Self-Configuring 3D Segmentation of Pediatric Dentition

Tullissi E,[1,2] Gaydamour A,[1,2] Prieto JC,[3] Mattos C,[4] Braga BMR,[5] Rosa RR,[4] Tinawi S,[1] Keener DJ,[1]Aliaga Del Castillo A,[1] Caleme E,[6] Larsson B,[7] de Oliveira Ruellas AC,[8] Arriola-Guillén LE,9 Bianchi J,[10] Oh H,[10] Lima Gurgel M,[11] Benavides E,[1] Soki F, Rodríguez-Cárdenas YA, Ruíz-Mora GA, [1] Cevidanes LH, [1] Sanna A, [5] de Silva I

1University of Michigan, Ann Arbor, MI, USA, 2 CPE Lyon, Lyon, France, 3University of North Carolina, Chapel Hill, NC, USA, 4Universidade Federal Fluminense, Brazil, 5Hospital for Rehabilitation of Craniofacial Anomalies, University of São Paulo 6Universidade positivo, Curitiba, Paraná,Brazil, 7University of Minnesota, Minneapolis,MN,USA,8Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, 9Universidad Científica del Sur, Lima, Perú, 10University of the Pacific, Arthur A. Dugoni, San Francisco, CA, USA, 11Federal University of Ceara, Fortaleza, Brazil, 12Universidad Nacional de Colombia, Bogotá, Colombia.

ABSTRACT

Robust 3D segmentation of primary and permanent teeth in cone-beam CT (CBCT) is critical for pediatric and orthodontic care. We propose a fully automatic deep-learning pipeline built on the self configuring nnU-Net v2 framework, tailored for high-fidelity dental shape modeling. Our approach learns fine-scale tooth geometries directly from volumetric data, eliminating manual tuning. On a pediatric CBCT cohort (369 training, 93 validation, 55 test scans), our model attains a mean Dice score of 0.87 across 55 dental and supporting anatomical structures. Key components include adaptive preprocessing (isotropic resampling, automatic craniofacial cropping, intensity normalization), on-the-fly 3D augmentations, and lightweight postprocessing to remove spurious segment. The resulting segmentations are consistent and clinically action able, supporting advanced 3D morphometric analysis and digital treatment planning. By extending state-of-the-art volumetric segmentation to mixed dentition CBCT data, our work facilitates integration of AI-driven geometric learning into routine pediatric dentistry workflows.

Keywords: 3Dsegmentation ·pediatric dentistry · Cone-beam CT (CBCT) · deep learning· nnU-Net · dental morphometry

METHODS

Dataset and Annotation Strategy

We assembled a retrospective dataset of 517 pediatric CBCTs covering full maxilla and mandible with mixed dentition (primary + permanent) across varied eruption stages (including unerupted buds, resorbing roots, impactions). Segmentations were initialized with DentalSegmentator and manually refined in ITK-SNAP to produce voxel-precise labels for 55 classes (52 teeth + mandible, maxilla, mandibular canal), capturing detailed morphology (curved roots, crown contours, interproximal spacing) and including identifiable atypical/unerupted structures. Class prevalence (Fig. 2) shows permanent teeth/supporting structures are near-universal (most premolars/molars ≥ 92%, e.g., UR first premolar 517/517), while age-dependent classes vary: third molars ~56–62%, primary incisors 2.7–13.3%, primary canines/molars 34–51%. The cohort is split 369/93/55 (train/val/test), supporting generalizable 3D learning.

