Cross-surface Representation Learning for Accurate Gingiva Prediction in CBCT

Lanzhuju Mei[1], Yuan Li[2], Han Wu[1], Wen Gu[2], Chong Wang[5], Maurizio Tonetti[2], Zhiming Cui[1(✉)], and Dinggang Shen[1,3,4(✉)]

[1] School of Biomedical Engineering, State & Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai, China
cuizm.neu.edu@gmail.com, Dinggang.Shen@gmail.com
[2] Shanghai Ninth People's Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, China
[3] Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China
[4] Shanghai Clinical Research and Trial Center, Shanghai, China
[5] Department of Radiology, Stanford University, Stanford, CA, USA

Abstract. Accurate gingival segmentation from Cone-Beam Computed Tomography (CBCT) is critical for periodontitis management and implantology, yet remains challenging due to low soft tissue contrast and the adhesion of adjacent tissues such as the cheeks and the tongue. In standard CBCT, the gingival margin is essentially indistinguishable from the surrounding mucosa, so existing methods either rely on additional intra-oral scan (IOS) data or apply voxel-based deep-learning methods that struggle with boundary precision. To address these limitations, this paper proposes a protocol-and-method co-design. First, we introduce a chair-side scanning protocol that uses a mouth retractor and an impression tray to physically isolate the gingiva from buccal and lingual tissues during acquisition, creating a continuous high-contrast air interface that makes the gingival margin imageable in CBCT in the first place; the protocol requires no scanner modification and adds less than one minute to the standard CBCT workflow. Second, building on this controlled input, we propose a cross-surface representation (CSR) that recasts the 3D volumetric segmentation as 1D radiodensity profile regression along tooth surface vertex normals, together with a band-aware Gaussian regression (BGR) module that produces a continuous, topologically consistent confidence field on the tooth manifold. Experiments demonstrate that the method outperforms strong 3D voxel baselines, achieving expert-level accuracy without auxiliary IOS data.

Keywords: Gingiva · CBCT · Periodontal Assessment.

1 Introduction

The gingival margin is a primary anatomical barrier and a fundamental determinant of periodontal stability [22, 1, 3]; its accurate delineation is critical for clinical measurement such as pocket depth and for assessing periodontal phenotypes [9, 10, 16, 17, 12, 20], with the 2017 World Workshop classification emphasizing that a thin gingival phenotype (< 1 mm) significantly heightens the risk of recession and attachment loss [15, 7]. Yet conventional transgingival probing remains invasive, uncomfortable, and susceptible to measurement errors of approximately 0.5 mm [8]. Digital alternatives such as IOS provide high-resolution superficial contours but inherently fail to capture the underlying alveolar bone and internal supracrestal attachments; clinicians therefore have to register IOS with CBCT to evaluate the complete periodontal complex, a workflow that is time-consuming, costly, and prone to registration errors.

While CBCT is the gold standard for three-dimensional hard tissue imaging, its application in periodontal soft tissue evaluation is severely restricted by low contrast resolution and scatter radiation [6]. The negligible density difference among the oral mucosa, lips, and tongue results in tissue adhesion on scans, making the gingival boundaries virtually indistinguishable, as shown in the top row of Fig. 1.

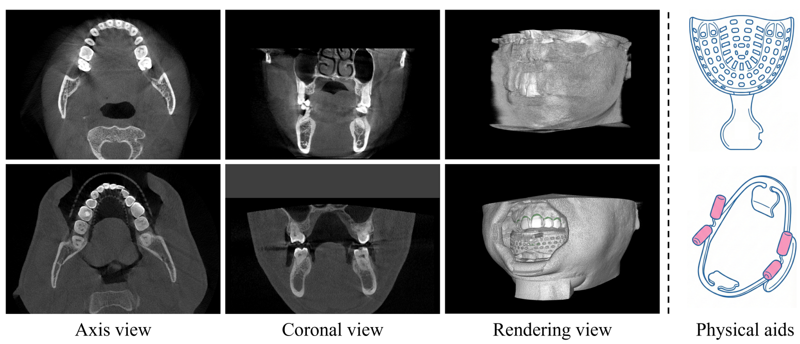

Fig. 1. Comparison of CBCT imaging protocols. (Top) Standard scans exhibit severe tissue adhesion and indistinguishable boundaries. (Bottom) Our standardized protocol with physical aids establishes a high-contrast air interface, providing distinct physical boundaries for precise digital processing.

