nnUNet for Semi-supervised Tooth and Pulp Root Canal Segmentation in CBCT

Ajo Babu George[1][0009−0005−3026−0959] and Sadhvik Bathini[2][0009−0007−1011−3761]

[1] DiceMed, Odisha, India.
[2] Indian Institute of Technology Kharagpur, West Bengal, India
drajo_george@dicemed.in, sadhvik.ini@gmail.com

Abstract. A solution for the Semi-supervised Teeth Segmentation and Registration (STSR) 2025 Challenge, which focused on the precise segmentation of teeth and pulp root canals in 3D Cone Beam Computed Tomography (CBCT) scans is presented in this paper. Accurate segmentation of the pulp root canal is crucial for clinical visualization and treatment planning, but manual annotation is extremely labor-intensive. The presented approach uses a semi-supervised framework powered by nnU-Net, leveraging a small labeled dataset of 30 scans alongside a much larger unlabeled dataset of 300 scans. To effectively utilize the unlabeled data, pseudo-labeling was employed to generate annotations, and the model was subsequently trained. The results for both tooth and pulp structures yield a Dice score of 0.8088 and an mIoU of 0.9638 in the all-data track, while the Dice score in the coreset track is 0.69. These metrics highlight the model's ability to accurately identify and delineate the target structures.

Keywords: Teeth Segmentation · Semi-supervised learning · nnUNet · CBCT · Pulp · Root canals

1 Introduction

The field of dentistry is increasingly benefiting from computer-aided diagnosis tools, particularly for treatment planning and prognosis evaluation. Precise segmentation of teeth and especially the root pulp canal from 3D Cone-Beam Computed Tomography (CBCT) scans is a crucial pre-processing step for many of these applications. This enables clearer visualization of dental anatomy, which in turn helps in developing more refined treatment strategies. However, manual annotation of these regions is an extremely labor-intensive task, requiring a substantial investment of time and human resources. This makes acquiring large, labeled datasets a significant challenge for training robust deep learning models.

Recent developments in deep learning have shown great promise in dental image analysis, with models capable of high-accuracy segmentation and disease classification[2] [4]. Networks like U-Net and its derivatives, including nnU-Net, have become standard for medical image segmentation tasks [1]. nnU-Net, in particular, stands out for its ability to automatically adapt to various 3D and 2D medical imaging tasks without extensive manual tuning of hyperparameters [7].

Despite these advances, a major limitation common to all deep learning-based methods is their reliance on a large quantity of high-quality training data, which is difficult and expensive to obtain for medical imaging. The manual annotation of 3D volume data, for example, requires experts to label each 2D slice, making the process even more challenging. To address this, semi-supervised learning has emerged as a highly practical approach, allowing models to benefit from a large quantity of readily available unlabeled data alongside a small set of labeled data [5].

This is the very purpose behind the Semi-supervised Teeth Segmentation (STS) Challenge. The challenge, a pioneering event in tooth segmentation, aimed to stimulate the development of effective semi-supervised algorithms for both 2D PXI and 3D CBCT volumes. The STS 2023 Challenge [8] [9] [10] focused on tooth instance segmentation. A significant research gap remains in the precise, automated segmentation of the intricate pulp root canal. This task is more complex due to the high variability in pulp canal morphology and the need for fine-grained annotation consistency.

A primary motivation is to advance the field of dental image analysis by developing a robust methodology for segmenting both teeth and pulp root canals using a semi-supervised learning strategy. Public dental imaging resources primarily focus on 2D panoramic radiographs rather than volumetric CBCT, especially for pediatric and mixed-dentition cohorts, highlighting the scarcity of large annotated 3D datasets and motivating semi-supervised learning for this task [12]. This approach is essential because the manual annotation of pulp canals is particularly labor-intensive due to their complex and variable morphology. The challenge dataset provided a scenario typical of real clinical settings: a limited number of labeled scans (30) and a large pool of unlabeled data (300). The following are the contibutions:

- A pseudo-labeling technique using the nnU-Net framework was adopted to leverage the large-scale unlabeled dataset. This approach allows a model to learn from a much larger volume of data, improving its robustness and generalization capabilities.
- The pipeline includes comprehensive pre-processing steps, such as resampling, cropping, and normalization.
- By combining these techniques, a model achieved robust performance across both teeth and pulp segmentation tasks, demonstrating its scalability and potential for real-world clinical applications where annotated data is scarce.

