# **1. 分割任务的整体流程**

## 1.1.有标签数据训练

### 1.1.1 数据准备

为标记的数据提供一个TaskID（例如Task001），并根据nnUNet的要求对其进行组织，文件结构如下：

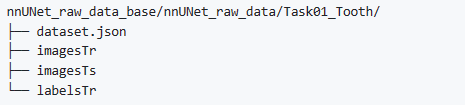


图1：训练数据的文件结构

imagesTr文件夹存放训练的原始图像数据，imagesTs文件夹存放测试的原始图像数据，labelsTr文件夹存放训练的像数据的标签。

### 1.1.2.使用nnUNet进行预处理

这里我们使用默认设置，命令如下：

nnUNet\_plan\_and\_preprocess -t 1 --verify\_dataset\_integrity

### 1.1.3.通过5折交叉验证进行训练

for FOLD in 0 1 2 3 4

do

CUDA\_VISIBLE\_DEVICES=0,1 nnUNet\_train\_DP 3d\_fullres nnUNetTrainerV2\_DP 1 $FOLD -gpus 2 -c --npz

done

## 1.2.为无标签数据生成伪标签

### 1.2.1生成伪标签

nnUNet\_predict -i $INPUTS\_FOLDER -o $OUTPUTS\_FOLDER -t 2 -m 3d\_fullres --save\_npz

### 1.2.2.迭代训练模型并生成伪标签

* 给出一个新的TaskID（例如Task002），并如上所述组织标记数据和伪标记数据。
* 使用nnUNet进行自动预处理：

nnUNet\_plan\_and\_preprocess -t 2 --verify\_dataset\_integrity

* 通过所有训练数据训练新的nnUNet模型

for FOLD in 0 1 2 3 4

do

CUDA\_VISIBLE\_DEVICES=0,1 nnUNet\_train\_DP 3d\_fullres nnUNetTrainerV2\_DP 2 $FOLD -gpus 2 -c --npz

done

* 为无标签数据生成新的伪标签。

### 1.2.3.过滤低质量伪标签。

通过select\_pseudo\_label.ipynb脚本，我们比较不同轮次中的伪标签，并筛选出具有高变体的标签。

## 1.3.通过5折交叉验证对nnUNet-att进行训练

### 1.3.1.将此repo中的以下文件复制到nnUNet环境中。

为了提高模型的性能，我们修改了nnunet源代码的generic\_UNet.py，构建了nnUNet-att模型。

./nnunet/network\_architecture/generic\_UNet.py

### 1.3.2.通过5折交叉验证对nnUNet-att进行培训

for FOLD in 0 1 2 3 4

do

CUDA\_VISIBLE\_DEVICES=0,1 nnUNet\_train\_DP 3d\_fullres nnUNetTrainerV2\_DP 3 $FOLD -gpus 2 -c --npz

done

### 1.3.3.将“\_000”后缀添加到测试集数据的文件名中

pyhton add\_suffix.py

### 1.3.4.使用nnUNet-att进行高效推理

nnUNet\_predict -i $INPUTS\_FOLDER -o $OUTPUTS\_FOLDER -t 3 -m 3d\_fullres --save\_npz

## 1.4.通过位置校正模块对异常分割点进行校正

python position\_correction.py

# **2. 部署实验的细节**

## 2.1 模型架构

### 2.1.1 整体框架



图1：我们的整体框架图。总体架构遵循U-Net架构。

在我们的研究中，我们精心选择了nnUNet[1]作为我们的backbone，这是因为nnUNet在医学图像分割任务中表现出了卓越的性能。然而，我们并不满足于简单地采用它，而是对其进行了深入的改进，以更好地适应我们的特定任务。我们的改进焦点主要集中在nnUNet的第二层解码器上，因为我们深知这一部分对于分割性能的重要性。我们决定在第二层解码器中引入了3D轴向注意力机制[2]，这是一项重要的创新。这个决策的核心动机在于，我们希望模型能够更加聚焦于牙齿结构的关键细节。为了实现这一目标，我们利用了3D轴向注意力机制，它在处理高维数据时具有出色的计算和存储效率，并能够捕捉局部和全局依赖关系。这意味着我们的模型能够更精确地识别和定位复杂的牙齿结构，如牙根和牙冠，从而实现更高精度的分割结果。图1展示了我们整体框架的结构。

