残差连接和引入变分自动编码器 (VAE) 分支以正则化训练，进一步改进了 UNet 架构，特别是在数据有限的情况下尤为有益。然而，由于 CNN 的感受野有限，固有地难以建模长距离依赖关系，使其在分割那些扩散的、不规则的或在体积扫描中跨越多个切片的复杂解剖结构时效果较差 [4, 5]。为了解决这个问题，基于序列的模型，如 transformers [6] 和 mamba [7] 因其通过注意力机制和状态空间表示捕捉全局上下文的能力而受到关注 [8]。在许多真实世界场景中，单个数据点的解释高度依赖于上下文。例如，在自然语言处理领域，单词的含义受其周围符号的影响，而在医学影像处理中，解剖结构通常跨越多个切片，需要跨切片的连贯性以实现准确的分割 [9]。序列模型专门设计用于学习局部和全局依赖关系，使其能够构建上下文感知的表示。这种能力对分割任务尤其有利，因为准确的边界划定不仅依赖于孤立特征，还依赖于嵌入在整个序列中的连续性和结构。transformer [6] 相较于传统的递归神经网络 (RNN [10]) 和长短期记忆网络 (LSTM [11]) 代表了一项重大进展，通过用一个完全并行的多头注意力机制来取代它们固有的顺序处理。这一操作在方程式 1 中被正式定义：

$$\text{Attn}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

其中  $Q, K, V$  分别表示查询、键和值，而  $d_k$  是查询的维度。另一种变体称为差分 Transformer [12]，通过信号差分去除共模信号来减少上下文噪声。该差分注意机制在公式 2 中表达：

$$\text{DiffAttn}(X) = \left(\text{softmax}\left(\frac{Q_1K_1^T}{\sqrt{d}}\right) - \lambda \text{softmax}\left(\frac{Q_2K_2^T}{\sqrt{d}}\right)\right)V \quad (2)$$

其中  $Q_1, K_1$  表示用于捕获相关信号组件的主要注意力的查询和键矩阵，而  $Q_2, K_2$  对应于用于估计和减去共模噪声模式的查询和键。参数  $\lambda$  控制噪声抑制的程度。基于变换器的架构在 3D 分割中显示了潜力。UNETR [13] 在非重叠的 3D 补丁上使用视觉变换器 (ViT) 编码器。变换器的多层特征被融合到一个卷积解码器中。然而，UNETR 存在高计算成本、缺乏层次化表示和缺乏局部归纳偏差的问题。SwinUNETR [14] 使用分层 swin 变换

器 [15] 解决了这些问题。它将自注意力限制在局部窗口中，并引入了用于跨窗口交互的平移窗口。

Mamba [7] 是另一种序列模型，它引入了一系列选择性状态空间模型 (SSMs)，在保持变压器的长程上下文建模能力的同时，实现了  $\mathcal{O}(T)$  的时间和内存复杂度。它通过隐藏状态  $h(t) \in \mathbb{R}^N$  将一维输入序列  $x(t) \in \mathbb{R}$  处理为输出  $y(t) \in \mathbb{R}$ ，通过线性常微分方程 (ODEs) 捕捉长期依赖关系：

$$h'(t) = Ah(t) + Bx(t), \quad y(t) = Ch(t), \quad (3)$$

其中  $A \in \mathbb{R}^{N \times N}$  是状态转移矩阵， $B \in \mathbb{R}^{N \times 1}$  和  $C \in \mathbb{R}^{1 \times N}$  是投影矩阵。通过使用时间尺度  $\Delta$ ，连续值使用零阶保持 (ZOH) 离散化：

在训练过程中，mamba 利用全局卷积来高效并行地处理整个序列。卷积核被定义为

$$\bar{K} = (C\bar{B}, C\bar{A}\bar{B}, \dots, C\bar{A}^{M-1}\bar{B}), \quad (4)$$

，其中  $M$  是输入长度。这个结构化核  $\bar{K} \in \mathbb{R}^M$  通过线性变换实现了高效的长程依赖建模。UMamba [16] 将状态空间 mamba 模块融入到 UNet 中，以实现高效的 3D 医学分割。他们提出了两种变体，UMamba-Bot（基线）在瓶颈处使用 mamba，而 UMamba-Enc 则在每个编码器层中使用 mamba 模块。与基于 Transformer 的模型相比，UMamba 在降低计算成本的同时保持或提升了性能，使其适合于资源受限的环境。

然而，序列模型的成功依赖于获取大规模标注数据集 ( $> 800$  卷)，然而由于生成像素级标签所需的高成本和努力，这在医学领域是稀缺的。这种数据稀缺带来了一个主要挑战：当在小数据集上训练时，基于序列的模型往往会过拟合，学习到虚假的噪声模式而不是临床相关的特征 [17, 18, 19, 20]。为了解决这些限制，最近的工作探索了各种将序列模型适应于数据受限环境的策略。Swin-UMamba [21] 引入了一种混合架构，利用在大规模数据集（如 ImageNet）上的预训练，将一般的视觉先验知识转移到较小的目标 2D 域，从而有效地减少了过拟合。然而，专门为解决小数据集挑战而设计的 3D 网络较少。因此，我们提出了一种架构，该架构包括一个专门的模块，用于在潜在空间中以非指导的方式学习和抑制噪声模式。该模块与主要的 UNet 主干共同训练，通过减少虚假的激活来帮助模型专注于任务相关特征。本文的贡献如下。

- 我们提出了一种用于肿瘤分割的差分 UMamba 架构，该架构整合了一种新颖的降噪模块 (NRM)，以减轻在小规模数据集上的过拟合。据我们所知，这是首次引入医学图像分割的差分网络，该网络包含旨在抑制噪声的额外参数。

Table 1: 我们研究中使用的数据集概述。

| Dataset             | Modality                  | Targets                          | Number of Volumes |
|---------------------|---------------------------|----------------------------------|-------------------|
| MSD Task06-Lung     | CT                        | Tumor                            | 63                |
| MSD Task07-Pancreas | Contrast-enhanced CT      | Pancreas & Tumor                 | 281               |
| Internal Dataset    | Planning CT & CBCT        | GTV                              | 564               |
| BraTS               | MRI (T1, T1ce, T2, FLAIR) | Enhancing tumor, edema, necrosis | 1251              |
| AIIB23              | HRCT                      | Airways (no tumor labels)        | 120               |

- 我们分析了在不同数据集规模上训练的模型的潜在空间，并展示了过拟合会产生与注入噪声所造成的类似模式。
- 该方法在多个小规模数据集上进行了广泛评估，并在性能上持续优于现有的最先进方法。
- 实现了一种基于深度学习的肿瘤轮廓传播流程，使模型能够利用先前的轮廓信息并提高分割性能。该方法是从以前用于分割危及器官 (OAR) 的方法 [22] 中采纳的。

论文的组织结构如下。在第 2 节中，我们描述了所使用的数据集和我们提出的方法。第 3 节介绍了实现细节以及探索性研究，并提供了定性和定量的结果。最后，在第 4 节中，我们讨论了局限性和未来工作的潜在方向。

