technology files

Document name: 3D-the 7th name-single combat-technical description

catalogue

[I. Overall Process 2 2](#_Toc8710)

[1.1 Inference 2 2](#_Toc17693)

[1.2, post-processing 2 2](#_Toc2849)

[2. Deployment Details 4 4](#_Toc8558)

[2.1, Preprocessing 4 4](#_Toc747)

[2.2 Network Model 4 4](#_Toc6385)

[2.3 Algorithm description and loss function setting 5 5](#_Toc28759)

[2.4 Training Strategy 6 6](#_Toc4409)

[2.5, post-processing 7 7](#_Toc8913)

[Results presentation 8 8](#_Toc19615)

[3.1 Preview of forecast results 8 8](#_Toc14251)

[3.2 Evaluation index results 9 9](#_Toc28820)

[References 11 11](#_Toc32094)

1. The whole process

The whole process is divided into two parts: reasoning and post-processing, which takes about 100s/.nii.gz in total, as shown in Figure 1-1.

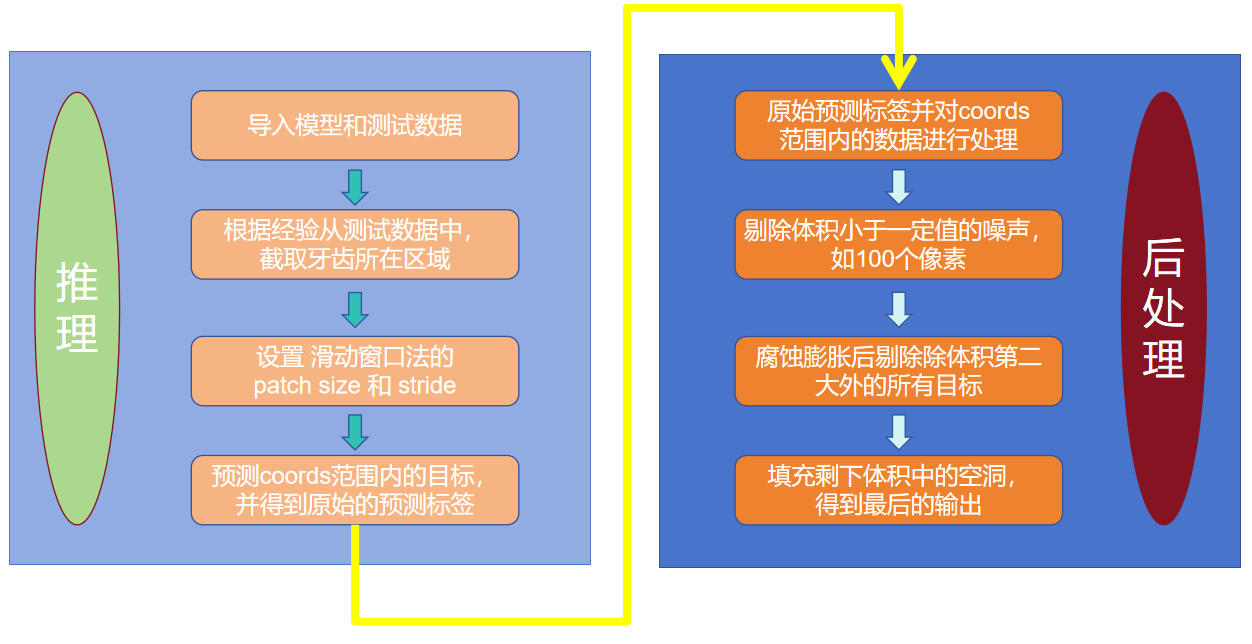


Figure 1-1 Schematic diagram of the overall process

1.1 Inference

First, import the trained model and test data. Second, perform operations such as scale to spacing=0.3, using 2500 and 500 as pixel upper and lower bounds, and normalization of maximum and minimum values. Third, set the window size and step size for the sliding window method to patch\_size= (112,112,80), stride\_xy = 48, and stride\_z = 48, meaning a prediction is made with a window of (112,112,80) and a step size of 48 pixels. Fourth, based on experience, set the ROI coordinates, coords= (100,440,0,350,10,320), corresponding to the minimum and maximum values of x-axis, y-axis, and z-axis, respectively; and predict within the ROI region according to the set sliding parameters, obtaining the original predicted labels.