### 2.1.2 backbone:

nnUNet是一种专为医学图像分割任务设计的神经网络架构。该模型继承了U-Net的基本结构，包括下采样编码器和上采样解码器，但引入了多项创新和优化以提高性能和适应性。其中最显著的特点是其自适应性：nnUNet能够根据不同的数据集和任务需求自动调整其网络架构和超参数。这一特性大大减少了模型调优的复杂性，使其能够在多个医学图像分割挑战中实现最先进的性能。除了自适应性，nnUNet还采用了一系列精心设计的训练策略，包括数据增强、损失函数选择和优化算法，进一步提高了模型的鲁棒性和准确性。由于这些优点，nnUNet已经成为医学图像分割领域的一个重要基准，为研究人员提供了一个高效、可靠的解决方案。

### 2.1.3 3D轴向注意力机制

3D轴向注意力机制是我们创新思路的关键组成部分。这个机制允许模型在进行分割时更好地关注图像中的水平和垂直结构，这对于3D牙齿分割尤为重要[3]。3D轴向注意力机制通过考虑不同方向的特征来调整解码器中的特征图。这有助于提高模型对细微结构的捕捉能力，从而改善分割的准确性。

我们对解码器上采样部分添加3D轴向注意力组件，具体如图2所示。3D轴向注意力机制的工作原理是通过分解3D空间为多个单一维度或“轴”来降低计算复杂性。因此，模型可以分别对每一个轴进行注意力计算，而不是在整个3D空间中进行。对于一个形状为的输入图像, 具体算法的迭代过程如下：

1. **轴选择与切片提取**：首先，我们选择一个轴（如深度轴）。然后，我们提取该轴上的所有切片，每一个切片都是一个 形状的2D张量。
2. **注意力权重计算**：对于每一个2D张量切片，我们使用标准的多头注意力机制[4]来计算查询（Query）、键（Key）和值（Value）矩阵。注意力权重通过以下公式得出：
3. **特征图更新：**使用计算出的注意力权重，更新每一个2D张量切片的表示。
4. **轴合并：**完成一个轴的注意力计算后，重新组合更新后的切片以形成一个新的3D张量。
5. 对其他轴重复1-3操作。

请注意，我们尝试过使用通道-空间注意力（Convolutional Block Attention Module，CBAM）[4]来增强模型的表征能力，但是通道-空间注意力机制会在整个空间和通道上计算注意力权重，这在处理高维数据时会导致更高的计算和存储需求。其次，CBAM主要关注通道和空间的全局依赖，可能会忽视不同轴之间的局部依赖关系。由于牙齿区域是集中且形状不规则，仅仅是全局依赖不能精细化处理边缘和凹陷等区域，所以我们需要关注更广泛的轴维度信息。



图2：上采样图示。我们采用3D轴向注意力与残差连接，增强表征能力同时增强梯度流。

### 2.1.3 位置修正模块

3D CBCT图像扫描范围的特殊性是我们研究中需要仔细处理的关键问题。这些图像包括了牙齿附近的骨性结构，其中一些结构的形态和密度与牙齿非常相似，例如枢锥齿状突[6]。这种相似性会导致分割模型容易错误地将这些非牙齿结构误认为是牙齿并进行分割，从而产生错误的分割结果。图3中展示了一个实际情况，其中枢锥齿状突也被误分割出来。



图3：枢锥齿状突被模型错误分割（蓝色框部分）

为了解决这个问题，我们进行了数据集分析，并咨询了相关的牙科专家。我们了解到，在进行CT扫描时，被试的头部位置是固定的，并且非牙齿的骨性结构通常位于牙齿较远的位置。因此，精确检测牙齿的范围对于排除错误的分割项至关重要。为了确保准确性，我们进行了统计分析，研究了训练集中已标记数据中牙齿的分割范围。在训练集中，我们发现牙齿的最大宽度（即，x轴方向的长度）为305个单位。通过与牙科医生的咨询，我们了解到牙齿的大小存在个体差异。为了保守起见，我们将牙齿的检测宽度w设置为320个单位，而长度l（即，y轴方向的长度）设置为宽度的0.8倍，即256个单位。