## 2. 材料与方法

### 2.1. 数据集

我们在三个以肿瘤为重点的数据集上进行了实验：MSD [23]（肺和胰腺）、BraTS-21 [24, 25, 26] 和一个内部数据集，以及一个用于气道分割的数据集 (AIIB23 [27])。本节提供了每个数据集的详细信息，摘要见表 1。

医学分割十项全能 [23] (MSD) 挑战赛提供了一套全面的 3D 医学影像数据集，涵盖了各种解剖部位，旨在促进可推广的分割算法的发展。

肺部分割：此任务涉及从胸部 CT 扫描中分割非小细胞肺癌 (NSCLC) 病变。该数据集由 63 个带注释的 3D CT 卷组成，作为医学分割竞赛的一部分发布 [23] (Task06-Lung)。使用单一的前景标签来识别肿瘤区域。由于肿瘤形态和密度的广泛异质性，以及胸腔内混杂的解剖结构的存在，该任务构成了一个重大挑战。我们将数据集分为 50 个用于训练的卷和 13 个用于测试的卷。

胰腺分割：这个任务 (Task07-Pancreas) 关注的是在增强对比的腹部 CT 扫描中分割胰腺及胰腺肿瘤。数据集包括注释的三维扫描图像，每个图像包含两个目标结构：胰腺及任何现存的肿瘤病灶。该任务的复杂性源于胰腺的体积小、不规则形状，以及与周围腹部器官的对比度较低。此外，肿

瘤外观通常稀疏且细微，这在训练过程中处理严重类别不平衡时带来了更多挑战。该数据集包括 281 例增强对比的腹部 CT 扫描，其中胰腺和胰腺肿瘤均有人工标注。数据集被分为 250 个用于训练的扫描和 31 个用于测试的扫描。

#### 2.1.1. BRaTS-21 数据集

这个数据集 [24, 25, 26] 主要集中在胶质瘤的分割上，胶质瘤是最常见且侵袭性最强的脑肿瘤之一。它由 1,251 个三维脑磁共振成像 (MRI) 扫描组成，每个扫描包含四种不同的成像模态：T1 加权 (T1)、增强型 T1 加权 (T1CE)、T2 加权 (T2) 以及 T2 液体衰减反转恢复 (T2-FLAIR)。它们提供逐体素标注的真实标签。图像具有固定的空间分辨率为  $240 \times 240 \times 155$  像素，来自使用各种 MRI 扫描仪的多个机构，确保采集设置的多样性。数据集包括三个关键肿瘤子区域的分割标签：

非增强型肿瘤核心 (NCR/NET) 增强型肿瘤 (ET) 肿瘤周围水肿 (ED)。

这些标签用于优化网络，并在最后结合以生成全肿瘤 (WT)、肿瘤核心 (TC) 和增强肿瘤 (ET) 的分割结果。数据集被随机分割，分配 150 张图像到测试集，其余图像用于训练和验证。

#### 2.1.2. 内部数据集：GTV 分割

这个数据集是匿名化并私有保存的，由 82 位诊断为未手术非小细胞肺癌患者的数据组成。他们接受了剂量在 60 至 70 Gy 的放射治疗，所有患者都同意使用他们的数据。成像使用了一种 3D free-breathing 技术，利用含碘对比剂、用于计划的西门子 CT 扫描仪，以及用于机载成像的 Varian CBCT 扫描仪和直线加速器进行。每位患者通常有 6 到 7 次 CBCT 扫描和一次计划 CT，依次将 CBCT 与计划 CT 首次对齐进行配准。对于标签轮廓，之前的 GTV ( $GTV_{n-1}$ ) 被应用于当前的 CBCT ( $CBCT_n$ )，使用  $[-400, +175 \text{ HU}]$  的阈值来分离 GTV，并对显著的解剖变化进行手动调整。数据被分为子集，61 名患者被分配到训练集，14 名到验证集，剩下的 7 名到测试集，并确保各子集之间患者数据不重叠。这分别对应 476、48 和 40 次扫描。

#### 2.1.3. AIIB23 数据集

AIIB23 数据集 [27] 作为“纤维化肺病的气道知情定量 CT 成像生物标志物-2023”挑战的一部分在 MICCAI 发布。该数据集旨在促进在间质性肺病患者中进行稳健的气道树分割和预测性生物标志物的发现，因为晚期纤维化的扭曲结构使得气道提取异常困难。虽然这一数据集中没有提供肿瘤

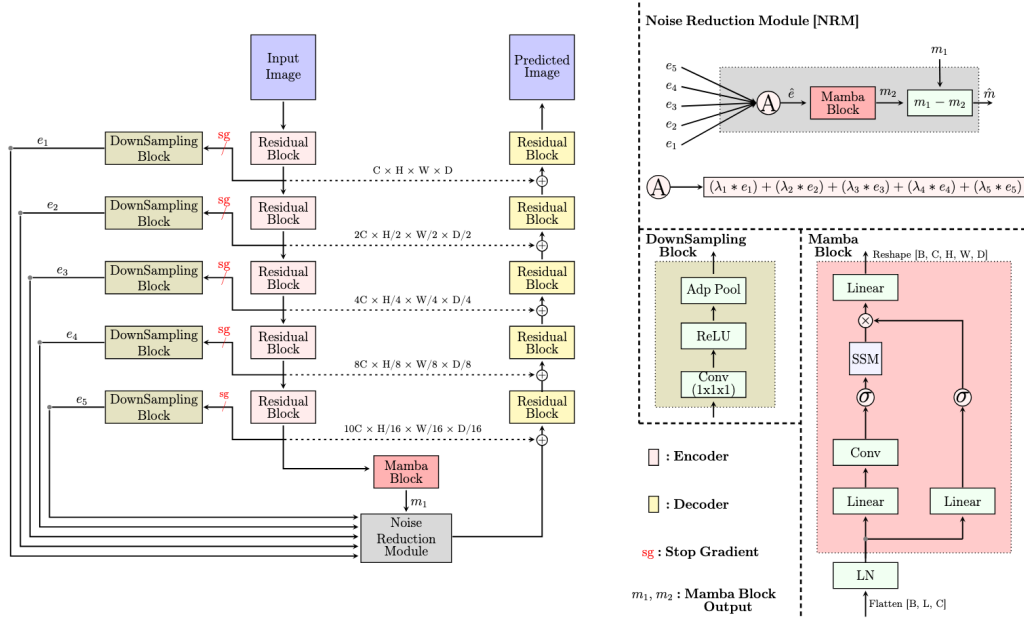


Figure 1: 提出的 Diff-UMamba 的概述。独立层用于检测噪声模式，随后通过噪声减少模块 (NRM) 进行过滤。(  $\lambda_1 - \lambda_5$  是可学习的参数。)  $e_1$  到  $e_5$  的特征形状，与  $m_1$  和  $m_2$  的形状，与最后的跳跃连接相同。

分割，但我们将其包括在内，以增加数据集池的多样性。挑战中的训练部分由 120 个高分辨率胸部 CT 体积组成，这些体积是从患有进展性纤维化间质性肺病的患者（主要是特发性肺纤维化）中通过全吸高分辨率临床 CT (HRCT) 协议获取的。所有扫描的层间距为  $\leq 1$  毫米，平面内像素大小约为 0.7 毫米，反映了常规 HRCT。数据集随机分割，分配 99 个体积到训练集，其余体积用于测试集。

