Prior knowledge preprocessing before tooth segmentation

Abstract: In recent years, nnU-Net has become a popular baseline for medical image segmentation. However, in practical clinical applications, it is also necessary to consider how to effectively utilize unlabeled medical image data. This paper proposes a semi-supervised method based on the original nnU-Net network structure. The method achieves precise segmentation of teeth by combining prior knowledge preprocessing and iterative optimization of pseudo-labels. Our research analysis shows that morphological opening and closing operations, as prior knowledge, played a crucial role in the preprocessing of the dataset during the competition. In the preliminary stage, we used 12 labeled data points that had been preprocessed to train the network and used their weights to predict unlabeled data, followed by retraining the network. This strategy enabled us to achieve a score of 0.9381 in the preliminary stage. In the semi-final stage, we first used the weights obtained from the preliminary stage to define ROIs for newly added unlabeled data, then continued to train with the same network structure and iterative method as in the preliminary stage. Ultimately, we achieved a score of 0.8427.

Key words: segmentation, prior knowledge, semi-supervision, pseudo-labeling.

1. method

In our research, we chose nnU-Net[1] as our backbone netwo-rk, and the entire training framework is shown in Figure 1. In previous work in medical image segmentation, it is usually necessary to adjust the network structure and parameters according to different datasets. This process not only relies on the experience of researchers but also typically requires extensive experimentation, with results often being difficult to replicate successfully. Previous studies have often focused on designing new network archit-ectures, but these structures are often empirically adjusted based on specific training data, lacking generalization ability and prone to overfitting. However, the key innovat-ion in our study lies in our rational design of the entire training process. Before network training, we analyze the data and add prior knowledge through preprocessing. This method reduces excessive adjustments to the network struct-ure and parameters, making the model more generalized. By designing a reasonable training framework, we can fully leverage the potential of nnU-Net, providing a more robust and reliable solution for medical image segmentation tasks. This is of great significance for research and application in the field of medical image segmentation.

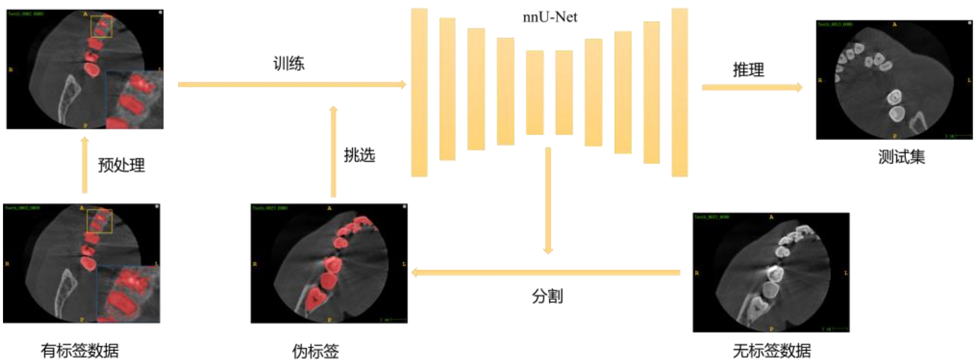


Figure 1. Overall framework

1.1 pretreatment

In our research, in addition to the network architecture based on nnU-Net, we also introduced prior knowledge such as morphological opening and closing operations and connected domain analysis, because these prior knowledge have extremely important value in the field of medical image segmentation. The application of these knowledge is reflected in many aspects, One is to improve the accuracy of the algorithm. Given that medical images typically contain complex anatomical structures and tissues, prior knowledge such as morphological operations and connected domain analysis can guide the algorithm to capture target structures more accurately, thereby significantly reducing the risk of segmentation errors. Another aspect is to reduce reliance on large-scale labeled data. In the medical field, obtaining high-quality labeled data is often very expensive and time-consuming. However, fully utilizing prior knowledge from morphological opening and closing operations and connected domain analysis can partially replace the need for large-scale labeled data, thus significantly reducing training costs. By optimizing model-generated pseudo-labels based on this prior knowledge, we can more effectively utilize limited labeled data and better adapt to different medical images

The scenario further improves the performance and robustness of the model. This strategy not only reduces the cost, but also conforms to the data-driven principle of deep learning, making the model more feasible for practical application.

1.2 Related configurations

To improve the segmentation accuracy of the network, we have designed a new set of configurations based on the default 3D_fullres configuration of nn-UNet. These adjustments include using larger input slice sizes and increasing batch sizes. Considering that this competition is divided into two stages: preliminary and semi-final, and the datasets used in these stages differ significantly, we have made different configurations for the datasets of the preliminary and semi-final stages, as shown in Table 1.

Table 1. Specific Settings for input data

	Preliminary contest		Intermediary heat	
Configure	**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 Fake label generation

We adopted a simple and effective pseudo-label generation method to facilitate the use of unlabeled data to train our model. The specific method of generating pseudo-labels is as follows:

1. A 5-fold cross-validation model is trained on the prepr-ocessed tag data.

2. The network model trained by the first step is used to predict unlabeled data.

3. A network model is retrained with labeled data and unlabeled data.

In the training process, we used a composite loss function of Dice Loss and Cross Entropy Loss.

2.experiment

2.1 The preliminary stage

Solution 1: First, only 12 labeled data were used for training, and the test result was 0.9229.

In the second scheme, it was found through analysis that the label data showed severe sharpening and noise, which did not conform to the medical prior knowledge that the surface of teeth should be smooth. Therefore, we considered using morphological opening and closing as well as removing connected domains to fine-tune and retrain the 12 labeled training sets, and got the result of 0.9300.