为了解决错误分割的问题，我们特意设计了一个位置修正模块，该模块经过精心构建，包括以下4个关键步骤，以确保我们的分割结果更加准确和可靠：

1. **点A和点B的确定**：首先，在横断面图像中，我们需要准确定位已分割出的牙齿结构的位置。我们在图像中找到最左侧的点A（坐标为：(x0, y0, z)）和最上方的点B（坐标为：(x1, y1, z)），如图4所示。
2. **点C的计算**：接下来，我们进行点C的计算，以便准确修正牙齿的位置。点C位于点A和点B连线的交点，其中点A连线朝上方延伸，点B连线朝左侧延伸。点C的坐标可以由点A和点B的坐标计算得出，通常为(x0, y1, z)。
3. **牙齿区域框的构建**：基于我们预定义的牙齿宽度和长度参数，我们从点C开始沿着x轴和y轴方向构建牙齿的区域框，如图4所示。这个框将准确地界定出牙齿的范围。
4. **数据筛选**：最后一步是将位于牙齿区域框内的分割数据保留下来，而将不在框内的数据丢弃。这个步骤确保了只有真正属于牙齿结构的数据被保留，从而有效地解决了错误分割的问题。

这个详细的位置修正模块的设计和操作流程，充分考虑了牙齿结构的特殊性，确保了我们的分割模型能够准确地定位和分割出牙齿，同时排除了误分割非牙齿结构的可能性。

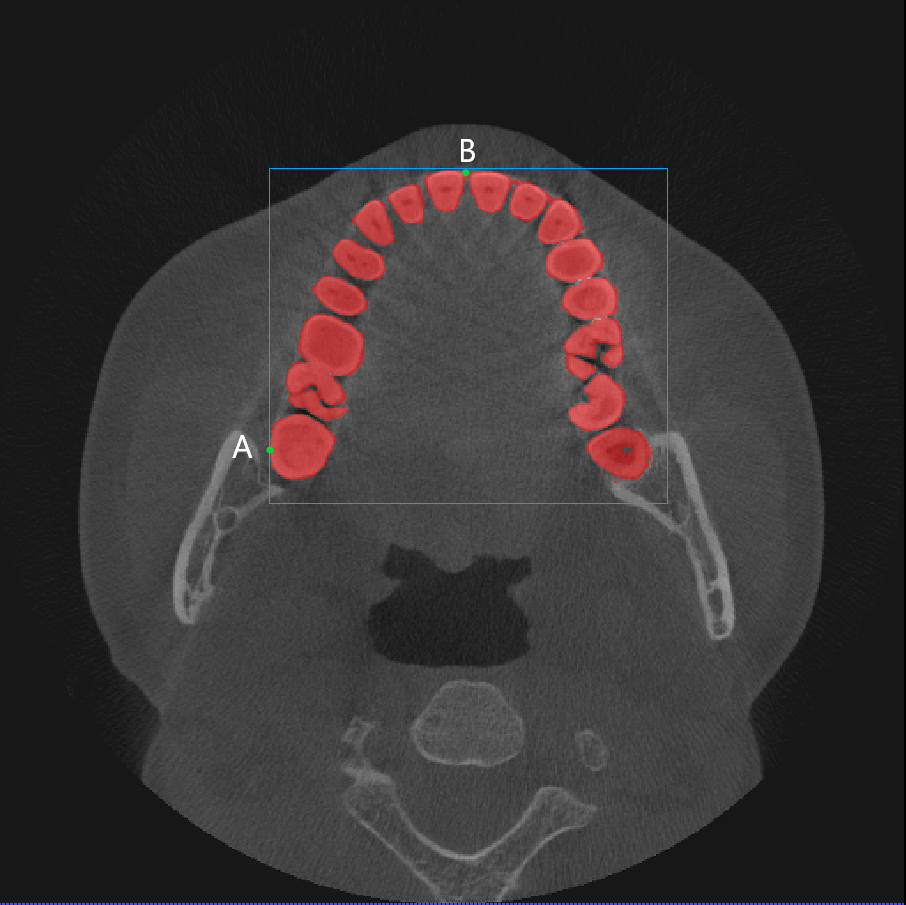


图4：构建的牙齿区域框

## 2.2 伪标签方法

为了最大程度地利用未标记数据来训练模型，我们采用了一种复杂而高效的伪标签生成策略[7][8]，以确保伪标签的质量和模型性能的提高。以下是我们的伪标签策略的详细步骤，这些步骤被精心设计，以充分利用未标记数据来提高模型性能：

1. **基于5折交叉验证的训练**：首先，我们对已标记的数据进行5折交叉验证，以训练5个独立的nnU-Net模型，以确保模型具有多样性。
2. **高质量伪标签的生成**：在伪标签生成阶段，我们使用我们精心设计的5折nnU-Net集合进行精确的推理设置，以预测未标记数据上的一个伪标签。这个步骤确保了伪标签的高质量和准确性。
3. **迭代训练和伪标签更新**：接下来，我们在带有伪标签的标记数据和未标记数据的并集上进行迭代训练一个nnU-Net模型，并在每轮训练后生成新的一个伪标签。这个过程有助于不断提升伪标签的质量和模型性能。
4. **伪标签的稳定性选择**：在伪标签生成的过程中，我们根据伪标签在不同训练轮次中的稳定性来选择伪标签。
5. **最终模型训练**：最终，我们在标记数据和选择的带有伪标签的未标记数据的并集上进行训练，采用了3D轴向注意力的nnUNet（nnUNet-Att）。引入3D轴向注意力的目的是更好地区分伪标签的干扰，以确保模型在最终评估中能够准确地分割牙齿结构。