## 2.2. 提出的架构

小型数据集由于缺乏足够的数据点来保证泛化能力，可能会因为过拟合或训练的记忆化导致模型引入不必要的噪声模式。所提出的 Diff-UMamba 通过在每个编码器层集成专门的模块来解决这一问题，这些模块学习数据中的噪声模式，并最终通过噪声减少模块 (NRM) 在瓶颈层中将其过滤掉，如图 1 所示。

这种过滤掉不重要特征的方法受到差分变压器 [12] 的启发，它在注意力层中使用否定技术来消除上下文的词元级噪声。

Diff-UMamba 的剩余块和 mamba 块与 UMamba [16] 中使用的结构相同。每个剩余块由一个卷积层组成，后接实例归一化 (IN [28]) 和 Leaky

ReLU [29] 激活函数。然后将原始输入加到处理输出上，使得该模块能够学习残差映射。mamba 模块通过两条并行路径处理输入。在两者中，序列通过线性层投影到更高的分辨率  $(B, 2L, C)$ 。第一路径应用 1D 卷积和 SiLU [30] 激活，以及结构化状态空间 (SSM) 操作。第二路径则应用另一个线性投影后接 SiLU [30] 激活。两条路径的输出通过 Hadamard 积逐元素结合。最后，产生的特征投影回  $(B, L, C)$  并重新整形到原始的 3D 布局。方程 5 概述了 Diff-UMamba 的完整前向传播过程。

$$\hat{y} = \mathcal{D} \left( \mathcal{M}_1(\mathcal{E}(x)) - \mathcal{M}_2 \left( \sum_{i=1}^l \lambda_i e_i \right) \right) \quad (5)$$

其中  $\hat{y}$  是网络预测输出， $\mathcal{E}$  是从输入  $x$  提取特征表示的主要编码器， $\mathcal{D}$  是重构最终输出的解码器， $\mathcal{M}_1$  是生成中间表示  $m_1$  的瓶颈蛇块，而  $\mathcal{M}_2$  是通过对外部嵌入  $e_i$  的加权和生成  $m_2$  的第二个蛇块。

### 2.2.1. 下采样块

来自编码器层的每个输出都经过一个专用的下采样块，以将特征图投影到统一的空间分辨率。每个块由一个  $1 \times 1 \times 1$  卷积、一个 ReLU [29] 激活以及一个自适应平均池化操作组成，该操作将特征图调整为预定义的瓶颈尺寸。给定来自  $i$ -th 编码器层的编码器输出特征图  $F_i \in \mathbb{R}^{C_i \times D_i \times H_i \times W_i}$ ，下采样输出  $\hat{F}_i$  按照方程 6 计算：

$$\hat{F}_i = \text{AdapPool}(\text{ReLU}(\text{Conv}_{1 \times 1 \times 1}(F_i))), \quad \hat{F}_i \in \mathbb{R}^{C' \times D' \times H' \times W'} \quad (6)$$

这里， $C'$  是目标通道数， $(D', H', W')$  表示瓶颈的空间期望尺寸。对于每个编码器层的输出，分别应用降采样块，以确保在 NRM 中进行特征融合或聚合之前的一致维度性。

该模块在图 1 中进行了说明，清楚地描绘了其输入和输出。它将  $e_1$  到  $e_5$  和  $m_1$  作为输入，并产生  $\hat{m}$  作为输出。模块汇总来自多个降采样层的噪声模式，如方程 7 中所定义：

$$\hat{e} = \sum_{i=1}^l \lambda_i e_i \quad (7)$$

这里， $\lambda_i$  是可训练的参数， $e_i$  表示专用下采样块的第 1 层的输出。在聚合完输出后，它们会被处理，通过 mamba 块  $\mathcal{M}_2$  [7]，然后被滤出主要瓶颈，如方程 8 所示。

$$\hat{m} = m_1 - m_2 \quad (8)$$

在这里， $m_1$  对应提取出的包含有用信息和噪声的特征，而  $m_2$  代表估计的噪声。与经典信号处理系统（例如，卡尔曼滤波器）一样，状态空间模型以估计隐含动态并滤除测量噪声而闻名于信号去噪 [31]。在 NRM 中使用 mamba 的灵感来源于这一原理。从  $m_1$  中减去  $m_2$  有效地发挥了滤波器的作用，以隔离有意义的特征并抑制噪声。

---

**Algorithm 1** 将 NRM 集成到基于 UNet 的架构中

---

**Require:** Encoder outputs  $\{F_1, F_2, \dots, F_l\}$ , bottleneck feature  $m_1$   
**Require:** Trainable parameters  $\{\lambda_1, \lambda_2, \dots, \lambda_l\}$ , Mamba block  $\mathcal{M}$   
**Ensure:** Denoised output  $\hat{m}$  for decoding

```

1: for  $i = 1$  to  $l$  do
2:    $e_i \leftarrow \text{AdapPool}(\text{ReLU}(\text{Conv}_{1 \times 1 \times 1}(F_i)))$   $\triangleright$  Downsample  $F_i$  to produce  $e_i$  as in Equation 6
3: end for
4:  $\hat{e} \leftarrow \sum_{i=1}^l \lambda_i \cdot e_i$   $\triangleright$  Aggregate encoder outputs using Equation (7)
5:  $m_2 \leftarrow \mathcal{M}(\hat{e})$   $\triangleright$  Estimate noise using a Mamba block
6:  $\hat{m} \leftarrow m_1 - m_2$   $\triangleright$  Filter noise as per Equation (8)
7: return  $\hat{m}$   $\triangleright$  Pass  $\hat{m}$  to decoder

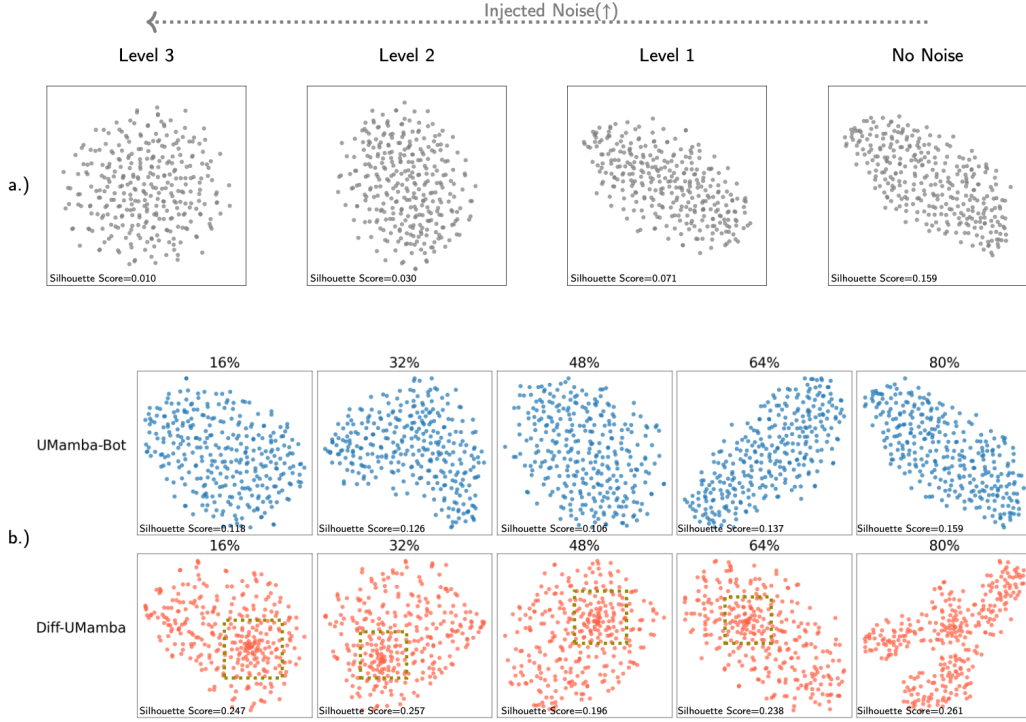
```

