Self-Configuring 3D Segmentation of Pediatric Dentition

Enzo Tulissi[1,2], Alban Gaydamour[1,2], Juan C. Prieto[3], Claudia Mattos[4], Renata R. Rosa[4], Sara Tinawi[1], Dylan J. Keener[1], Aron Aliaga Del Castillo[1], Eduardo Caleme[6], Brent Larson[7], Antonio C. de Oliveira Ruellas[8], Luis E. Arriola-Guillén[9], Jonas Bianchi[10], Heesoo Oh[10], Marcela Lima Gurgel[11], Erika Benavides[1], Fabiana Soki[1], Yalil A. Rodríguez-Cárdenas[12], Gustavo A. Ruíz-Mora[12], Bruno M. R. Braga[5], Ana B. Teodoro[13], Selene Barone[14], Martin Styner[3], Roberto Bespalez-Neto[15], and Lucia H. Cevidanes[1]

[1] University of Michigan, Ann Arbor, MI, USA
[2] CPE Lyon, France
[3] University of North Carolina, Chapel Hill, NC, USA
[4] Universidade Federal Fluminense, Brazil
[5] Hospital for Rehabilitation of Craniofacial Anomalies, University of São Paulo, Brazil
[6] Universidade Positivo, Curitiba, Paraná, Brazil
[7] University of Minnesota, Minneapolis, MN, USA
[8] Federal University of Rio de Janeiro, Brazil
[9] Universidad Científica del Sur, Lima, Perú
[10] University of the Pacific, San Francisco, CA, USA
[11] Federal University of Ceara, Fortaleza, Brazil
[12] Universidad Nacional de Colombia, Bogotá, Colombia
[13] Federal University of Goiás, Goiânia, Brazil
[14] Magna Graecia University, Catanzaro, Italy
[15] University Anhanguera/Uniderp, Campo Grande, MS, Brazil

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 actionable, 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: 3D segmentation · pediatric dentistry · Cone-beam CT (CBCT) · deep learning· nnU-Net · dental morphometry.

1 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 [1, 2]. 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 application 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.

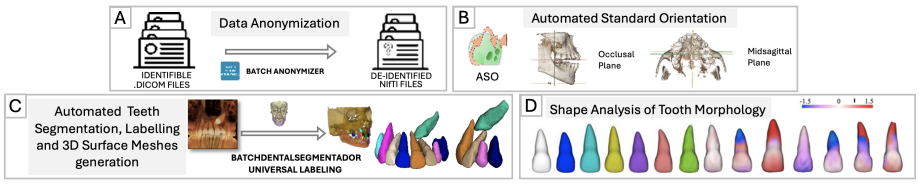


Fig. 1. End-to-end automated dental analysis , deployed as modules of 3D Slicer software [5] (A) *Batch Anonymizer* converts identifiable DICOM image stacks into de-identified NIfTI volumes. **(B)** *Automated Standard Orientation* (ASO) aligns each volume to the occlusal and midsagittal planes, ensuring consistent orientation across patients. **(C)** The *BatchDentalSegmentador* module performs fully automatic tooth segmentation. **(D)** Individual tooth meshes undergo statistical shape analysis to quantify morphological variability.

2 Methods

2.1 Dataset and Annotation Strategy

To enable robust shape modeling of the mixed dentition (Fig.1), we curated a retrospective dataset of 517 pediatric CBCT scans exhibiting both primary and permanent teeth [6]. Each volume encompasses complete maxillary and mandibular arches and captures a broad spectrum of tooth development stages, including unerupted buds, resorbing roots, and impactions, each presenting unique geometric challenges. On average, each scan includes 6–8 primary teeth and over 28 permanent teeth in varying states of eruption. Initial segmentations were generated using the DentalSegmentator module in 3D Slicer [7] and refined manually in ITK-SNAP [8] to assign anatomically precise voxel-wise labels for each discernible tooth. This resulted in a dense multi-class segmentation with 55 total unique anatomical structures, each representing 52 distinct dental shapes, the upper jaw/cranium, lower jaw and the mandibular canal. Importantly, our labeling captures detailed morphologies, including curved roots, crown morphology, and interproximal spacing, thereby supporting downstream geometric analysis. Unerupted or malformed structures were included when identifiable, ensuring anatomical completeness. The dataset was split into 369 scans for training, 93 for validation, and 55 for testing. This diverse set provides a strong basis for generalizable 3D learning across a spectrum of pediatric dental geometries.