这一复杂的伪标签生成策略结合了多个关键步骤，旨在提高模型性能并确保伪标签的质量。通过精心设计和实施这些步骤，我们能够更有效地利用未标记数据，以提高医学影像分割模型的准确性和鲁棒性。

## 2.3. 实现细节

我们的开发环境详细列在表1中，以确保工作的可重复性和可验证性。为了训练我们的模型，我们在4个不同的服务器上执行了不同折的交叉验证。在模型构建方面，我们严格遵循了nnU-Net的标准训练方法，每个模型都经过了5折交叉验证的训练。在整个训练过程中，我们还采用了动态的数据增强策略，以增加模型的泛化性能。这些数据增强技巧包括随机旋转和缩放、弹性变形、加性亮度增强以及伽玛缩放。

**表1** 开发环境

|  |  |
| --- | --- |
| 系统版本 | Ubuntu 20.04.2 LTS | Ubuntu 20.04.2 LTS | Ubuntu 18.04.6 LTS | Ubuntu 18.04.6 LTS |
| GPU(数量和类型) | 1NVIDIA GTX 3090Ti 24GB|2NVIDIA V100 24GB|2NVIDIA GTX 3090Ti 24GB |1NVIDIA 3090 24GB |
| CUDA版本 | 11.4|11.4|11.4|11.3 |
| 编程语言 | Python 3.9.16 |
| 深度学习框架 | Pytorch (Torch 1.11.0) |
| 特定的依赖关系 | nnU-Net 1.7.1 |
| 代码 | https://github.com/qpuchen/nnUNet\_att\_position\_correction/ |

我们选择了将Dice损失和交叉熵损失结合使用作为总损失函数，因为复合损失函数在多种医学图像分割任务中被证明具有鲁棒性。不同于样本级别的骰子损失，我们采用了批次级别的骰子损失，将整个批次视为一个样本，而不是对每个minibatch中的每个样本计算骰子分数的平均值。这种批处理骰子损失有助于稳定训练，通过减少对带有少量注释样本的样本的误差。在优化方面，我们使用随机梯度下降进行网络的训练，采用Nesterov动量，动量值为0.99。初始学习率设置为0.01，并按照多项式调度进行衰减。每次训练持续3000个epoch，每个epoch包含250个小批次。我们使用当前折交叉验证集上的骰子分数来实时监控训练进度，以确保模型在训练过程中达到期望的性能。我们的研究中开发了三种不同的模型，它们分别是：

