brief introduction

With the widespread application of cone-beam computed tomography (CBCT) technology in dentistry, especially in oral imaging, precise and reliable tooth structure segmentation has become a key challenge in the field of dental medical image analysis. Moreover, as socioeconomic levels rise and the trend of population aging intensifies, the development of dentistry becomes increasingly important and urgent. To address this challenge, we not only need highly accurate methods for tooth segmentation but also effective datasets and algorithms to handle common issues in CBCT images.

In this MICCAI 3D-CBCT semi-supervised dental segmentation competition, our research team leveraged deep learning techniques, particularly the nnUNet architecture, to address this issue. nnUNet, renowned for its outstanding performance and flexibility, provided a powerful tool for our study to tackle complex dental image segmentation tasks. However, despite significant advancements in deep learning methods in image segmentation, the quality and quantity of labeled data remain a major challenge for CBCT images.

In this competition, we face the issue of noise in labeled data, particularly due to artifacts from dental restorations and implants, as well as the complexity of diverse tooth shapes and restoration of dental defects. Moreover, the insufficiency of labeled data has also plunged us into the dilemma of data scarcity. Therefore, this report will detail our research methods, including our nnUNet-based model architecture, and the innovative techniques we have adopted to address the challenges of noisy labeled data and data scarcity.

Methodology

nnUNet Expansion

In this study, we constructed a nnUNet-based reinforcement model to address the complexity of dental CBCT image segmentation tasks. Our focus was mainly on the decoder part of the model, and we enhanced the expression ability of the model by introducing more decoding layers and expanding the neural network properties (number of channels).

Decoder structure

We increase the depth of the model by introducing additional layers in the nnUNet decoder. These additional layers help extract higher-level feature representations to more accurately capture complex structures in CBCT images, such as teeth and surrounding tissues.

Attribute enhancement

In order to enhance the perception ability of the model, we extended the properties of nnUNet and increased the number of channels in the neural network. This can provide more feature information and help the model better distinguish different anatomical structures, thus improving the segmentation accuracy.

Two-stage model

Our research also introduces a two-stage model training strategy to address the challenges of CBCT image segmentation. The key to this strategy is to break down the task into two subtasks and train the model step by step to better capture the contextual information of the image.

Stage 1: prediction of maxilla and mandible (Maxilla and Mandible)

In the first stage, we trained a model aimed at predicting the position of the jawbone in CBCT images. The purpose of this sub-task is to extract features related to the jawbone, thereby better understanding the overall structure of the image. This prediction task helps the model better understand the contextual information of dental CBCT images.

Stage 2: Tooth segmentation

In the second stage, we built a model that combined CBCT images and the prediction results of jawbone position to realize the segmentation of teeth. By introducing the information of jawbone position, the model could locate the teeth more accurately, making the segmentation task more feasible.

The results show that the difficulty of tooth segmentation task is significantly reduced by introducing the prediction of jawbone position, which improves the performance and segmentation accuracy of the model.

data augmentation

The importance of data processing

In the relatively simple task of dental segmentation, we believe that overly complex model structures are not necessary, as complex models can easily lead to overfitting and noise. Instead, we consider data processing to be key to solving the problem. Our data faces two main challenges: the noise in labeled data and the limited availability of labeled data.

Data processing method

Noise marked data processing

The primary challenge lies in the noise of labeled data, with unclear boundaries and a large number of noisy labels. To address this issue, we used tools like 3DSlicer to smooth the boundaries of labeled data and reduce isolated segmentation areas, thereby improving label accuracy. This preprocessing step helps enhance the robustness and performance of the model.

Data augmentation

Secondly, the challenge we face is limited labeled data. To make full use of this limited labeled data, we employed data augmentation techniques to increase the diversity of training data. Our decision was to leverage nnUNet's built-in data augmentation feature, which has been widely validated and can effectively enhance the model's generalization performance.

Utilization of whole face data

We noticed that the training data consists of hemimandibular images, while the prediction data includes complete mandibular and maxillary images. This mismatch in data distribution can lead to issues with tooth recognition in the model. Therefore, we decided to remove all hemimandibular data and use only complete mandibular and maxillary facial images for training. This decision helps improve the model's accuracy in recognizing teeth.

The introduction of false labels

In order to make better use of unlabeled data, we adopted the method of pseudo-labeling. By using the prediction results of the model on unlabeled data as pseudo-labels, we were able to expand the training data and improve the diversity of the data. These pseudo-labels became as reliable as the real label data through repeated optimization.

