**Configuration**

为了支持模型训练的高计算需求，我们使用了4张NVIDIA RTX 4090 GPU进行实验。每张GPU具有24GB显存，总计提供96GB显存容量，能够高效处理高分辨率的3D腰椎MRI图像数据。训练过程采用分布式并行策略，以加速收敛并优化模型性能。

**Method**

DrSAM3D结合了一个基于U-Net的深度学习模型（LumoDiscNet, LDN）和一个预训练的图像分割模型（SAM 3D），以提高脊椎间盘分割的准确性。LDN提供初步的分割结果，作为SAM 3D的提示，引导SAM 3D进行更精确的分割。

具体流程如下图所示，首先，输入脊柱的MRI图像及其对应的标注信息。随后，这些图像被送入LDN网络进行处理，LDN基于改进的U-Net架构，采用多尺度卷积模块和残差连接，旨在捕捉脊椎间盘的局部特征和全局上下文信息，从而生成一个初步的粗糙分割图像。接下来，该初步分割图像与原始MRI图像作为提示信息（Prompt）输入到Segment Anything Model 3D (SAM 3D）模型中，SAM 3D利用其预训练的庞大三维语义知识库和自适应注意力机制，进一步优化分割边界，特别针对脊椎间盘的非均匀形态和复杂解剖结构进行精细调整。最终，SAM 3D输出高精度的三维分割结果，并通过体素级可视化技术清晰呈现每个脊椎间盘的位置、形状及空间分布。这种结合LDN的粗分割能力和SAM 3D的精细优化优势的混合策略，有效提升了模型的鲁棒性和准确性，尤其在处理MRI图像中的噪声和伪影时表现出色。

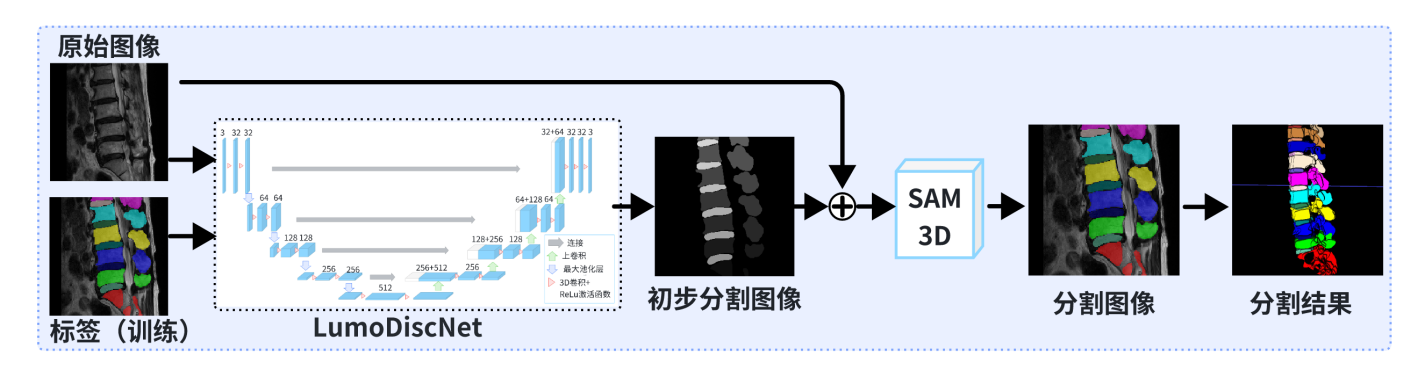


图1 DRSAM架构图

**Model**

LumoDiscNet (LDN) 网络由一个编码器（下采样路径）和一个解码器（上采样路径）组成，并且通过跳跃连接将编码器和解码器中对应的层连接起来。

具体流程如图2所示，基于U-Net架构的图像分割网络采用对称的编码器-解码器结构，通过跳跃连接实现多尺度特征融合。编码器利用卷积和最大池化层提取特征，特征图数量从32增至512；解码器通过上采样和卷积重建图像，特征图数量从512减至32，跳跃连接如256+512拼接保留细节。参数配置包括3个输入通道、32初始特征图、32-512编码器特征图、256-32解码器特征图及3个输出通道。该网络在医学影像分割中展现广泛应用潜力。