2.2 Self-Configuring Segmentation Pipeline

We implemented the nnU-Net v2 pipeline [3, 4], 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)[9]. This design is especially suited for medical shape analysis tasks, where anatomical variability and image heterogeneity demand robust adaptation.

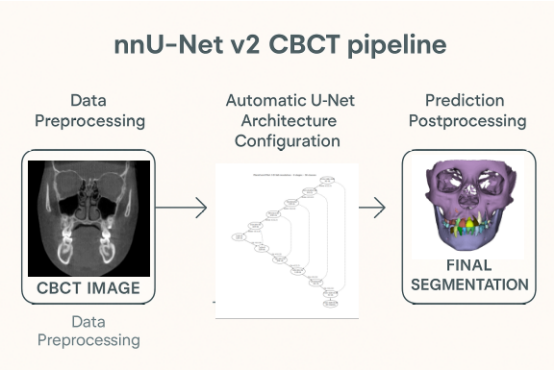


Fig. 2. Illustration of the nnU-Net v2 segmentation pipeline. Left: coronal slice from the original pediatric CBCT volume. Right: voxel-wise prediction showing all 55 teeth individually.

Preprocessing : To prepare the CBCT volumes for geometric learning Fig 3, nnU-Net performs: (1) Resolution normalization. All scans are resampled to a consistent isotropic spacing, preserving geometric proportions and enabling scale-invariant feature learning. (2) Spatial cropping. The field-of-view is restricted to the craniofacial region via automatic bounding box detection, focusing the learning process on relevant anatomical context. (3) CT-based intensity normalization. Voxel intensities are clipped and z-normalized to homogenize grayscale representations of bone and soft tissue, preserving contrast critical for shape boundaries. (4) On-the-fly 3D augmentation. During training, randomized transformations such as elastic deformations, rotations, and intensity shifts are applied to simulate anatomical variation and imaging noise without distorting the underlying geometry. These steps produce standardized yet anatomically diverse inputs that allow the network to learn generalizable shape features.

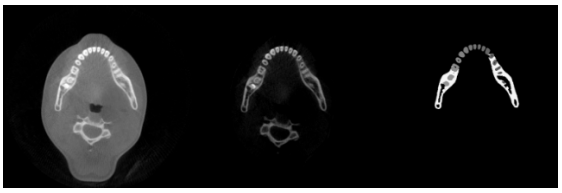


Fig. 3. nnU-Net v2 preprocessing. (Left) Original axial pediatric CBCT. (Center) After automatic cropping and intensity normalization. (Right) Segmentation of 55 anatomical structure (52 dental and 3 skeletal), plus background (label 0), in the preprocessed space.

Architecture Configuration : nnU-Net v2 automatically selected a full-resolution 3D U-Net with six resolution levels (about 30 M parameters), omitting low-resolution cascades since our volumes fit in GPU memory. Training patches measured 112 x 128 x 128 voxels, covering several teeth while respecting memory limits Fig.4. The encoder applies successive strides to downsample, and a symmetric decoder with skip-connections reconstructs spatial detail. The network concludes with a $1 \times 1 \times 1$ convolution followed by softmax activation and a voxel-wise argmax over 56 classes. This full-resolution setup proved both accurate and efficient, without needing 2D or multi-resolution variants.

Automated training schedule : We trained using the default nnU-Net v2 optimization, which automatically configures hyperparameters, learning rate scheduling, and optimizer selection. A combined soft Dice and cross-entropy loss handled the severe class imbalance across 56 labels. Each iteration sampled two random 112x128x128 patches, with on-the-fly augmentations (rotations, elastic deformations, zooms, noise) to improve generalization. We trained for 150 epochs (3000 iterations per epoch) per fold using 5-fold cross-validation and ensembled the softmax outputs, to produce consistent, anatomically robust segmentations.