Model training (Model Training)

Training parameter setting

During the model training process, we selected a set of carefully tuned hyperparameters to achieve the best performance. Our parameter Settings are as follows: epoch = 500, learning rate = 0.01, and the optimizer is stochastic gradient descent (SGD) with Nesterov momentum.

epoch select

Considering the limited labeled data, we conducted detailed experiments comparing epoch settings of 100,200,500, and 1000. Our study shows that setting the epoch to 500 yields the best model performance. Larger epochs may lead to performance degradation, so we chose 500 as the appropriate training period.

Learning rate selection

Learning rate is another important hyperparameter in the training process. We believe that a slightly larger learning rate can benefit the convergence speed and performance of the model. However, the exact learning rate value may vary depending on the dataset and model structure, so it can be fine-tuned according to specific circumstances.

Optimizer selection

We chose the SGD optimizer and enabled Nesterov momentum to help the model converge better and avoid local minima. may consider RMSProp which better align with unbalanced data.

Training strategies

We adopted a series of training strategies to make full use of the limited labeled data and improve the performance of the model.

Data smoothing and removal of low quality data

First of all, we smooth the labeled data to reduce boundary noise and improve the quality of labels. At the same time, we remove the poor quality labeled data to ensure the high quality of training data.

The introduction of false labels

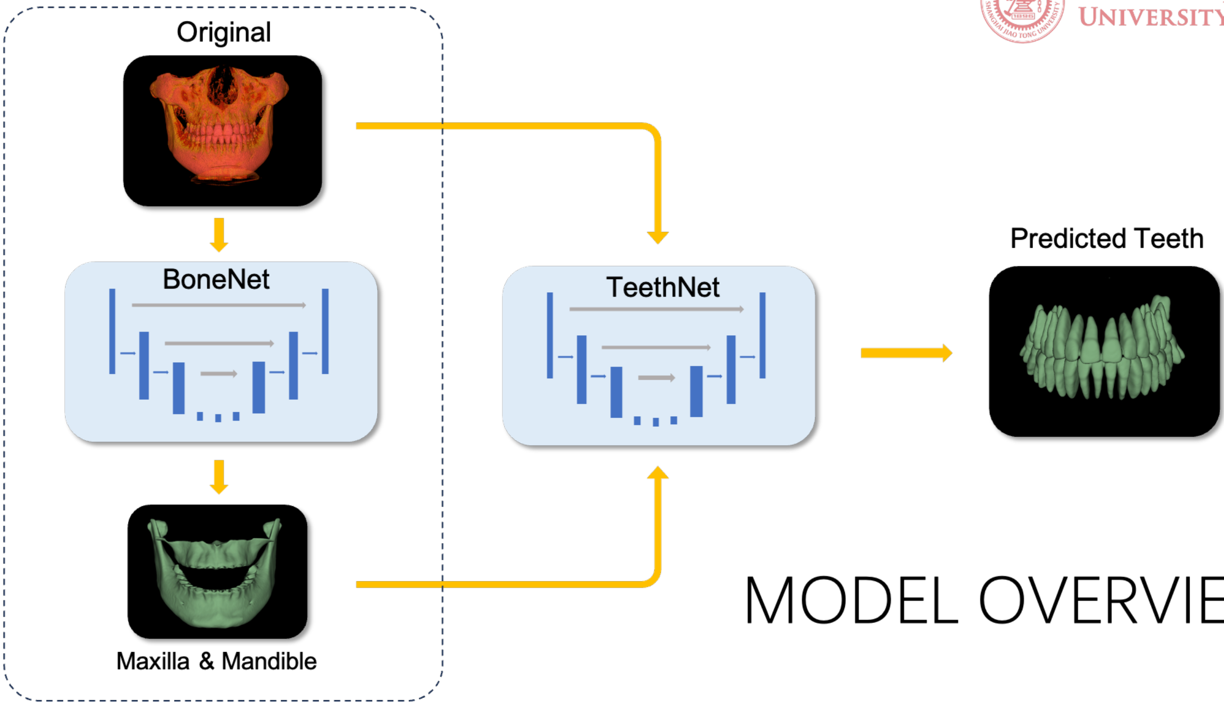
Next, we employed a semi-supervised learning strategy, training the baseline model on smoothed labeled data and using the model's predictions for unlabeled data as pseudo-labels. We iterated this process repeatedly until the pseudo-labels stabilized. This approach helps expand our training data and enhances the model's generalization performance.