2 Method

The methodology employs a semi-supervised learning strategy to address the challenge of limited labeled data. Figure 1 describes the overall pipeline involves pre-processing the 3D CBCT scans, followed by a pseudo-labeling technique and extensive model training.

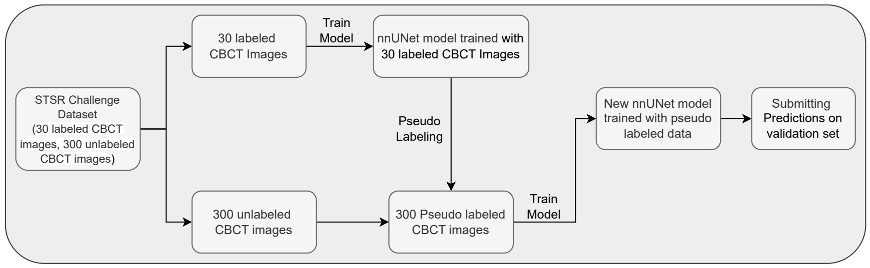

Fig. 1. Overall Pipeline of the proposed methodology.

2.1 nnU-Net Architecture

The nnU-Net framework [6] is employed for the network architecture, a self-contained python package that automatically optimizes network structure and training strategies for medical image segmentation. The architecture has a characteristic U-shape, where 3D arrays representing each input image undergo a series of convolutions, maximum pooling, up-convolutions, and concatenation steps. The initial half of the network is responsible for feature extraction, while the second half synthesizes the segmentation output. The design includes horizontal concatenation steps that pass early network information to later stages, a key feature of U-Net-based architectures.

3 Experiments

3.1 Dataset and evaluation metrics

The dataset for this task consists of 3D CBCT images for semi-supervised segmentation of teeth and pulp root canals. The training dataset consists of two parts: a labeled set of 30 images with fine-grained segmentation masks for teeth, wisdom teeth, and various root canal structures, and a much larger unlabeled set of 300 images. For public validation, 40 labeled images are provided, with segmentation results submitted to the Codabench platform for evaluation.

3.2 Implementation details

Preprocessing The images are first resampled and cropped to a default size that the nnUNet model configures while maintaining key anatomical features. A per-scan z-score normalization is applied to each image, standardizing the data with a mean of 0 and a standard deviation of 1 across the non-zero voxels.

Environment settings The development environments and requirements are presented in Table 1.

Table 1. Development environments and requirements.

System	CentOS 7.6
CPU	Intel Xeon SKL G-6148 CPU@2.4GHz
RAM	384 GB
GPU (number and type)	NVIDIA V100
CUDA version	11.0
Programming language	Python 3.20
Deep learning framework	torch 2.0, torchvision 0.2.2

3.3 Semi-Supervised Pseudo-Labeling Scheme

A core component of the methodology is the use of a pseudo-labeling technique to leverage the large amount of unlabeled data. An initial nnU-Net model is trained for 750 epochs on the 30 available labeled scans. This trained model is then used to generate pseudo-labels for the 300 unlabeled scans. The pseudo-labeled data is then used to retrain the model for an extended period of 500 epochs. This approach effectively expands the training set, allowing the model to learn from a much larger volume of data than the labeled set alone. Extensive training improves the model's ability to accurately segment variable pulp canal morphology.

Table 2. Training protocols of the nnUNet model trained with 30 labeled scans

Pre-trained Model	None (training from scratch)
Batch size	2
Patch size	128×128×128
Total epochs	750
Optimizer	SGD with Nesterov momentum (0.99)
Initial learning rate (lr)	0.01
Lr decay schedule	Polynomial decay ($lr = lr_0(1 - \frac{epoch}{max\_epoch})^{0.9}$)
Training time	∼12 hours
Loss function	Cross-entropy + Dice loss (sum)
Number of model parameters	∼30–35M[3]
Number of flops	∼250–300G[4]

Table 3. Training protocols of the nnUNet model trained with 300 pseudo labeled scans

Pre-trained Model	None (training from scratch)
Batch size	2
Patch size	128×128×128
Total epochs	500
Optimizer	SGD with Nesterov momentum (0.99)
Initial learning rate (lr)	0.01
Lr decay schedule	Polynomial decay ($lr = lr_0(1 - \frac{epoch}{max\_epoch})^{0.9}$)
Training time	∼36 hours
Loss function	Cross-entropy + Dice loss (sum)
Number of model parameters	∼30–35M[5]
Number of flops	∼250–300G[6]

4 Results and discussion

The proposed methodology demonstrates robust performance in both teeth and pulp segmentation, successfully addressing the challenges of a limited labeled dataset. The semi-supervised approach, which combines pseudo-labeling with the nnU-Net framework, proved to be scalable to large CBCT datasets.