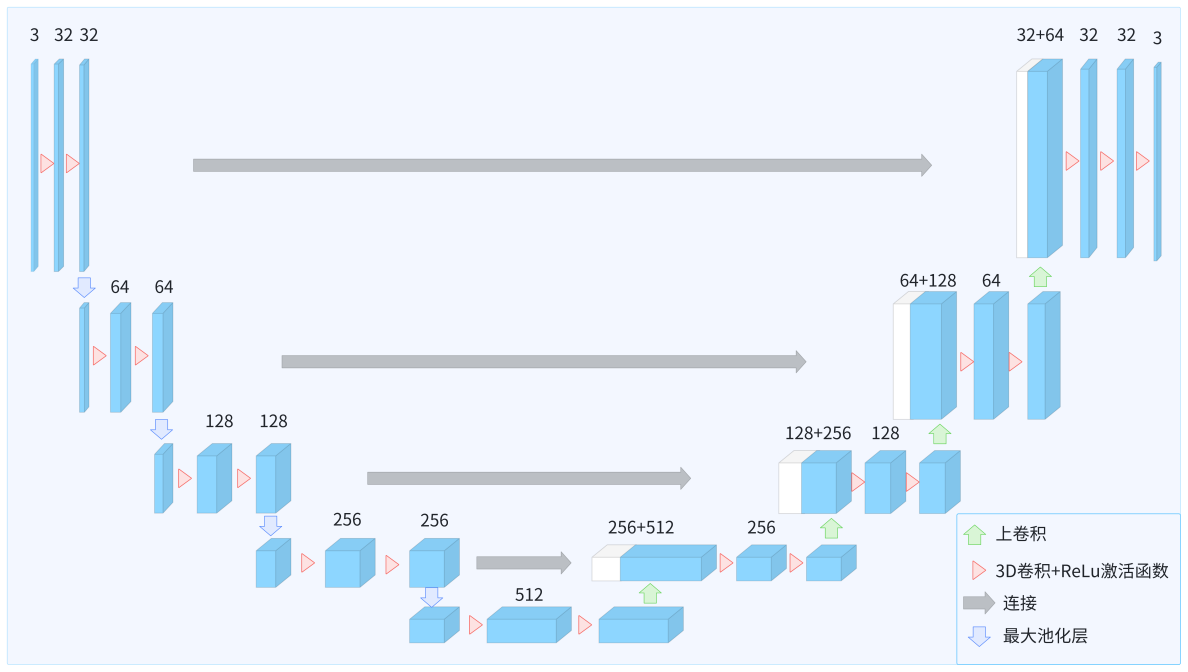


图2 LumoDiscNet结构图

SAM3D的网络结构如图3所示:

该方法的核心在于融合图像特征、用户提示和注意力机制。首先，ViT图像编码器分别从原始与预分割图像中提取全局上下文特征，为后续分割提供基础信息。图中Ori\_img (Original Image)表示原始输入图像；Preseg\_img (Pre-segmentation Image)代表预分割图像。同时，Prompt编码器将用户指定的点提示转换为特征向量，作为分割的引导信号。为了增强特征表示，模型引入了3D块注意力机制，通过计算图像不同区域之间的相关性，突出与分割任务相关的区域，并对特征进行加权。最终，Mask解码器将图像特征和Prompt信息融合，生成精确的分割Mask，实现对图像中目标对象的准确分割。这种方法结合了ViT的全局特征提取能力、Prompt的交互式引导以及注意力机制的特征增强能力，能够有效地提高图像分割的精度和鲁棒性。