(Note: roi coordinates refer to the coordinate of the largest connected domain selected after nii.gz, which is extracted from all predicted test set labels after multiple tests and superimposed into one. It can cover 50 teeth regions of test data.)

1.2, post-processing

First, read the original prediction labels and set the coords range. Second, based on the skimage.measure.label() function, obtain the connected region information of the original prediction labels (including volume and number), and remove targets with a volume less than 100 pixels from the original prediction labels. Third, use the scipy.ndimage.gray\_dilation() function with size= (2,2,2) for erosion, and then apply the scipy.ndimage.gray\_erosion() function with size= (4,4,4) to dilate the eroded data; again use the skimage.measure.label() function on the eroded and dilated data to select the second largest volume as the tooth target. Fourth, use the scipy.ndimage.binary\_fill\_holes() function to fill all cavities within the tooth target, which serves as the final prediction output.

(Note: For test22.nii.gz, test23.nii.gz, test.24.nii.gz, after many observations and experiments, the set coords ranges are (120,440,0,200,0,150), (120,440,0,200,0,130) and (120,440,0,345,50,280).)

1. Deploy details

2.1, Pre-treatment

For the retest data, there are 12 sets of labeled data and 300 sets of unlabeled data, but most shapes are above (266,266,200), with the spacing of most data being 0.3. Observations show that tooth shapes generally fall between (100,150,100) and (150,200,150). Therefore, random cropping can be used to increase sample randomness while reducing the training burden.

The preprocessing steps are as follows:

First, use the trilinear or nearest method to interpolate the original data into the same spacing (i.e., 0.3).

Then, with 500 and 2500 as the upper and lower bounds respectively, the maximum and minimum values are normalized. [1]

Finally, the patch (128,128,100) is randomly cropped 20 times in the set coords region and saved to the corresponding h5 file.

After preprocessing, we get (12+300)\* 20 = 6240 pieces of data.

The schematic diagram is shown in Figure 2-1.

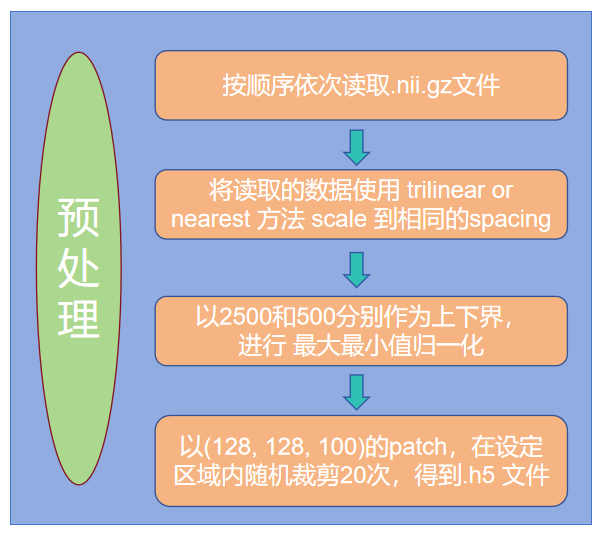


Figure 2-1 Schematic diagram of preprocessing

2.2 Network model

Referring to MCNet+[2] and CBAM[3] algorithms, the network structure of the adopted method is shown in Figure 2-2.