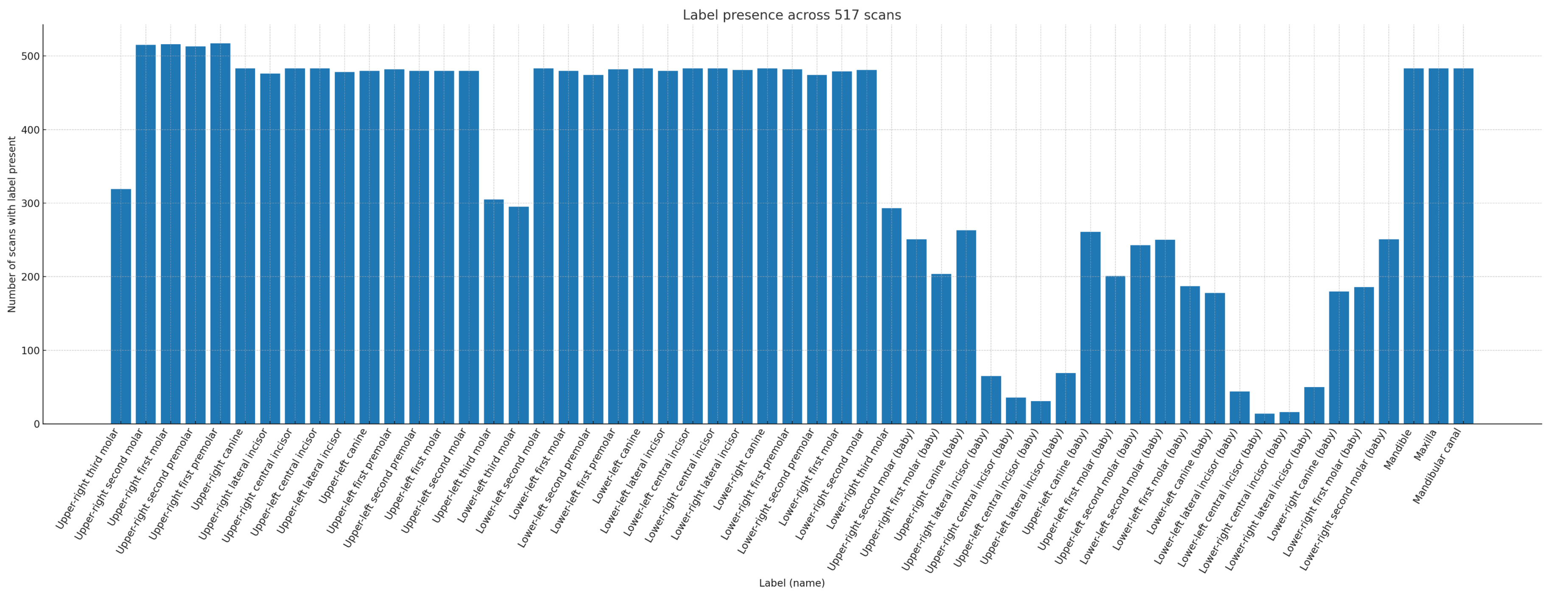


Fig.2. Per-label prevalence across the cohort (n=517).

Self-Configuring Segmentation Pipeline

We implemented the nnU-Net v2 pipeline, which offers a fully automated segmentation framework that configures all aspects of preprocessing, network architecture, and training strategy based on data-driven heuristics (Fig. 2). This design is especially suited for medical shape analysis tasks, where anatomical variability and image heterogeneity demand robust adaptation.

Preprocessing. Isotropic resampling, craniofacial ROI cropping, CT intensity clipping + z-normalization, and on-the-fly 3D augmentations (elastic, rotations, intensity shifts) to standardize yet diversify inputs.

Architecture. Auto-configured full-resolution 3D U-Net (6 levels, ~30M params); patches 112×128×128; softmax over 56 classes; no low-res cascades.

Training. nnU-Net v2 defaults with Dice+CE loss; 2 patches/iteration; 150 epochs, 5-fold cross-validation; fold ensembling at inference with overlapping tiles.

Post-processing. Per-class connected components: remove small islands, keep one anatomically plausible instance; export watertight meshes; Laplacian smoothing for surface regularization..