Predicting the jawbone

To better understand the contextual information of CBCT images, we attempted to construct a jawbone prediction model. By selecting appropriate regions of interest (ROIs) and segmenting with fixed background intensity, we successfully removed the background from the image, including water and air, thus obtaining the predicted results for the jawbone. Due to the high noise in jawbone images, smoothing and selection were initially used to improve the quality of the jawbone images.

Treatment of dental restorations and their artifacts

Dental restoration artifacts and post-restoration tooth defects are highly challenging issues. However, due to the limited competition time, after reviewing the annotated data from the participants, we decided to treat all teeth with post-restoration defects as normal teeth and remove the data of dental restorations with artifacts. We believe that focusing on accurately predicting healthy, normal teeth rather than dealing with dental restorations and their associated artifacts will lead to better performance. Although the number of restored teeth is limited, we think this strategy has the potential to improve the model's scores.



MODEL OVERVIEW

discuss

In our research, although some success has been achieved, there are still some problems that need to be further addressed. The following are some of the key issues discussed:

Training and inference speed

In model training and inference, time cost is crucial. To apply our model more broadly, we need to explore ways to accelerate the training and inference processes, thereby reducing the consumption of time and computational resources. This may involve research into model optimization, hardware acceleration, parallel computing, and other areas.

Use the restored tooth image

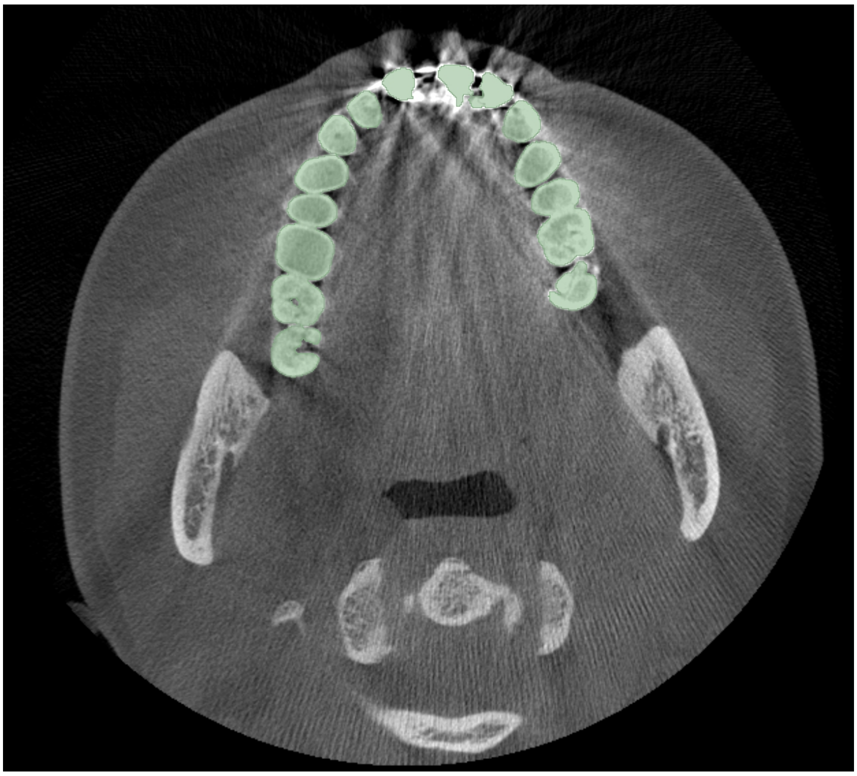
Although we default to treating the restored teeth as normal teeth for image segmentation in training to improve performance, these images of dental restorations may have significant medical value. Future research should explore how to better utilize these images to provide more information for medical diagnosis and treatment. This may require developing specialized models to process dental restoration images or adopting multi-task learning methods.

Remove dental restoration artifacts

Dental restoration artifacts appear as radiating light and dark stripes, significantly impacting the clarity and accuracy of surrounding images. To better address this issue, we need to explore new segmentation methods and artifact reduction techniques to improve the quality of dental CBCT images. This may require interdisciplinary collaboration, integrating expertise from both image processing and medical fields.

In summary, although our research has achieved some success in the segmentation task of dental CBCT images, we still face a series of challenges. By continuing to delve deeper into research and innovation, we hope to address these issues and provide more accurate and efficient solutions for medical image analysis, thereby improving patients' diagnostic and treatment experiences.

Results shown



	SCORE	DICE	IOU	HAUSDORFF DISTANCE
ROUND 1	0.9359	0.9348	0.9374	0.0642
ROUND 2	0.8497	0.8442	0.8661	0.1595