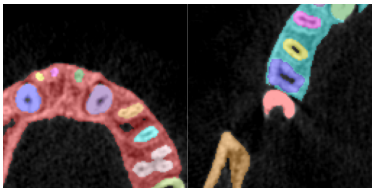


Fig. 4. 112x128^2 voxel training patches: (A) deciduous incisors and permanent buds; (B) deciduous molars with an erupting first permanent molar. Colored overlays show tooth labels on the CBCT.

Postprocessing : After inference, connected components are extracted per class and small isolated false positives are removed, retaining the largest anatomically plausible shape per label. This preserves the integrity of each tooth's morphology and ensures separation of closely spaced structures. This light-weight postprocessing step helps preserve geometric consistency without requiring external priors or templates. Post-segmentation , surface meshes are generated and made available for shape analysis. We enforced shape smoothness via Laplacian regularization.

3 Results

3.1 Quantitative Segmentation Performance

We tested our nnU-Net v2 on 55 pediatric CBCT scans, reporting Dice (DSC) and Intersection-over-Union (IoU) across all 55 labels. As Figure 5 and Table 1 show, permanent teeth average DSC about 0.90, while smaller primary teeth average about 0.85. All annotated structures were recovered (no missing labels), matching prior CBCT segmentation benchmarks [10].

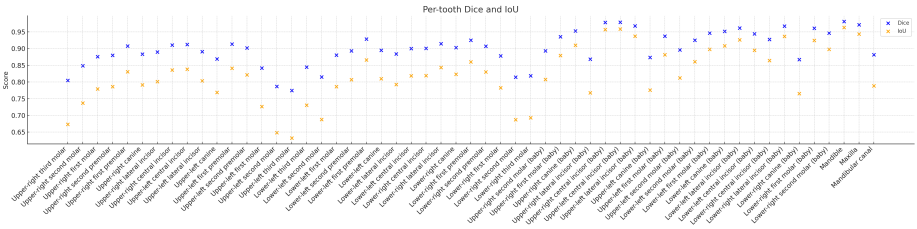
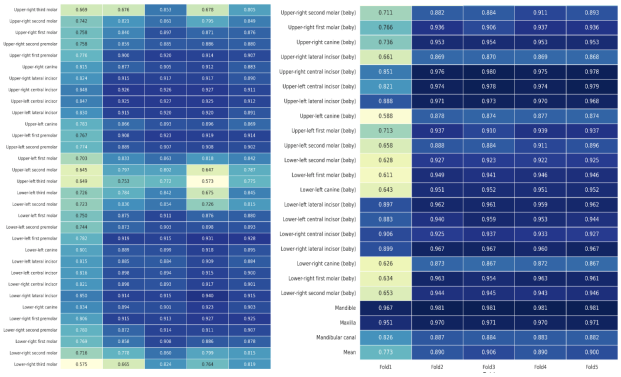


Fig. 5. Per-class accuracy on the test set. Blue circles = Dice, orange triangles = IoU, for the 55 labels ordered from the upper right third molar . Most permanent teeth score Dice $\geq$ 0.90, primary teeth around 0.90. Lower values appear only for the rare third molars.

Table 1. Results demonstrate the model's capacity to robustly segment both primary and permanent teeth across all regions of the dentition. Lower Dice and higher variability are observed for third molars and primary incisors, reflecting anatomical variation and limited sample representation.



To further characterize the segmentation performance at the voxel level, we computed per-class confusion matrices grouped by dentition type. Figure 6 shows the normalized confusion matrices separately for permanent and deciduous teeth, divided into upper and lower arches. The matrices highlight that misclassifications, when present, are mostly limited to adjacent teeth or homologous structures, especially among primary teeth with similar shapes.

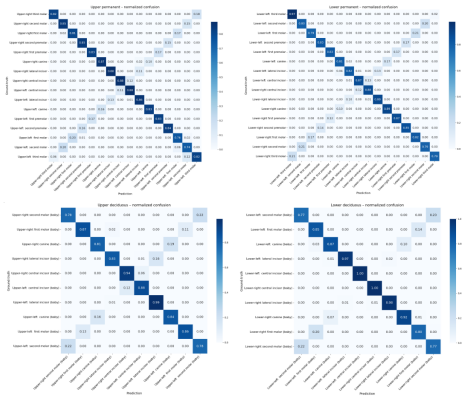


Fig. 6. Normalized confusion matrices for upper/lower permanent and upper/lower primary teeth. Notably, errors are rare and usually limited to neighboring classes or to the same tooth on the opposite side (i.e., right–left confusion), especially among primary incisors and molars.