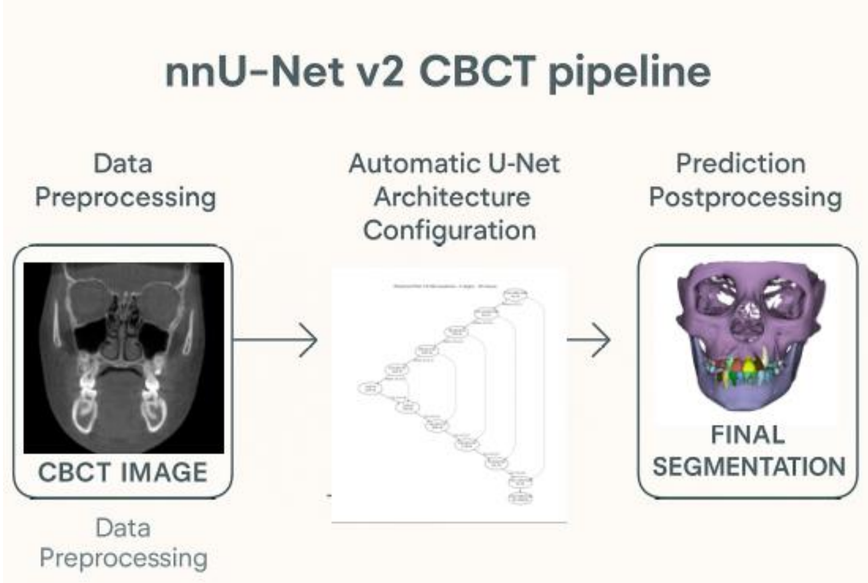


Fig.3. Illustration of the nnU-Net v2 segmentation pipeline. Left: coronal slice from the original pediatric CBCT volume. Right: voxel-wise prediction showing all 55 teeth individually.

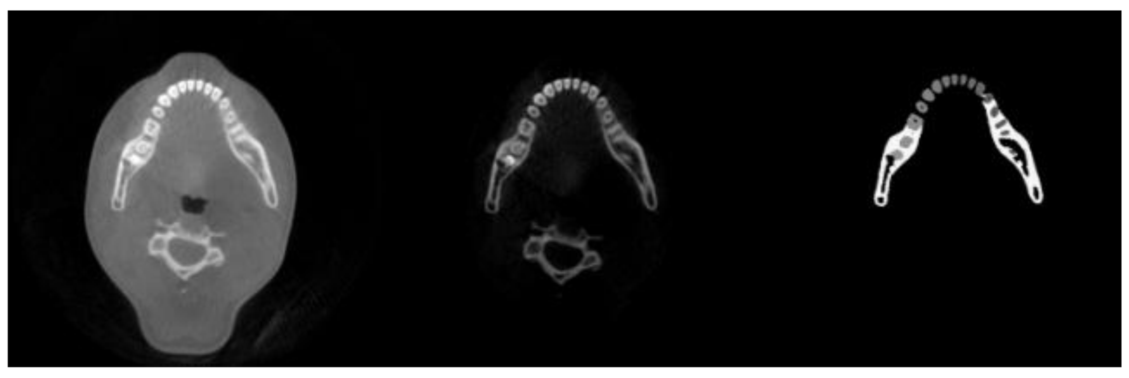


Fig.4. nnU-Net v2 preprocessing. (Left) Original axial pediatric CBCT. (Cen ter) After automatic cropping and intensity normalization. (Right) Segmentation of 55 anatomical structure (52 dental and 3 skeletal), plus background (label 0), in the preprocessed space.

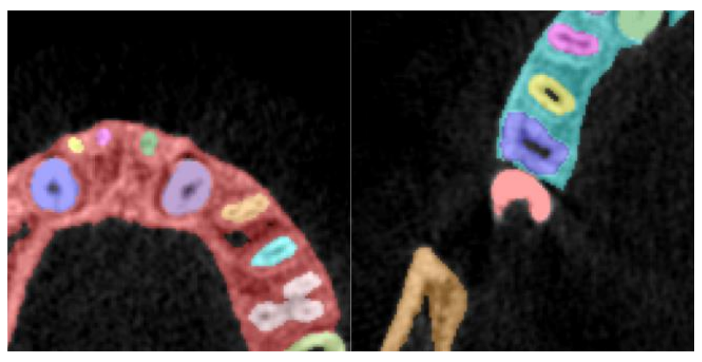


Fig.5. 112x1282 voxel training patches: (A) deciduous incisors and permanent buds; (B) deciduous molars with an erupting first permanent molar. Colored overlays show tooth labels on the CBCT.

