

牙齿分割前的先验知识预处理

摘要：近年来，nnU-Net 成为备受欢迎的医学图像分割基线，然而，在实际临床应用中也需要考虑如何有效利用未标记的医学图像数据。本文基于原始的 nnU-Net 网络结构，提出了一种半监督方法，该方法通过结合先验知识的预处理和对伪标签的迭代优化，以实现牙齿的精确分割。我们的研究分析表明，形态学的开闭操作作为先验知识最为数据集的预处理在比赛中发挥了关键作用。在初赛阶段，我们利用经过预处理的 12 个带标签的数据进行了网络训练，并用其权重进行无标签数据的预测，随后再对网络进行重新训练。这一策略使我们在初赛中取得了 0.9381 分数。在复赛阶段，我们首先利用初赛获得的权重对新增的无标签数据进行 ROI 限定，然后沿用初赛的网络结构和迭代方法进行训练。最终，我们取得了 0.8427 的分数。

关键词：分割、先验知识、半监督、伪标签。

1.方法

我们在研究中选择了 nnU-Net^[1]作为我们的骨干网络，整个训练框架如图 1 所示。在医学图像分割领域的以往工作中，通常需要根据不同的数据集来调整网络结构和参数。这一过程不仅依赖于研究者的经验，而且通常需要进行大量的实验，结果往往难以成功复现。以前的研究往往将重点放在新的网络结构设计上，但这些结构往往是基于特定的训练数据经验性调整的，缺乏泛化能力并容易导致过拟合。然而，我们的研究中的关键创新在于，我们合理设计了整个训练过程，在网络训练前分析数据并添加先验知识的预处理，这种方法减少了对网络结构和参数的过多调整，使得模型更具泛化性能。通过合理的训练框架设计可以充分发挥 nnU-Net 的潜力，为医学图像分割任务提供了一种更为稳健和可靠的解决方案。这对于医学图像分割领域的研究和应用具有重要意义。

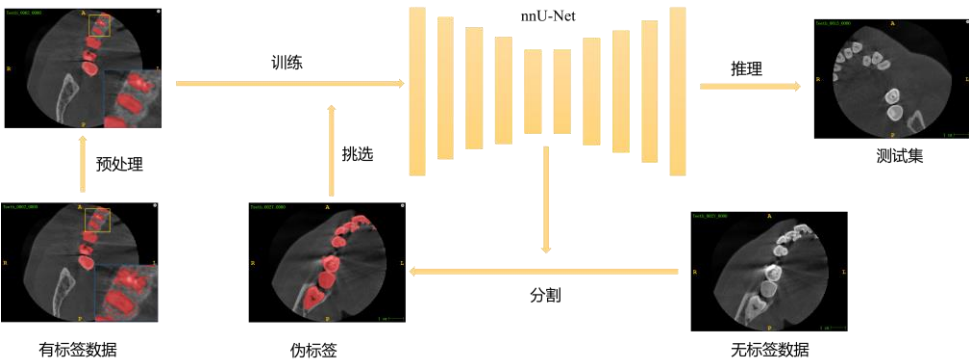


图 1.整体框架

1.1 预处理

在我们的研究中，除了基于 nnU-Net 的网络架构，我们还引入了形态学的开闭操作和连通域分析等先验知识，因为在医学图像分割领域，这些先验知识具有极其重要的价值。这些知识的应用体现在多个方面，

其中之一是提高算法的准确性。由于医学图像通常包含复杂的解剖结构和组织，形态学操作和连通域分析等先验知识可以指导算法更精确地捕捉目标结构，从而显著降低分割错误的风险。另一方面是减少对大规模标注数据的依赖。在医学领域，获取高质量的标注数据通常非常昂贵且耗时。然而，充分利用形态学的开闭操作和连通域分析等先验知识可以在一定程度上替代对大规模标注数据的需求，从而显著降低了训练成本。通过基于这些先验知识优化模型生成的伪标签，我们能够更有效地利用有限的标记数据，并且更好地适应不同的医学图像情境，进一步提高了模型的性能和鲁棒性。这一策略不仅降低了成本，还符合深度学习的数据驱动原则，使得模型更具实际应用的可行性。

1.2 相关配置

为了提高网络的分割精度，我们在 nn-UNet 的 3D_fullres 默认配置基础上进行了一组新的配置设计。这些配置的调整包括使用更大的输入切片尺寸以及增加批处理大小。考虑到本次比赛分为初赛和复赛两个阶段，而这两个阶段所使用的数据集具有较大的差异，因此我们为初赛和复赛的数据集分别进行了不同的配置，具体如表 1 所示。