3.2 Qualitative visual assessment

Figure 7 shows a representative test-set CBCT with the nnU-Net v2 segmentation overlaid. The automatic masks align closely with true tooth boundaries, isolating each tooth in a distinct color and accurately separating adjacent structures even unerupted buds and resorbing roots. No major label swaps were observed.

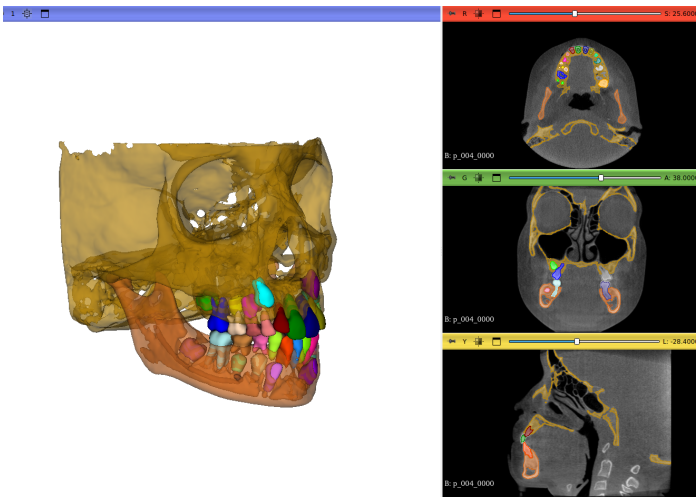


Fig. 7. Qualitative example of the automatic segmentation. **Left:** 3-D surface rendering of the maxilla, mandible and teeth generated from the predicted labels. **Right:** axial, coronal and sagittal CBCT slices with the same labels over-laid in semi-transparency.

Clinically, our open-source method, deployed and available in the 3D Slicer platform delivers a patient specific digital twin within minutes: each tooth becomes a separate 3D object for measurement or virtual extraction. Segmentation takes 2 min on GPU or $<$7 min on CPU, cutting manual effort by $>$ 90 percent and enabling chair-side deployment [10].

4 Discussion

Our results confirm that nnU-Net v2 accurately segments primary and permanent teeth in pediatric CBCT, achieving a mean Dice of 0.87, performance that approaches expert level [10]. We discuss clinical implications, compare with prior dental-segmentation work, and outline limitations and future directions.

Precise 3D tooth segmentation enables advanced geometric analyses in pediatric dentistry. Mixed dentition presents variable anatomy unerupted buds, erupting premolars, resorbing roots while 2D imaging fails to capture these shapes and manual segmentation is slow and inconsistent. We reconstruct the entire dentition volumetrically, each tooth as a separate 3D object. These models

support applications like extraction simulation, eruption-path prediction, surgical guidance, and 3D orthodontic planning. Our pipeline is fast, reliable, and maximizes clinical CBCT value. While we emphasize enhancing the utility of already prescribed scans, we also stress the importance of judicious CBCT use in pediatric populations [11].

We demonstrate how nnU-Net v2, a self-configuring framework, extends DentalSegmentator for pediatric tooth shape modeling [1, 12], segmenting 55 classes in pediatric CBCT while preserving inter-class 3D geometry. A single nnU-Net generalizes across tooth types and stages without custom tuning, leveraging shape priors during augmentation and outputting semantic labels directly, avoiding heuristic post-processing. Unlike watershed or atlas methods [13], our model infers tooth morphology directly while maintaining anatomical coherence.

Where classical methods fail near weak separation or artifacts, our network remains robust through targeted data augmentation. Although Metal artifacts continue to limit segmentation accuracy [14], training with bracketed scans may help [15]. Reported per-tooth Dice ranges (0.75–0.93) [14] place our mean of 0.87 in the mid-range: better than early CNNs but below top permanent tooth only pipelines, yet we uniquely handle mixed dentition in one pass. We disabled left–right flips to avoid laterality errors (e.g., mis-labelling contralateral teeth) and achieved 56-class segmentation in a single forward pass streamlining over previous binary or hierarchical approaches while preserving anatomical fidelity.