DISCUSSION

Our nnU-Net v2 model segments primary and permanent teeth in pediatric CBCT with a mean Dice of 0.87, approaching expert performance. The output is a complete, volumetric reconstruction where each tooth is an individual 3D object, enabling extraction simulation, eruption-path prediction, surgical guidance, and 3D orthodontic planning while emphasizing judicious CBCT use in pediatrics.

Methodologically, a single-head 55-class network handles mixed dentition in one pass, preserving inter-class geometry without heuristic pipelines. Compared with watershed/atlas approaches, the model infers morphology directly and maintains anatomical coherence. Disabling left–right flips reduces contralateral mislabeling and stabilizes side-specific predictions.

Performance sits in the mid-range of reported per-tooth Dice (0.75–0.93): better than early CNNs and below the very best permanent-tooth-only systems, but unique in addressing mixed dentition comprehensively. Targeted augmentations help in regions with weak separations or artifacts, though metal artifacts remain a limiting factor.

Limitations include underrepresented anatomical outliers (e.g., syndromic cases) and cross-scanner variability. Future work will expand rare anatomies, explore shape-aware losses, fuse CBCT with intraoral scans, and integrate geometric learning tools to further boost fidelity and clinical utility.

INTRODUCTION

Understanding and analyzing the 3D shape of anatomical structures is a cornerstone of medical image computing. In dentistry, tooth morphology including crown and root geometries directly impacts diagnosis, treatment planning, and monitoring. Recent advances in artificial intelligence, particularly convolutional neural networks (CNNs) have significantly improved the segmentation of anatomical structures in medical images. They now learn shape representations from voxel data, overcoming the limitations of thresholding or region growing caused by partial volume effects and anatomical variability. Accurate, automatic segmentation of individual teeth in cone-beam CT (CBCT) is especially critical for pediatric mixed dentition, which involves coexisting primary and erupting permanent teeth, unerupted buds, resorbing roots, and impacted teeth all of which introduce significant shape variability. While previous methods have shown strong results for permanent teeth [1], few have addressed the challenge of comprehensive shape segmentation across all dentition types in pediatric CBCTs. This paper explores a fully automatic segmentation approach using the self-configuring nnU-Net v2 framework [3,4]. Unlike traditional pipelines that require manual network tuning or rule-based preprocessing, nnUNet v2 adapts its architecture and training plan to the geometry of the input data. This makes it particularly well-suited for tasks involving complex, irregular, and densely packed structures such as mixed dentition. Our study represents the first ap plication of nnUNet v2 to the joint segmentation of both primary and permanent teeth, including all 52 dental and 3 skeletal supporting structures, treated as independent classes in a high-resolution 3D domain. Our contributions are twofold: (1) to design and evaluate a self-configuring deep learning pipeline tailored to pediatric dental anatomy in CBCT; and (2) to demonstrate how this approach enables high-fidelity geometric modeling of the entire dentition, facilitating downstream applications such as orthodontic planning, eruption tracking, and surgical guidance.