表 1.输入数据具体设置

| 配置 | 初赛 | | 复赛 | |
|------------------|---------------|---------------|---------------|---------------|
| | default | our | default | our |
| input patch size | (112,160,128) | (112,160,160) | (160,96,160) | (160,96,160) |
| input batch size | 2 | 4 | 2 | 4 |
| input spacing | (0.3,0.3,0.3) | (0.3,0.3,0.3) | (0.3,0.3,0.3) | (0.3,0.3,0.3) |

1.3 伪标签生成

我们采用了简单有效的伪标签生成方法，方便利用无标签的数据来训练我们的模型。生成伪标签的具体方法如下：

1. 在预处理后的标签数据基础上训练一个五折交叉验证模型。

2. 使用第一步训练好的网络模型对无标签数据进行预测。

3. 结合有标签数据与无标签数据重新训练一个网络模型。

在训练过程中我们使用了 Dice Loss 和 Cross Entropy Loss 的复合损失函数。

2. 实验

2.1 初赛阶段

方案一：首先仅使用有标签的 12 个数据进行训练，得到的 test 结果为 0.9229。

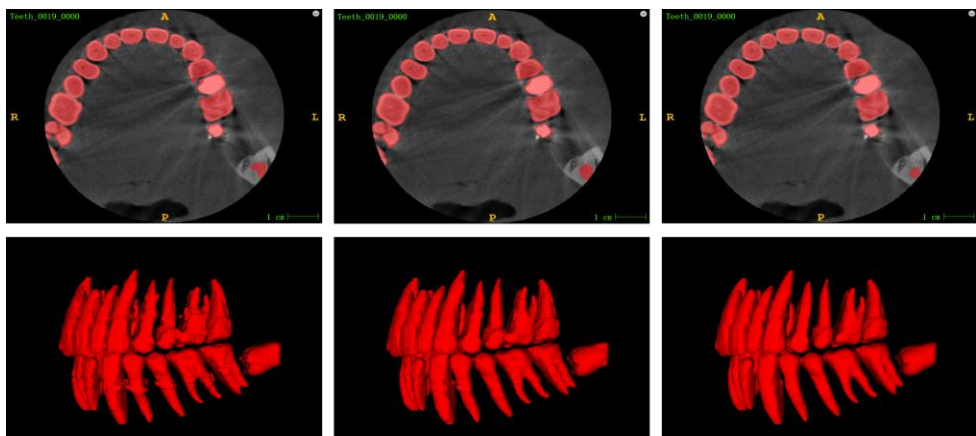
方案二：通过分析发现标签数据表明锐化严重，且存在噪点，并不符合牙齿表面应该是光滑的医学先验知识，因此我们考虑使用形态学开闭的以及去除连通域的操作对 12 个有标签训练集进行微调并重新训练，得到了 0.9300 的结果。

方案三：后续使用该模型生成了无标签数据的伪标签，同时参考^[2]对伪标签进行了不确定性检测的筛选。将所有选择的无标签数据预测为一个新的数据集进行迭代训练，最终结果是 0.9381。具体结果如表 2 所示。

表 2.初赛阶段不同策略的结果

| 评价指标 | 方案一 | 方案二 | 方案三 |
|-------|--------|--------|--------|
| dice | 0.9081 | 0.9191 | 0.9300 |
| iou | 0.9136 | 0.9235 | 0.9330 |
| hd | 0.0479 | 0.0488 | 0.0462 |
| score | 0.9229 | 0.9300 | 0.9381 |

为了更加直观的看到各个方案之间的比较，我们将分割结果进行了三维可视化并展示出来，结果如图 2 所示。



方案一

方案二

方案三

图 2.不同方案结果对比

2.2 复赛阶段

方案一：首先使用初赛 score 最高的模型直接对复赛的 test 数据进行预测，得分为 0.7694。

方案二：通过分析发现，初赛与复赛的 test 数据集差异较大，初赛数据主要是 $[200,266,400] \times [200,266,400] \times [200,300]$ ，spacing 为 $[0.3,0.3,0.3]$ 而复赛数据则均为 $512 \times 512 \times 400$ ，spacing 则为 $[0.25,0.25,0.25]$ 。我们对复赛的 test 数据进行了限定 roi 区域的后处理，得到的 0.8259 的分数。

方案三：后续我们对复赛提供的无标签数据的伪标签进行了 roi 区域的限定，然后和初赛选出的无标签数据一起进行迭代训练，最终得分是 0.8427。结果如表 3 所示。

