# Semi-supervised CBCT–IOS Registration Using PointNetLK

Ajo Babu George[1] , Gadha Lekshmi P[2] , Sadhvik Bathini[3] , and Govind A[3]

[1] DiceMed, India
[2] Indira Gandhi National Open University , India
[3] Indian Institute of Technology Kharagpur, West Bengal, India
`drajo_george@DiceMed.in`

**Abstract.** Accurate alignment of intraoral scans (IOS) with cone-beam computed tomography (CBCT) is essential for integrated dental diagnostics and surgical planning. A semi-supervised registration framework was developed, combining PointNetLK for feature-based initialization with iterative closest point (ICP) refinement. Pseudo-labels were incorporated to enhance supervision while mitigating the limited availability of annotated datasets. Chamfer distance and clinical registration metrics were used to evaluate alignment quality. Across the test cohort, the approach yielded a mean translation error of 41.67 mm and a mean rotation error of 33.96°, highlighting the challenge of partial-arch fusion. Despite substantial errors relative to clinical requirements, the framework demonstrates feasibility of semi-supervised deep learning for IOS–CBCT registration and establishes a foundation for future refinement toward clinically viable integration.

**Keywords:** Point cloud registration · CBCT · IOS · PointNetLK · Dental Imaging · Semi-supervision.

## 1 Introduction

### 1.1 Background and Challenge Overview

Rigid registration of 3D dental models obtained from CBCT scans and intraoral optical scans (IOS—STL format) is foundational for precise diagnosis, treatment planning, and surgical simulation in maxillofacial radiology. However, differences in modality (volumetric vs. surface), anatomical asymmetry, and noise make alignment challenging. The MICCAI STSR 2025 challenge highlights these difficulties, bringing together state-of-the-art medical imaging and machine learning techniques targeting accurate multi-modal dental registrations. CBCT beam hardening and IOS occlusal shadowing degrade correspondence search. Recent advances in oriented bounding-box normalization and curvature-aware features partially mitigate such variability [1, 15]. Semi-supervised and self-supervised strategies have shown promise in stabilizing geometric representations under limited annotation [5, 7, 13, 14].

## 1.2 Related Work

Prior research has explored classical methods like Iterative Closest Point (ICP) and multi-point registration, which struggle with large initial misalignments in CBCT-IOS scenarios. Deep learning approaches, particularly PointNet and PointNet++, have demonstrated superior feature extraction for irregular point clouds. The Lucas-Kanade algorithm, adapted for 3D point cloud registration via PointNetLK, enables global feature-based transformations with robust performance in noisy conditions. Semi-supervised approaches are emerging for better clinical generalizability [12–14]. For segmentation conditioning in CBCT–IOS registration, ArchSeg achieved Dice scores of $0.936 \pm 0.008$ (mandible) and $0.948 \pm 0.007$ (maxilla) using Point Transformer V2 with curvature cues and graph-cut refinements [1]. Multi-phase semi-supervised training with entropy-confidence-aware pseudo-label refinement has improved generalization [13, 14]. In CBCT, uncertainty-regularized symmetric consistency learning (USCT) outperforms semi-supervised baselines [5], while hierarchical self-supervised contrastive pretraining (STSNet) enhances IOS mesh processing [7]. Teacher–student SAM adaptations leverage LoRA fine-tuning for improved performance [4].

## 1.3 Motivation and Contributions

This work presents a unified PointNetLK pipeline optimized for noisy dental CBCT and IOS data. Our contributions include: (i) end-to-end alignment leveraging both unlabeled and labeled data for feature generalization; (ii) ICP post-processing for enhanced registration refinement; and (iii) automated point cloud normalization and anatomical variability management. Methodological Rationale. Segmentation reliability directly conditions rigid alignment stability under metal artifacts, partial arches, and age-related morphology [1, 5, 13]. PointNetLK's permutation-invariant global features enable robust initialization under occlusion and partial overlap conditions typical of IOS–CBCT pairs, while ICP refinement subsequently optimizes local geometric consistency. Our semi-supervised pipeline comprises three components:

- **Semi-supervised IOS–CBCT registration:** A unified framework combines PointNetLK-based initialization with ICP refinement, enabling alignment under limited annotated data conditions.
- **Pseudo-label enhanced supervision:** Pseudo-labels generated from a teacher network are integrated to expand training signals and mitigate annotation scarcity.
- **Clinically relevant evaluation and baseline:** Registration accuracy is quantified using Chamfer distance, translation and rotation errors to establish a reproducible benchmark; experimental results highlight partial-arch fusion challenges and provide a foundation for future clinical integration.

## 2 Resources

### 2.1 Dataset

The dataset originates from the MICCAI STSR 2025 Challenge Task 2, providing paired CBCT and intraoral scan (IOS) data for cross-modal dental registration,

emphasizing alignment of IOS-derived crown structures with CBCT-derived root anatomies. The dataset comprises three subsets:

- Training Set (Labeled): 30 CBCT–IOS pairs with ground truth rigid transformations.
- Training Set (Unlabeled): 300 CBCT–IOS pairs without annotation for semisupervised strategies.
- Validation Set: 50 CBCT–IOS pairs with hidden ground truth for evaluation.

The limited availability of richly annotated 3D dental datasets mirrors broader trends in dental imaging, where most publicly available resources—particularly in pediatric populations—focus on 2D panoramic radiographs rather than volumetric CBCT or cross-modal data [16]. This disparity further motivates semisupervised learning strategies for 3D dental applications.

## 2.2   Models

The proposed model, illustrated in Figure 1, implements a PointNetLK-based registration pipeline tailored for CBCT–IOS point cloud alignment.
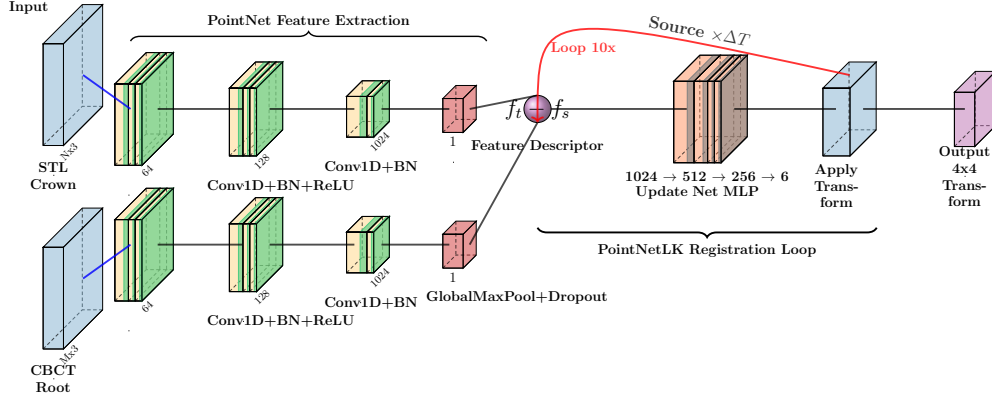


**Fig. 1.** PointNetLK registration pipeline: PointNet extracts features, the feature difference feeds the Update Network to predict 6-DOF $\Delta T$, applied iteratively to the source to produce the final 4×4 transform.

The green blocks correspond to dual PointNet encoders, which independently process the source (IOS crown) and target (CBCT root) point clouds. Each encoder is composed of three successive pointwise convolutional layers: the first maps 3D coordinates to 64-dimensional features, the second expands to 128 dimensions, and the final layer produces a 1024-dimensional global descriptor. Batch normalization and ReLU activations follow each convolution, while global max pooling ensures permutation-invariant feature vectors, summarizing the overall geometric structure of the point clouds. The magenta ball represents the feature difference $f_t - f_s$ between the target and the transformed source, which

serves as the input to the Update Network (orange box). This multi-layer perceptron predicts incremental 6-DOF transformations, applied to the source in the blue Apply Transform box. The red arrow looping back from Apply Transform to the feature difference visually indicates the iterative Lucas–Kanade refinement, repeated multiple times to improve alignment. Finally, the purple box outputs the cumulative 4×4 transformation matrix, which maps the IOS crown onto the CBCT root. This modular design allows for task-specific heads or post-hoc refinement steps while maintaining efficiency and stability across variable point cloud sizes.