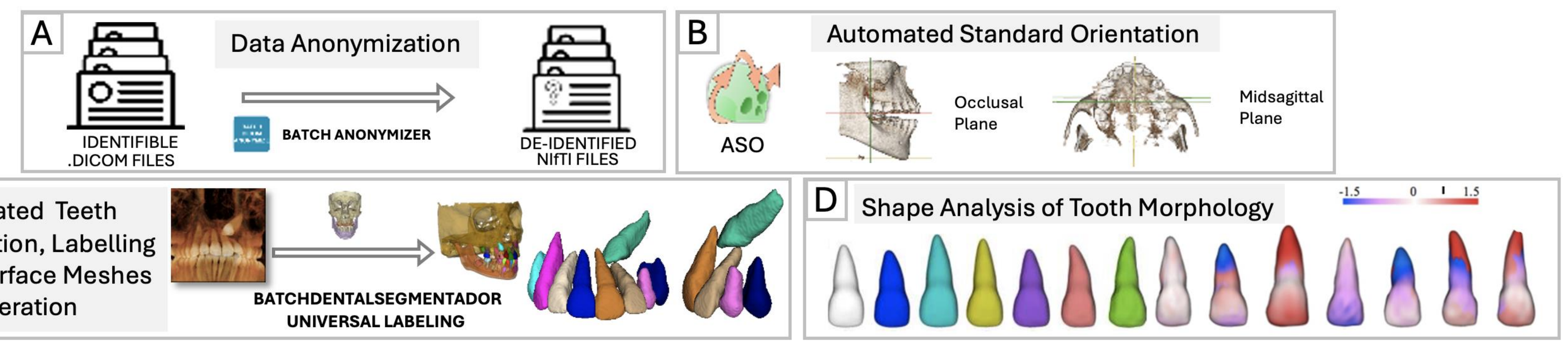


Fig.1. End-to-end automated dental analysis , deployed as modules of 3D Slicer software (A) Batch Anonymizer converts identifiable DICOM image stacks into de-identified NIfTI volumes. (B) Automated Standard Orientation (ASO) aligns each volume to the occlusal and midsagittal planes, ensuring consistent orientation across patients. (C) The BatchDentalSegmentador module performs fully automatic tooth segmentation. (D) Individual tooth meshes undergo statistical shape analysis to quan tify morphological variability.

RESULTS

We evaluated the model on 55 pediatric CBCT scans and reported Dice (DSC) and IoU for all 55 classes. The evaluation covers every individual tooth plus mandible, maxilla, and mandibular canal, ensuring a complete, label-by-label assessment of mixed dentition.

Accuracy. Permanent teeth achieve an average Dice of ~0.90, while primary teeth average ~0.85. Importantly, all annotated structures were recovered no missing labels indicating robust coverage across the full anatomy.

Per-class behavior. Lower scores concentrate in anatomically variable or sparsely represented classes, notably third molars and primary incisors. This pattern is consistent with the expected variability in eruption stage and morphology in pediatric cohorts.