* nnUNet-base：这是原始的nnUNet模型，没有任何特殊的修改或增强。
* nnUNet-Att：在nnUNet-base的基础上，生成伪标签，并在nnUNet的解码器的第二层后添加了3D轴向注意力机制。该模型在标记数据和选择的带有伪标签的未标记数据的并集上进行训练。
* nnUNet-Att-p：在nnUNet-Att的基础上，进一步添加了位置修正模块，以对预测的分割异常点进行修正。

# **3. 结果和讨论**

我们的方法在MICCAI Workshop的3D牙齿分割任务中取得了显著的成绩。我们的模型在测试集上表现出色，与其他竞争方法相比，具有很好的分割准确性和鲁棒性。

**表2** 不同模型的的预测性能比较

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Dice | IOU | Hausdorff distance | Score |
| nnUNet-base | 0.6669 | 0.7497 | 0.2921 | 0.7041 |
| nnUNet-Att | 0.7556 | 0.8045 | 0.2521 | 0.7680 |
| nnUNet-Att-p | **0.7959** | **0.8307** | **0.1922** | **0.8099** |

表2显示了交叉验证的三种模型的在3D牙齿分割的预测性能。模型nnUNet-base在整个训练数据集（1折）上进行训练。从表2中可以看出， 相对于nnUNet-base，nnUNet-Att在所有指标上都取得了显著的提升，一方面和nnUNet-base只进行了一折的训练有关，当然主要是伪标签和3D轴向注意力机制的添加发挥了重要作用。如图5所示，因为比赛给出的无标签数据是手工裁剪的仅包含部分牙齿的数据，对这些数据生成伪标签，能够很大的扩充样本量，对提取牙齿的轮廓、纹理和形状是十分重要的。然而，伪标签是一种简单而常规的半监督学习方法，但即使在基于不确定性的伪标签选择之后，伪标签仍然可能是有噪声的。因此我们添加了3D轴向注意力机制使得模型在处理垂直和水平结构时更关注牙齿的特定结构，忽略噪声信息。伪标签和3D轴向注意力的结合使得模型的性能得到了很大的提高。