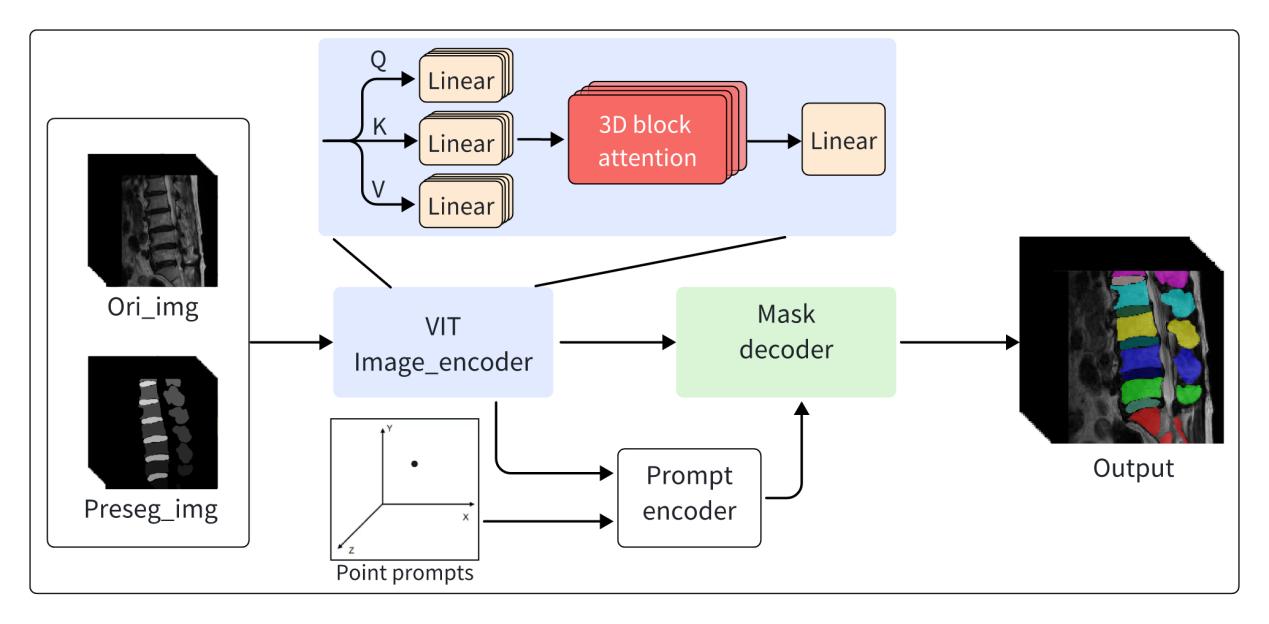


图3 SAM3D结构图

**Dataset**

本研究所使用的数据集来源于脊柱磁共振成像（MR）检查，共包含215例受试者的T2加权三维体数据[1]。所有影像均由临床常规检查采集，覆盖腰椎区域，具有较高的空间分辨率和对比度，能够清晰展示椎体与椎间盘等关键解剖结构。每个体数据均经过标准化预处理，包括强度归一化、体素重采样以及空间对齐，以确保不同个体间的图像具有可比性和一致性。

在数据标注方面，采用经验丰富的影像科医生进行逐层手工标注，生成椎体与椎间盘的精细分割掩膜。最终标注结果涵盖10个椎体和9个椎间盘，包括背景共20个类别，为多类别分割任务提供了可靠的参考标准。为了保证标注的准确性，所有数据均经过多位专家的交叉验证和一致性评估，从而最大程度降低主观差异带来的影响。

该数据集不仅涵盖了正常解剖结构，也包含一定数量的退变、畸形及病理变化病例，具有较高的多样性和代表性。这一特点使得模型在训练过程中能够学习到不同形态下的脊柱特征，从而具备更好的泛化能力和临床应用价值。

**Results**

表1展示了DRSAM3D模型在20个类别上的分割性能。整体而言，模型在多标签三维分割任务中表现良好，平均DICE系数为0.7849，平均IoU为0.6706，说明模型能够在大多数类别中实现较高的空间重叠度和区域一致性。

在类别层面，Class 0取得了最高的分割精度（DICE = 0.9864，IoU = 0.9731），几乎实现了完美分割，表明该类别的几何特征显著且易于区分。Class 1-5的DICE值均在0.83以上，分割性能稳定且较高。部分中等难度类别（如Class 6–8）的DICE值在0.68–0.80之间，存在一定预测偏差。值得注意的是，Class 9的表现显著低于其他类别（DICE = 0.3398，IoU = 0.2147），导致该类别分割效果较差的原因是因为该标签数量较少，模型学习到的知识不足。

表1 DrSAM3D分割指标

|  |  |  |
| --- | --- | --- |
| Class | DICE | IOU |
| Class 0 | **0.9864** | **0.9731** |
| Class 1 | 0.8371 | 0.7204 |
| Class 2 | 0.8449 | 0.7336 |
| Class 3 | **0.8508** | 0.7420 |
| Class 4 | 0.8408 | 0.7269 |
| Class 5 | 0.8427 | 0.7293 |
| Class 6 | 0.8030 | 0.6885 |
| Class 7 | 0.7703 | 0.6614 |
| Class 8 | 0.6803 | 0.5359 |
| Class 9 | 0.3398 | 0.2147 |
| Class 10 | 0.7099 | 0.5503 |
| Class 11 | 0.8195 | 0.6994 |
| Class 12 | 0.8388 | 0.7346 |
| Class 13 | **0.8693** | **0.7706** |
| Class 14 | **0.8724** | **0.7753** |
| Class 15 | **0.8618** | **0.7621** |
| Class 16 | 0.7896 | 0.6922 |
| Class 17 | 0.7642 | 0.6429 |
| Class 18 | 0.7287 | 0.5800 |
| Class 19 | 0.6484 | 0.4798 |
| **Average\_All** | **0.7849** | **0.6706** |