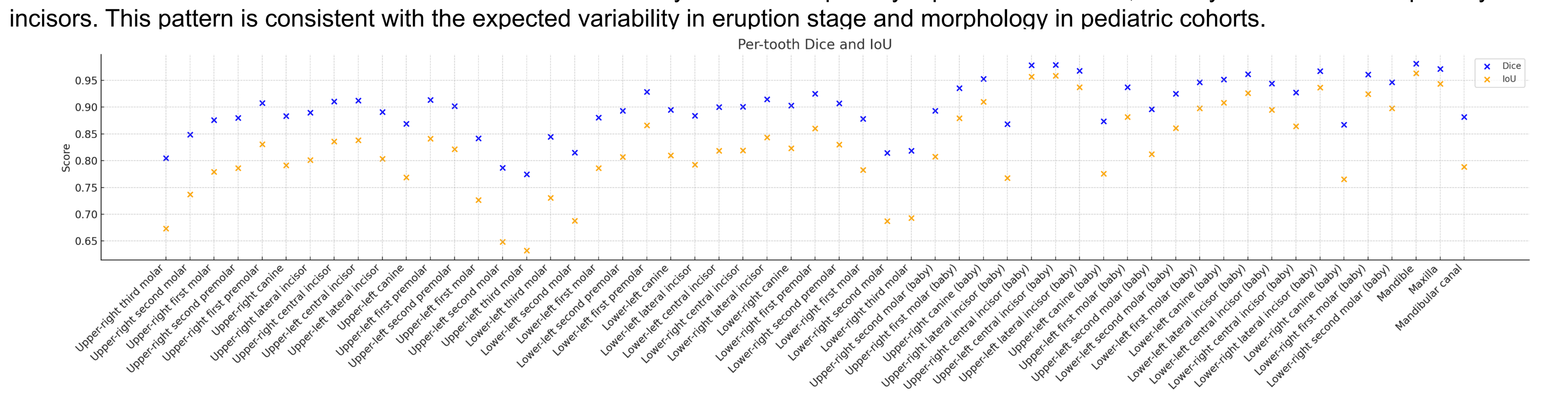


Fig.6. Per-class accuracy on the test set. Blue circles=Dice, orange triangles=IoU, for the 55 labels ordered from the upper right third molar . Most permanent teeth score Dice ≥ 0.90, primary teeth around 0.90. Lower values appear only for the rare third molars.

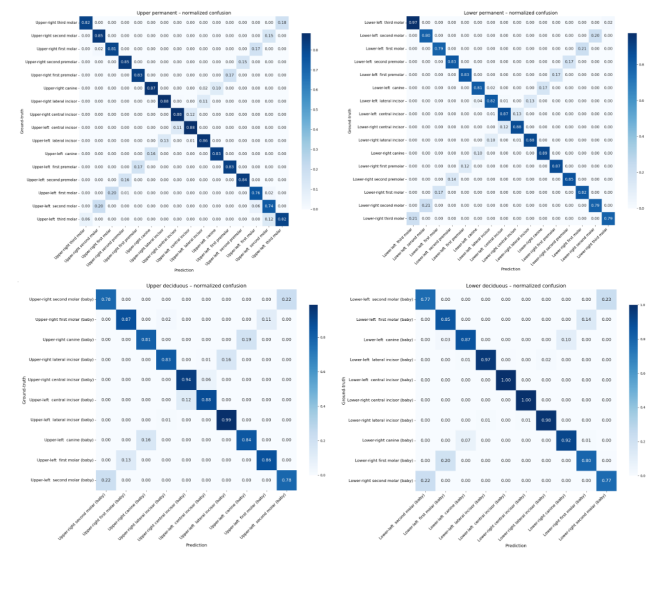
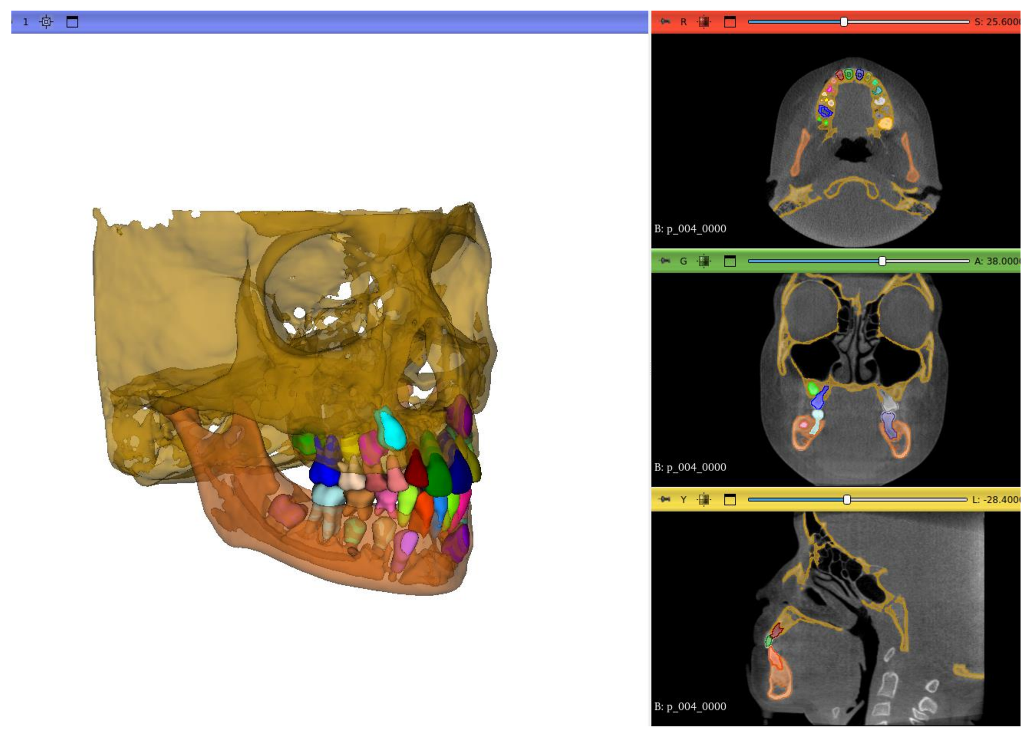