Figure 2 and Table 2 detail the training curves and protocols, respectively, for the initial model utilizing the limited dataset of 30 labeled images. Another nnUNet model trained on the 300 pseudo-labeled images is illustrated by the training curves and protocols presented in Figure 3 and Table 3. These figures provide a comprehensive overview of the training progression and the impact of the semi-supervised approach on the model's final performance.

Table 4. Mean Dice and IoU Scores for Valid Dental Anatomy Labels.

Label ID	**Anatomical Structure**	**Mean Dice**	**Mean IoU**
1	Dental Hard Tissues	0.9591	0.9215
2	Pulp Chamber	0.8298	0.7100
4	Palatal Root	0.7449	0.5971
5	Mesial Root Canal	0.4976	0.3730
6	Distal Root Canal	0.6700	0.5116
7	Mesiobuccal Root Canal	0.6986	0.5390
8	Mesiolingual Root Canal	0.5978	0.4366
9	Distobuccal Root Canal	0.7150	0.5583
12	Impacted Tooth	0.9565	0.9173

Fig. 2. Training and validation curves from nnUNet with 30 labeled scans

Fig. 3. Training and validation curves from nnUNet with 300 pseudo labeled scans

Table 5. Per-case Dice scores for all non-background labels.

Case	D1	D2	D4	D5	D6	D7	D8	D9	D12
ToothPulp_004	0.964	0.831	0.695	0.675	0.742	0.671	0.662	0.691	0.966
ToothPulp_009	0.955	0.872	0.848	0.248	0.776	0.780	0.704	0.808	0.961
ToothPulp_013	0.961	0.852	0.751	0.791	0.475	0.746	0.658	0.668	0.977
ToothPulp_016	0.953	0.797	0.756	0.000	0.679	0.686	0.337	0.706	0.923
ToothPulp_019	0.968	0.830	0.758	0.619	0.718	0.662	0.670	0.713	0.977
ToothPulp_027	0.952	0.798	0.660	0.653	0.630	0.647	0.555	0.704	0.935

4.1 Quantitative results on validation set

Across the six validation cases, the model shows a clear separation in performance between large dental structures and the smaller, more complex root canal pathways. As shown in Table 5, the model achieves consistently high Dice scores for large anatomical structures (labels 1 and 12), while the fine-grained canal structures exhibit substantial variability. Table 4 further highlights this contrast, demonstrating strong segmentation performance for Dental Hard Tissues and the Impacted Tooth compared to the smaller canal structures. Dice scores for these major structures consistently fall within the 0.92–0.97 range, whereas the root canals show markedly lower and more unstable performance, ranging from moderate scores (approximately 0.79–0.81) to complete failures in the most challenging cases.

4.2 Qualitative results on validation set

Qualitative analysis of the validation set indicates that the model's performance, which achieved an overall accuracy of 80.88%, is highly dependent on anatomical characteristics. The model consistently demonstrates robust segmentation of large, well-defined structures but exhibits limitations in regions with low image contrast and complex micro-anatomy as noted in the varous sections in Figures 4 and 5.The visualization of the segmentations was done using 3D Slicer. [3] A successful segmentation is shown in the top row of the Figure 6 STS25_Validation_0007 where the model accurately delineates the main bodies of the dental hard tissues and the pulp chambers. The resulting boundaries are clear, and the 3D reconstruction is anatomically cohesive, reflecting the model's strength in identifying structures with distinct intensity gradients.

STS25_Validation_0025 is highlighted in the bottom row of the Figure 6 as it exemplifies the model's primary weakness: the inability to completely segment the apical third of the tooth roots. The segmentation is visibly incomplete, resulting in fragmented masks and a disjointed 3D reconstruction where the teeth appear to lack proper root structures.

The primary reason for these failures is the low contrast-to-noise ratio (CNR) between the root apex and the surrounding trabecular bone. This ambiguity is exacerbated by partial volume averaging artifacts, which are common in thin structures, and the complex anatomy of the root tip, which often includes lateral canals. Consequently, while the model reliably segments the bulk of the tooth structures, its accuracy is diminished by its inability to resolve these challenging but clinically significant apical regions.

