Technical Report

1. Development process

We use self-training method with filtering mechanism to utilize label-free data. Specifically, it is divided into three steps:

Step 1: We use labeled data ( ) for supervised training, iteratively train E times, and save checkpoints at three epoch points (, ,). Use these three similar models to predict the segmentation graph of labelless data respectively, and record it as . Then calculate the average Dice of , and , and then obtain the average Dice of the three models through arithmetic average. The expression is shown in Equation 1:

X_{nl}, M_{nl} 3 E 2 3 E E $\langle \langle M_1$ M_2 M_3 M_1 M_2 M_3

$$MeanDice = \frac{Dice(M_1, M_3) + Dice(M_2, M_3)}{2} \qquad (1)$$

We believe that *MeanDice* can be used as an evaluation indicator for evaluation label prediction model for label-free data. The higher the value of *MeanDice*, the better the model predicts label-free data and obtains low-noise pseudo-label effects. According to this idea, we sort the unlabeled data from high to low according to the *MeanDice* value, and then divide the data into two groups: easy prediction group and difficult prediction group according to the proportion. And use the best saved model to predict easily predicted data to get pseudo-labeled data ().

$1:1$ X_{easy}, M_{easy}

Step 2: Oversample labeled data to ensure that its number is consistent with the amount of pseudo-label data, prevent the model from overfitting due to unbalanced number of tags, and perform MixUp and CutMix data enhancement on the data with pseudo-labeled before the model is trained. Finally, use the easy-to-predicted data group with pseudo-labeled () and labeled data ( ), and then use the trained model to predict the remaining difficult-to-predicted data to obtain

X_{easy}, M_{easy} X_{hard} (M_{hard}) (,

Step 3: Retrain the model with labeled data, repeat step 2 operations, and use three sets of data to train the model together.

X_{easy} (X_{hard}) M_{hard})

2. Training skills

1. Preprocessing part: The original image size is used as input (3,640,320), and is standardized accor

2. Data enhancement: divided into weak data enhancement and strong data enhancement.

Weak data enhancement: Random flips the width and height dimensions.

Strong data enhancement: flip, rotation, elastic transformation (Elastic), Gaussian smoothing, Gauss

Sharpening, intensity change

Use weak data enhancement for labeled data, and use MixUp and CutMix for pseudo-labeled data to perform strong data enhancement, to prevent the model from overfitting pseudo-labeled data.

3. Model and training part:

The model mainly adopts the Unet architecture, backbone is efficientnet-b7, and a denoiser is also introduced to reduce noise. The denoiser also uses the Unet architecture, backbone is resnet-10t

Training section:

The loss function is BCELoss+DiceLoss, the optimizer is AdamW, the learning rate is lr=1e-4, the learning rate is adjusted with cosine annealing, and iterates 50 times

TTA: The strategy is horizontal flip and vertical flip.

4. Innovative ideas

A denoiser is introduced to reduce the noise of the segmentation graph predicted by the model, reduce the noise in the pseudo-label, and improve the quality of the pseudo-label. The input of the denoiser is a segmentation diagram, and the output is noise. The segmentation diagram subtracts the noise to obtain the final segmentation diagram. The model diagram is shown in Figure 1.