Studies on direct soft-tissue segmentation from CBCT remain limited and are typically tackled with a standard 3D CNN backbone [24], inheriting the voxel-based recipe of V-Net [18], DeepLabV3+ [2], and nnU-Net [5] that also dominates CBCT tooth and bone segmentation [25]; none of these methods address the acquisition-side contrast problem. A parallel surface-based line of work, including arch-guided tooth arrangement [23] and mesh or point-cloud backbones such as PointNet++ [19] and PointMLP [14], has not yet been applied to gingival margin localization. While voxel models demonstrate excellent performance in general medical segmentation, their application to fine-grained gingival identification from CBCT remains limited: the partial volume effect and resolution limitations blur the already thin gingival band, and these voxel-wise approaches treat the gingival margin as an independent volumetric mask rather than a boundary strictly constrained by the tooth surface. The absence of explicit anatomical prior knowledge limits their precision and motivates a more robust formulation that integrates geometric constraints with deep learning.

To address these issues, we introduce a standardized physical protocol for creating a continuous, high-contrast air interface across the dental arch, shown in the bottom row of Fig. 1. By reconstructing the tooth surface and using it as a geometric prior, we propose a cross-surface representation to transform 3D volumetric segmentation into structured analysis along vertex normals. This manifold-based approach exposes distinct bone–gingiva–air transition patterns, enabling continuous scalar-field regression instead of discrete classification. Supervised by a band-aware hybrid loss, our framework ensures topologically smooth and clinically accurate margins, providing a robust tool for gingival margin assessment without auxiliary scans.

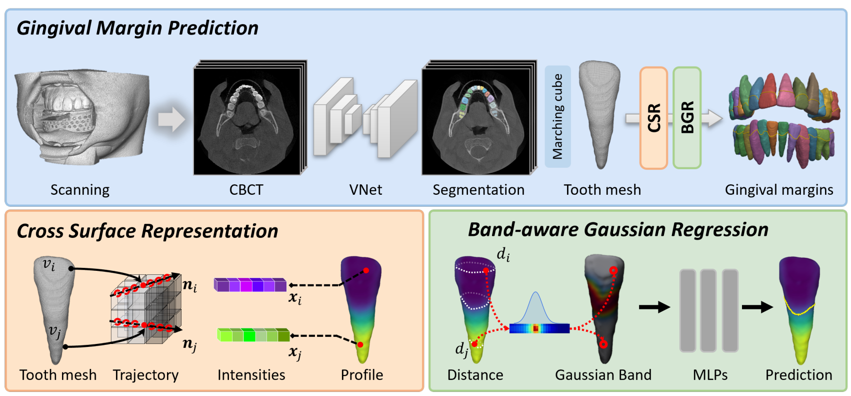

Fig. 2. Overview of the proposed framework. Geometric Prior Construction (tooth mesh) via Physical Separation and Instance Segmentation with Marching Cubes; the Cross-Surface Representation (CSR) encodes radiodensity along vertex normals; the Band-Aware Gaussian Regression (BGR) module then regresses a continuous, topologically consistent confidence field for the gingival margin.

2 Methodology

2.1 Geometric Prior Construction

Physical Scanning Protocol: To ensure optimal visibility of the gingival margin, as illustrated in Fig. 1, we introduce a physical scanning protocol utilizing commonly available chair-side clinical aids, specifically a mouth retractor and an impression tray. The retractor is used to displace the buccal mucosa, while the impression tray provides a stable platform for tongue placement, ensuring isolation of the lingual and palatal gingiva. Collectively, these tools establish a continuous, high-contrast air interface that physically separates the gingiva from adjacent soft tissues, providing the deterministic physical boundary necessary for the automated pipeline. The aids are routinely available, add under one minute to the standard CBCT workflow, and admit drop-in variants (e.g. smaller retractors or trimmed trays) for patients with severe gag reflex, restricted mouth opening, or extensive tooth loss. The present scope is CBCT acquired under this protocol; routine or metal-artifact-dominated CBCT is left as future work.