表2展示了各算法的平均DICE和IoU值。实验结果表明，DrSAM-3D模型在整体分割性能上显著优于其余对比算法，包括SAMadapter [2]、MedLSAM [3]、Memorizing SAM [4]、Fastsam3D [5]、3DSAM adapter [6]和Promise [7]。

我们的模型取得了平均DICE 0.7849和IoU 0.6706，超越了所有对比算法。例如，DRSAM的DICE仅为0.2801，IoU为0.2800；SAMadapter的DICE为0.5000，IoU为0.3333；即使与性能较佳的Promise（DICE 0.5956，IoU 0.3342）相比，我们的模型仍表现出约19%的DICE提升，34%的IOU提升，表明我们的模型在综合性能上更平衡。

表2 对比试验结果

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | DrSAM3D | Medical  SAMadapter | MedLSAM | Memorizing  SAM | Fastsam3D | 3DSAM  adapter | Promise |
| Loss | **0.0001** | 0.1007 | 0.7887 | 8.6485 | 8.6997 | 0.4881 | 0.0161 |
| Dice | **0.7849** | 0.2801 | 0.5000 | 0.5332 | 0.5956 | 0.4734 | 0.5009 |
| IoU | **0.6706** | 0.2800 | 0.3333 | 0.3653 | 0.4255 | 0.5540 | 0.3342 |

DrSAM3D实验结果

训练损失曲线如图4所示，从初始值约2.0迅速下降至0.5，并在50个epoch后趋于平稳，最终在200个epoch时接近0，显示出DrSAM3D在前期高效学习数据特征，后期参数收敛的特性。该曲线凸显了模型优越的性能，尤其在分割精度和鲁棒性方面表现突出。

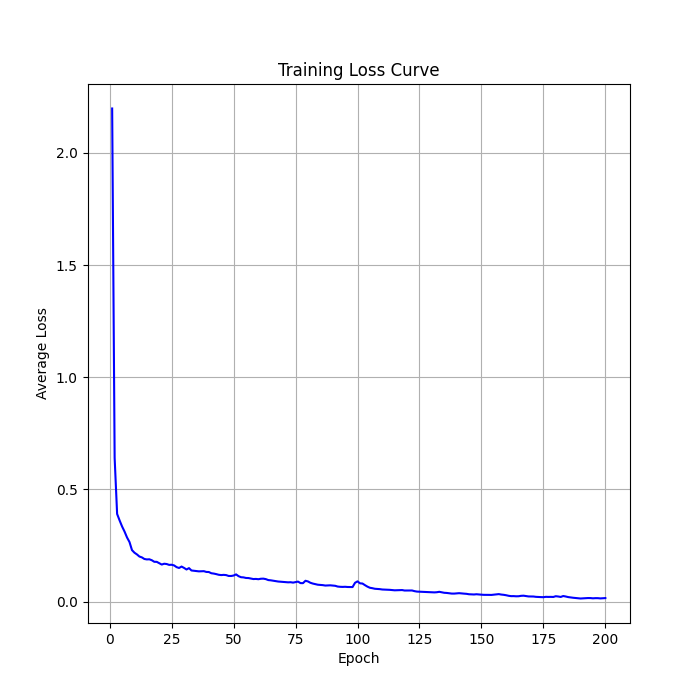


图4 训练损失随epoch变化的曲线图

我们基于验证结果分析了模型性能。Validation Average Dice Curve显示平均DICE值在训练初期快速上升，稳定在约0.7-0.8范围内；Validation Average IoU Curve表明平均IoU值稳定在约0.6-0.7范围内。两者在100 epoch后趋于收敛，反映模型对数据分布的良好适应性。