表 3.复赛阶段不同策略的结果

| 评价指标 | 方案一 | 方案二 | 方案三 |
|-------|--------|--------|--------|
| dice | 0.7659 | 0.8162 | 0.8343 |
| iou | 0.8115 | 0.8455 | 0.8583 |
| hd | 0.2679 | 0.1806 | 0.1615 |
| score | 0.7694 | 0.8259 | 0.8427 |

复赛阶段不同方案的三维可视化结果如图 3 所示。

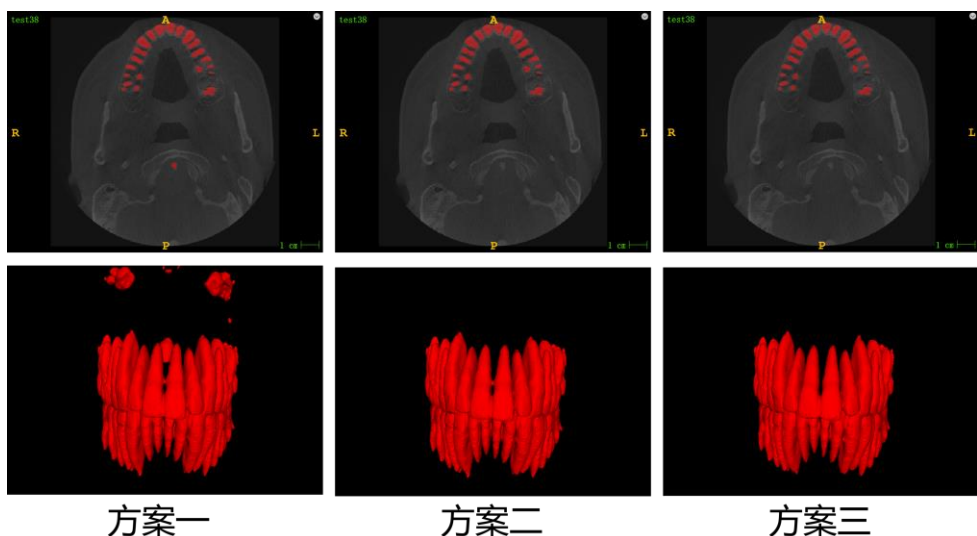


图 3.不同方案结果对比

2.3 实现细节

环境设置: 我们基于 pytorch 对 nnU-Net 进行的二次开发。所有的训练过程都是在 Linux 系统中进行的，所有模型都是从头开始训练的。有关开发环境的详细信息如表 4 所示。

表 4.实验环境配置

| 参数 | |
|----------------------|--|
| Linux version | CentOS Linux release 7.9.2009 |
| CPU | Intel(R) Xeon(R) Gold 6338 CPU @ 2.00GHz |
| RAM | 16GB;3200MT/s |
| GPU | A100-PCIE-40GB |
| CUDA version | 11.7 |
| Programming language | Python 3.9.16 |
| Framework | Pytorch(torch2.0.1,torchvision0.15.2) |

训练方案: 所有的训练数据均基于其相应的数据集进行各自的预处理，预处理用到的方法都是 nnU-Net 默认的处理方法。我们的训练策略结合了 Dice 损失和交叉熵损失函数，初始学习率设置为 0.01，训练轮数为 1000，适用于所有实验。详细设置如表 5 所示。

表 5.训练方案

| | 参数 |
|------------------------|---|
| Network initialization | nnU-Net default |
| Optimizer | SGD with nesterov momentum ($\mu = 0.99$) |
| Total epochs | 1000 |
| Initial learning rate | 0.01 |
| Training time | 24/48 hours |

3 总结

在本研究中，我们在 nnU-Net 的基础上巧妙地融合了医学领域的先验知识，有效地利用了未标记数据参与模型训练，从而显著提升了模型在 MICCAI 2023 挑战赛的 3D CBCT 牙齿分割任务中的性能。通过充分利用医学领域的专业知识和图像处理技术，我们的方法不仅提高了牙齿分割的准确性，还减轻了对大规模标记数据的依赖。这不仅有助于提高工作效率，降低了成本，还有望加速牙科 CAD 技术的进步，从而为口腔医学领域的诊断和治疗提供更可靠和高效的工具。

参考文献

- [1] Isensee F, Jaeger P F, Kohl S A A, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature methods, 2021, 18(2): 203-211.
- [2] Huang Z, Wang H. Revisiting nnUNet for Iterative Pseudo Labeling and Efficient Sliding Window Inference in Abdominal Organ Segmentation[J].