Tooth Surface Reconstruction: Building upon this standardized imaging, we further establish a deterministic geometric prior for gingival prediction by reconstructing the tooth surface, to which the gingival margin is anatomically attached. We deploy an instance segmentation network to extract individual teeth from the CBCT volume V. Let $T = \{1, \ldots, 32\}$ denote the set of universal tooth IDs. The reconstruction of the surface mesh $\mathcal{M}_t$ for a specific tooth $t \in T$ is defined by:

$$\mathcal{M}_t = \mathrm{MC}\left(\mathbf{1}_{>\tau}\left([S_{\Phi}(V)]_t\right)\right) = (\mathcal{V}_t, \mathcal{F}_t), \tag{1}$$

where S_{Φ} denotes the segmentation network parameterized by Φ that generates a multi-channel probability tensor, $[\cdot]_t$ isolates the channel corresponding to tooth ID t, $\mathbf{1}_{>\tau}$ performs binary thresholding at level τ, and $\mathrm{MC}(\cdot)$ represents the Marching Cubes algorithm followed by Laplacian smoothing to ensure surface manifold regularity. The refined mesh $\mathcal{M}_t$, defined by its vertices $\mathcal{V}_t$ and faces $\mathcal{F}_t$, serves as a deterministic geometric prior that constrains the search space strictly to the tooth surface. This anatomical constraint filters out extrinsic imaging noise and adjacent structures, ensuring topological consistency and clinical plausibility across the dental arch.

2.2 Cross-surface Representation

To capture the subtle intensity changes at the gingival margin, we construct the cross-surface representation. This formulation shifts the computational focus from the unconstrained 3D volume to a targeted surface-normal domain anchored by the geometric prior $\mathcal{M}_t$. By sampling intensities perpendicular to the tooth surface, the cross-surface representation explicitly models the anisotropic radiodensity transitions where the contrast between hard tissue, soft tissue, and air is most discriminative. Specifically, for each vertex $v_i \in \mathcal{V}_t$ on the tooth mesh, we define a sampling trajectory $\mathcal{T}_i$ extending outwards along the unit normal vector $\mathbf{n}_i$. This trajectory is discretized into a sequence of K spatial coordinates:

$$\mathcal{T}_{i,k} = v_i + k \cdot \delta \cdot \mathbf{n}_i, \quad k \in \{0, 1, \ldots, K-1\}, \tag{2}$$

where δ denotes the sampling stride and K determines the radial receptive field. The cross-surface representation feature vector $\mathbf{x}_i \in \mathbb{R}^K$ is constructed by projecting the volumetric CBCT intensities onto this trajectory via a trilinear interpolation operator Ψ:

$$\mathbf{x}_i = \Psi(V, \mathcal{T}_i) = \left(V[\mathcal{T}_{i,0}], V[\mathcal{T}_{i,1}], \ldots, V[\mathcal{T}_{i,K-1}]\right)^{\top}, \tag{3}$$

where $V[\cdot]$ represents the continuous intensity function interpolated from the discrete voxel grid. The resulting structured profile $\mathbf{x}_i$ serves as a radiodensity signature. Specifically, the cross-surface representation profile characterizes three critical anatomical zones: (1) a supra-gingival zone with low intensity, reflecting radiolucent air; (2) a gingival zone with an intermediate intensity plateau, corresponding to soft tissue encapsulated between bone and external air; (3) a sub-gingival zone with a high-intensity continuum where the root is closely opposed to the alveolar bone. By capturing these distinct signatures, the cross-surface representation effectively disentangles the gingival margin from neighboring anatomical noise, reformulating the complex 3D segmentation task into a robust and computationally efficient surface-normal analysis. Notably, for multi-rooted teeth, the gingival margin typically remains coronal to the furcation; even under advanced periodontal recession, the sampling band rarely reaches the divergent roots and is therefore unaffected by root multiplicity.

2.3 Band-aware Gaussian Regression

Formulation: Building on the tooth surface $\mathcal{M}_t$ and cross-surface representation, we reformulate gingival margin extraction as a continuous scalar-field regression task. Specifically, for each vertex $v_i \in \mathcal{V}_t$, we construct a fused feature vector $C_i = [\mathbf{n}_i \oplus \mathbf{x}_i] \in \mathbb{R}^{24}$ that fuses the local geometric orientation $\mathbf{n}_i$ with the cross-surface representation radiodensity profile $\mathbf{x}_i$. While these features capture local anatomical signatures, the raw spatial coordinates v_i are also fed into a point cloud encoder $\mathcal{E}$ to facilitate hierarchical grouping and neighborhood aggregation. The network learns to predict a continuous confidence field $\hat{y}_i$, representing the probabilistic proximity of each vertex to the ground-truth gingival margin:

$$\hat{y}_i = \mathcal{E}\left(v_i, C_i\right) \in [0, 1]. \tag{4}$$

To generate the ground-truth tooth-wise heatmap $\mathcal{Y}_t$, we first compute the Euclidean distance $d(v_i, \mathcal{G})$ from each vertex to the expert-annotated gingival curve $\mathcal{G}$ using a KD-Tree spatial index. This distance field is transformed into a smooth Gaussian response via the kernel:

$$y_i = \exp\left(-\frac{d(v_i, \mathcal{G})^2}{2\sigma^2}\right) \in \mathcal{Y}_t, \tag{5}$$

where σ controls the spatial variance. This formulation creates a soft gingival band $\mathcal{B}_t = \{v_i \mid y_i > \tau\}$, which softens the hard annotation margin and lets the model learn the intensity gradients characteristic of soft-tissue transitions.

Loss Function: To supervise this regression task while ensuring topological consistency, we use a hybrid objective $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{WBCE}} + \mathcal{L}_{\text{Dice}}$ that balances per-vertex fidelity with structural overlap. The weighted binary cross entropy loss $\mathcal{L}_{\text{WBCE}}$ treats the continuous heatmap values as soft probabilistic targets. Since the gingival band occupies only a minute fraction of the tooth surface, we implement an adaptive spatial weighting scheme to counteract this severe class imbalance:

$$\mathcal{L}_{\mathrm{WBCE}} = -\frac{1}{N}\sum_{i=1}^{N} w_i \left[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i) \right], \tag{6}$$

where w_i provides a positive emphasis factor to prioritize sparse gingival signals. Because cross entropy treats vertices independently and may yield spatially fragmented predictions, we add a soft Dice loss $\mathcal{L}_{\mathrm{Dice}}$ to explicitly optimize the global structural overlap between the predicted distribution $\hat{\mathcal{Y}}_t$ and the ground-truth $\mathcal{Y}_t$ in a differentiable manner:

$$\mathcal{L}_{\mathrm{Dice}} = 1 - \frac{2\sum_{i=1}^{N} \hat{y}_i y_i + \epsilon}{\sum_{i=1}^{N} \hat{y}_i + \sum_{i=1}^{N} y_i + \epsilon}. \tag{7}$$

This integrated loss ensures the resulting gingival margin is locally accurate and globally smooth across the manifold.

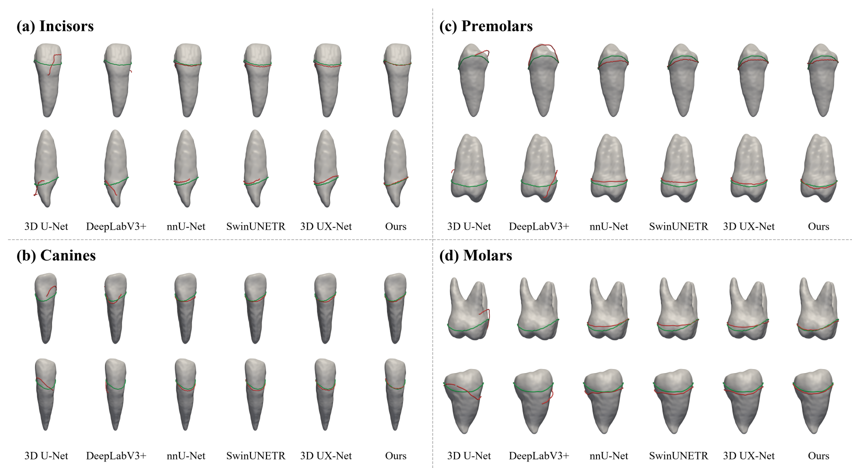

Fig. 3. Qualitative comparison across tooth categories: predicted (red) vs. ground-truth (green) margins for (a) Incisors, (b) Canines, (c) Premolars, and (d) Molars, illustrating the bone–gingiva–air transition exploited by CSR and the smoothness imposed by BGR.