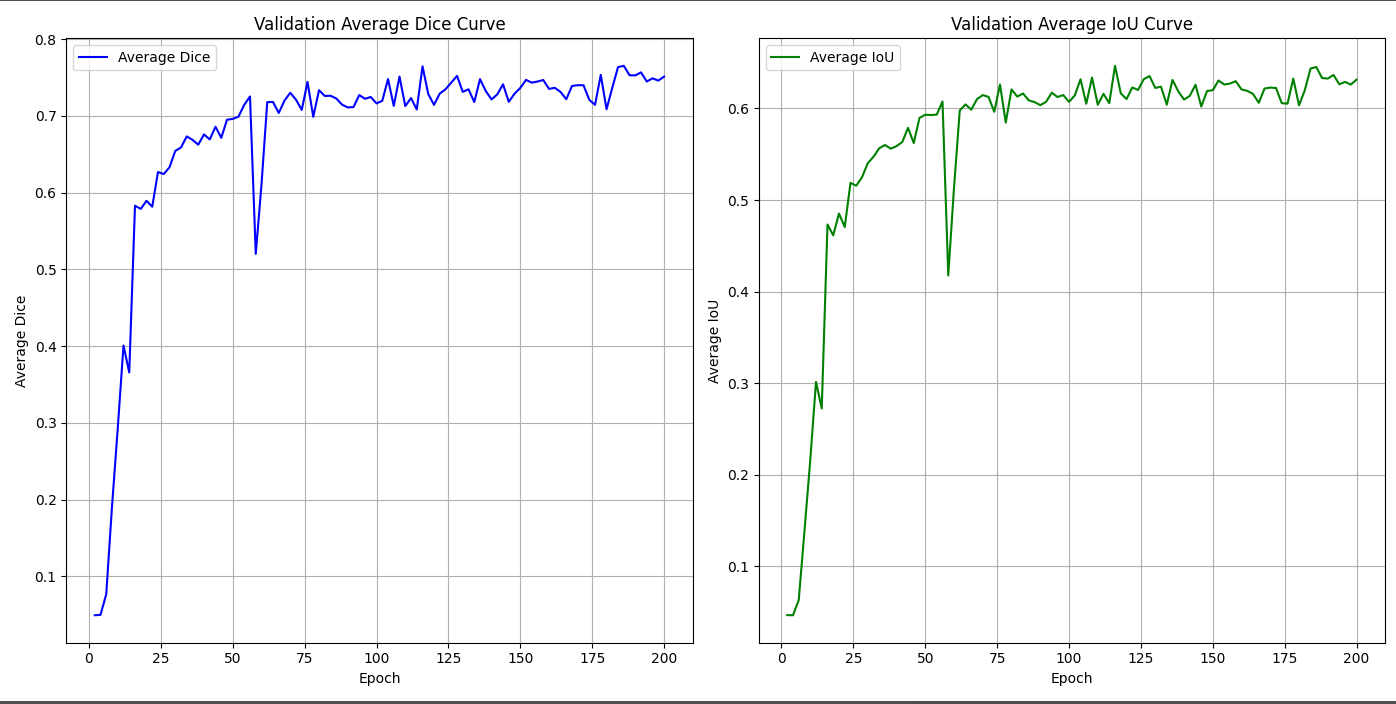


图5 实验平均指标验证曲线

从图5中图像分割结果来看，DrSAM3D的分割结果最接近groundtruth，这得益于其在三维场景中对复杂边界和细微结构的精准捕捉，颜色分布与groundtruth高度一致，边界重合度显著优于其他模型（如Promise、Fastsam3D和Memorizing SAM）。

从实验结果对比来看，DrSAM3D的分割效果整体最接近真实标注（groundtruth）。在脊柱MRI的分割中，DrSAM3D能够较好地保持椎体与椎间盘的完整性和空间连续性，边界清晰，几乎没有明显的断裂或粘连问题。相比之下，Promise模型在分割时存在边界模糊和错分的情况，部分椎体与椎间盘的分隔不够准确，整体形态不如DrSAM3D规整。FastSAM3D的结果则出现了明显的碎片化问题，生成了许多孤立的小块区域，空间一致性较差，导致与真实标注存在较大偏差。Memorizing SAM的表现介于二者之间，整体轮廓相对完整，但仍存在粘连和局部分割不清的问题，尤其在上段椎体和椎间盘的边界部分，精度不及DrSAM3D。综上，DrSAM3D在脊柱MRI分割任务中展现了最优的表现，能够更准确地贴合真实解剖结构，说明其在保持三维结构连续性和提升分割精度方面具有明显优势。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| DrSAM3D | Promise | Fastsam3D | Memorizing SAM | groundtruth |