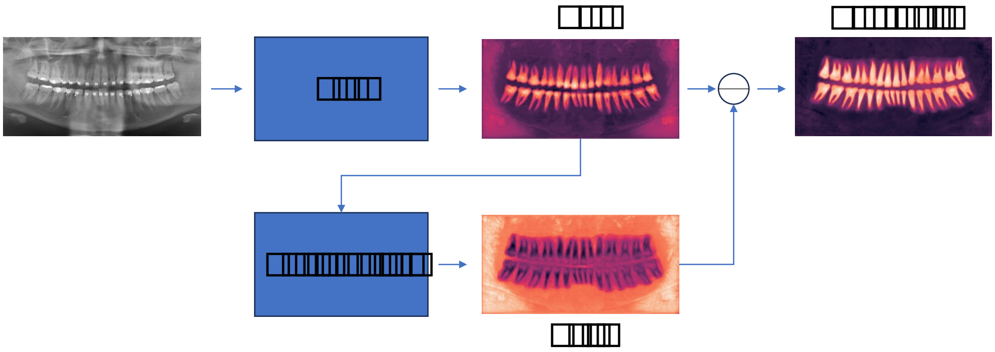


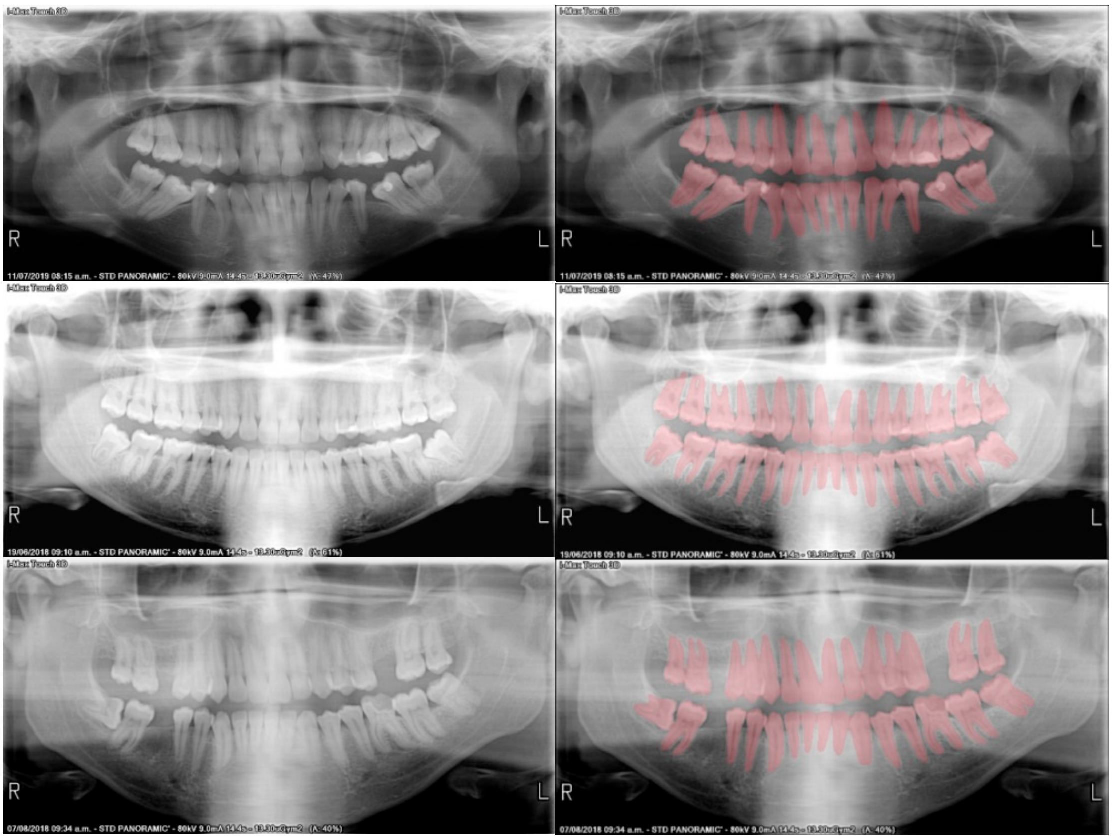
Figure 1 Segmentation model diagram

Pseudo-tag retraining:

The key point in improving the model's effect is to strengthen the data enhancement +MixUp+CutMix, which can effectively prevent the model from overfitting the pseudo-label. Both tagged data do not use MixUp and CutMix. Labeled data does not use MixUp+CutMix, and the test set score is 95.98. Labeled data uses MixUp+CutMix, with a score of 95.90 in the test. When training mixed with pseudo-labels, the ease of learning of label data should be ensured.

Prevent the model from being biased by pseudo-labels.

Results:


I-Max Touch 3D
R
L
11/07/2019 08:15 a.m. - STD PANORAMIC' - 80kV 9.0mA 14.4s - 13.30mGym2 (A: 47%)
19/06/2018 09:10 a.m. - STD PANORAMIC' - 80kV 9.0mA 14.4s - 13.30mGym2 (A: 61%)
07/08/2018 09:34 a.m. - STD PANORAMIC' - 80kV 9.0mA 14.4s - 13.30mGym2 (A: 40%)