Limitations include underrepresented anatomical outliers (e.g., syndromic cases) and unmodeled cross-scanner variability. Future work will aim to: (i) expand training to include rare anatomies, (ii) incorporate shape-aware loss functions, (iii) fuse with intraoral scans, (iv) integrate geometric learning tools [16].

5 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 vol- umes. 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.

Acknowledgments. This work was funded by NIH, grant number R01-DE024450.

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

References

1. Lahoud, P., EzEldeen, M., Beznik, T., et al.: Artificial intelligence for fast and accurate 3-dimensional tooth segmentation on cone-beam computed tomography. *Journal of Endodontics* **47**(5), 827–835 (2021)
2. Tarce, M., Zhou, Y., Antonelli, A., Becker, K.: The Application of Artificial Intelligence for Tooth Segmentation in CBCT Images: A Systematic Review. *Applied Sciences* **14**(14), 6298 (2024)
3. Isensee, F., Jäger, P.F., Kohl, S.A.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)
4. Isensee, F., Kirchhoff, Y., Kraemer, L., Rokuss, M., Ulrich, C., Maier-Hein, K.H.: Scaling nnU-Net for CBCT Segmentation. arXiv:2411.17213 [cs.CV] (2024)
5. Slicer Automated Dental Tools github : `https://github.com/DCBIA-OrthoLab/SlicerAutomatedDentalTools`, last accessed 2025/07/22
6. Cui, W., Wang, Y., Zhang, Q., et al.: CTooth: A Fully Annotated 3D Dataset and Benchmark for Tooth Volume Segmentation on Cone Beam Computed Tomography Images. arXiv:2206.08778 [cs.CV] (2022)
7. 3D Slicer page : `https://www.slicer.org/`, last accessed 2025/07/22
8. ITK-SNAP page : `https://www.itksnap.orgto`, last accessed 2025/07/22
9. Beser, B., Reis, T., Berber, M.N., et al.: YOLO-V5 based deep learning approach for tooth detection and segmentation on pediatric panoramic radiographs in mixed dentition. *BMC Medical Imaging* **24**, 172 (2024)
10. Elsonbaty, S., Elgarba, B.M., Fontenele, R.C., Swaity, A., Jacobs, R.: Novel AI-based tool for primary tooth segmentation on CBCT using convolutional neural networks: a validation study. *International Journal of Paediatric Dentistry* **35**(1), 31–40 (2025)
11. Oenning, A.C., Jacobs, R., Pauwels, R., Stratis, A., Hedesiu, M., Salmon, B.: Cone-beam CT in paediatric dentistry: DIMITRA project position statement. *Pediatric Radiology* **48**(3), 308–316 (2018)
12. Shaheen, E., Leite, A., Alqahtani, K.A., et al.: A novel deep learning system for multi-class tooth segmentation and classification on cone beam computed tomography: A validation study. *Journal of Dentistry* **115**, 103865 (2021)
13. Galibourg, A., Dumoncel, J., Telmon, N., Calvet, A., Michetti, J., Maret, D.: Assessment of automatic segmentation of teeth using a watershed-based method. *Dentomaxillofacial Radiology* **47**(3), 20170220 (2018)
14. Polizzi, A., Quinzi, V., Ronsivalle, V., et al.: Tooth automatic segmentation from CBCT images: a systematic review. *Clinical Oral Investigations* **27**(6), 3363–3378 (2023)
15. Alqahtani, K.A., Jacobs, R., Smolders, A., et al.: Deep convolutional neural network-based automated segmentation and classification of teeth with orthodontic brackets on cone-beam computed tomographic images: a validation study. *European Journal of Orthodontics* **45**(2), 169–174 (2023)
16. Zhang, Y., Liu, Z., Feng, Y., Xu, R.: 3D-U-SAM Network For Few-shot Tooth Segmentation in CBCT Images. arXiv:2309.11015 [cs.CV] (2023)