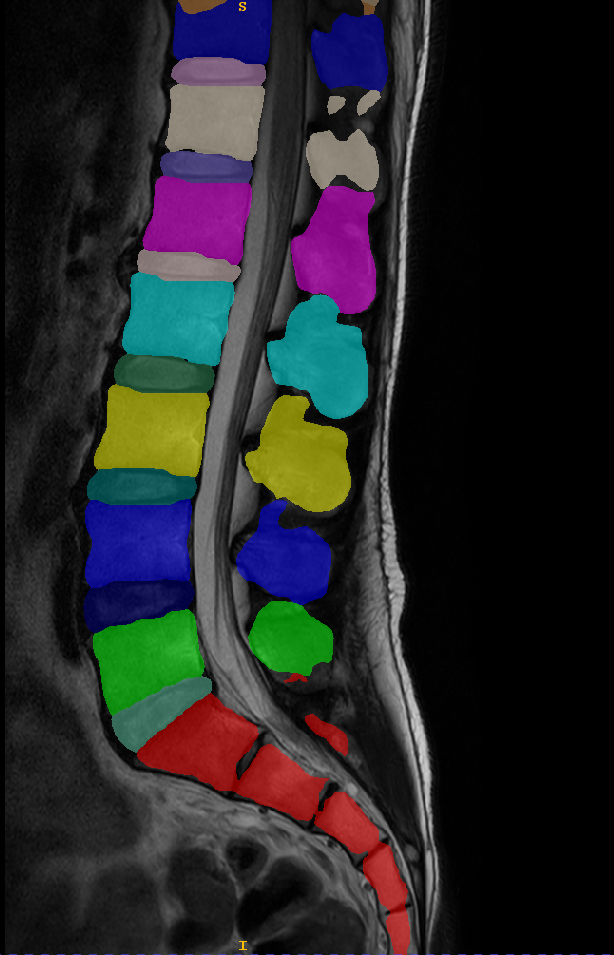
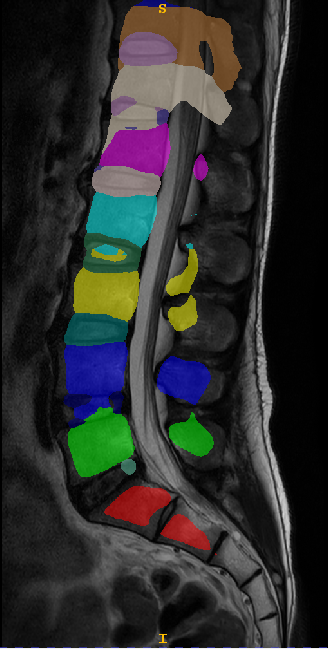
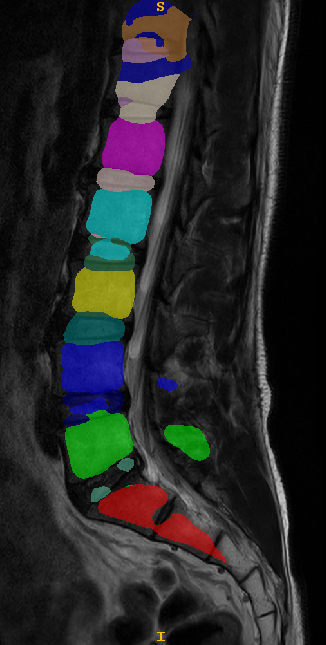
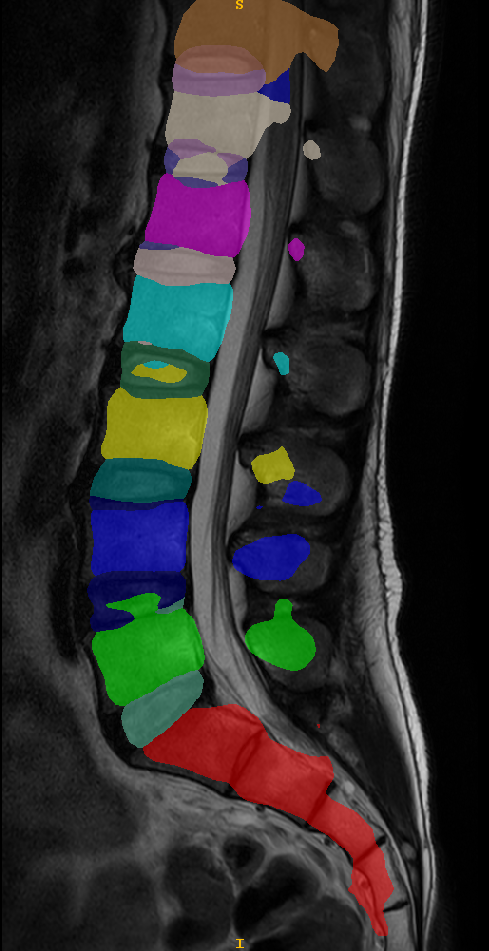
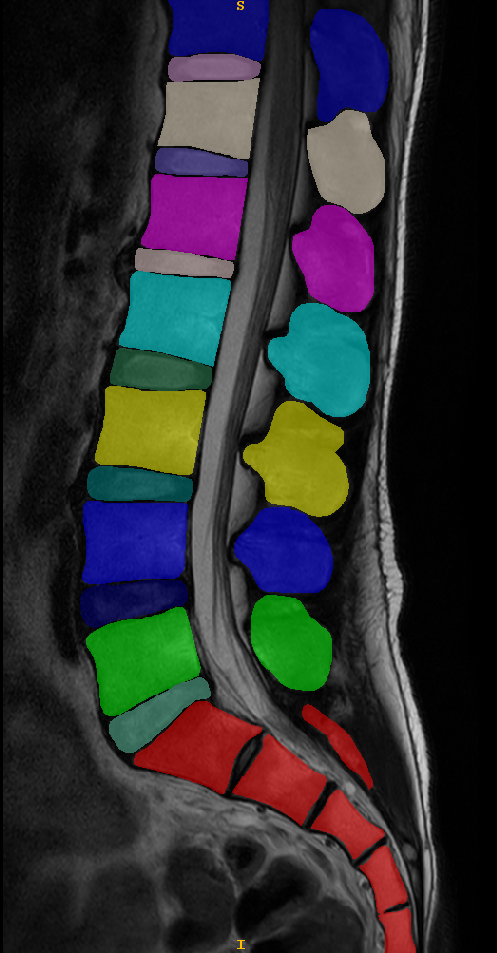
    