3 Experiment

Dataset: Our framework is validated on a clinical dataset of 28 patient CBCT scans with 684 teeth (voxel size: $0.15 \times 0.15 \times 0.15$ mm^3). The cohort spans the full clinical range from gingivitis to Stage IV periodontitis. The dataset was partitioned at the patient level for five-fold cross-validation: each fold contains

Table 1. Comparison of gingival margin prediction methods across tooth categories (MAE in mm, IoU@0.3 mm in %). All baselines use the same protocol-acquired data and the same patient-level five-fold split.

Method	Incisors		Canines		Premolars		Molars		Overall	
	MAE	IoU	MAE	IoU	MAE	IoU	MAE	IoU	MAE	IoU
UNet3D	2.17	4.4	2.66	3.3	3.61	2.7	3.22	1.5	3.00	2.9
DeepLabV3+	2.11	5.2	2.69	2.2	2.86	3.6	2.68	3.6	2.56	3.9
nnUNet	0.50	33.8	0.44	35.5	0.50	32.3	0.54	37.4	0.50	34.6
SwinUNETR	0.49	36.1	0.46	37.6	0.49	36.0	0.51	40.0	0.48	37.1
3D UX-Net	0.47	37.7	0.44	39.0	0.47	37.6	0.49	41.6	0.46	38.8
Ours	**0.36**	**57.4**	**0.35**	**57.7**	**0.39**	**54.6**	**0.39**	**55.6**	**0.38**	**56.2**

disjoint patients, so no tooth from the same patient ever appears in both training and evaluation. The study was approved by the institutional review board under protocol `SH9H-2026-T98-1`; written informed consent was obtained from all participants, and all CBCT volumes were anonymized prior to processing. All scans were acquired using the standardized protocol described in Section 2.1, in which a generally acceptable mouth retractor and impression tray were employed to establish a consistent, high-contrast air gap. For ground-truth annotation only, intra-oral scans (IOS) were acquired and registered to the corresponding CBCT volumes via ICP; the proposed framework itself requires no IOS at inference, consistent with the abstract. Three experienced periodontists then refined the gingival margins on these references. Target teeth with full crowns or in-tooth implants were excluded, while minor restorations were retained.

Implementation Details: The pipeline runs on a single NVIDIA RTX 4090. The tooth mesh $\mathcal{M}_t$ is reconstructed by nnU-Net [5] for instance segmentation, followed by Marching Cubes [13] and Laplacian smoothing [4]; the upstream nnU-Net uses the standard 3d_fullres recipe on the same patient-level split and reaches 96% Dice on teeth, sufficient for stable downstream prediction on the retained restoration cohort. For the CSR we sample $K = 21$ points along each normal trajectory in $[-0.3, 1.7]$ mm with stride 0.1 mm. The point cloud network is PointMLP [14], trained for 200 epochs with batch size 8 using AdamW and learning rate 10^{-3}; the Gaussian target uses $\sigma = 0.2$ mm and positive weight $\omega^+ = 20$. We report MAE, RMSE, F1 score and IoU within 0.3 mm. Baselines include 3D U-Net and DeepLabV3+ trained on CBCT, nnU-Net with tooth instance segmentation, and the recent SwinUNETR [21] and 3D UX-Net [11]; all baselines share the same protocol-acquired data and patient-level split. Code and configurations will be released upon publication.

3.1 Comparison on Gingival Margin Prediction

As shown in Tab. 1, our framework reaches an overall MAE of 0.38 mm and IoU@0.3 mm of 56.2%, significantly outperforming all voxel baselines. Older 3D U-Net and DeepLabV3+ [24] yield MAE > 2.5 mm with IoU below 5%,

Table 2. Ablation study on the effectiveness of Cross-surface Representation (CSR) and Band-aware Gaussian Regression (BGR).

Method	MAE (mm)	RMSE (mm)	F1@0.3 mm (%)	IoU@0.3 mm (%)
w/o Cross-surface	0.55 ± 0.20	0.65 ± 0.23	45.7 ± 19.1	35.7 ± 17.5
w/o Gaussian Band	0.40 ± 0.12	0.52 ± 0.17	63.5 ± 14.1	53.2 ± 14.8
Ours (Full)	$\mathbf{0.38 \pm 0.13}$	$\mathbf{0.49 \pm 0.18}$	$\mathbf{66.2 \pm 14.5}$	$\mathbf{56.2 \pm 14.9}$