---

降噪模块（NRM）是一个灵活的组件，可以集成到任何基于 U 型网络的架构中。将 NRM 整合到现有架构的过程在算法 1 中概述。注意：在 Diff-UMamba 中有 2 个 mamba 模块。我们建议在较小的数据集上使用这个模型，在这种情况下，尽管 mamba 模型 [7] 通常比 transformers [6] 表现更好，但常常难以有效地表现。相比之下，较大数据集通常提供更广泛多样的训练样本，使模型能够更有效地泛化，并且在基于序列的模型中通常表现更好。

### 2.3. 潜在空间分析

我们发现，在有限数据集上训练的模型往往会捕捉到虚假的或过拟合的模式，这些模式在特征空间中表现为类似噪声的扭曲。为探讨这一现象，我们可视化了训练模型（UMamba-Bot [16]）的瓶颈特征。在图 2 a 中，我们展示了瓶颈特征的按通道 t-SNE 投影，其中人为地加入了不同级别的高斯噪声。随着噪声水平的增加，特征的结构化组织逐渐恶化，表明潜在空间对噪声的敏感性。有趣的是，当模型在数据集的缩减子集上进行训练时，也观察到了类似的特征结构退化。为了证明这一点，我们将 Figure 2 b 中的 BraTS-21 数据集上使用不同大小的训练数据进行训练的模型的瓶颈特征进行了可视化。随着训练集规模的减小，特征嵌入变得越来越分散且不



连贯，这表明有限数据会在潜在空间中引发类似噪声的伪影。为了量化这一行为，我们采用了轮廓系数 [32]，对于数据点  $i$ ，定义在公式 9 中：

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (9)$$

其中  $a_i$  表示簇内的平均距离（即  $i$  与同一簇内所有其他点之间的平均距离），而  $b_i$  表示  $i$  与最近邻簇中所有点的最小平均距离。利用 k-means 进行聚类，聚类数量为  $k$ ，并使用最大轮廓分数来评估聚类质量和选择最佳的  $k$ 。观察到 UMamba-Bot (16 %, 32 %, 48 %) 之间的聚类行为减弱，这类似于嵌入空间中噪声水平的增加。然而，在分析不同训练样本量的 Diff-UMamba 模型的瓶颈特征时，我们观察到某些通道始终表现出良好的分离簇，即（**强调**）。Diff-UMamba 拥有更高的轮廓分数，表明其特征表示更加紧凑且具有区别性。虽然我们不能确定聚类瓶颈特征与分割指标改

进之间的直接因果关系（详见第 ?? 节），但我们观察到这种一致的聚类行为出现在使用不同数据集大小训练的 Diff-UMamba 模型中。此外，这个数据集中有轻微的过拟合迹象，因为使用四种输入模式有助于减小训练集 80 % 和 16 % 之间的性能差距。进一步的调查在第 ?? 节中对 Diff-UMamba 中的 NRM 模块进行了评估。

### 3. 实验和结果

我们将 Diff-UMamba 集成到 UMamba-Bot 架构中，该架构基于 nnUNet [2] 框架自动管理预处理、装配和网络架构的选择，为医学分割任务提供了强大而适应性强的设置。优化是通过 SGD 进行的，使用 dice 损失和交叉熵损失的未加权和。实验在内部数据集中进行了 1000 个 epoch，使用单个输入 (CBCT) 或两个输入 (见第 ?? 节)。对于公共 BRaTS-21 数据集 [24, 25, 26]，使用了四种输入模态 [T1, T1CE, T2, FLAIR]。相比之下，MSD [23] 和 AIIB23 [27] 数据集使用单个 CT 体积作为输入。所有实验均在 Tesla A100 GPU 上进行。

在本节中，我们在不同的数据集上评估和分析各种最新的分割网络。Diff-UMamba 与各种架构进行了比较，包括基于 CNN 的方法（如 nnUNetv2、SegResNet）、基于 transformer 的方法 (UNETR、SwinUNETR)，以及基于 mamba 的模型 (如 LKM-UNet、UMamba-Bot、UMamba-Enc 和 SegMamba) 用于内部数据集。由于肿瘤形状和大小的多样性，以及它们经常出现在低对比度区域中的特性，训练用于 GTV 分割的深度学习网络带来了很大的挑战。为了解决这一问题，我们结合使用了刚性注册的轮廓 (r-CT) 作为先验信息，以指导网络精确描绘 GTV，正如图中所示，这类似于之前在 CBCT 体积上的 OAR 分割中展示的方法。该过程首先将计划 CT（传统放射治疗时间线如图所示）与 CBCT 体积进行配准。

使用梯度下降优化转换参数以对齐这两种模态。为最小化人工干预，对所有患者使用相同的超参数。注意：该方法仅应用于内部数据集。表 ?? 显示了结果，分为两个组：一个排除了 r-CT 轮廓，另一个在输入中包含了它们。评估指标由骰子系数 (DSC  $\uparrow$ )、交并比 (IOU  $\uparrow$ ) 和豪斯多夫距离 (HD95  $\downarrow$ ) 组成，我们观察到包含 r-CT 轮廓为所有模型提供了显著好处。此外，Diff-UMamba 在两种情况下的结果均优于其余情况。与 Umamba-Bot 相比，在没有 r-CT contours 的情况下，DSC 增加了 5.02% ( $pvalue = 0.03$ ，配对 t 检验)；在使用 r-CT contours 的情况下，DSC 增加了 4.44% ( $pvalue = 0.002$ ，配对 t 检验)。有趣的是，当包含额外数据 (r-CT 轮廓) 时，改进更小。NRM 的增加使总参数数量增加了 2 %，同时在潜在特征空间中提供了降噪。

Table 2: 轮廓传播方法的 DSC 比较。

|     | Translation | Rigid | DIR-T | DIR-L | Ours  |
|-----|-------------|-------|-------|-------|-------|
| DSC | 55.19       | 56.89 | 63.20 | 66.59 | 76.91 |

