Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Q-Bone System: an intelligent quantitative system for alveolar bone loss to assist the diagnosis of periodontitis – model development and validation

Wei Li

dentiwei1998@163.com

Beijing Tongren Hospital CMU

Jingyi Liu

Nankai University

Gepeng Ji

Nankai University

Zhuotao Yao

Beijing Tongren Hospital CMU

Dengping Fan

Nankai University

Jiang Lin

Beijing Tongren Hospital CMU https://orcid.org/0000-0002-6170-7566

Research Article

Keywords: Deep Learning, Multitask Learning, Anatomically-Driven, Alveolar Bone Loss, Periodontitis Diagnosis, Multicenter Imaging

Posted Date: January 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8425991/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Journal of Translational Medicine on June 25th, 2026. See the published version at https://doi.org/10.1186/s12967-026-08444-z.

Q-Bone System: an intelligent quantitative system for alveolar bone loss to assist the diagnosis of periodontitis – model development and validation

Wei Li[a], Jingyi Liu[b], Gepeng Ji[b], Zhuotao Yao[a], Dengping Fan[b], Jiang Lin[a]*

Affiliations

[a] Department of Stomatology, Beijing Tongren Hospital, Capital Medical University, Beijing, China

[b] College of Computer Science & Visual Computing and Intelligent Perception (VCIP) Lab, Nankai University, Tianjin, China

* **Corresponding author**: Jiang Lin, Department of Stomatology, Beijing Tongren Hospital, Capital Medical University, No. 1, Dongjiaominxiang, Dongcheng District, Beijing 100730, China. E-mail address: kelvinperio@163.com (J. Lin).

Abstract

Objectives: To develop and validate the Q-Bone system for precise alveolar bone loss quantification and intelligent periodontitis diagnosis across multiple clinical centers and imaging devices.

Methods: This study included 1,273 periodontitis cases from four clinical centers using diverse imaging devices. A multitask deep learning model, DGNet, was employed for tooth segmentation and anatomical key points localization, integrated with an anatomically-driven quantification algorithm. Performance was assessed using several validation datasets.

Results: The Q-Bone system demonstrated strong performance: tooth segmentation achieved an S-measure of 0.929, and key point localization reached a PRCK@0.5 of 0.994 in internal validation. The system showed high consistency with expert measurements, with an ICC of 0.973. It quantified alveolar bone loss with minimal bias (-0.238%) and assisted in periodontitis diagnosis, achieving a Kappa value of 0.955 for tooth-level diagnosis.

Conclusions: The Q-Bone system provides accurate, automated alveolar bone loss quantification and intelligent periodontitis diagnosis. It showed excellent generalization across multicenter and cross-device settings, making it a reliable tool for periodontitis diagnosis.

Keywords: Deep Learning; Multitask Learning; Anatomically-Driven; Alveolar Bone Loss; Periodontitis Diagnosis; Multicenter Imaging

1.Introduction

The Global Burden of Disease study shows that periodontitis affects approximately 1.1 billion people worldwide [1]. Its pathological characteristics include the loss of periodontal connective tissue attachment and progressive alveolar bone loss[2], making it the leading cause of tooth loss and damage in middle-aged and elderly populations[3]. Periodontitis represents one of the significant challenges to global public health[4] , and accurate diagnosis, treatment, and prognosis analysis are crucial for its prevention and control [5].

Alveolar bone loss is one of the core criteria for evaluating the prognosis of teeth affected by periodontitis [6]. In 2018, the AAP/EFP classification of periodontal disease incorporated radiographic bone loss ratio as a key criterion for staging, which raises higher demands for the quantitative accuracy of imaging assessment[7]. While

existing imaging technologies, such as Cone Beam Computed Tomography (CBCT), can clearly present the anatomical details of hard tissues, such as the spatial relationship between the tooth root and alveolar bone, the extent of periapical lesions, and the jawbone structure, the quantification process still faces challenges such as complexity, low automation, and interference from subjective factors[8].

Research shows[9] that the effective radiation dose of CBCT is significantly higher than that of Orthopantomogram (OPG), with a wide range of effective doses, from 17.2 μSv to 396 μSv. At the lowest exposure settings, the average dose is 31.6 μSv, while at the highest exposure settings, it reaches 209 μSv [10]. In contrast, the single exposure dose of OPG is only 9–24 μSv [11]. Long-term, large-scale use of CBCT will significantly increase the radiation exposure risk for patients [12].

In contrast, OPG has advantages such as low radiation, ease of operation, and relatively low cost, making it the mainstream imaging tool for periodontitis screening [13]. However, there are still significant limitations in the clinical application of OPG. Firstly, the evaluation by clinicians relies on visual judgment, and there is a lack of standardized operational methods in clinical practice, which makes subjective assessment prone to errors in identifying the crest of the alveolar ridge, thus affecting the accuracy of determining the level of alveolar bone loss [14]. Secondly, the existing auxiliary measurement tools only provide linear measurement functions, which cannot accommodate complex anatomical structures such as curved roots [15]. This method, which neglects the natural curvature of the root, leads to quantification bias [16], making it difficult to precisely calculate the alveolar bone loss percentage and provide a basis for accurate periodontitis staging, thus limiting its application value in individualized periodontal treatment and becoming a bottleneck in clinical use.

Artificial intelligence-assisted imaging diagnosis has recently made significant progress in identifying oral lesions [17]. Deep learning-based methods, particularly architectures such as U-Net and its variants, have been widely applied to tasks such as periapical lesion detection [18], tooth segmentation [19,20], and gingival inflammation area identification [21] significantly improving the accuracy of tissue recognition.

Despite these advancements, critical limitations remain in current AI systems regarding precision medicine and early diagnosis.

First, existing AI-assisted periodontitis research primarily focuses on screening and classification tasks. For instance, the HC-Net+ model, which integrates clinical diagnostic standards into its algorithm design, has achieved or even surpassed specialist-level performance in classifying Stages II–IV periodontitis[22]. However, such classification-oriented models cannot provide precise quantitative outputs for the extent of Alveolar Bone Loss (ABL), whereas the quantitative assessment of ABL is a crucial indicator for monitoring periodontitis severity and assisting diagnosis. Second, the current data foundation is suboptimal: most studies still rely on conventional panoramic radiographs for annotation [23]. Due to the two-dimensional projection characteristics of panoramic images, overlapping or blurring of anatomical structures such as the cementoenamel junction (CEJ), the alveolar bone crest (ABC), and the apical point (AP) often occurs, leading to visual biases from clinicians. This bias directly affects the accuracy of

annotations, which in turn distorts the gold standard for model training, reducing both the model's recognition accuracy [24] and the precision of ABL quantification. Third, ABL quantification is inherently a cascade data processing workflow, involving fine segmentation of the tooth, anatomical key points localization, root curvature mapping, and staging conversion. However, most existing studies focus on single tasks and lack effective multitask collaborative solutions, preventing the complete translation from image recognition to clinical staging. As a result, these methods fail to meet the clinical diagnostic needs for precise periodontitis staging.

To address the aforementioned clinical issues and technical challenges, this study constructed a DGNet multitask deep learning model for tooth segmentation and anatomical key point localization based on panoramic surface tomography images and 2D-reconstructed CBCT (2D-CBCT) panoramic images. The model was combined with a local coordinate system and curve measurement method to achieve precise quantification of alveolar bone loss and intelligent staging of periodontitis at both single-tooth and full-mouth levels. Finally, the performance and clinical applicability of the Q-Bone AI-assisted system were validated in real-world clinical settings across multiple centers and imaging devices.

2. Materials and methods

2.1 Study design

This multicenter diagnostic trial was approved by the Ethics Committee, registered with the Chinese Clinical Trial Registry, and conducted in accordance with the Helsinki Declaration[25] and the dental AI research guidelines checklist[26]. It collaborates with four clinical hospitals to validate the model’s adaptability across multiple centers and heterogeneous imaging devices. Beijing Tongren Hospital served as the internal training center, and Beijing Daxing District People’s Hospital, the Fourth Affiliated Hospital of Harbin Medical University, and Jining People’s Hospital acted as external validation centers, providing independent data to assess the model's generalizability.