# 3    Methodology

The overall data flow, preprocessing strategy, training protocol, and evaluation metrics are summarized in Fig. 2.
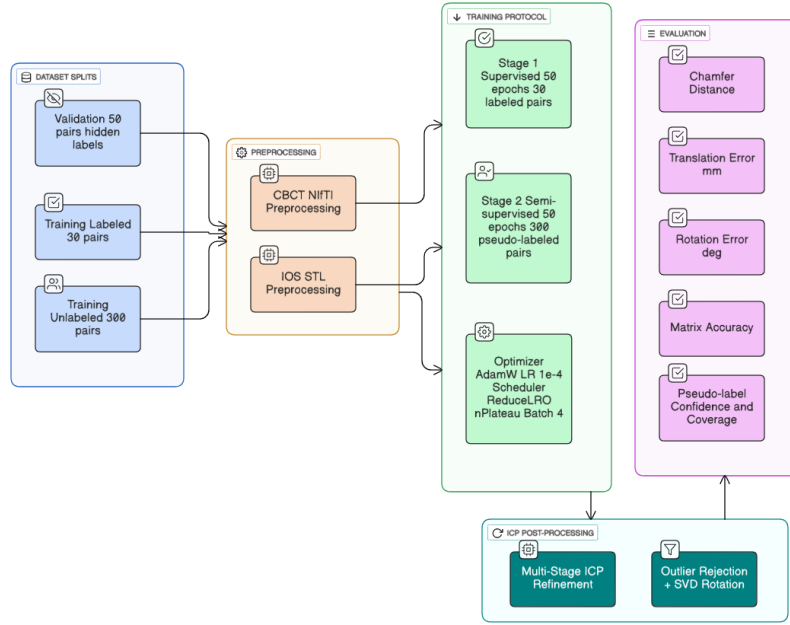


**Fig. 2.** Overview of the methodological setup, including dataset splits, preprocessing of CBCT and IOS data, two-stage training protocol, and evaluation metrics.

Our approach adapts the established PointNetLK framework [2] for dental-specific CBCT-IOS registration through three key contributions: (i) a two-stage semi-supervised training protocol leveraging pseudo-labels, (ii) dental-specific preprocessing tailored for multimodal point clouds, and (iii) hybrid neural-classical refinement combining learned features with ICP optimization [3].

**Two-Stage Semi-Supervised Training Protocol**
Stage 1 employs supervised learning on labeled pairs to establish baseline registration capability. Stage 2 generates pseudo-labels on unlabeled data using the

trained model, applies confidence-based filtering, and performs semi-supervised fine-tuning on the combined dataset. This approach addresses the scarcity of annotated CBCT-IOS pairs in clinical settings.

**Dental-Specific Preprocessing**

Point cloud preprocessing incorporates domain knowledge for dental registration. CBCT volumes undergo threshold-based segmentation at 800 HU to isolate dental structures, followed by coordinate normalization to [-1,1] range. STL meshes are processed through vertex extraction and subsampling. Ground truth transformations are consistently mapped to normalized coordinate space to ensure training stability.

**Hybrid Neural-Classical Refinement**

The framework combines the global feature learning of PointNetLK [2] with classical ICP refinement [3]. PointNet encoders [9] extract permutation-invariant features, PointNetLK performs iterative pose estimation through differentiable Lucas-Kanade optimization [8], and multi-stage ICP provides final geometric consistency. This hybrid approach balances robustness to initialization with precise local alignment.