confirming that unconstrained voxel classification is insufficient for fine-grained gingival prediction. nnU-Net improves to 0.50 mm by using tooth masks as a spatial prior, and the recent SwinUNETR (0.48 ± 0.16 mm MAE, $37.1 \pm 17.6\%$ IoU) and 3D UX-Net (0.46 ± 0.15 mm MAE, $38.8 \pm 17.4\%$ IoU) further narrow the gap. However, all voxel backbones remain limited by the discrete voxel grid: when the gingival band is thinner than the 0.15 mm scan resolution they show staircase artifacts that the surface-domain formulation avoids. Since all baselines share the same protocol-acquired data and patient-level split, the gains isolate the algorithmic contribution of CSR+BGR. Fig. 3 reinforces this qualitatively. Per-fold overall MAE ranges from 0.36 to 0.40 mm and IoU@0.3 mm from 54.7% to 57.9%, indicating that performance is stable across folds.

3.2 Ablation Study

We isolate our two design choices on the same patient-level split, as summarized in Tab. 2. (i) **Effect of CSR.** "w/o Cross-surface" ($K = 1$) keeps the geometric prior and the same PointMLP encoder, using only vertex coordinates and normals; this is a direct surface-domain classification baseline. IoU@0.3 mm drops from 56.2 to 35.7% (MAE 0.55 mm, F1@0.3 mm 45.7%), a 20.5-pt gap that isolates the contribution of the localized intensity profile for exposing the bone–gingiva–air signature. (ii) **Effect of BGR.** "w/o Gaussian Band" keeps CSR and the encoder but replaces the soft Gaussian target with a hard binary band label; F1@0.3 mm drops to 63.5% and IoU@0.3 mm to 53.2% (MAE 0.40 mm), showing that continuous soft-target regression is superior to hard classification for modeling fuzzy CBCT interfaces. The two ablations decouple the gain of operating on the tooth manifold (CSR) from the gain of soft Gaussian regression (BGR); both are necessary.

4 Conclusion

We presented a protocol-and-method co-design for gingival margin extraction in CBCT. The chair-side scanning protocol renders the gingival margin imageable, while the cross-surface representation with band-aware Gaussian regression converts volumetric segmentation into a surface-anchored 1D regression task. The framework attains expert-level accuracy, consistently outperforming strong 3D voxel baselines and providing a practical, automated tool for periodontal assessment.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (grant numbers U23A20295, 62131015, 82441023, 82394432, 82301157), the STI2030 Major Projects of the China Ministry of Science and Technology (grant numbers 2022ZD0209000, 2022ZD0213100), the Key R&D Program of Guangdong Province (grant number 2023B0303040001), the Clinical Research Program of the 9th People's Hospital, Shanghai Jiao Tong University School of Medicine (grant number JYLJ2025015), and the HPC Platform of ShanghaiTech University.

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

References

1. Caton, J.G., Armitage, G., Berglundh, T., Chapple, I.L., Jepsen, S., Kornman, K.S., Mealey, B.L., Papapanou, P.N., Sanz, M., Tonetti, M.S.: A new classification scheme for periodontal and peri-implant diseases and conditions–introduction and key changes from the 1999 classification. Journal of periodontology **89**, S1–S8 (2018)
2. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV). pp. 801–818 (2018)
3. Dutzan, N., Konkel, J.E., Greenwell-Wild, T., Moutsopoulos, N.M.: Characterization of the human immune cell network at the gingival barrier. Mucosal immunology **9**(5), 1163–1172 (2016)
4. Field, D.A.: Laplacian smoothing and delaunay triangulations. Communications in applied numerical methods **4**(6), 709–712 (1988)
5. 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)
6. Jain, S., Choudhary, K., Nagi, R., Shukla, S., Kaur, N., Grover, D.: New evolution of cone-beam computed tomography in dentistry: Combining digital technologies. Imaging science in dentistry **49**(3), 179 (2019)
7. Jepsen, S., Caton, J.G., Albandar, J.M., Bissada, N.F., Bouchard, P., Cortellini, P., Demirel, K., de Sanctis, M., Ercoli, C., Fan, J., et al.: Periodontal manifestations of systemic diseases and developmental and acquired conditions: Consensus report of workgroup 3 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. Journal of clinical periodontology **45**, S219–S229 (2018)
8. Kim, Y.S., Park, J.S., Jang, Y.H., Son, J.H., Kim, W.K., Lee, Y.K., Kim, S.H.: Accuracy of periodontal probe visibility in the assessment of gingival thickness. Journal of Periodontal & Implant Science **51**(1), 30 (2020)
9. Ko, T.J., Byrd, K.M., Kim, S.A.: The chairside periodontal diagnostic toolkit: Past, present, and future. Diagnostics **11**(6), 932 (2021)
10. Lang, N.P., Bartold, P.M.: Periodontal health. Journal of periodontology **89**, S9–S16 (2018)
11. Lee, H.H., Bao, S., Huo, Y., Landman, B.A.: 3D UX-Net: A large kernel volumetric ConvNet modernizing hierarchical transformer for medical image segmentation. In: International Conference on Learning Representations (ICLR) (2023)