图5 不同模型对比推理图

**Discussion**

本研究提出的DrSAM3D模型在多类别三维医学图像分割任务中展现了显著的优势。实验结果表明，该模型在20个类别上的平均DICE系数为0.7849，平均IoU为0.6706，整体性能优于现有的多种主流方法，包括SAMadapter、MedLSAM、Memorizing SAM、Fastsam3D、3DSAM adapter和Promise。尤其值得关注的是，与表现较好的Promise相比，DrSAM3D在DICE上实现了约19%的提升，在IoU上实现了约34%的提升，表明本模型在保持分割精度与鲁棒性方面具有更为均衡的表现。

在类别层面的分析显示，DrSAM3D在大多数类别上均能实现较高的分割精度，其中 Class 0 的分割效果几乎达到完美水平（DICE = 0.9864，IoU = 0.9731），表明该类别具有明显的几何特征，易于模型捕捉。而Class 1–5的DICE值均在0.83以上，分割稳定性良好。然而，对于部分中等难度类别（如Class 6–8），模型的分割精度有所下降，DICE处于0.68–0.80之间，提示这些类别可能存在结构复杂或边界模糊的问题。特别地，Class 9的分割性能显著低于其他类别（DICE = 0.3398，IoU = 0.2147），主要原因在于该类别样本数量不足，导致模型难以充分学习其特征。这一结果揭示了数据分布不均衡对模型分割性能的影响，也为后续研究在小样本类别上的提升提供了方向，例如通过数据增强、类别重加权或少样本学习方法进行优化。

训练曲线进一步验证了DrSAM3D的学习效率和稳定性。损失曲线在前期迅速下降并逐渐收敛，表明模型能够高效提取特征并实现参数优化；验证集上的平均DICE和IoU曲线在 100个epoch后趋于稳定，反映了模型在面对复杂数据分布时的良好适应性和泛化能力。这些结果共同说明，DrSAM3D不仅在整体性能上优于对比算法，而且在训练收敛速度和鲁棒性方面同样具备优势。