表 2 对比了包括 DIR 在内的非深度学习轮廓传播方法的两种变体：仅肿瘤变形的 DIR-T 和整个肺变形的 DIR-L。我们的方法在所有其他方法上始终至少领先 10%。图 A.5 提供了两种可变形配准变体的更多细节。我们在 MSD [23] 数据集上进行了实验，任务包括肺和胰腺，以及 AIIB23 [27] 数据集的气道分割。结果如表 ?? 所示。在所有任务和数据集中，与 UMamba-Bot [16] 相比，Diff-UMamba 的 IOU 始终增加 1 到 3%。这种稳定的改进表明，集成到 Diff-UMamba 中的降噪模块 (NRM) 在增强特征表示和模型泛化方面发挥了重要作用，特别是在训练数据有限的情况下。

Table 3: 在 BRaTS-21 数据集上对最先进的分割模型进行比较，使用不同比例 (16 %, 32 %, 48 %) 的训练样本。表格报告了 DSC、IoU 和 HD95，表示为 (平均值  $\pm$  标准差)。

| Models          | Avg (WT, TC, ET)  |                   |                      |                   |                   |                      |                   |                   |                      |
|-----------------|-------------------|-------------------|----------------------|-------------------|-------------------|----------------------|-------------------|-------------------|----------------------|
|                 | 16 %              |                   |                      | 32 %              |                   |                      | 48 %              |                   |                      |
|                 | DSC( $\uparrow$ ) | IOU( $\uparrow$ ) | HD95( $\downarrow$ ) | DSC( $\uparrow$ ) | IOU( $\uparrow$ ) | HD95( $\downarrow$ ) | DSC( $\uparrow$ ) | IOU( $\uparrow$ ) | HD95( $\downarrow$ ) |
| nnUNetv2 [2]    | 89.87 $\pm$ 15.11 | 83.97 $\pm$ 17.63 | 5.00 $\pm$ 16.60     | 91.43 $\pm$ 12.70 | 85.96 $\pm$ 15.30 | 4.36 $\pm$ 15.55     | 91.25 $\pm$ 13.50 | 86.08 $\pm$ 15.47 | 3.61 $\pm$ 13.40     |
| SegResNet [3]   | 89.31 $\pm$ 16.73 | 83.90 $\pm$ 17.87 | 4.95 $\pm$ 18.24     | 90.74 $\pm$ 14.43 | 85.43 $\pm$ 16.29 | 3.86 $\pm$ 14.16     | 90.83 $\pm$ 14.47 | 85.39 $\pm$ 16.79 | 4.19 $\pm$ 15.18     |
| UMamba-Bot [16] | 89.74 $\pm$ 15.38 | 84.03 $\pm$ 17.24 | 5.01 $\pm$ 16.89     | 90.59 $\pm$ 13.44 | 84.70 $\pm$ 15.79 | 4.79 $\pm$ 16.10     | 91.18 $\pm$ 13.74 | 86.21 $\pm$ 14.98 | 3.55 $\pm$ 11.01     |
| UMamba-Enc [16] | 89.17 $\pm$ 16.12 | 83.28 $\pm$ 17.91 | 5.16 $\pm$ 16.95     | 90.54 $\pm$ 14.09 | 84.80 $\pm$ 16.54 | 5.14 $\pm$ 16.95     | 91.14 $\pm$ 13.65 | 86.14 $\pm$ 15.09 | 3.51 $\pm$ 10.78     |
| LKM-UNet [33]   | 89.23 $\pm$ 16.45 | 83.49 $\pm$ 18.24 | 5.58 $\pm$ 18.63     | 90.94 $\pm$ 13.60 | 85.34 $\pm$ 15.97 | 4.63 $\pm$ 15.90     | 90.77 $\pm$ 14.07 | 85.19 $\pm$ 16.59 | 4.64 $\pm$ 15.82     |
| Diff-UMamba     | 89.93 $\pm$ 14.83 | 84.21 $\pm$ 16.82 | 4.85 $\pm$ 16.15     | 91.24 $\pm$ 12.67 | 86.22 $\pm$ 14.56 | 3.41 $\pm$ 14.56     | 91.30 $\pm$ 13.10 | 86.27 $\pm$ 14.51 | 3.19 $\pm$ 9.78      |

此外，表 3 和 ?? 显示了使用不同比例的训练样本 (16 %、32 %、48 %、64 %、80 %) 在 BRaTS-21 数据集 [24, 25, 26] 中的网络性能。我们比较了 nnUNetv2 [2]、SegResNet [3] 和 mamba 架构 [33, 16] 与 Diff-UMamba。我们观察到，当数据集的 32 % 进行训练时，Diff-UMamba 显示出了最显著的改进，在三项指标中显著优于其他方法 (基于与  $pvalue < 0.05$  进行配对 t 检验)。由于存在多个输入模式，训练数据的 16 % 和 80 % 之间的骰子分数差异相对较小。此外，在这种情况下引入 NRM 的好处并不明显，因为多模态输入本身有助于减少过拟合。因此，在所有较小的数据集规模 (16 % 和 48 %) 中，性能提升并不总是显著的。表 ?? 显示，随着数据集规模增大，UMamba-Bot [16] 和 UMamba-Enc [16] 的性能优于 Diff-UMamba 和 nnUNetv2 [2]，反映出 mamba 在建模长程依赖并减轻对 NRM 需求方面的优势。这表明，随着数据集规模的扩大，NRM 模块的优势被削弱，而基础模型可以更有效地进行泛化。我们重复了实验，使用单一模态，即来自 BraTS-21 [24, 25, 26] 的 T1 加权 MRI，包含训练样本的 16 %，以证明在

真正小数据环境中的改进。图 ?? 提出了 Diff-UMamba、UMamba-bot [16] 和 nnUNetv2 [2] 在三个肿瘤子区域中的比较。尤其是，Diff-UMamba 在 TC、ET 和平均性能上表现出显著改进 ( $p < 0.05$ , 配对 t 检验)。这些结果突出显示了 Diff-UMamba 的能力，尤其是在数据受限的环境中可以学习到更有意义的表示。

### 3.0.1. 消融研究

进行了一项针对内部数据集（使用 r-CT 轮廓线）的消融研究，以评估不同 NRM 配置的影响，其中包括排除  $e_1$  到  $e_5$ ，移除 mamba 块  $\mathcal{M}_2$  以及更改  $\lambda$  初始值 ( $\lambda_{init}$ )，这些在表 ?? 中进行了总结。所有消融变体在比较 UMamba-Bot 时一致地达到了更高的 DSC。发现最佳配置（具有最大 IOU）包括 ( $e_1$  到  $e_5$ )，保留 mamba 块  $\mathcal{M}_2$ ，并将  $\lambda_{init}$  设为 0.5。结果表明，移除  $e_1$  或  $e_2$  显著降低性能，而排除更深层的影响相对较小。这支持了 NRM 有效过滤主要在早期编码器层中捕获的噪声的假设。另外，在表 ?? 中还包括标准 NRM 模块装备的 SegResNet [3] 和 UNETR [13] 的结果，相较于非 NRM 版本（见表 ??），有适度的改善，SegResNet 的 dice 得分提高了 1.01%，UNETR 提高了 2.85%。相比之下，Diff-UMamba 获得了显著更高的 4.44% 的提升，可能是因为 mamba 架构提供了更好的归纳偏置以整合 NRM，能够更好地建模序列模式并在数据少的情况下抑制噪声，而 transformers 对数据稀缺更敏感，较难有效利用 NRM 捕获的差异化信号。