The experimental process is shown in Fig. 1, which includes the following steps: multicenter, cross-device imaging and clinical data collection, construction of the DGNet multitask deep learning model framework[27], model training and explainability analysis, development of the alveolar bone loss quantification algorithm based on root anatomical curvature, and application deployment and validation in real multicenter clinical scenarios.

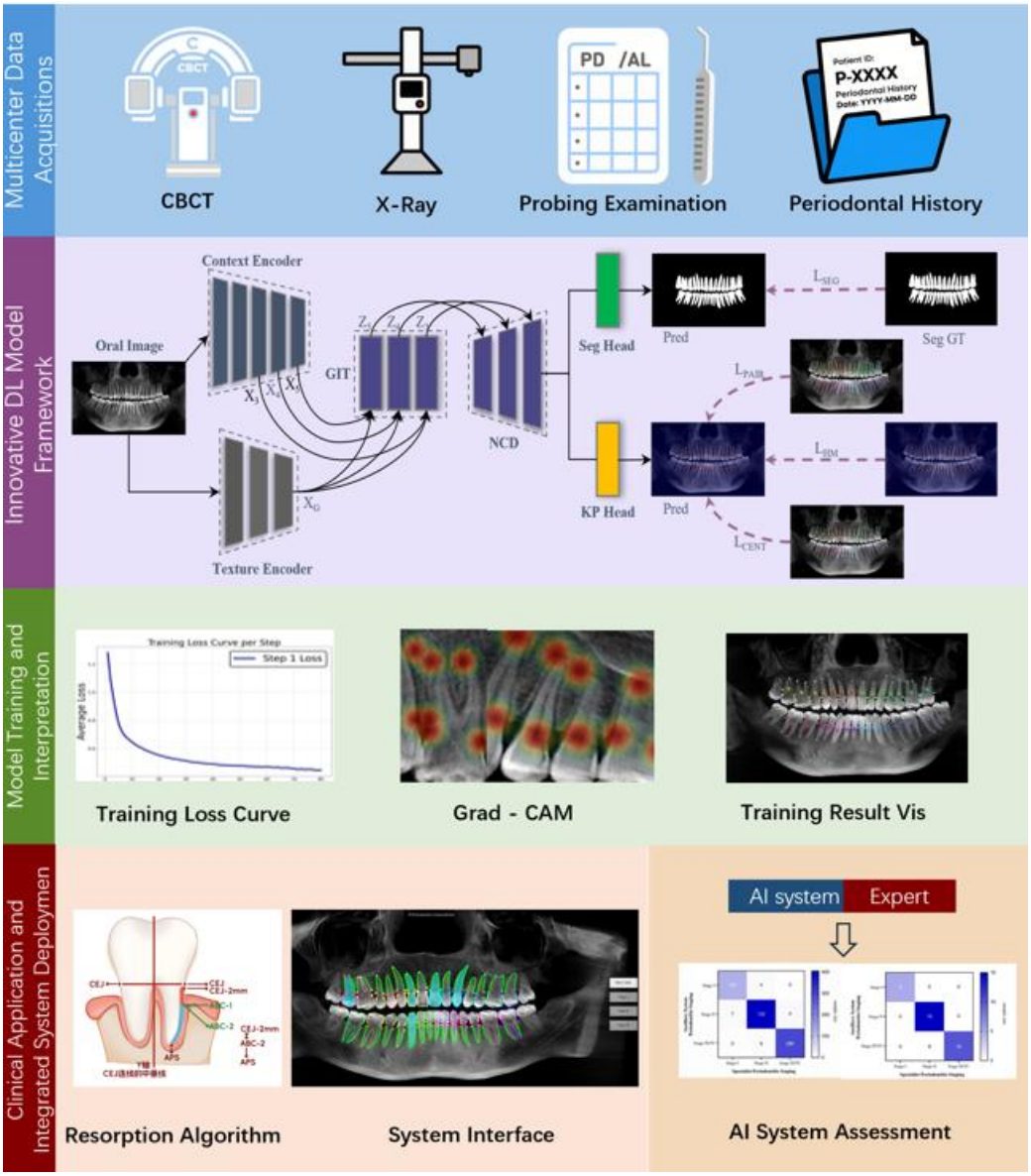


Fig. 1. Integrated Imaging Analysis and Clinical Assessment Workflow for Smart Periodontitis Diagnosis. Overview of the Q-Bone workflow, from multicenter periodontal data acquisition and DGNet-based image analysis to anatomically-driven alveolar bone loss quantification and clinical evaluation of AI–expert agreement.

2.2 Study subjects

This retrospective multicenter study collected de-identified clinical and imaging data from patients diagnosed with periodontitis. The dataset included demographic information, systemic and periodontal history, oral hygiene habits, and core periodontal indices (probing depth, clinical attachment loss, and gingival bleeding index), alongside corresponding 2D-CBCT or OPG images. Inclusion criteria were: (1) Age ≥ 18 years; (2) Diagnosis of periodontitis according to the 2018 AAP/EFP guidelines; and (3) Availability of complete clinical records and diagnostic-quality OPG or 2D-CBCT imaging data.

2.3 Multicenter imaging dataset construction and sample size determination

To ensure robustness and generalizability, imaging data from four clinical centers were grouped into an internal training/development dataset and an external validation dataset

(Table 1). The internal dataset comprised 500 cases with 2D-CBCT images from Beijing Tongren Hospital and was used for model training and development. The external dataset comprised 773 cases (OPG or 2D-CBCT) from three independent centers and was used exclusively for external validation to assess performance across different devices and populations. Baseline demographic characteristics of the multicenter cohort are summarized in Table 2.

Table 1. Sample size and imaging modality distribution across clinical Centers

Clinical Center	Quantity	Imaging Type
Center 1: Beijing Tongren Hospital	500	2D-CBCT
Center 2: Beijing Daxing District People's Hospital	300	2D-CBCT/OPG
Center 3: The Fourth Affiliated Hospital of Harbin Medical University	273	OPG
Center 4: Jining People's Hospital	200	OPG
Total	1273	

Note: 2D-CBCT, two-dimensional reconstructed cone-beam computed tomography; OPG, orthopantomogram.

Table 2. Baseline demographic characteristics of the multicenter cohort

Feature	Statistic
Number of Patients	1273
Male/Female	618/655
Age Distribution	
30-40	276
40-60	424
60-80	573

Note: Values are presented as counts of patients.

The overall training sample size (n = 1,273) was determined based on preliminary experiments using publicly available datasets [28] , and was considered sufficient to cover diverse imaging modalities and periodontitis stages, thereby enhancing model generalisation and reducing the risk of overfitting. For the multicenter clinical consistency validation, the required sample size was calculated using Fleiss' kappa statistics [29] , with a significance level of $\alpha = 0.05$ (two-sided) and a statistical power of $1–\beta = 80\%$, yielding a minimum of 26 cases. To ensure robustness, 30 subjects were ultimately planned for inclusion via stratified sampling by periodontitis stage.

2.4 Multicenter imaging devices and imaging data processing

OPG devices were used in three centers: the Dentsply Sirona system at Beijing Daxing District People's Hospital, the SOREDEX unit at the Fourth Affiliated Hospital of Harbin Medical University, and the Vatech PaX-400C system at Jining People's Hospital. The panoramic radiographic images retained the original grayscale information without additional preprocessing and were exported in PNG format.

CBCT devices were used in two centers: the Dentsply Sirona Axeos CBCT system at Beijing Tongren Hospital, and the CEFLA S.C. NewTom VGi CBCT system at Beijing Daxing District People's Hospital. The original CBCT DICOM data were

processed using the Sidexis4 workstation. Two periodontics specialists manually defined the panoramic curve of the dental arch to match the tooth morphology. The extracted 2D-CBCT images had a reconstruction layer thickness of 0.5mm to ensure clear display of the alveolar ridge and root contours, which were then exported in PNG format.