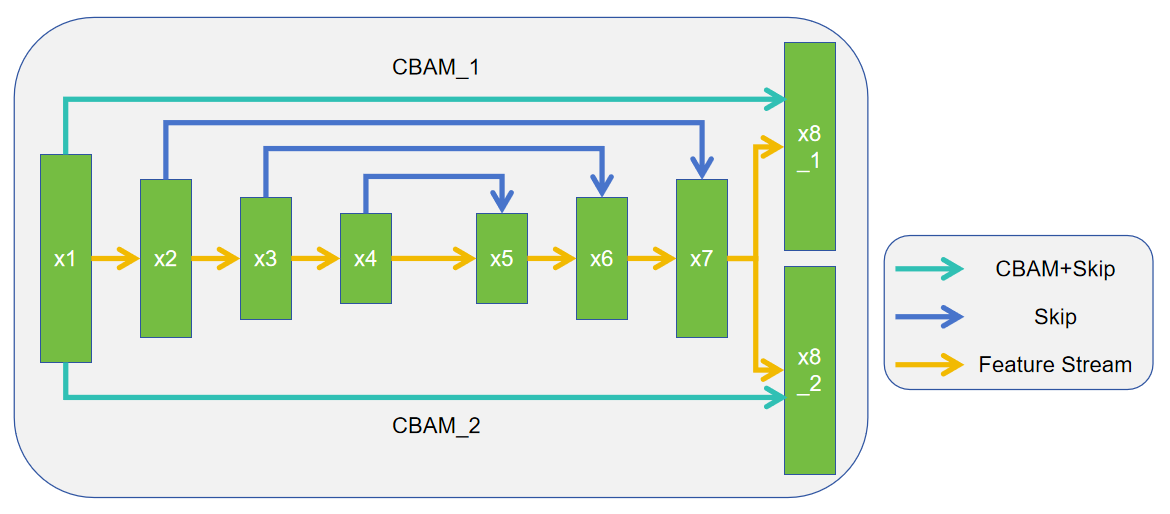


Figure 2-2 Schematic diagram of network model

Among them, referring to MCNet+[2], two different upsampling branches (x8\_1 and x8\_2) are used in the last upsample of the Decoder at VNet[4], aiming to learn different regions of the teeth. Referring to the spatial attention module of CBAM[3], the feature x1 obtained after the first downsample at VNet[4] is sequentially passed through two CBAM[3] modules skip to different upsampling branches (x8\_1 and x8\_2), resulting in the final outputs output1 and output2.

2.3 Algorithm description and loss function setting

The algorithm diagram is shown in Figure 2-3.