12. Li, Y., Cui, Z., Mei, L., Xie, Y., Marini, L., Pelekos, G., Gu, W., Yu, X., Wu, X., Wei, X., Tao, L., Deng, K., Pilloni, A., Shen, D., Tonetti, M.S.: A novel AI-powered radiographic analysis surpasses specialists in stage II–IV periodontitis detection: a multicenter diagnostic study. npj Digital Medicine **8**, 691 (2025)
13. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. In: Seminal graphics: pioneering efforts that shaped the field, pp. 347–353. ACM Press (1998)
14. Ma, X., Qin, C., You, H., Ran, H., Fu, Y.: Rethinking network design and local geometry in point cloud: A simple residual MLP framework. In: International Conference on Learning Representations (ICLR) (2022)
15. Malpartida-Carrillo, V., Tinedo-Lopez, P.L., Guerrero, M.E., Amaya-Pajares, S.P., Özcan, M., Rösing, C.K.: Periodontal phenotype: A review of historical and current classifications evaluating different methods and characteristics. Journal of Esthetic and Restorative Dentistry **33**(3), 432–445 (2021)
16. Mei, L., Deng, K., Cui, Z., Fang, Y., Liu, Y., Lai, H., Tonetti, M., Shen, D.: Clinical knowledge-guided hybrid classification network for automatic periodontal disease diagnosis in X-ray image. Medical Image Analysis **99**, 103376 (2025)
17. Mei, L., Fang, Y., Cui, Z., Deng, K., Wang, N., He, X., Zhan, Y., Zhou, X., Tonetti, M., Shen, D.: HC-Net: Hybrid classification network for automatic periodontal disease diagnosis. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. Lecture Notes in Computer Science, vol. 14225, pp. 54–63. Springer (2023)
18. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). pp. 565–571. Ieee (2016)
19. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: PointNet++: Deep hierarchical feature learning on point sets in a metric space. In: Advances in Neural Information Processing Systems (NeurIPS). pp. 5099–5108 (2017)
20. Tan, M., Cui, Z., Li, Y., Fang, Y., Mei, L., Zhao, Y., Wu, X., Lai, H., Tonetti, M.S., Shen, D.: PerioAI: A digital system for periodontal disease diagnosis from an intra-oral scan and cone-beam CT image. Cell Reports Medicine **6**(6), 102186 (2025)
21. Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V., Hatamizadeh, A.: Self-supervised pre-training of swin transformers for 3D medical image analysis. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 20730–20740 (2022)
22. Vitkov, L., Singh, J., Schauer, C., Minnich, B., Krunić, J., Oberthaler, H., Gamsjaeger, S., Herrmann, M., Knopf, J., Hannig, M.: Breaking the gingival barrier in periodontitis. International Journal of Molecular Sciences **24**(5), 4544 (2023)
23. Wang, C., He, Z., Chen, L., Yang, S., Sun, G., Duan, F., Wang, S., Zhou, Y.: ITMatch: Arch-guided semi-supervised tooth arrangement via iterative confidence evaluation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. Lecture Notes in Computer Science, vol. 15967, pp. 268–277. Springer (2025)
24. Yang, L., Zhu, Z., Li, Y., Huang, J., Wang, X., Zheng, H., Chen, J.: Clinical-oriented 3d visualization and quantitative analysis of gingival thickness using convolutional neural networks and cbct. Frontiers in Dental Medicine **6**, 1635155 (2025)
25. Zhou, Z., Chen, Y., He, A., Que, X., Wang, K., Yao, R., Li, T.: NKUT: Dataset and benchmark for pediatric mandibular wisdom teeth segmentation. IEEE Journal of Biomedical and Health Informatics **28**(6), 3523–3533 (2024)