2.5 Dataset annotation and division

Data annotation was performed by three periodontists, with the final review conducted by one periodontics expert. All annotators underwent standardized training. The intraclass correlation coefficient (ICC) was 0.92, indicating excellent inter-observer agreement among the annotators.For images with unclear anatomical structures, the annotators referred to the original three-dimensional data and performed difference calibration based on the 3D morphological features of the alveolar bone and tooth roots to ensure accuracy.

To construct the gold standard for model training, we selected 110 cases from 500 2D CBCT images at Beijing Tongren Hospital, annotating a total of 2,800 teeth. The specific annotations included: (1) pixel-level segmentation masks of the teeth; (2) tooth position identification based on the FDI two-digit coding system [30]; (3) Localization of six anatomical key points per tooth, including the CEJ, AP, and mesial and distal points of the ABC. This work resulted in the annotation of 16,800 key points.

The initial tooth segmentation was performed using MetaAI's SAM-ViT-Huge model [31]. Three periodontics specialists manually refined the segmentation and reached expert consensus, ultimately forming the annotated dataset.

The dataset was stratified by individual patients and divided into 90 cases for training, 10 for validation, and 10 for internal testing. Additionally, independent data from other centers were used as external test sets to evaluate the model's generalizability across different clinical centers.

2.6 Algorithm design and multitask recognition model construction

To achieve collaborative optimization of single-tooth segmentation and periodontal anatomical key points localization, we adopted a three-step training strategy for the DGNet multitask model.

Step 1: Using PVTv2-B3[32] as the backbone encoder, segmentation pre-training is first performed on 300 relevant images selected from the MICCAI STS23 dataset [28]. Afterward, segmentation tasks are further optimized using the internal training dataset from Beijing Tongren Hospital, with a focus on learning the overall anatomical contours and boundary features of the teeth and alveolar bone to obtain stable global feature representations.

Step 2: Based on Step 1, the model undergoes pre-training for the sparse anatomical key points adaptation task. The internal dataset from Beijing Tongren Hospital is used to further enhance the localization capability of key point anatomical landmarks such as the CEJ, ABC, and AP.

Step 3: Using the internal dataset from Tongren Hospital, joint optimization of the tasks is performed to further improve segmentation accuracy and the anatomical correctness of key points, achieving collaborative optimization of both tasks.

Finally, post-processing steps are applied to refine the segmentation results and

automatically correct the key point positions.

The model is built on the PyTorch framework, utilizing the Adam optimizer with cosine annealing learning rate scheduling. The parameters for each stage are dynamically adjusted based on task requirements, with specific settings as follows: learning rate of $5\times10^{-5} \rightarrow 1\times10^{-4} \rightarrow 1\times10^{-6}$, batch size of 16 → 16 → 12, and training epochs of 80 → 150 → 100.

2.7 DGNet multitask model architecture and loss function design

DGNet adopts a dual-encoder architecture, with PVTv2-B3 used for deep semantic feature extraction, while another lightweight shallow convolutional network captures texture information. The features from the two encoders are fused using the Gradient-Induced Transition module to generate a shared feature map. Finally, parallel 1×1 convolutional layers are employed to perform segmentation (seg) and key point tasks, outputting a single-channel segmentation mask and a 168-channel heatmap.

The total loss of the model is calculated by weighting the losses from multiple tasks. Specifically, for the segmentation task, a weighted cross-entropy loss and Intersection over Union (IoU) loss are used to reinforce the edge regions. For key point detection, a dynamically weighted binary cross-entropy (BCE) loss is applied to address the sample sparsity issue. Additionally, two Smooth L1 losses are used to constrain the overall centroid and the relative distance between key points groups, ensuring the correctness of the anatomical structures.

2.8 Alveolar bone resorption quantification algorithm based on root anatomical curvature

In response to the natural curvature of the root's true anatomical morphology, this study developed an anatomically-driven alveolar bone resorption ratio (ABRR) quantification algorithm (Fig. 2).

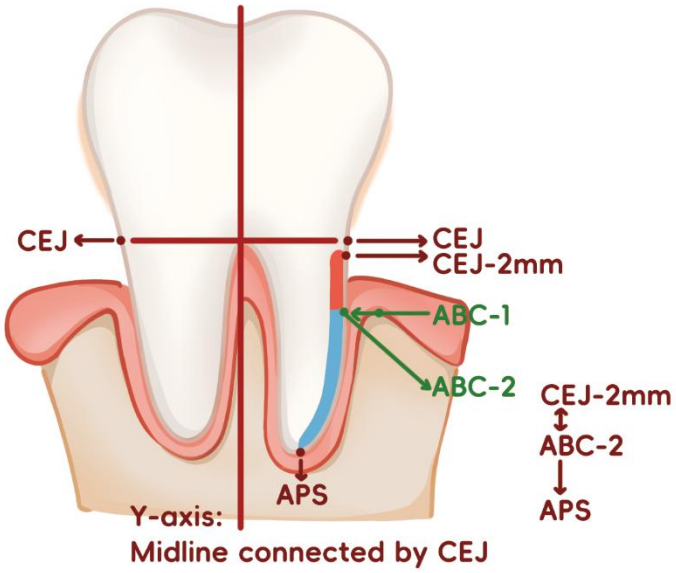


Fig. 2. Schematic illustration of the anatomically-driven, curvature-based algorithm for single-tooth alveolar bone resorption quantification, showing the local CEJ-based coordinate system, correction of the alveolar bone crest point,

and definition of root length and bone loss along the root surface.

The algorithm mainly includes the following three steps:

(1) Local Coordinate System Construction: The X-axis is established by connecting the mesial-distal points of the CEJ, and the Y-axis is formed by a perpendicular line drawn from the midpoint of the CEJ. This creates a local coordinate system for the tooth, eliminating the impact of positional and orientation differences in panoramic images.

(2) Anatomical key points Correction : The original alveolar crest point (ABC-1) is first obtained, then projected along the normal direction and fitted to the external curve of the root. The corrected alveolar crest point (ABC-2) is derived, ensuring that the crest point lies on the anatomical edge of the root.

(3) ABRR Calculation: Within the local coordinate system, two arc lengths are defined along the true edge of the root:

1.Root Length (RL): The curve distance from CEJ to the AP.

2.Bone Loss (BL): Along the same root surface curve, 2 mm is measured from CEJ towards the apex to define a CEJ-2mm reference point. The curve distance from this point to the corrected alveolar crest point (ABC-2) is then calculated.

The ABRR is calculated based on Formula (1):

$$\mathrm{ABRR} = \frac{\mathrm{BL}}{\mathrm{RL}} \times 100\% \text{(Formula 1)}$$

The periodontitis staging is then performed based on ABRR:Stage I: ABRR ≤ 15%;Stage II: 15% < ABRR ≤ 33%;Stage III/IV: ABRR > 33%[8].

2.9 Development of alveolar bone loss visualization and interactive system

To visualize alveolar bone loss measurement results and facilitate clinical application, this study developed a clinical interactive system based on OpenCV and Matplotlib, incorporating AI-assisted periodontal imaging recognition functions to enhance system usability. The main features of the system include:

(1) Core Information Visualization: Simultaneously displays tooth position, alveolar bone loss ratio, and periodontitis staging, with quantified data embedded in the image interface as structured tables to ensure intuitive presentation of information.

(2) Anatomical Feature Visual Presentation: Differentiates tooth segmentation boundaries using color coding and marks key anatomical points with anatomical labels to ensure traceability of measurement references.

(3) Staging Highlight Function: Supports highlighting corresponding tooth positions based on staging, offers toggle functionality for the visibility of quantified tables, and allows users to view the original image alongside AI-identified quantified results, assisting periodontists in accurately assessing alveolar bone loss levels.

2.10 Evaluation metrics

2.10.1 Model recognition performance metrics

Given the dual-task characteristics of the model, the following core evaluation metrics were selected[33]: for the segmentation task: S-Measure[34], Weighted F-Measure[35], Eφ[36], and Mean Absolute Error (MAE) [37] were used to assess the segmentation accuracy of tooth contours and edges ;for key point localization: Percentage of Correctly Recognized Key points (PRCK) [38] was adopted. Using the mean value of the diagonals of the tooth bounding box as the normalization factor (L),

the correct localization rate was calculated with d-thresh = 0.05, 0.25, and 0.5, respectively, to quantify the localization accuracy of sparse anatomical landmarks such as root apices and alveolar crests.