尽管DrSAM3D模型在大多数类别上表现出色，但仍存在一些不足。首先，模型在处理某些具有复杂形态或较弱几何特征的类别时，其分割精度相对较低。例如，对于Class 6–8的分割，模型在细节处理上存在一定的欠缺，可能是由于这些类别的解剖结构较为复杂，导致模型难以准确捕捉。其次，Class 9的分割性能差，表明在少样本类别上，模型的学习能力受到限制，可能需要在训练过程中采用更多的数据增强和类别平衡策略。此外，尽管模型在大多数类别上的表现相对稳定，但在面对极为复杂或不规则的结构时，仍可能出现不精确的边界和小区域的断裂。未来的研究可以探索进一步改进模型的精细化分割能力，尤其是在面对低对比度或结构不清晰的类别时。

未来的改进方向

针对上述不足，未来的工作可以从以下几个方面进行改进：

1. 小样本学习：针对Class 9等少样本类别，模型的学习能力有所限制。为了解决这一问题，可以引入少样本学习方法，如迁移学习、数据增强（例如生成对抗网络生成更多样本）、或者使用类别重加权策略，来加强模型在少样本类别上的表现。
2. 改进数据预处理和增强技术：对于具有复杂几何结构的类别（如Class 6–8），可以采用更复杂的数据增强策略，来提高模型的鲁棒性。此外，采用更先进的图像预处理技术，可以帮助模型更好地捕捉细节，减少误分割。
3. 模型结构优化：为了应对复杂解剖结构，未来可以考虑引入更精细的网络结构，例如基于多尺度特征提取的深度学习架构或改进的注意力机制，这可以帮助模型更好地聚焦于不同类别的关键区域，提高分割精度。
4. 集成学习与后处理技术：可以通过集成多种模型的预测结果来提升整体分割性能，特别是在复杂或边界不清晰的类别上。此外，后处理步骤（如形态学操作）可以进一步优化分割边界，减少小区域的断裂现象。

通过上述方法，可以进一步提升DrSAM3D模型在医学图像分割任务中的表现，尤其在面对小样本类别、复杂形态及不规则结构时的处理能力。

**Conclusion**

本文提出的DrSAM3D模型在三维医学图像多类别分割任务中展现出优越性能。通过在20个类别上的系统实验，我们验证了模型在整体精度、空间一致性和鲁棒性上的优势，平均DICE达到0.7849，平均IoU达到0.6706，显著超越了多种现有算法。在类别层面，DrSAM3D能够在大多数类别上保持较高的分割精度，尤其是在几何特征清晰的类别上接近完美分割，但在数据稀缺类别上表现不足，凸显了未来在小样本学习与不平衡数据处理上的研究价值。

综合来看，DrSAM3D在医学图像分割中提供了一种高效、稳定且鲁棒的解决方案，为临床应用中的自动化脊柱结构识别与定量分析提供了技术支撑。未来的工作将进一步聚焦于如何提升在少样本类别上的分割表现，以及探索更具普适性的模型结构，以拓展DrSAM3D 在更多医学影像任务中的应用潜力。

**Reference**

1. Pang S, Pang C, Zhao L, et al. SpineParseNet: spine parsing for volumetric MR image by a two-stage segmentation framework with semantic image representation[J]. IEEE Transactions on Medical Imaging, 2020, 40(1): 262-273.
2. Wu J, Wang Z, Hong M, et al. Medical sam adapter: Adapting segment anything model for medical image segmentation[J]. Medical image analysis, 2025, 102: 103547.
3. Lei W, Xu W, Li K, et al. MedLSAM: Localize and segment anything model for 3D CT images[J]. Medical Image Analysis, 2025, 99: 103370.
4. Shao X, Shen Y, Unberath M. Memorizing SAM: 3D medical Segment Anything Model with memorizing transformer[C]//Medical Imaging 2025: Image Processing. SPIE, 2025, 13406: 7-11.
5. Shen Y, Li J, Shao X, et al. Fastsam3d: An efficient segment anything model for 3d volumetric medical images[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2024: 542-552.
6. Gong S, Zhong Y, Ma W, et al. 3dsam-adapter: Holistic adaptation of sam from 2d to 3d for promptable medical image segmentation[J]. arXiv e-prints, 2023: arXiv: 2306.13465.