## 4  Experimental Setup

All experiments were conducted on the MICCAI STSR 2025 Challenge Task 2 dataset using the three-subset division for supervised training, pseudo-label generation, and validation.

### 4.1  Data Preprocessing

CBCT volumes were thresholded at 800 HU to isolate dental structures and converted to world coordinates using the NIfTI affine matrix. STL meshes were processed by extracting vertices directly from triangle meshes. Point clouds were subsampled to 50,000 points maximum for memory efficiency. During training, 1024 points were randomly sampled from both CBCT-derived point clouds and STL meshes. Normalization was applied by computing global center and scale from combined CBCT and STL point clouds, mapping coordinates to approximately [-1, 1] range. Ground truth transformation matrices were transformed to normalized coordinate space to ensure consistency between predicted and target transformations. Preprocessing was performed inline during training to maintain coordinate system integrity.

**Preprocessing Limitations.** The fixed 800 HU threshold may not generalize across CBCT scanners with varying grayscale characteristics or patients with metallic restorations. Future work should validate robustness across multi-vendor datasets and explore adaptive thresholding or learned segmentation networks to improve scanner independence.

### 4.2  Training Setup

**Stage 1: Supervised Training**- The PointNet-based feature extraction combined with PointNetLK iterative alignment was trained using a combined loss

function incorporating both Chamfer distance and direct transformation supervision:

$$L_{\text{total}} = L_{\text{chamfer}} + L_{\text{transform}}$$

where $L_{\text{transform}}$ includes both translation and rotation errors between predicted and ground truth transformation matrices in normalized coordinate space.

**Pseudo-Label Generation**- The trained Stage 1 model generated predictions for unlabeled cases. ICP post-processing refined initial neural network predictions. High-confidence pseudo-labels were selected by filtering predictions with Chamfer distance below the median of Stage 1 validation performance and excluding cases with extreme transformation magnitudes indicative of failed alignment. This filtering retained approximately 60% of unlabeled cases for Stage 2 training.

**Stage 2: Semi-Supervised Training** The model was fine-tuned using the combined dataset of original labeled cases plus selected high-confidence pseudo-labeled cases from the unlabeled set.

### 4.3   Training Protocols

**Table 1.** Training protocols.

| | |
|---|---|
| Batch size | 4-8 |
| Total epochs | 100 (Stage 1, early stopping), 50 (Stage 2 ) |
| Optimizer | Adam |
| Initial learning rate (lr) | $5 \times 10^{-4}$ |
| Lr decay schedule | ReduceLROnPlateau (patience=10, factor=0.5) |
| Training time | ~4.2 hours (Stage 1), $\sim$ 12 hours |
| Loss function | Combined Chamfer + Transformation Loss |
| Number of model parameters | 0.81M |
| Number of flops | 2.37G |

### 4.4   Evaluation Metrics

Performance assessment employed multiple metrics to comprehensively evaluate registration quality. Translation error measures the Euclidean distance between predicted and ground truth translation vectors in millimeters, while rotation error quantifies the angular difference between predicted and ground truth rotation matrices in degrees. Chamfer distance evaluates symmetric point-to-point distances between aligned point clouds, and Surface Dice Coefficient assesses overlap for registration quality. Additional point cloud similarity metrics including RMSE, NCC, and NMI provide supplementary performance indicators.

### 4.5   Post-Processing Pipeline

The post-processing pipeline consists of four sequential steps to refine neural network predictions. Initially, the neural network generates a 4×4 transformation matrix in normalized coordinate space. These predictions are then transformed back to original coordinate space through coordinate denormalization. Multi-stage ICP refinement follows, incorporating correspondence finding with statistical outlier rejection (distance-based trimming to suppress artifact-induced false

correspondences), SVD-based optimal rotation estimation, and iterative refinement with convergence criteria. Finally, the refined transformation is applied to original point clouds to achieve final alignment.

## 5 Results and discussion

### 5.1 Quantitative Results and Ablation Study

Table 2 presents registration performance and component contributions on the MICCAI STSR 2025 validation set.