2.10.2 Clinical validation metrics

Statistical analyses were performed using MedCalc 20.0 software. A validation protocol was designed for the core clinical output metrics as follows. For the ABRR, the ICC and Bland–Altman plots were used to assess the agreement between the Q-Bone AI-assisted system and expert manual measurements. For periodontitis staging, the kappa statistic was employed to evaluate consistency with clinical diagnoses, and a consistency heatmap was generated. The statistical significance level was set at a two-sided $P < 0.05$.

3.Results

3.1Training process and loss curves

Fig. 3 shows the loss curves of the three training stages. The loss steadily decreases and eventually converges throughout the entire training process, demonstrating the effectiveness of the training strategy. In Step 1 (Fig. 3a), the loss drops rapidly, indicating that the model quickly learns the basic features. In Step 2 (Fig. 3b), the loss exhibits a staged decline, suggesting that the model gradually adapts to the new key point task. Finally, in Step 3 (Fig. 3c), the loss converges slowly in a smooth manner, indicating that the model has entered a stable fine-tuning phase, further improving its accuracy.

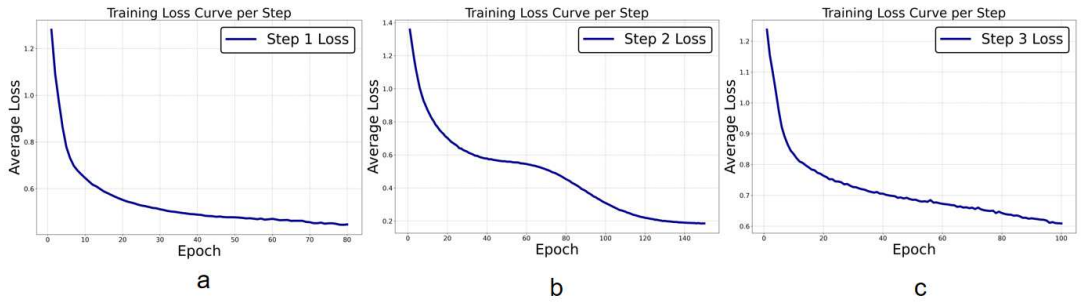


Fig. 3. Training loss curves for the three stages: (a) Step 1 segmentation pre-training, (b) Step 2 key point adaptation, and (c) Step 3 joint fine-tuning.

3.2 Grad-CAM heatmap interpretability analysis

To assess whether the decision logic of DGNet is consistent with anatomical structures, Grad-CAM was applied to generate attention heatmaps. In the key point detection task (Fig. 4), the activation was mainly concentrated around the CEJ, ABC and AP, with minimal responses in surrounding areas. In the tooth segmentation task (Fig. 5), the heatmaps closely followed the contours of all teeth and showed different activation patterns for anterior and posterior teeth, while the background regions exhibited low activation.

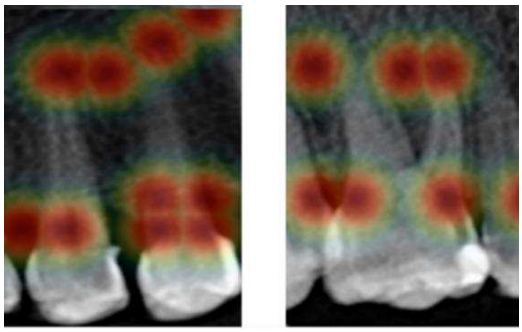


Fig. 4. Grad-CAM visualizations for the key point detection task, showing focused model attention on the CEJ, ABC, and AP with minimal activation in non-target regions.

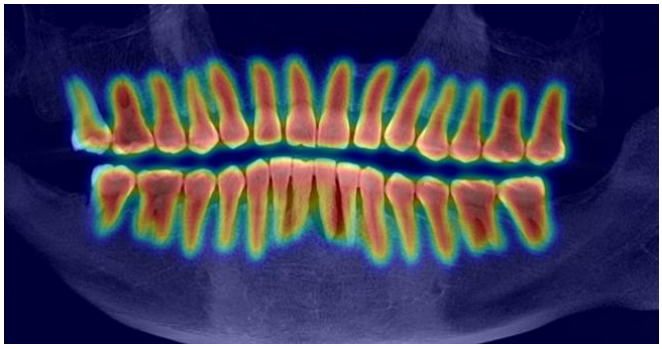


Fig. 5. Grad-CAM visualizations for the tooth segmentation task, illustrating that model attention covers the full-mouth tooth contours and differentiates anterior and posterior tooth morphology.

3.3 Multicenter recognition performance of the multitask deep learning model

The performance of the multitask model was evaluated across four clinical centers. For the tooth segmentation task (Table 3), the S-measure ranged from 0.893 to 0.929, with the highest value at Beijing Tongren Hospital and all external centers remaining above 0.89. Other segmentation metrics also remained consistently high, while MAE stayed within a low range (0.033–0.061), indicating robust and stable tooth-contour segmentation across different centers and imaging devices.

Table 3. Multicenter performance evaluation of tooth segmentation

Evaluation Metric \ Centers	Beijing Tongren Hospital	Daxing District People's Hospital	Harbin Medical University Fourth Hospital	Jining People's Hospital
Tooth Segmentation				
Smeasure ↑	0.929	0.893	0.906	0.910
wFmeasure ↑	0.832	0.783	0.747	0.763
MAE ↓	0.033	0.061	0.046	0.040
adpEm ↑	0.978	0.958	0.969	0.976
meanEm ↑	0.935	0.905	0.920	0.927
maxEm ↑	0.982	0.957	0.975	0.976
adpFm ↑	0.894	0.865	0.845	0.872
meanFm ↑	0.874	0.838	0.835	0.854
maxFm ↑	0.916	0.887	0.885	0.896

Note: S-measure, Structure-Measure; wF-measure, weighted F-measure; MAE, mean absolute error; Em, E-measure (enhanced-alignment measure); Fm, F-measure. The arrows (↑/↓) indicate whether higher or lower values correspond to better performance. The data summarise segmentation performance across the four clinical centers.

For the key point localization task (Table 4), PRCK was used as the evaluation metric. PRCK@0.5 exceeded 0.86 in all four centers, demonstrating high localization accuracy for sparse anatomical landmarks such as the root apex and alveolar crest, and confirming that the model is suitable for subsequent anatomical measurements.

Table 4. Multicenter performance evaluation of anatomical key points detection

Evaluation Metric \ Centers	Beijing Tongren Hospital	Daxing District People's Hospital	Harbin Medical University Fourth Hospital	Jining People's Hospital
Key point Detection				
PRCK@0.5	0.994	0.905	0.863	0.920
PRCK@0.25	0.911	0.661	0.657	0.616
PRCK@0.05	0.354	0.238	0.202	0.137

Note: PRCK, Percentage of Correctly Recognized key point. The thresholds @0.5, @0.25, and @0.05 represent the strictness of the evaluation, calculated based on the normalization factor (L, the diagonal length of the tooth bounding box). Higher PRCK values indicate greater precision in anatomical landmark localization.

3.4 Tooth segmentation and key point recognition visualization

Representative examples of tooth segmentation and anatomical key point recognition are shown in Fig. 6. Compared with the ground-truth annotations, the Q-Bone predictions produce binary and instance masks that closely follow the true tooth contours and individual tooth shapes. The predicted anatomical key points are well aligned with the reference landmarks along the dental arch in both anterior and posterior regions, illustrating the ability of the multitask model to simultaneously achieve accurate tooth segmentation and key point localization.