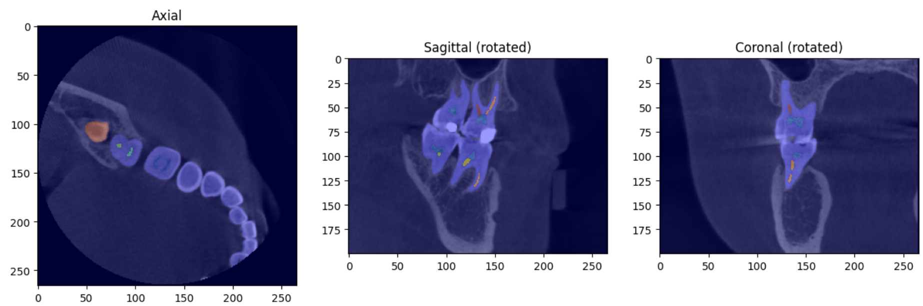

Fig. 4. Segmentation result visualization on validation case 001

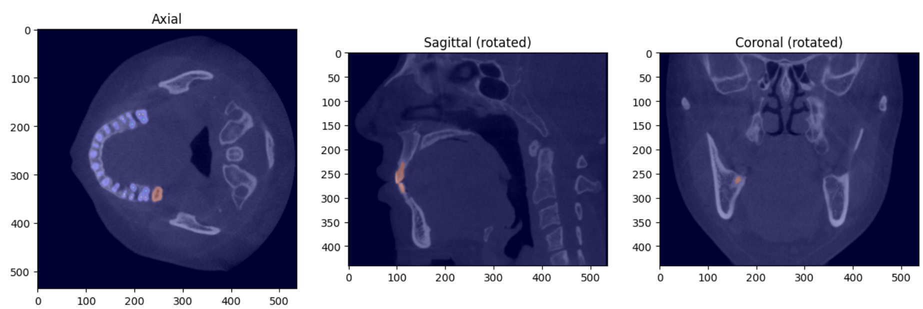

Fig. 5. Segmentation result visualization on validation case 003

5 Conclusion

The proposed methodology overcame the challenges of teeth and pulp root canal segmentation in 3D CBCT scans, particularly the scarcity of labeled data. The semi-supervised approach, which combines pseudo-labeling with the nnU-Net framework. By generating pseudo-labels for the large-scale unlabeled dataset, the model was able to learn from a much larger data pool, improving its generalization capabilities beyond what would be possible with the limited labeled data alone. The use of nnU-Net was also crucial in handling the high variability in pulp canal morphology and ensuring fine-grained annotation consistency. The achieved validation dice score of 0.8088 on the validation set demonstrates the effectiveness of this approach.

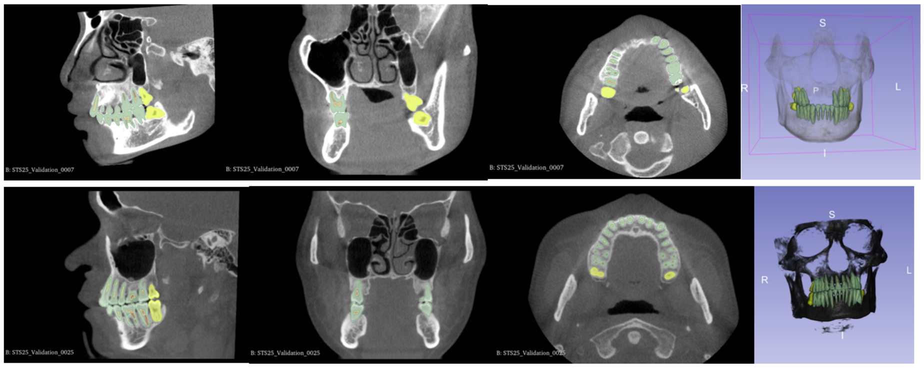

Fig. 6. Segmentation result visualization using 3D Slicer for cases 007 and 025

6 Limitations and Future Work

The proposed method, while effective, still faces challenges in segmenting the apical third of the roots due to low contrast, complex anatomy, and partial-volume effects, which lead to incomplete or fragmented predictions. The reliance on pseudo-labels also introduces potential noise that may propagate during training, limiting consistency in difficult regions. Future work can focus on improving thin-structure representation through super-resolution or topology-aware modules, incorporating uncertainty-guided pseudo-label refinement, and expanding the diversity of labeled data through active learning.