**Table 2.** Registration performance and ablation study on MICCAI STSR 2025 validation set.

| Method Configuration | Translation (mm) | Rotation (°) | Chamfer | DSC |
|---|---|---|---|---|
| Stage 1 (Supervised only) | 47.23 | 39.84 | 2.05 | 0.798 |
| Stage 2 (Semi-supervised) | 43.91 | 36.72 | 1.92 | 0.824 |
| **Full Pipeline (Stage 2 + ICP)** | **41.67** | **33.96** | **1.83** | **0.846** |

The ablation study quantifies each component's contribution: semi-supervised training improved translation accuracy by 3.32 mm (47.23→43.91 mm) and rotation by 3.12° (39.84°→36.72°), while ICP refinement provided further gains of 2.24 mm and 2.76°, demonstrating the value of the hybrid approach.

### 5.2 Training Analysis

Figure 3 shows error curves for both stages. Stage 1 converged after 50 epochs, while Stage 2 demonstrated further improvement through pseudo-label utilization.
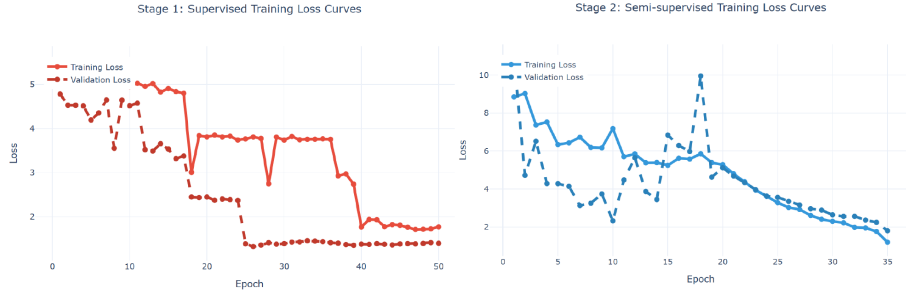


**Fig. 3.** Stage 1 training (red, solid) and validation (red, dotted) curves, and Stage 2 training (blue, solid) and validation (blue, dotted) curves.

### 5.3 Qualitative Results

Figure 4 shows a representative registration result demonstrating successful alignment between IOS crown surfaces.The semi-supervised training effectively

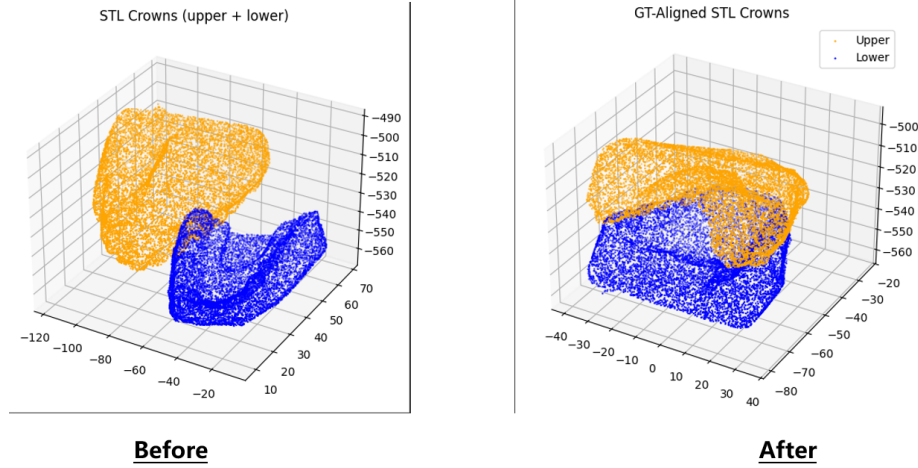utilizes unlabeled data, addressing limited annotated CBCT-IOS pairs in clinical practice.