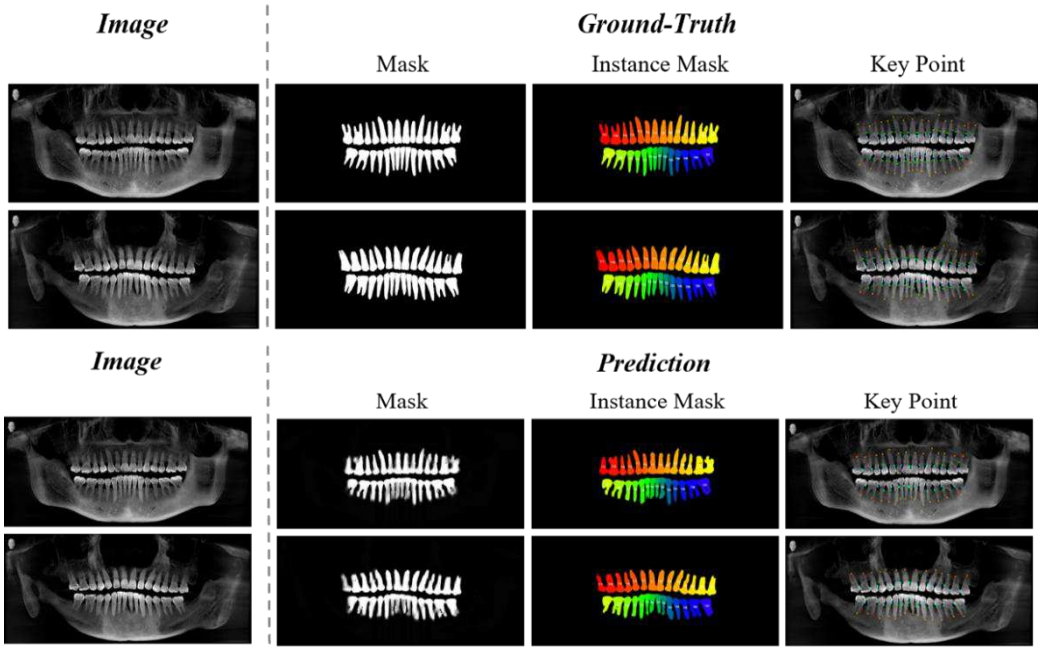


Fig. 6. Visual comparison of ground-truth annotations and predictions for full-mouth tooth segmentation and anatomical key points. Ground-truth annotations: original panoramic image, binary tooth mask, color-coded instance mask (different colors represent different teeth), and overlaid anatomical key point. Predictions: corresponding

predicted binary masks, instance masks, and key point localization results.

3.5 Interactive visualization system for alveolar bone resorption quantification

This study further developed the Q-Bone interactive visualization system for alveolar bone loss quantification, which presents the deep learning model's detection, quantification, and staging outputs in an intuitive clinical interface (Fig. 7). In the central panoramic radiograph, color-coded overlays display the segmentation contours of each tooth and the severity of alveolar bone loss. On the left side of the interface, a structured table lists the alveolar bone resorption ratio (ABRR; reported to two decimal places) and the radiographic stage for each tooth (FDI tooth codes). The bottom information bar dynamically shows metrics for the currently selected tooth, while interactive controls on the right allow users to filter teeth by stage and review information for individual teeth. By combining oral imaging with AI-based quantitative analysis, the system provides a clear and structured visual representation of alveolar bone loss and periodontitis staging results.

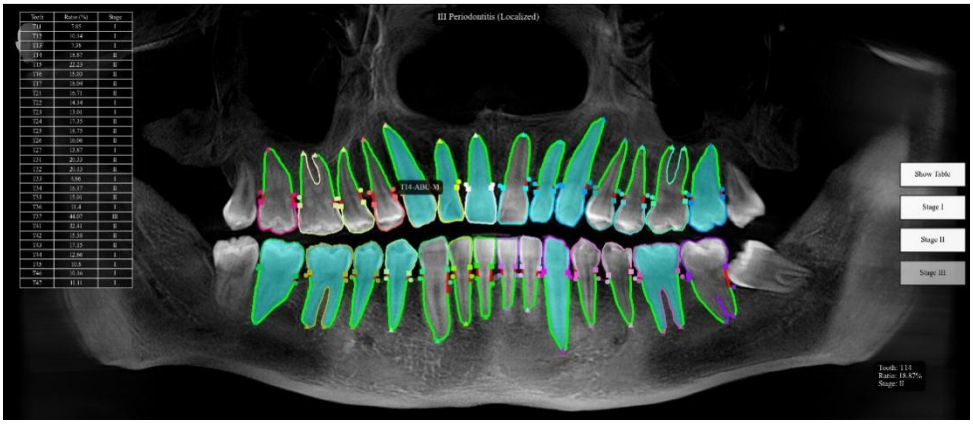


Fig. 7. Q-Bone interactive visualization interface for full-mouth alveolar bone loss quantification and periodontitis staging. Color-coded tooth segmentation overlays, anatomical keypoints, and tooth-wise ABRR and staging tables are displayed in a single view, with interactive controls for filtering and inspecting individual teeth.

3.6 Multicenter imaging dataset construction clinical Validation

In the multicenter diagnostic trial, 30 patients were included in the clinical validation set, yielding 840 evaluable teeth. Using the expert consensus diagnosis as the reference standard, we assessed the agreement between Q-Bone and specialists for tooth-level ABRR measurements and periodontitis staging, as detailed below.

3.6.1 Consistency analysis of alveolar bone loss ratio

The consistency between the specialist's manual measurements and the Q-Bone AI-assisted system for tooth-level alveolar bone resorption ratio (ABRR) was evaluated for 840 tooth positions. As summarized in Table 5, the mean ABRR was 31.85% ± 18.59% for the specialist and 32.08% ± 18.41% for the Q-Bone system. The mean difference (Q-Bone – specialist) was −0.24% (95% CI: −8.66% to 8.19%; P = 0.110), and the ICC was 0.973, indicating excellent agreement between the two methods and no statistically significant systematic bias.

Table 5. Statistical comparison of tooth-level alveolar bone resorption ratio (ABRR) between specialist and Q-Bone AI-assisted system

Test	n	Mean ± SD	Median (IQR)	Mean Difference and 95% CI	ICC	P
Specialist	840	31.85±18.59	27.00(18.52,43.91)	-0.238(-8.663, 8.188)	0.973	0.110
Q-Bone system	840	32.08±18.41	26.98(19.15,44.40)			

Note: Data are presented as mean ± standard deviation (SD) or median (interquartile range [IQR]). CI, confidence interval; ICC, intraclass correlation coefficient. P values are from paired t-tests comparing specialist and Q-Bone system measurements. The mean difference is calculated as "Q-Bone – specialist".

The Bland–Altman plot (Fig. 8) further illustrates this agreement. The average bias was −0.24%, with 95% limits of agreement from approximately −8.7% to 8.2%, and most data points lay within these limits without an obvious trend across the measurement range. This pattern suggests minimal systematic error and clinically acceptable random variation between the AI-assisted system and specialist measurements.

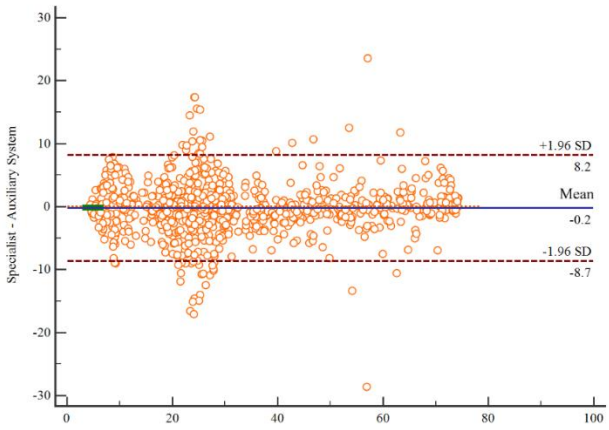


Fig. 8. Bland–Altman plot comparing tooth-level alveolar bone resorption ratio (ABRR) between the AI-assisted system and specialist measurements. The solid line represents the mean difference (−0.24%), and the dashed lines indicate the 95% limits of agreement (−8.7% to 8.2%).