Acknowledgements We thank all the data owners for making the medical images publicly available and Codabench [11] for hosting the challenge platform.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Azad, R., Aghdam, E.K., Rauland, A., Jia, Y., Avval, A.H., Bozorgpour, A., Karimijafarbigloo, S., Cohen, J.P., Adeli, E., Merhof, D.: Medical image segmentation review: The success of u-net. IEEE Transactions on Pattern Analysis and Machine Intelligence (2024)
2. Chen, X., Wang, X., Zhang, K., Fung, K.M., Thai, T.C., Moore, K., Mannel, R.S., Liu, H., Zheng, B., Qiu, Y.: Recent advances and clinical applications of deep learning in medical image analysis. Medical image analysis **79**, 102444 (2022)
3. Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al.: 3d slicer as an image computing platform for the quantitative imaging network. Magnetic resonance imaging **30**(9), 1323–1341 (2012)

4. George, A.B., Bathini, S., et al.: Grad-cam & grad-cam++ for explainable oral squamous cell carcinoma detection using deep learning on orthopantomograms. In: 2025 International Conference on Sensors and Related Networks (SENNET) Special Focus on Digital Healthcare (64220). pp. 1–6. IEEE (2025)
5. Han, K., Sheng, V.S., Song, Y., Liu, Y., Qiu, C., Ma, S., Liu, Z.: Deep semi-supervised learning for medical image segmentation: A review. Expert Systems with Applications **245**, 123052 (2024)
6. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
7. Pettit, R.W., Marlatt, B.B., Corr, S.J., Havelka, J., Rana, A.: nnu-net deep learning method for segmenting parenchyma and determining liver volume from computed tomography images. Annals of Surgery Open **3**(2), e155 (2022)
8. Wang, Y., Chen, X., Qian, D., Ye, F., Wang, S., Zhang, H. (eds.): Semi-supervised Tooth Segmentation: First MICCAI Challenge, SemiToothSeg 2023, Held in Conjunction with MICCAI 2023, Vancouver, BC, Canada, October 8, 2023, Proceedings. Lecture Notes in Computer Science, Springer, Cham (Oct 2024). https://doi.org/10.1007/978-3-031-72396-4, https://link.springer.com/book/10.1007/978-3-031-72396-4
9. Wang, Y., Zhang, Y., Chen, X., Wang, S., Qian, D., Ye, F., Xu, F., Zhang, H., Dan, R., Zhang, Q., Huang, X., Huang, Z., Liu, J., Zheng, Z., Wu, C., Li, Y., Li, Z., Ma, Z., Cui, W., Luo, S., Wang, C., Chen, Y., Li, T., Liu, Y., Feng, X., Ni, J., Liu, D., Wang, Q., Lin, Z., Song, W., Li, Y., Wang, B., Wang, C., Chen, Q., Li, M., Zhou, H., Jin, Q.: MICCAI 2023 STS Challenge: A retrospective study of semi-supervised approaches for teeth segmentation (7 2025), https://figshare.le.ac.uk/articles/journal_contribution/MICCAI_2023_STS_Challenge_A_retrospective_study_of_semi-supervised_approaches_for_teeth_segmentation/29512100
10. Wang, Y., Zhang, Y., Chen, X., Wang, S., Qian, D., Ye, F., Xu, F., Zhang, H., Zhang, Q., Wu, C., Li, Y., Cui, W., Luo, S., Wang, C., Li, T., Liu, Y., Feng, X., Zhou, H., Liu, D.L., Wang, Q., Lin, Z., Song, W., Li, Y., Wang, B., Wang, C., Chen, Q., Li, M.: Sts miccai 2023 challenge: Grand challenge on 2d and 3d semi-supervised tooth segmentation. ArXiv **abs/2407.13246** (2024), https://api.semanticscholar.org/CorpusID:271270695
11. Xu, Z., Escalera, S., Pavão, A., Richard, M., Tu, W.W., Yao, Q., Zhao, H., Guyon, I.: Codabench: Flexible, easy-to-use, and reproducible meta-benchmark platform. Patterns **3**(7), 100543 (2022)
12. Zhang, Y., Ye, F., Chen, L., Xu, F., Chen, X., Wu, H., Cao, M., Li, Y., Wang, Y., Huang, X.: Children's dental panoramic radiographs dataset for caries segmentation and dental disease detection. Scientific Data **10**(1), 380 (2023)