在这一部分中，我们在内部数据集上进行探索性实验，以分析 NRM 的行为并将 Diff-UMamba 的性能与基线模型进行比较。

图 ?? 展示了通过 Diff-UMamba 编码器处理的测试集中的 1,280 个小块中  $m_1$  和  $m_2$  之间的皮尔逊相关性。结果一致显示这两种嵌入类型之间的相关性较低，这表明它们捕捉到了不同的特征表示。这暗示着  $m_1$  和  $m_2$  关注输入数据的不同方面。此外，我们通过在推理过程中将各种类型的噪声（高斯、散斑、周期性和椒盐）以不同的强度引入第一个残差块来评估模型的鲁棒性。如图 ?? 所示，Diff-UMamba 比 UMamba-Bot [16] 更有效地适应，这表明 NRM 能够在 UMamba 瓶颈中减轻噪声。噪声水平与其对应参数之间的映射在附录 A.4 中给出。

### 3.0.2. 参数的演变

图 ?? 展示了在 BRaTS-21 [24, 25, 26] 数据集的 32% 和 80% 上训练的模型的  $\lambda$  可训练参数的演变对比。当在 32% 的数据上训练时，我们观察到  $\lambda_1$  和  $\lambda_2$  达到了显著高于其余参数的值。这表明在数据有限的条件下，模型对早期特征（即来自较浅的编码器阶段）赋予了更大的重要性，这支持了微分操作通过从早期表示中放大更多信息丰富的信号来减弱噪声的作用。

与之相对，当在数据集的 80 % 上进行训练时，模型给所有参数  $\lambda$  指派了一致较低的值，且它们之间的变化极小。这表明对微分操作的噪声减弱效应的依赖性降低。通过获取足够的数据，模型可以直接从输入中学习更加具有表现力的表示，而不必强调可微信号。结果是对 NRM 的需求减少，模型通过在训练期间抑制数值  $\lambda$  来隐式地调节其影响。这种行为突显了模型基于数据可用性自适应调节其架构机制的能力。另一个观察到的现象是  $\lambda$  参数的稳定化行为。在数据集的 32 % 和 80 % 的设置中， $\lambda$  值最初表现出波动，因为模型正在寻找噪声特征的最佳组合。然而，在数据集的 80 % 上训练的模型中，稳定化发生得更快。这种较早的稳定化支持了这样一种观点，即在有足够的训练样本时，模型对特征微分的依赖性较少，从而能够更快速地稳定到有效的已学得表示中。

### 3.1. 与现有最先进方法的视觉比较

图 3 展示了在四个公共数据集中，Diff-UMamba 与几种最先进的分割方法之间的定性比较：(a) BRaTS-21 [24, 25, 26]，(b) MSD [23]（胰腺），(c) MSD [23]（肺部），和 (d) AIIB23 [27]。这些比较突出了 Diff-UMamba 在不同模态和器官中准确描绘复杂解剖和病理结构的有效性。

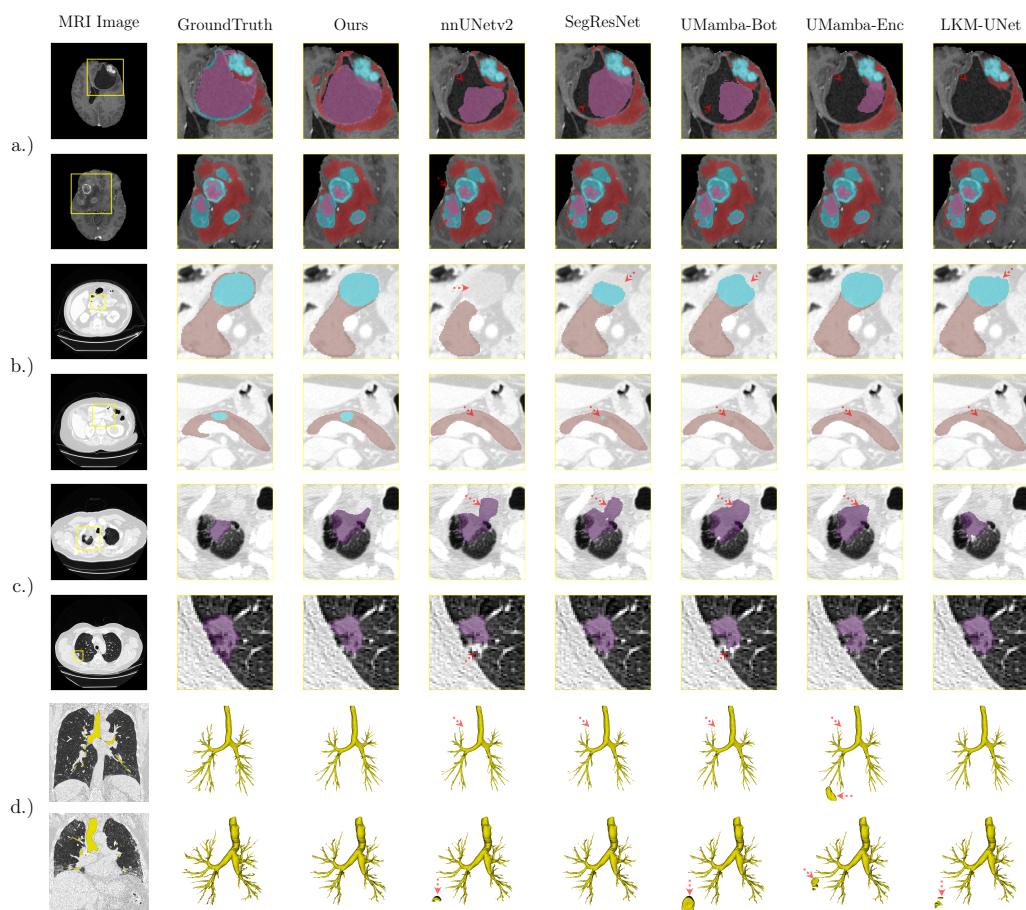


Figure 3: 四个数据集分割结果的视觉比较：a) BRaTS-21 数据集 [24, 25, 26]（在可用训练集的 32 % 上训练），显示的分割区域：NCR, ED 和 ET。b) 医学分割十项全能 [23] 胰腺任务，突出显示 pancreas 和 tumor 区域。c) 肺任务，显示 tumor 区域。d) AIIB23 数据集 [27]，突出显示 airway 区域。

在图 3 a 的第一个例子中，Diff-UMamba 在准确描绘非增强核心 (NCR) 区域上明显优于其他方法。在第二个例子中，虽然大多数方法的表现大致相似，Diff-UMamba 在某些细微区域提供了更精确的描绘，如图中红色箭头所示。在图 3 b (MSD [23]：胰腺) 中，第二个例子清楚地展示了 Diff-UMamba 的优势：它是唯一一种成功检测出肿瘤区域的方法，而大多数其他模型错误地将其分类为胰腺的一部分。这突出显示了模型对小病理区域的敏感性及其在应对类别混淆方面的鲁棒性，这在胰腺肿瘤分割中由于肿瘤的细微表现和边界不规则性而是一个常见问题。对于 MSD [23]