3.6.2 Single-tooth level periodontitis staging consistency validation

Agreement between the AI-assisted system and specialists for single-tooth periodontitis staging was evaluated using Cohen's kappa statistic and visualised with a confusion matrix heatmap (Table 6 and Fig. 9). The overall kappa value was 0.955 (P = 0.001), indicating near-perfect agreement. Most teeth lay on the diagonal of the confusion matrix, with only a small number of discrepancies between adjacent stages (Stage I vs. Stage II and Stage II vs. Stage III/IV), and no misclassification between Stage I and Stage III/IV.

Table 6. Single-tooth level periodontitis staging: confusion matrix and kappa agreement between the AI-assisted system and specialist reference standard

AI System Periodontitis Staging	Specialist Periodontitis Staging			Kappa	P
	Stage I	Stage II	Stage III/IV		
Stage I	122	4	0		
Stage II	7	405	4	0.955	0.001
Stage III/IV	0	8	290		

Note: Agreement between the AI-assisted system and specialist staging was assessed using Cohen's kappa. Kappa values > 0.90 indicate near-perfect agreement, and P values < 0.05 were considered statistically significant.

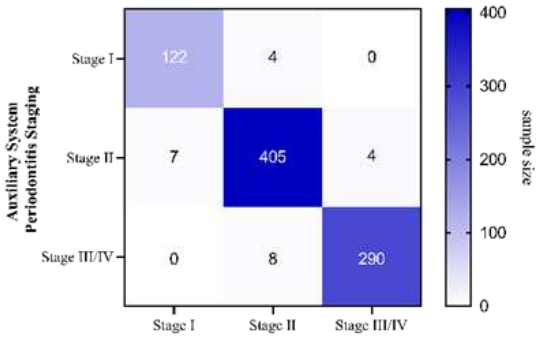


Fig. 9. Confusion matrix of single-tooth periodontitis staging between the AI-assisted system and specialist reference standard. Diagonal cells indicate concordant staging, whereas off-diagonal cells represent discrepancies between stages.

3.6.3 Full-mouth agreement in periodontitis staging

Agreement between the AI-assisted system and specialists for full-mouth periodontitis staging was evaluated using Cohen's kappa (Table 7 and Fig. 10). The Q-Bone system assigned exactly the same stage as the specialists for all 30 patients, yielding a kappa value of 1.000 (P = 0.001), indicating perfect agreement with no misclassified cases.

Table 7. Full-mouth level periodontitis staging: confusion matrix and kappa agreement between the AI-assisted system and specialist reference standard

AI System Periodontitis Staging	Specialist Periodontitis Staging			Kappa	P
	Stage I	Stage II	Stage III/IV		
Stage I	5	0	0		
Stage II	0	15	0	1.000	0.001
Stage III/IV	0	0	10		

Note: Agreement between the AI-assisted system and specialist staging was assessed using Cohen's kappa. A kappa value of 1.000 indicates perfect agreement, and P values < 0.05 were considered statistically significant.

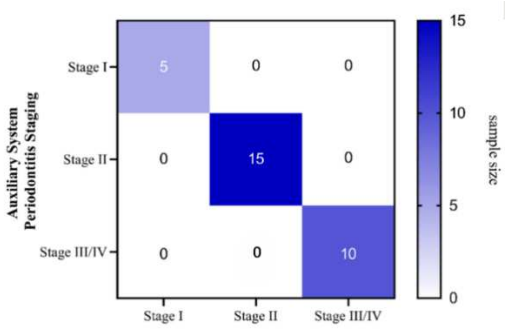


Fig. 10. Confusion matrix of full-mouth periodontitis staging between the AI-assisted system and the specialist reference standard. All 30 cases lie on the diagonal, indicating 100% agreement with no discrepancies between stages.

4. Discussion

This study developed the Q-Bone AI-assisted system, which integrates the DGNet multitask deep learning model with an anatomically-driven quantification algorithm, enabling automated and precise alveolar bone loss quantification and auxiliary diagnosis of periodontitis across different centers and imaging devices. In clinical settings, the system showed stable and reliable performance in tooth segmentation,

anatomical key point localization, and staging consistency, suggesting good generalization across heterogeneous imaging conditions and patient populations.

A progressive three-step training strategy was applied to the DGNet multitask model, introducing an effective approach to task synergy. By gradually refining the model through segmentation pre-training, subsequent key point localization, and final joint fine-tuning, the strategy allowed each task to benefit from the representations learned in the previous stage. This progressive optimization helped to balance the objectives of segmentation and localization and facilitated robust performance across multiple centers without apparent task interference.

Several previous AI-assisted periodontal diagnostic studies have reported strong results on single-center, single-device datasets[39]. In contrast, the present study adopted a multicenter, cross-device design and still maintained high performance in both tooth segmentation and anatomical key point localization[40] , This indicates that Q-Bone is not only accurate under controlled conditions, but also robust to variations in imaging devices and patient compositions, providing a solid basis for its potential deployment in large-scale, multiregional clinical applications.

Moreover, Li et al.'s work has made important progress in staging and classifying periodontitis; however, classification-based models typically output only categorical labels and do not provide precise quantitative information on bone destruction[22]. In clinical practice, there is an urgent need for accurate quantification of ABRR in periodontitis patients. The Q-Bone system, with its anatomically-driven curve-based quantification pathway, generates high-precision, interpretable numerical outputs. This offers a more refined diagnostic basis than simple categorical staging alone and provides a foundation for future dynamic and precise monitoring of periodontal treatment responses.

Traditional linear measurement methods are easily affected by the natural curvature of the tooth root, making it difficult to accurately reflect the true extent of alveolar bone loss[15] .To address this technical challenge, Q-Bone employs a local coordinate system and curve fitting approach: using the CEJ as a reference, an independent local coordinate system is constructed for each tooth, and the alveolar crest point is projected onto the true root surface. The bone resorption ratio is then calculated along the actual curved trajectory of the root, rather than using straight-line distances. This anatomically aligned measurement strategy reduces systematic errors introduced by root curvature and improves the precision and clinical reliability of alveolar bone loss assessment.

Clinical validation across multiple centers further confirmed the feasibility and reliability of the Q-Bone system for quantifying alveolar bone loss under real-world clinical conditions. Using manual measurements by senior periodontists as the reference standard, the AI-derived ABRR values showed high agreement with expert measurements, with a high intraclass correlation coefficient and no evidence of systematic bias. In addition, Q-Bone avoids the inherent subjective visual variability of manual measurements during large-scale assessments[14] and provides stable, efficient quantitative outputs across different devices and clinical environments. Taken together, these results indicate that Q-Bone is not only methodologically robust, but also holds substantial potential, in real-world clinical practice, as an AI-based quantitative

assistance tool for large-scale screening of periodontitis, monitoring disease progression, and dynamically evaluating periodontal prognosis.

One of the core challenges in the clinical translation of AI technology is enabling clinicians to effectively understand, build trust in, and apply AI-generated results[41]. Q-Bone was therefore designed to integrate into the periodontal workflow at three levels. First, color-coded overlays display tooth boundaries and anatomical key point directly on the images, making the model output intuitive to interpret. Second, stage-based highlighting and switchable quantitative tables enable rapid identification of severely affected teeth and quick access to key metrics. Third, the interface presents the original image alongside the AI result, allowing clinicians to verify doubtful cases and retain final control over staging and diagnosis in a human–AI collaborative mode. This design is consistent with the AAP/EFP classification standards, which emphasise quantitative assessment of alveolar bone loss for precise evaluation of periodontitis, and highlights the potential practical value of Q-Bone for baseline assessment and follow-up evaluation in periodontal care.

Grad-CAM visualizations (Figs. 4 and 5) further support the anatomical plausibility of Q-Bone's decisions. Rather than relying on background artefacts or irrelevant regions, DGNet consistently allocates attention to periodontal landmarks that clinicians routinely use for diagnosis, such as the CEJ, ABC and AP. This convergence between model focus and human diagnostic cues suggests that Q-Bone's decision process is largely driven by clinically meaningful anatomy, which may facilitate clinician trust and the integration of the system into routine periodontal imaging workflows.