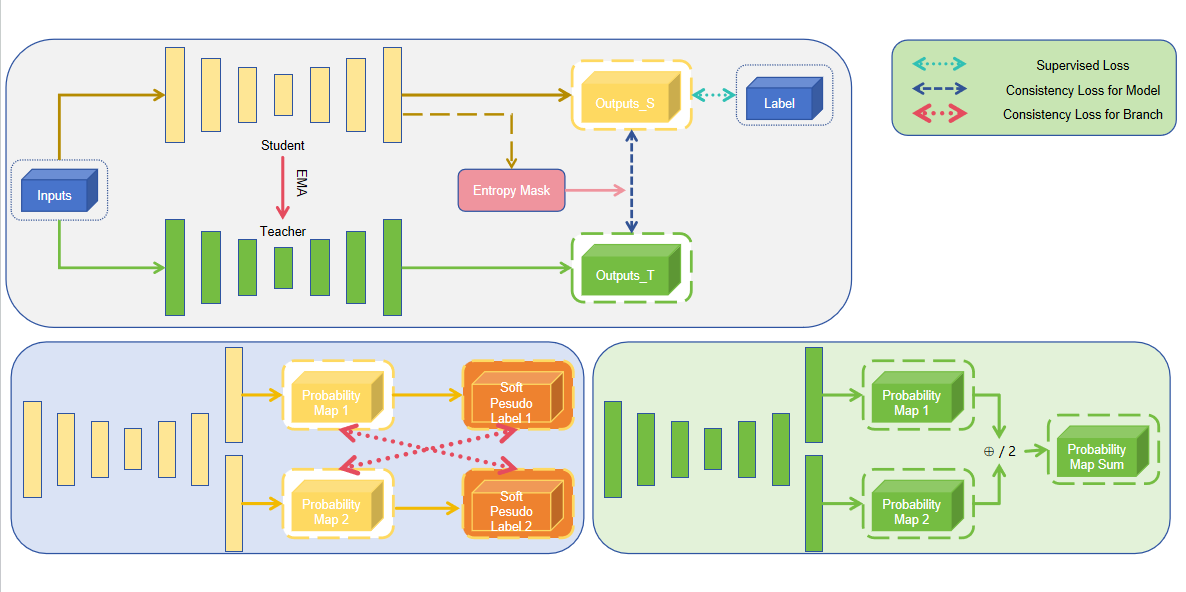


Figure 2-3 Schematic diagram of the algorithm

Referring to the Ent strategy in AC-MT[5] and the soft\_shapen\_pseudo method in MCNet+[2], this method also sets up the Student and Teacher models. It uses the soft\_shapen\_pseudo method to obtain pseudo-labels for the two outputs of the Student model, then calculates the MSE Loss between them as part of the consistency loss, Lc1. After merging the two outputs of Student and dividing by 2, it obtains the average output of the Student model, and only for the unlabeled part, calculates the MSE Loss with the average output of the Teacher model as the other half of the consistency loss, Lc2.

In addition, the entropy is calculated according to the output of Student model, and the mask of consistency loss is obtained according to the threshold and applied to Lc2.

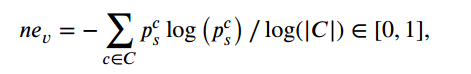


Figure 2-4 Entropy calculation formula [5]

The final loss function is:

Among them, C is the category of classification, c is the number of categories, and i and j are the number of outputs of the network.

2.4 Training strategy

During the training process, the basic parameters are set according to the training framework of AC-MT[5] as follows:

Max iterations: 20000

Batch size: 4

Patch size: (112, 112, 80)

Base lr: 0.01

Optimizer: SGD, lr=0.01, decay=0.0001, momentum=0.9, and every 2500 iter, lr = base lr \* 0.1 \*\* (iter\_num // 2500)

EMA decay: 0.99

Num workers: 4

Seed: 42

In addition, when importing data, the data enhancement operation of RandomRotFlip () and RandomCrop (patch\_size) will be performed. During training, the model is saved every 1000 iter.

2.5, post-processing

See I. Overall process.

1. Results shown

3.1 Preview of forecast results

Some of the forecast results are selected for display, namely test0.nii.gz, test20.nii.gz and test34.nii.gz, as shown in Figure 3-1 to 3-3.