(肺)数据集(图 3 c), 在第一个例子中, 大多数方法表现出分割缺陷, 尤其是在边界区域。然而, Diff-UMamba 显示出更清晰且连贯的描绘, 具有更少的误报和不连续性, 而 LKM-UNet 与其他模型相比也表现良好。在图 3 d 中的 AIIB23 [27] 数据集, 专注于气道和周围结构的分割, 许多模型在两个例子中都出现了显著的伪影。然而, Diff-UMamba 表现出更大的抗性, 提供了更符合解剖学的结果, 具有更少的气道段断开或碎片。使用内部数据集的比较如图 A.6 所示, 其中子图 (a)、(b) 和 (c) 展示了 r-CT 轮廓的引入显著提高了肿瘤边界的可见性, 即使在对比度较低的区域也是如此。然而, 一些示例如 c 和 d 显示出当使用 r-CT 轮廓时, 性能由于不准确的配准而下降的情况。

#### 4. 局限性

噪声减少模块 (NRM) 显示出令人鼓舞的结果, 尤其是在小规模数据集上。在这项研究中, NRM 被集成到 UMamba-Bot 架构中, 以利用 mamba 在小规模数据集上的长程依赖建模能力。替代的集成方案, 例如在每个编码器层中嵌入 NRM 的 UMamba-Enc, 可能会提高去噪性能。然而, 早期的编码器层通常在具有大量标记的高分辨率特征图上运行, 使得这种集成在计算上代价高昂且非平凡。在未来的工作中解决这一挑战将是关键, 以便在整个架构中更广泛地使用 NRM。此外, 这项工作仅限于 3D 医学图像分割任务。尚未评估 NRM 对 2D 数据和非医疗领域的适用性。更广泛地说, 这项作为我们如何利用小数据集进行学习开启了一个新方向。传统上, 小数据集通常与轻量级模型配对以避免过拟合。然而, 如果像 NRM 这样的噪声感知组件允许较大的模型在数据稀缺的情况下具有更好的泛化能力, 这可能促使我们重新思考模型设计策略。与其简单地减小模型尺寸, 我们可能会开始探索通过选择性正则化的方法, 开发更大的架构, 即使在受限的数据环境中也能为更具表现力的模型铺平道路。未来工作的另一范围包括动态 NRM 的开发, 其可根据数据集的特征选择性激活。例如, NRM 仅在学到的噪声阈值 (例如  $\lambda_1$  或  $\lambda_2$  值) 超过某个水平时应用, 从而允许模型自适应地决定何时需要去噪。这种机制将使 NRM 的动态部署成为可能, 使得 Diff-UMamba 更具可扩展性。

#### 5. 结论

在这项工作中, 我们介绍了 Diff-UMamba, 这是一种新颖的模型, 通过检测和去除每个编码器层中的噪声来增强像 mamba 这样的序列模型在较小数据集上的性能。我们的方法基于 UMamba-Bot, 整合了一个噪声消除模块 (NRM), 以在数据有限的情况下提高性能。我们展示了该模型超越

了现有的 sota 模型，在内部数据集和公共数据集上均取得了优异的结果，即使训练样本有限。我们的研究表明，Diff-UMamba 适用于医学成像应用，其中获取大量高质量标注的挑战仍然是一个显著的限制。

## 6.

### 致谢

本研究得到了 MINMACS Normandie 地区卓越奖和 ANR LabCom L-Lisa ANR-20-LCV1-0009 的支持。我们感谢 CRIANN 的同事为本项目提供了必要的计算资源。我们还感谢 Su Ruan 和 Tsiry Mayet 的校对。

## 7.

利益披露 作者声明没有与本文内容相关的竞争利益。

## References

- [1] I. D. Mienye, T. G. Swart, G. Obaido, M. Jordan, P. Ilono, Deep Convolutional Neural Networks in Medical Image Analysis: A Review, *Information* 16 (3) (Mar. 2025).
- [2] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, K. H. Maier-Hein, nnu-net: a self-configuring method for deep learning-based biomedical image segmentation, *Nature Methods* 18 (2) (2021).
- [3] A. Myronenko, 3d mri brain tumor segmentation using autoencoder regularization, in: A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, T. van Walsum (Eds.), *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Springer International Publishing, Cham, 2019.
- [4] J. M. J. Valanarasu, P. Oza, I. Hacıhaliloglu, V. M. Patel, Medical transformer: Gated axial-attention for medical image segmentation, in: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Springer International Publishing, Cham, 2021.
- [5] Y. Gao, Y. Jiang, Y. Peng, F. Yuan, X. Zhang, J. Wang, Medical Image Segmentation: A Comprehensive Review of Deep Learning-Based Methods, *Tomography* 11 (5) (Apr. 2025).
- [6] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is All you Need, *Advances in neural information processing system* (2017).

- [7] A. Gu, T. Dao, Mamba: Linear-Time Sequence Modeling with Selective State Spaces (2024). [arXiv:2312.00752](#).
- [8] F. Shamshad, S. Khan, S. W. Zamir, M. H. Khan, M. Hayat, F. S. Khan, H. Fu, Transformers in medical imaging: A survey, *Medical Image Analysis* 88 (2023) 102802.
- [9] J. Chen, L. Yang, Y. Zhang, M. Alber, D. Chen, Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation (09 2016).
- [10] R. M. Schmidt, Recurrent Neural Networks (RNNs): A gentle Introduction and Overview (Nov. 2019). [arXiv:1912.05911](#).
- [11] R. C. Staudemeyer, E. R. Morris, Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks (Sep. 2019). [arXiv:1909.09586](#).
- [12] T. Ye, L. Dong, Y. Xia, Y. Sun, Y. Zhu, G. Huang, F. Wei, Differential Transformer, *International Conference on Learning Representations* (2025).
- [13] A. Hatamizadeh, Y. Tang, V. Nath, D. Yang, A. Myronenko, B. Landman, H. R. Roth, D. Xu, Unetr: Transformers for 3d medical image segmentation, in: *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2022.
- [14] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. Roth, D. Xu, Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images (2022). [arXiv:2201.01266](#).
- [15] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.
- [16] J. Ma, F. Li, B. Wang, U-Mamba: Enhancing Long-range Dependency for Biomedical Image Segmentation (2024). [arXiv:2401.04722](#).
- [17] D. Hussain, Y. Hyeon Gu, Exploring the impact of noise and image quality on deep learning performance in dxa images, *Diagnostics* 14 (13) (2024).
- [18] X. Li, D. Chang, Z. Ma, Z.-H. Tan, J.-H. Xue, J. Cao, J. Yu, J. Guo, OSLNet: Deep Small-Sample Classification with an Orthogonal Softmax Layer, *IEEE Transactions on Image Processing* 29 (2020).