Option three: The model was subsequently used to generate pseudo-labels for unlabeled data, and uncertainty detection screening was conducted on these pseudo-labels with reference to [2]. All selected unlabeled data were predicted as a new dataset for iterative training, resulting in a final accuracy of 0.9381. Specific results are shown in Table 2.

Table 2. Results of different strategies in the preliminary stage

Evaluating indicator	Option 1	Option 2	Option three
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

In order to see the comparison between each scheme more intuitively, we divided the results into three dimensions and showed them. The results are shown in Figure 2.

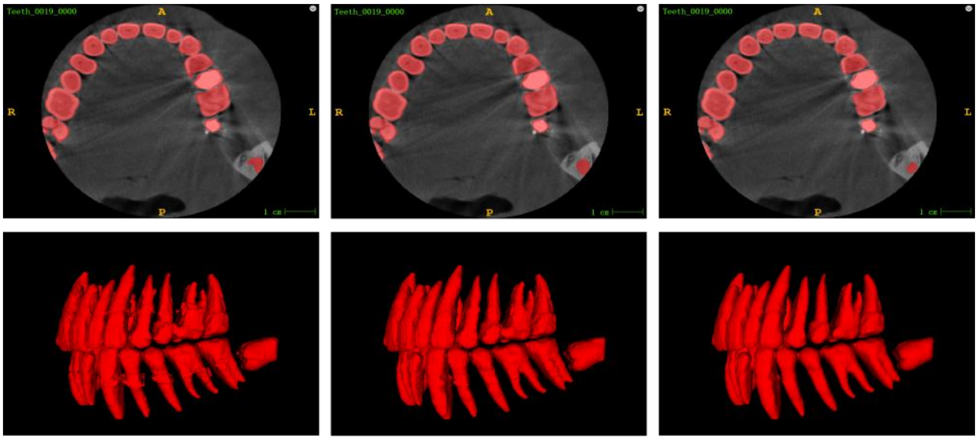


Figure 2. Comparison of results of different schemes

2.2 The second round

Plan 1: First, the model with the highest score in the preliminary round was used to directly predict the test data in the final round, and the score was 0.7694.

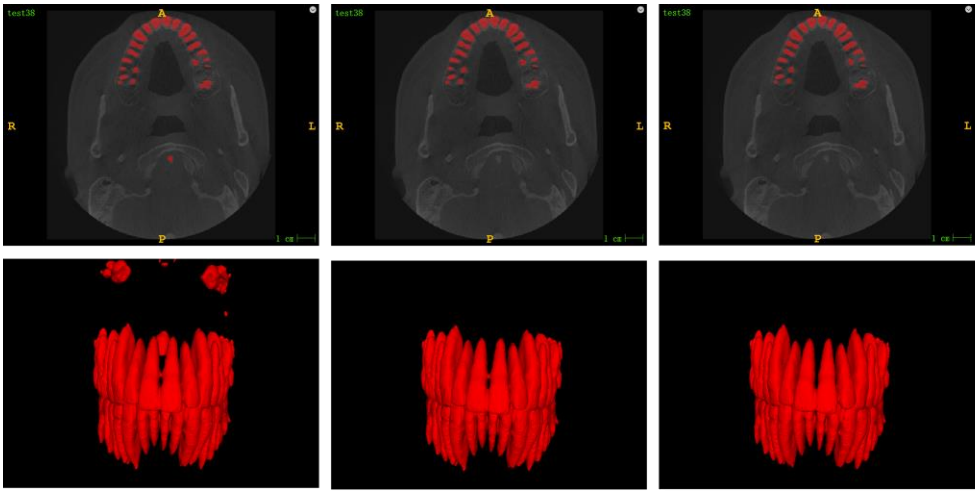
Option Two: Through analysis, it was found that the test datasets for the preliminary and final rounds had significant differences. The data for the preliminary round mainly consisted of [200,266,400]*[200,266,400]*[200,300] with spacing ranging from [0.3, 0.3, 0.3], while the data for the final round were all 512*512*400 with spacing ranging from [0.25,0.25,0.25]. We applied ROI limitations to the test data of the final round and performed post-processing on the regions, resulting in a score of 0.8259. Option Three: Subsequently, we applied ROI limitations to the pseudo-labels provided for the unlabeled data in the final round, then iteratively trained them together with the unlabeled data selected from the preliminary round. The final score was 0.8427. The results are shown in Table 3.

Table 3. Results of different strategies in the second round

Evaluating indicator	Option 1	Option 2	Option three

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

The three-dimensional visualization results of different schemes in the second round are shown in Figure 3.



方案一 方案二 方案三

Figure 3. Comparison of results of different schemes

2.3 Implement the details

Environment Settings: We have developed nnU-Net based on pytorch. All training processes are carried out in the Linux system, and all models are trained from scratch. Details of the development environment are shown in Table 4.

Table 4. Experimental environment configuration

Parameter	
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)

Training Plan: All training data undergoes preprocessing based on their respective datasets, using the default methods of nnU-Net. Our training strategy combines Dice loss and cross-entropy loss functions, with an initial learning rate set to 0.01 and 1000 training epochs, applicable to all experiments. Detailed settings are shown in Table 5.

Table 5. Training programme

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

3 sum up

In this study, we ingeniously integrated prior knowledge from the medical field into the nnU-Net, effectively utilizing unlabeled data in model training. This significantly enhanced the model's performance in the MICCAI 2023 and Challenge 3D CBCT dental segmentation tasks. By fully leveraging professional knowledge from the medical field and image processing techniques, our approach not only improved the accuracy of dental segmentation but also reduced reliance on large-scale labeled data. This not only helps to improve work efficiency and reduce costs but also holds promise for accelerating advancements in dental CAD technology, thereby providing more reliable and efficient tools for diagnosis and treatment in oral medicine.

reference documentation

[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].