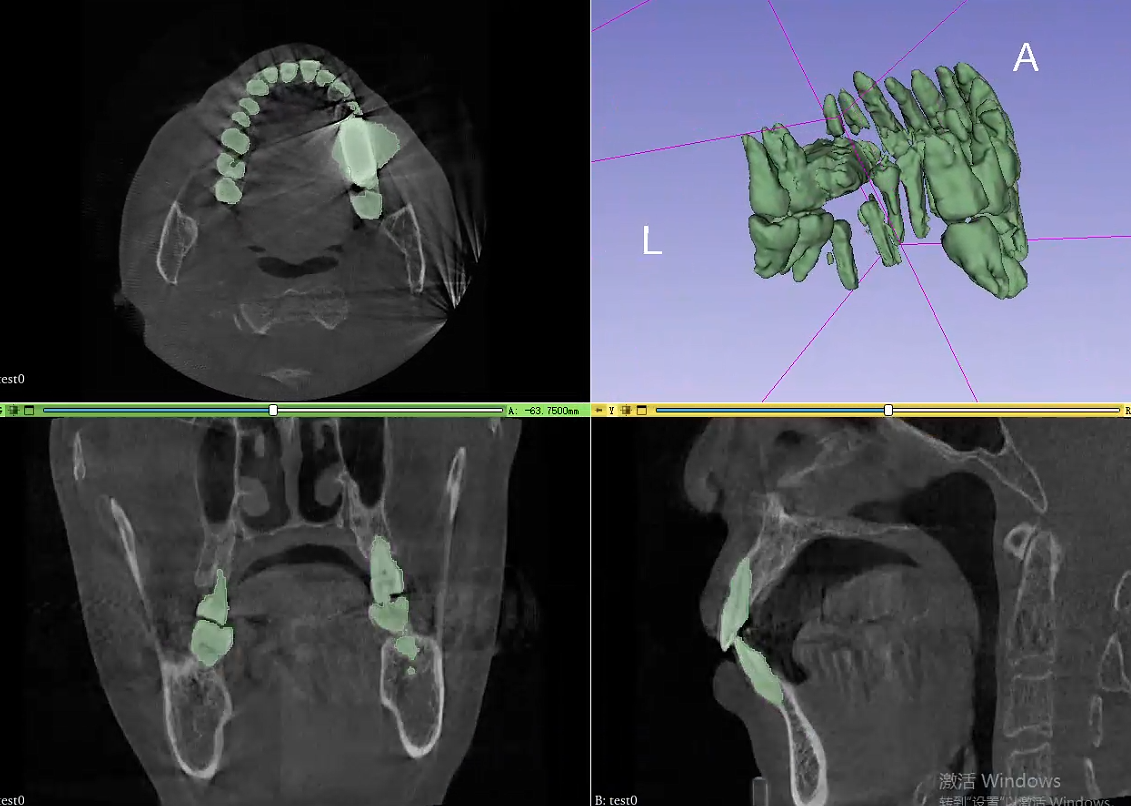


Figure 3-1 Prediction results of test0.nii.gz

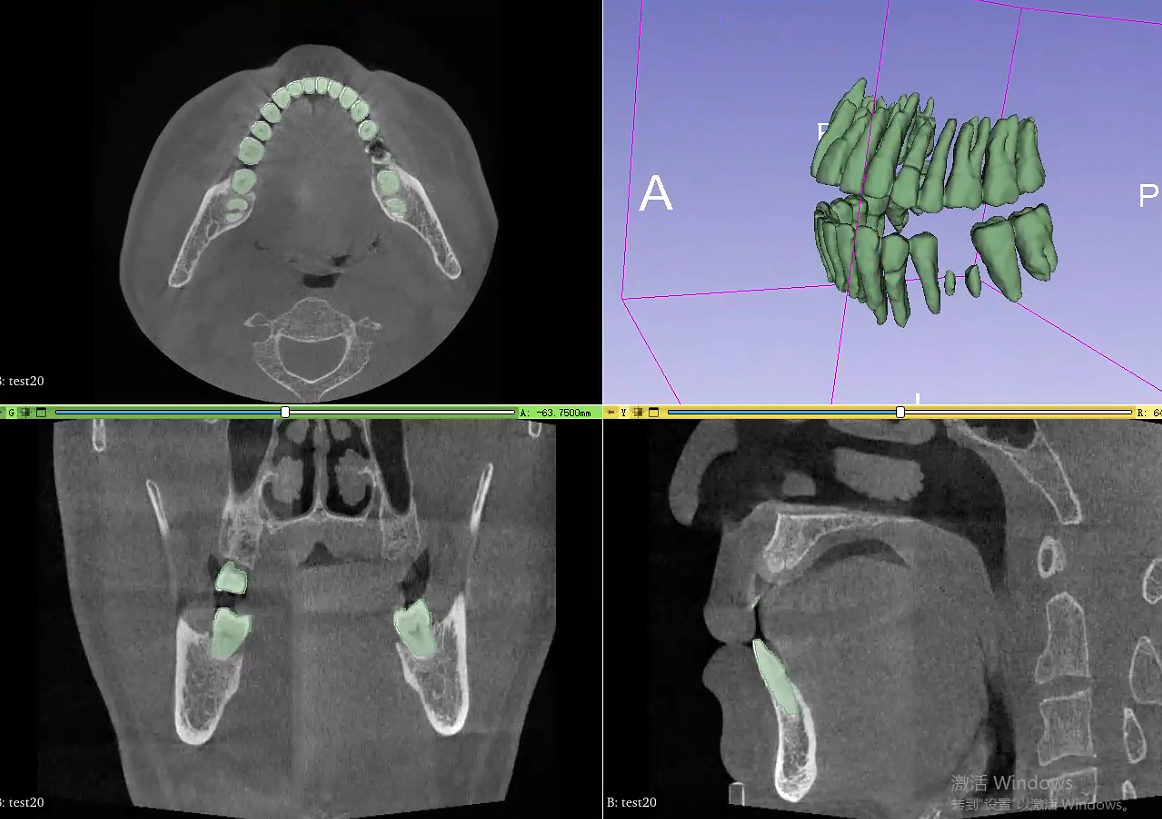


Figure 3-2 Prediction results of test20.nii.gz

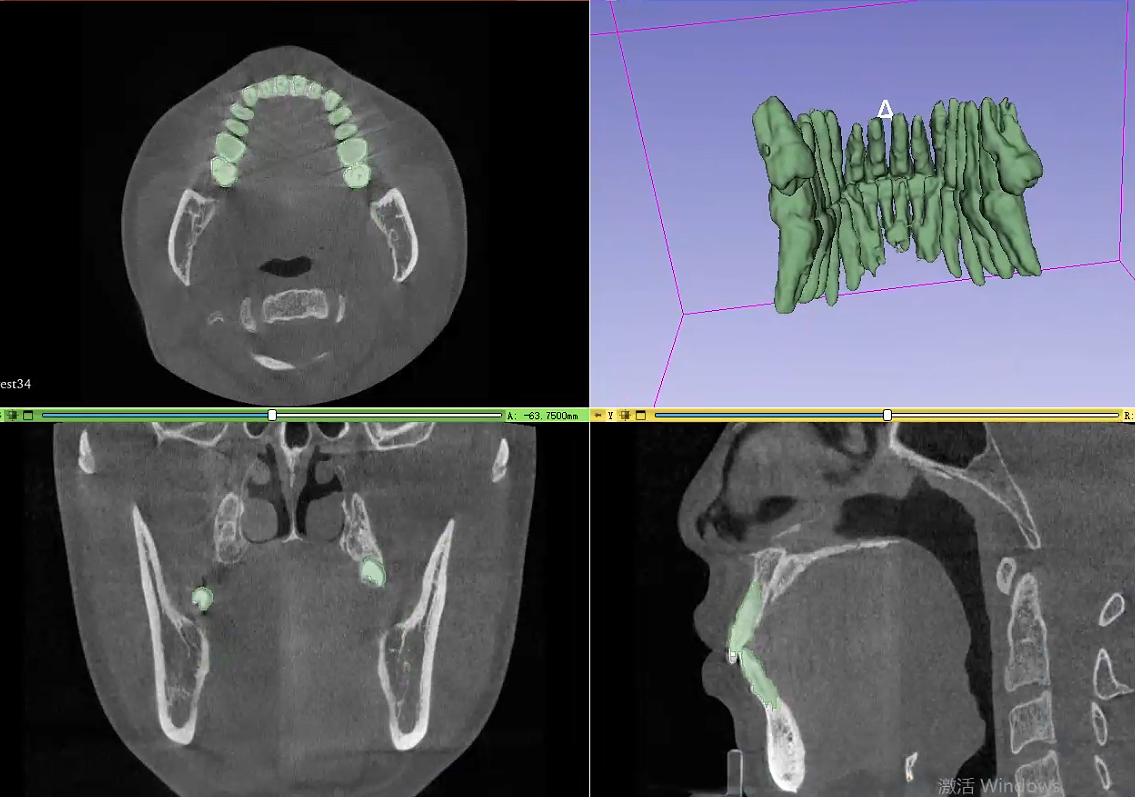


Figure 3-3 Prediction results of test34.nii.gz

3.2 Evaluation index results

Table 3-1 Evaluation index results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ranking** | **Dice** | **IOU** | **HD** | **SCORE** |
| **7** | **0.7749** | **0.8161** | **0.1580** | **0.8074** |

1. reference documentation
2. Zhiming C,Yu F,Lanzhuju M, Zhiming C,Yu F,Lanzhuju M, et al. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images[J]. Nature Communications,2022,13(1).
3. Yicheng W,Zongyuan G,Donghao Z, Yicheng W,Zongyuan G,Donghao Z, et al. Mutual consistency learning for semi-supervised medical image segmentation[J]. Medical Image Analysis,2022,81.
4. Woo S, Park J, Lee J Y, Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
5. Milletari F,Navab N,Ahmadi S. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation.[J]. CoRR,2016,abs/1606.04797.
6. Xu Z,Yixin W,Donghuan L, Xu Z,Yixin W,Donghuan L, et al. Ambiguity-selective consistency regularization for mean-teacher semi-supervised medical image segmentation[J]. Medical Image Analysis,2023,?.