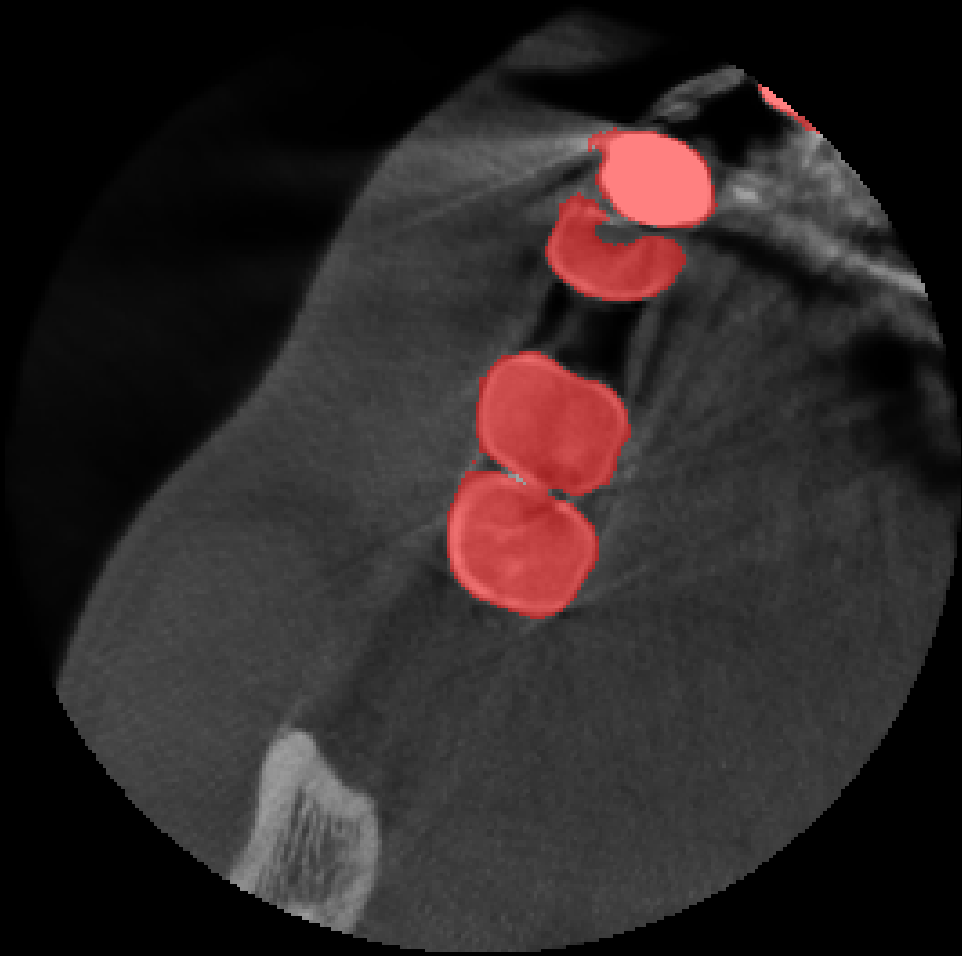


图5：对无标签的数据的分割样例

nnUNet-Att-p模型相对于nnUNet-Att模型在所有四个评估指标下都表现出明显的性能提升，Dice Score提升了约4.03%，IOU提升了约2.62%，Hausdorff距离改善了约5.99%，而总分数提升了约4.19%。正如我们2.1.3分析的，枢锥齿状突等的形态和密度与牙齿相似的部位使得模型错误地将它们误认为是牙齿而进行分割，我们设计的位置修正模块的添加纠正了这些错误的分割，使用位置修正模块前后的结果如图6所示。实验结果表明nnUNet-Att-p模型在3D牙齿分割任务中相对于nnUNet-Att模型具有更高的准确性和精确性，为分割任务提供了更优越的解决方案。



a) 使用位置修正模块前的分割示例 b) 使用位置修正模块后的分割示例

图6：使用位置修正模块前后对比图

我们设计了一个基于nnU-Net的框架，使用未标记的数据进行训练并有效地进行推理，通过修改nnUNet的第二层解码器并引入3D轴向注意力机制来提高分割性能。我们的训练技巧和创新思路在该任务中取得了显著的成功，并有望在医学图像分割的其他领域得到应用。我们相信我们提出的框架可以作为半监督学习和医学图像分割高效推理的良好基线。 我们将继续改进和优化我们的方法，以进一步提高医学图像分割的性能和效率。

本文作者声明，我们为参与MICCAI Workshop 2023挑战而实施的分割方法没有使用任何预训练模型，也没有使用除组织者提供的数据集之外的其他数据集。提出的解决方案是全自动的，不需要任何人工干预。

# **4. 参考文献**

1. Isensee, Fabian, et al. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation." *Nature methods* 18.2 (2021): 203-211.
2. Ho, Jonathan, et al. "Axial attention in multidimensional transformers." *arXiv preprint arXiv:1912.12180* (2019).
3. Luu, Huan Minh, and Sung-Hong Park. "Extending nn-UNet for brain tumor segmentation." *International MICCAI Brainlesion Workshop*. Cham: Springer International Publishing, 2021.
4. Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
5. Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." *Proceedings of the European conference on computer vision (ECCV)*. 2018.
6. Akobo, Seleipiri, et al. "The odontoid process: a comprehensive review of its anatomy, embryology, and variations." *Child's Nervous System* 31 (2015): 2025-2034.
7. Lee, Dong-Hyun. "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks." *Workshop on challenges in representation learning, ICML*. Vol. 3. No. 2. 2013.
8. Huang, Ziyan, et al. "Revisiting nnU-net for iterative pseudo labeling and efficient sliding window inference." *MICCAI Challenge on Fast and Low-Resource Semi-supervised Abdominal Organ Segmentation*. Cham: Springer Nature Switzerland, 2022. 178-189.