- [19] A. Power, Y. Burda, H. Edwards, I. Babuschkin, V. Misra, Grokking: Generalization Beyond Overfitting on Small Algorithmic Datasets, International Conference on Learning Representations Workshop (2022).
- [20] R. Shao, X.-J. Bi, Transformers Meet Small Datasets, IEEE Access 10 (2022).
- [21] J. Liu, H. Yang, H.-Y. Zhou, L. Yu, Y. Liang, Y. Yu, S. Zhang, H. Zheng, S. Wang, Swin-umamba: Adapting mamba-based vision foundation models for medical image segmentation, IEEE Transactions on Medical Imaging (2024) 1–1.
- [22] L. Ma, W. Chi, H. E. Morgan, M.-H. Lin, M. Chen, D. Sher, D. Moon, D. T. Vo, V. Avkshtol, W. Lu, X. Gu, Registration-guided deep learning image segmentation for cone beam ct-based online adaptive radiotherapy, Medical Physics (2022).
- [23] M. Antonelli, A. Reinke, Bakas, The Medical Segmentation Decathlon, Nature Communications 13 (1) (Jul. 2022).
- [24] B. H. Menze, A. Jakab, S. Bauer, The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS), IEEE Transactions on Medical Imaging 34 (10) (2015).
- [25] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, Crimi, Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge (2018).
- [26] S. Bakas, H. Akbari, A. Sotiras, Bilello, Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features, Scientific Data 4 (1) (2017).
- [27] Y. Nan, X. Xing, Wang, Hunting imaging biomarkers in pulmonary fibrosis: Benchmarks of the AIIB23 challenge, Medical Image Analysis 97 (Oct. 2024).
- [28] D. Ulyanov, A. Vedaldi, V. Lempitsky, Instance normalization: The missing ingredient for fast stylization (07 2016).
- [29] A. L. Maas, A. Y. Hannun, A. Y. Ng, Rectifier Nonlinearities Improve Neural Network Acoustic Models (2013).
- [30] D. Hendrycks, K. Gimpel, Gaussian Error Linear Units (GELUs) (Jun. 2023). [arXiv:1606.08415](https://arxiv.org/abs/1606.08415).

- [31] H. Nalatore, M. Ding, G. Rangarajan, Denoising neural data with state-space smoothing: Method and application, *Journal of Neuroscience Methods* 179 (1) (2009).
- [32] K. R. Shahapure, C. Nicholas, Cluster quality analysis using silhouette score, in: *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, 2020.
- [33] J. Wang, J. Chen, D. Chen, J. Wu, LKM-UNet: Large Kernel Vision Mamba UNet for Medical Image Segmentation, *Medical Image Computing and Computer-Assisted Intervention* 15008 (2024).

Appendix A. 实验的额外信息

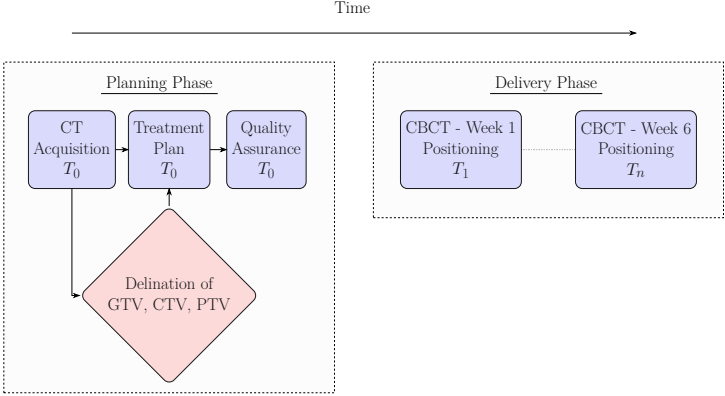


Figure A.4: 传统放射疗法展示了计划和实施阶段。在这段时间内，需要获取 CBCT 图像用于定位，整个治疗大约需要 30-45 天。

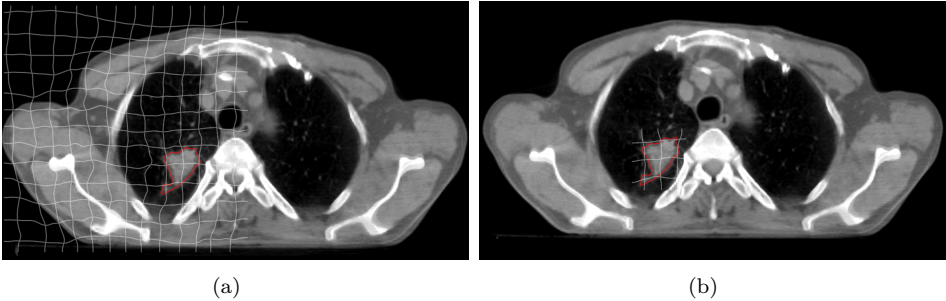


Figure A.5: 两种变形的可变形矢量场表示，a.) DIR\_L 利用肿瘤所在的肺部，b.) 而 DIR\_T 仅使用一个边界框来变形肿瘤周围的区域。

| Noise Level | Normal   |             |            | Speckle |             |            | Periodic  |             |            | Salt and Pepper |             |            |
|-------------|----------|-------------|------------|---------|-------------|------------|-----------|-------------|------------|-----------------|-------------|------------|
|             | Low/High | Diff-UMamba | UMamba-Bot | Scale   | Diff-UMamba | UMamba-Bot | Amplitude | Diff-UMamba | UMamba-Bot | Probability     | Diff-UMamba | UMamba-Bot |
| Level 1     | 0        | 76.91       | 72.64      | 0       | 76.91       | 72.64      | 0         | 76.91       | 72.64      | 0               | 76.91       | 72.64      |
| Level 2     | -2/2     | 76.83       | 70.18      | 0.3     | 76.81       | 73.04      | 0.5       | 76.98       | 71.20      | 0.002           | 74.94       | 15.27      |
| Level 3     | -5/5     | 68.82       | 60.52      | 0.5     | 76.39       | 68.88      | 1         | 76.27       | 66.80      | 0.005           | 68.47       | 0.91       |
| Level 4     | -8/8     | 64.25       | 29.53      | 0.7     | 74.78       | 59.32      | 2         | 70.77       | 20.18      | 0.008           | 63.47       | 0.69       |
| Level 5     | -10/10   | 55.69       | 17.45      | 0.9     | 71.05       | 30.86      | 3.5       | 65.55       | 0.073      | 0.01            | 56.79       | 0.56       |
| Level 6     | -12/12   | 35.25       | 12.36      | 1.1     | 66.03       | 0.07       | 5         | 56.96       | 0.036      | 0.02            | 25.34       | 0.25       |

Table A.4: 与图 ?? 对应的噪声映射表，说明不同噪声类型和强度对 DSC 的影响。



Figure A.6: 肿瘤分割在内部数据集上的视觉比较。所有肿瘤都位于低对比度区域，用 PET-CT 辅助在真实值中勾画出来。a) 一个位于肺外的异常肿瘤。b-e) 位于肺内的肿瘤。在案例 a、b 和 c 中，通过 r-CT 轮廓改善了分割，而在 d 和 e 中没有改善，可能是由于 r-CT 轮廓的错位。