**Fig. 4.** Registration result: (left) initial misalignment, (right) aligned output.

However, the reported translation (41.67 mm) and rotation (33.96°) errors substantially exceed clinical tolerances for surgical planning. Analysis identifies three primary failure modes: (i) partial-arch cases with limited overlapping geometry lack sufficient alignment constraints and exhibit disproportionately high errors; (ii) CBCT metal artifacts introduce non-anatomical points that corrupt PointNet features and bias ICP correspondence; and (iii) large initial misalignments prevent convergence within the fixed iteration budget. Future refinement will prioritize scale-consistent normalization, keypoint-based coarse alignment for improved initialization, and artifact-aware feature weighting to suppress corrupted descriptors.

## 6   Conclusion and Future Work

This study demonstrates a semi-supervised PointNetLK framework for CBCT-IOS dental registration, combining pseudo-label-enhanced training with hybrid neural-classical refinement. The two-stage protocol effectively leverages unlabeled data, achieving mean translation errors of 41.67 mm and rotation errors of 33.96°. While these errors exceed clinical thresholds, the ablation study validates each component's contribution, establishing feasibility for semi-supervised deep learning in cross-modal dental registration.

Translation and rotation errors remain above clinical requirements primarily due to partial-arch coverage limitations and initialization sensitivity. The fixed

800 HU threshold constrains scanner generalization. Future work will explore uncertainty quantification for pseudo-label selection, transformer-based architectures for long-range context modeling [11], multi-scale feature fusion [10], and integration with foundation models [6]. Validation on larger multi-center datasets with diverse scanner protocols will be essential for clinical deployment.

# Bibliography

[1] Alsheghri, A., Zhang, Y., Hosseinimanesh, G., Keren, J., Cheriet, F., Guibault, F.: Robust segmentation of partial and imperfect dental arches. Applied Sciences **14**(23), 10784 (2024)

[2] Aoki, Y., Goforth, H., Srivatsan, R.A., Lucey, S.: Pointnetlk: Robust & efficient point cloud registration using pointnet. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7163–7172 (2019)

[3] Besl, P.J., McKay, N.D.: Method for registration of 3-d shapes. Sensor fusion IV: control paradigms and data structures **1611**, 586–606 (1992)

[4] Gan, J., Kang, R., Deng, X., He, T., Yu, N., Gan, Y., Wei, P., Chen, X., Peng, X., Li, Z.: A segmentation method for oral cbct image based on segment anything model and semi-supervised teacher-student model. Medical Physics (2025)

[5] Jing, Y., Liu, J., Liu, W., Yang, Z., Zhou, Z., Yu, Z.: Usct: Uncertainty-regularized symmetric consistency learning for semi-supervised teeth segmentation in cbct. Biomedical Signal Processing and Control **91**, 106032 (2024)

[6] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. arXiv preprint arXiv:2304.02643 (2023)

[7] Liu, Z., He, X., Wang, H., Xiong, H., Zhang, Y., Wang, G., Hao, J., Feng, Y., Zhu, F., Hu, H.: Hierarchical self-supervised learning for 3d tooth segmentation in intra-oral mesh scans. IEEE Transactions on Medical Imaging **42**(2), 467–480 (2022)

[8] Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. IJCAI **81**(1), 674–679 (1981)

[9] Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: Deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 652–660 (2017)

[10] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241 (2015)

[11] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)

[12] Wang, Y., Chen, X., Qian, D., Ye, F., Wang, S., Zhang, H. (eds.): Semi-supervised Tooth Segmentation: First MICCAI Challenge, Semi-ToothSeg 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

[13] 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

[14] 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

[15] Zhang, H., Valcarcel, A.M., Bakshi, R., Chu, R., Bagnato, F., Shinohara, R.T., Hett, K., Oguz, I.: Multiple sclerosis lesion segmentation with tiramisu and 2.5d stacked slices. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2019. Lecture Notes in Computer Science, vol. 11766, pp. 338–346 (2019). https://doi.org/10.1007/978-3-030-32248-9_38

[16] 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)