This study does have certain limitations. First, the number and geographical scope of the external validation samples are still limited, and further expansion of the data to more regions, institutions, and devices is needed to consolidate the evidence for generalizability. Second, the current multicenter cohort does not include international populations, and potential differences in periodontal disease profiles across races and regions were not fully captured. Despite these limitations, the findings offer clear directions for future optimization without undermining the significance of the methodological innovation, clinical applicability, and interpretability demonstrated by Q-Bone in AI-assisted diagnosis of periodontitis.

In the future, we plan to further enhance the accuracy and precision of alveolar bone loss quantification, particularly for complex resorption patterns, through more detailed modeling of the three-dimensional anatomical structures of the oral cavity. In addition, we aim to integrate periodontal probing indices, longitudinal imaging data, and electronic medical records to develop a multimodal model for precise diagnosis and risk stratification of periodontitis.

5. Conclusion

The Q-Bone system integrates a DGNet multitask model, an anatomically-driven curvature-based quantification algorithm, and a clinical visual interactive interface, enabling automated and accurate quantification of alveolar bone loss and assisting radiographic diagnosis and staging of periodontitis. Multicenter clinical validation demonstrated strong agreement with specialist assessments and robust generalizability, providing standardized and interpretable quantitative outputs.

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee (Approval No. TREC2022-KY016) and registered with the Chinese Clinical Trial Registry (Registration No. ChiCTR2200058275), and all participants provided written informed consent prior to enrollment.

Consent for publication

All participants included in this study provided written informed consent for the publication of relevant clinical data and findings. No individual identifiable information is disclosed in this manuscript.

Availability of data and materials

Interested researchers may request the data by emailing the corresponding author at kelvinperio@163.com, with a statement of legitimate research purpose and a commitment to complying with relevant data-sharing and privacy protection protocols. The underlying code, pre-trained model weights, and complete detailed results of this study are publicly available on GitHub at the following link: https://github.com/AI4PeriodontalDisease/Intelligent-Periodontitis.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research was supported by the Capital's Funds for Health Improvement and Research (Grant No. CFH 2022-2-2056).

Authors' contributions

Jiang Lin: Conceptualization, Study design, Data analysis, Result interpretation, Writing – Review & Editing. Deng-ping Fan: Algorithm supervision, Model output interpretation, Writing – Review & Editing. Wei Li: Study design, Clinical data acquisition, Data annotation, Model development, Software design and development, Writing – Original draft. Gepeng Ji: Model algorithm development, Algorithm improvement, Training. Jingyi Liu: Algorithm development, Clinical software development. Zhuotao Yao: Clinical data collection, Data annotation, Data analysis. All authors: Writing – Review & Editing, Final manuscript approval.

Acknowledgements

We sincerely acknowledge Beijing Daxing District People's Hospital, The Fourth Affiliated Hospital of Harbin Medical University, and Jining People's Hospital for their valuable contributions to the multicenter data collection.

Declaration of use of generative AI

During the revision of this manuscript, the authors utilized ChatGPT-5.1 for English grammar checking. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

[1] L. Wu, C.-M. Huang, Q. Wang, J. Wei, L. Xie, C.-Y. Hu, Burden of severe periodontitis: new insights based on a systematic analysis from the Global Burden of Disease Study 2021, BMC Oral Health 25 (2025) 861.

https://doi.org/10.1186/s12903-025-06271-0.
[2] G. Hajishengallis, T. Chavakis, J.D. Lambris, Current understanding of periodontal disease pathogenesis and targets for host-modulation therapy, Periodontol. 2000 84 (2020) 14–34. https://doi.org/10.1111/prd.12331.
[3] G.G. Nascimento, S. Alves-Costa, M. Romandini, Burden of severe periodontitis and edentulism in 2021, with projections up to 2050: The Global Burden of Disease 2021 study, J. Periodontal Res. 59 (2024) 823–867. https://doi.org/10.1111/jre.13337.
[4] R.J. Genco, M. Sanz, Clinical and public health implications of periodontal and systemic diseases: An overview, Periodontol. 2000 83 (2020) 7–13. https://doi.org/10.1111/prd.12344.
[5] M.A. Peres, L.M.D. Macpherson, R.J. Weyant, B. Daly, R. Venturelli, M.R. Mathur, S. Listl, R.K. Celeste, C.C. Guarnizo-Herreño, C. Kearns, H. Benzian, P. Allison, R.G. Watt, Oral diseases: a global public health challenge, Lancet Lond. Engl. 394 (2019) 249–260. https://doi.org/10.1016/S0140-6736(19)31146-8.
[6] P.N. Papapanou, M. Sanz, N. Buduneli, T. Dietrich, M. Feres, D.H. Fine, T.F. Flemmig, R. Garcia, W.V. Giannobile, F. Graziani, H. Greenwell, D. Herrera, R.T. Kao, M. Kebschull, D.F. Kinane, K.L. Kirkwood, T. Kocher, K.S. Kornman, P.S. Kumar, B.G. Loos, E. Machtei, H. Meng, A. Mombelli, I. Needleman, S. Offenbacher, G.J. Seymour, R. Teles, M.S. Tonetti, Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions, J. Clin. Periodontol. 45 Suppl 20 (2018) S162–S170. https://doi.org/10.1111/jcpe.12946.
[7] A new classification scheme for periodontal and peri-implant diseases and conditions – Introduction and key changes from the 1999 classification - Caton - 2018 - Journal of Clinical Periodontology - Wiley Online Library, (n.d.). https://onlinelibrary.wiley.com/doi/10.1111/jcpe.12935 (accessed November 4, 2025).
[8] G.B. Ortigara, T. de G. Mário Ferreira, K.F. Tatsch, G.A. Romito, T.M. Ardenghi, C.S. Sfreddo, C.H.C. Moreira, The 2018 EFP/AAP periodontitis case classification demonstrates high agreement with the 2012 CDC/AAP criteria, J. Clin. Periodontol. 48 (2021) 886–895. https://doi.org/10.1111/jcpe.13462.
[9] J.B. Ludlow, R. Timothy, C. Walker, R. Hunter, E. Benavides, D.B. Samuelson, M.J. Scheske, Effective dose of dental CBCT-a meta analysis of published data and additional data for nine CBCT units, Dento Maxillo Facial Radiol. 44 (2015) 20140197. https://doi.org/10.1259/dmfr.20140197.
[10] D. Rottke, S. Patzelt, P. Poxleitner, D. Schulze, Effective dose span of ten different cone beam CT devices, Dento Maxillo Facial Radiol. 42 (2013) 20120417. https://doi.org/10.1259/dmfr.20120417.
[11] T.O. Gamba, F. Visioli, D.R. Bringmann, P.V. Rados, H.L.D. da Silveira, I.L. Flores, Impact of dental imaging on pregnant women and recommendations for fetal radiation safety: A systematic review, Imaging Sci. Dent. 54 (2024) 1–11. https://doi.org/10.5624/isd.20230177.
[12] D.M. Brasil, K. Merken, J. Binst, H. Bosmans, F. Haiter-Neto, R. Jacobs,

Monitoring cone-beam CT radiation dose levels in a University Hospital, Dento Maxillo Facial Radiol. 52 (2023) 20220213. https://doi.org/10.1259/dmfr.20220213.

[13] Evaluation of Panoramic Radiography Diagnostic Accuracy in the Assessment of Interdental Alveolar Bone Loss Using CBCT - Anbiaee - 2024 - Clinical and Experimental Dental Research - Wiley Online Library, (n.d.). https://onlinelibrary.wiley.com/doi/10.1002/cre2.70042 (accessed October 31, 2025).

[14] D. Clark-Perry, G.A. Van der Weijden, W.E.R. Berkhout, T. Wang, L. Levin, D.E. Slot, ACCURACY OF CLINICAL AND RADIOGRAPHIC MEASUREMENTS OF PERIODONTAL INFRABONY DEFECTS OF DIAGNOSTIC TEST ACCURACY (DTA) STUDIES: A SYSTEMATIC REVIEW AND META-ANALYSIS, J. Evid.-Based Dent. Pract. 22 (2022) 101665. https://doi.org/10.1016/j.jebdp.2021.101665.