Fig.7. Normalized confusion matrices for upper/lower permanent and upper/lower primary teeth. Notably, errors are rare and usually limited to neighboring classes or to the same tooth on the opposite side (i.e., right–left confusion), especially among primary incisors and molars.

Error profile (confusion matrices). Misclassifications, when they occur, are typically restricted to neighboring or homologous teeth (e.g., left–right counterparts). Right–left confusions are rare and mainly observed among primary incisors and molars, where shape similarity is highest

Qualitative assessment. Visual overlays show close adherence to true boundaries, with clear separation between adjacent teeth even in challenging scenarios such as unerupted buds and resorbing roots. We did not observe major label swaps, and individual tooth identities remain stable across planes.



Comparison to DentalSegmentator (nnU-Net v2). Our predictions align more closely with ground truth at inter-tooth contacts, yielding cleaner boundaries and fewer merge errors. The improvement is most visible around crowded regions where subtle gaps matter clinically.

Fig.8. Qualitative example of the automatic segmentation. Left: 3-D surface rendering of the maxilla, mandible and teeth generated from the predicted labels. Right: axial, coronal and sagittal CBCT slices with the same labels over-laid in semi-transparency.

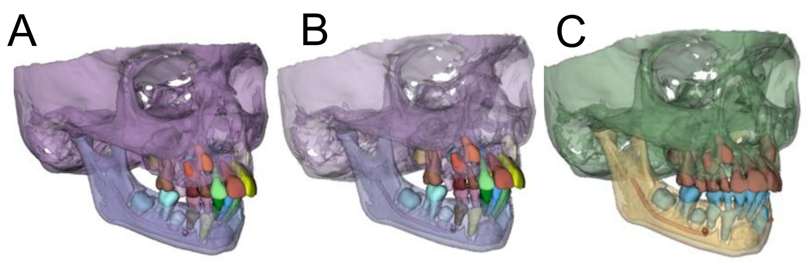


Fig.9. Outputs A. Ground Truth; **B..** DentalSegmentator

Runtime and usability. Inference runs in about 2 minutes on an RTX A6000 GPU and under 7 minutes on CPU, cutting manual effort by over 90%. The pipeline is packaged within 3D Slicer for chairside use, producing per-tooth 3D objects ready for measurement, planning, or virtual extraction.

CONCLUSION

We have presented a fully automatic segmentation pipeline that leverages the self- configuring nnU-Net v2 framework to model the complete mixed dentition in pediatric CBCT volumes. Our approach accurately segments all 55 anatomical structures, including both primary and permanent teeth, while preserving the geometric fidelity of the voxel level. The method achieves a mean Dice score of 0.87 and delivers anatomically coherent segmentations suitable for clinical and computational applications. By capturing the full spatial extent and inter class relationships of dental structures, our pipeline enables the generation of patient-specific digital twins in which each tooth is represented as a manipulable 3D object. These models support shape-based clinical tasks such as eruption assessment, surgical planning, and orthodontic treatment design. Critically, the workflow runs in near real-time, enabling seamless integration into diagnostic pipelines without manual post processing.

Supported by NIDCR R01DE024450 & CAPES