[15] E.D. Şeker, A.N. Dinçer, N. Kaya, Apical Root Resorption of Endodontically Treated Teeth after Orthodontic Treatment: A Split-mouth Study, Turk. J. Orthod. 36 (2023) 15–21. https://doi.org/10.4274/TurkJOrthod.2022.2022.48.

[16] R.C. Hartmann, E.S. Ferraz, T. Weissheimer, J.A. Poli de Figueiredo, G. Rossi-Fedele, M.S. Gomes, Comparative analysis of methods for measuring root canal curvature based on periapical radiography: A laboratory study, Int. Endod. J. 57 (2024) 1848–1857. https://doi.org/10.1111/iej.14142.

[17] S. Vinayahalingam, B. Berends, F. Baan, D.A. Moin, R. van Luijn, S. Bergé, T. Xi, Deep learning for automated segmentation of the temporomandibular joint, J. Dent. 132 (2023) 104475. https://doi.org/10.1016/j.jdent.2023.104475.

[18] M. Boztuna, M. Firincioglulari, N. Akkaya, K. Orhan, Segmentation of periapical lesions with automatic deep learning on panoramic radiographs: an artificial intelligence study, BMC Oral Health 24 (2024) 1332. https://doi.org/10.1186/s12903-024-05126-4.

[19] M. Yang, C. Li, W. Yang, C. Chen, C.-H. Chung, N. Tanna, Z. Zheng, Accurate gingival segmentation from 3D images with artificial intelligence: an animal pilot study, Prog. Orthod. 24 (2023) 14. https://doi.org/10.1186/s40510-023-00465-4.

[20] B. Silva, J. Fontinele, C.L.Z. Vieira, J.M.R.S. Tavares, P.R. Cury, L. Oliveira, A holistic approach for classifying dental conditions from textual reports and panoramic radiographs, Med. Image Anal. 105 (2025) 103709. https://doi.org/10.1016/j.media.2025.103709.

[21] W. Li, L. Li, W. Xu, Y. Guo, M. Xu, S. Huang, D. Dai, C. Lu, S. Li, J. Lin, Identification of Gingival Inflammation Surface Image Features Using Intraoral Scanning and Deep Learning, Int. Dent. J. 75 (2025) 2104–2114. https://doi.org/10.1016/j.identj.2025.01.002.

[22] Y. Li, Z. Cui, L. Mei, Y. Xie, L. Marini, G. Pelekos, W. Gu, X. Yu, X. Wu, X. Wei, L. Tao, K. Deng, A. Pilloni, D. Shen, M.S. Tonetti, A novel AI-powered radiographic analysis surpasses specialists in stage II-IV periodontitis detection: a multicenter diagnostic study, NPJ Digit. Med. 8 (2025) 691. https://doi.org/10.1038/s41746-025-02077-0.

[23] R. Jiao, Y. Zhang, L. Ding, B. Xue, J. Zhang, R. Cai, C. Jin, Learning with limited annotations: A survey on deep semi-supervised learning for medical image segmentation, Comput. Biol. Med. 169 (2024) 107840. https://doi.org/10.1016/j.compbiomed.2023.107840.
[24] K. Ertaş, I. Pence, M.S. Cesmeli, Z.Y. Ay, Determination of the stage and grade of periodontitis according to the current classification of periodontal and peri-implant diseases and conditions (2018) using machine learning algorithms, J. Periodontal Implant Sci. 53 (2023) 38–53. https://doi.org/10.5051/jpis.2201060053.
[25] World Medical Association, World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects, JAMA 310 (2013) 2191–2194. https://doi.org/10.1001/jama.2013.281053.
[26] F. Schwendicke, T. Singh, J.-H. Lee, R. Gaudin, A. Chaurasia, T. Wiegand, S. Uribe, J. Krois, Artificial intelligence in dental research: Checklist for authors, reviewers, readers, J. Dent. 107 (2021) 103610. https://doi.org/10.1016/j.jdent.2021.103610.
[27] G.-P. Ji, D.-P. Fan, Y.-C. Chou, D. Dai, A. Liniger, L. Van Gool, Deep Gradient Learning for Efficient Camouflaged Object Detection, Mach. Intell. Res. 20 (2023) 92–108. https://doi.org/10.1007/s11633-022-1365-9.
[28] Y. Wang, Y. Zhang, X. Chen, S. Wang, D. Qian, F. Ye, F. Xu, H. Zhang, Q. Zhang, C. Wu, Y. Li, W. Cui, S. Luo, C. Wang, T. Li, Y. Liu, X. Feng, H. Zhou, D. Liu, Q. Wang, Z. Lin, W. Song, Y. Li, B. Wang, C. Wang, Q. Chen, M. Li, STS MICCAI 2023 Challenge: Grand challenge on 2D and 3D semi-supervised tooth segmentation, (2024). https://doi.org/10.48550/arXiv.2407.13246.
[29] K.L. Gwet, Large-Sample Variance of Fleiss Generalized Kappa, Educ. Psychol. Meas. 81 (2021) 781–790. https://doi.org/10.1177/0013164420973080.
[30] A.B. Rajendra Santosh, T. Jones, Enhancing Precision: Proposed Revision of FDI's 2-Digit Dental Numbering System, Int. Dent. J. 74 (2024) 359–360. https://doi.org/10.1016/j.identj.2023.12.001.
[31] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W.-Y. Lo, P. Dollár, R. Girshick, Segment Anything, (2023). https://doi.org/10.48550/arXiv.2304.02643.
[32] W. Wang, E. Xie, X. Li, D.-P. Fan, K. Song, D. Liang, T. Lu, P. Luo, L. Shao, PVT v2: Improved baselines with Pyramid Vision Transformer, Comput. Vis. Media 8 (2022) 415–424. https://doi.org/10.1007/s41095-022-0274-8.
[33] Camouflaged Object Detection | IEEE Conference Publication | IEEE Xplore, (n.d.). https://ieeexplore.ieee.org/document/9156837 (accessed October 31, 2025).
[34] Structure-Measure: A New Way to Evaluate Foreground Maps | IEEE Conference Publication | IEEE Xplore, (n.d.). https://ieeexplore.ieee.org/document/8237749 (accessed October 31, 2025).
[35] Salient Object Detection: A Benchmark | IEEE Journals & Magazine | IEEE Xplore, (n.d.). https://ieeexplore.ieee.org/abstract/document/7293665 (accessed October 31, 2025).
[36] Salient object detection: A survey | TUP Journals & Magazine | IEEE Xplore,

(n.d.). https://ieeexplore.ieee.org/abstract/document/10897429 (accessed October 31, 2025).
[37] T.O. Hodson, Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not, Geosci. Model Dev. 15 (2022) 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022.
[38] R. Banks, V. Thengane, M.E. Guerrero, N.M. García-Madueño, Y. Li, H. Tang, A. Chaurasia, Periodontal Bone Loss Analysis via Keypoint Detection With Heuristic Post-Processing, (2025). https://doi.org/10.48550/ARXIV.2503.13477.
[39] C. Gao, L. Wu, W. Wu, Y. Huang, X. Wang, Z. Sun, M. Xu, C. Gao, Deep learning in pulmonary nodule detection and segmentation: a systematic review, Eur. Radiol. 35 (2025) 255–266. https://doi.org/10.1007/s00330-024-10907-0.
[40] T. Xue, L. Chen, Q. Sun, Deep learning method to automatically diagnose periodontal bone loss and periodontitis stage in dental panoramic radiograph, J. Dent. 150 (2024) 105373. https://doi.org/10.1016/j.jdent.2024.105373.
[41] N. Peek, D. Capurro, V. Rozova, S.N. van der Veer, Bridging the Gap: Challenges and Strategies for the Implementation of Artificial Intelligence-based Clinical Decision Support Systems in Clinical Practice, Yearb. Med. Inform. 33 (2024) 103–114. https://doi.org/10.1055/s-0044-1800729.