Automated Tooth Numbering and Periodontal Bone Loss Severity Detection Using Residual U-Net and Tooth Long Axis in Panoramic Radiography

Widi Arrohman[1], Arna Fariza[1,*], Muhammad Udin Harun Al Rasyid[1],
Eha Renwi Astuti[2], Ramadhan Hardani Putra[2]

[1]Department of Informatics and Computer Engineering, Politeknik Elektronika Negeri Surabaya, Surabaya, Indonesia
[2]Department of Dentomaxillofacial Radiology, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

Received 08 June 2026; received in revised form 06 August 2026; accepted 7 August 2026

DOI: https://doi.org/10.46604/aiti.2026.16511

Abstract

Periodontitis is a major cause of tooth loss, and its assessment from panoramic radiographs remains subjective. This study proposes an integrated framework for automatic tooth numbering and periodontal bone loss severity assessment using a residual U-Net architecture. The framework performs cemento-enamel junction (CEJ) segmentation, alveolar bone crest (ABC) segmentation, and tooth numbering based on the Fédération Dentaire Internationale (FDI) system. It estimates tooth long-axis using an oriented bounding box (OBB) for anatomically consistent bone loss measurement. Several residual backbones are evaluated, with ResNet34 achieving the best performance. Results obtain dice scores of 0.968 and 0.960 for CEJ and ABC segmentation, respectively, and 0.852 for tooth numbering. Bone loss severity evaluation on 265 teeth from 10 panoramic radiographs shows that the OBB-based approach outperforms axis-aligned bounding boxes (AABB) method, improving accuracy from 0.61 to 0.63 and macro-F1 score from 0.45 to 0.47. These results demonstrate the feasibility of automated and anatomically consistent periodontal assessment.

Keywords: residual U-net, tooth numbering, periodontal bone loss severity, panoramic radiographs, oriented bounding box

1. Introduction

More than one billion individuals worldwide suffer from periodontal disease, a prevalent chronic inflammatory condition that increases the risk of systemic health problems, causes considerable tooth loss, and lowers quality of life [1-2]. Bone loss is the primary clinical indicator of periodontal tissue damage and can affect a patient's chewing, speech, and quality of life. In addition, several studies have shown a link between periodontal disease and systemic conditions such as diabetes mellitus, heart disease, and metabolic disorders. Therefore, early detection and monitoring of bone loss are crucial in clinical practice [3].

Bone loss assessment is generally performed using panoramic radiography because it provides a comprehensive view of the jaw structure and tooth position in a single image. However, interpreting radiographic images still relies on the dentist's manual observations, which are subject to subjectivity and variation among examiners. In addition, bone loss in the early stages often does not show obvious symptoms, thereby increasing the risk of delayed detection [4-5].

* Corresponding author. E-mail address: arna@pens.ac.id

Accurate dental numbering is essential for identifying locations and effective clinical communication. Errors can lead to incorrect localization of bone loss and influence therapy. Thus, a system that integrates bone loss detection and tooth numbering automatically is necessary [6].

Previous research has developed a variety of methods to detect bone loss, including using periapical and bite radiography. However, these methods cover only a portion of the teeth, so they do not allow a thorough analysis in a single image [5,7-8]. In addition, you only look once (YOLO)-based models also have limitations in that they only produce a rough bounding box; therefore, they are not able to detect anatomical boundaries such as the cemento-enamel junction (CEJ) and alveolar bone crest (ABC) contours with precision, which is very important in bone loss analysis [9-10].

Based on this, the main gap in this study lies in the lack of an integrated system that combines precise segmentation of dental anatomical structures, automatic dental numbering, and analysis of the severity of periodontal bone loss within a comprehensive framework. Most previous studies still focused on a single aspect in isolation, without integrating the results into a broader clinical context [11-13].

Another challenge in periodontal bone loss assessment is accurately representing the anatomical orientation of individual teeth. Conventional axis-aligned bounding boxes (AABB) are constrained to the image axes and may not adequately reflect the natural inclination of teeth. To address this limitation, the proposed framework employs an oriented bounding box (OBB) to estimate the tooth long axis according to the orientation of each segmented tooth. The resulting tooth long axis serves as the anatomical reference for CEJ–ABC measurement because periodontal bone loss is clinically evaluated along the natural orientation of the tooth. This geometry-based approach improves the anatomical consistency of bone loss assessment and constitutes a key novelty of the proposed framework.

To overcome these challenges, this study proposes a deep learning-based approach using the U-Net architecture to perform CEJ and ABC, as well as dental numbering according to the Fédération Dentaire Internationale (FDI) system. Some previous studies have shown that the U-Net has excellent capabilities for segmenting medical images and dental radiographs with high accuracy [14]. U-Net has been successfully applied to tooth segmentation in panoramic radiographs with high performance, demonstrating its potential for automating dental diagnostic tasks [15]. To further improve anatomical measurement, the segmented tooth regions are subsequently processed using an OBB to estimate the tooth long axis, which serves as the anatomical reference for periodontal bone loss severity assessment.

Recent transformer-based and foundation-model approaches, including medical adaptations of the segment anything model (SAM), demonstrated promising performance in medical image segmentation by improving contextual understanding and generalization capability [16-17]. However, these methods generally require larger datasets and higher computational resources than conventional CNN-based architectures. Given the relatively limited size of the CEJ and ABC datasets, residual U-Net is selected in this study because of its proven effectiveness, training stability, and computational efficiency for medical image segmentation tasks [14]. The main contributions of this study are summarized as follows:

(1) To propose an integrated deep learning and geometry-based framework based on a residual U-Net with a ResNet34 backbone that combines CEJ and ABC segmentation, automatic tooth numbering, and periodontal bone loss severity classification within a unified pipeline.

(2) To introduce an anatomically oriented tooth long-axis framework based on an OBB for periodontal bone loss severity classification by measuring the geometric relationship between the CEJ and the ABC along the natural orientation of each tooth.

(3) To provide a comprehensive periodontal assessment system that improves diagnostic accuracy, objectivity, and efficiency for clinical decision support.

The structure of this paper is as follows: Section 2 outlines the methodology, detailing the system architecture and coordinate transformation procedure. Section 3 presents the experimental results and their discussion. Lastly, Section 4 provides the conclusion and highlights several directions for future research.

2. Proposed Methodology

The CEJ and ABC severity scales are based on those developed by Tonetti et al [18]. The proposed methodology is described in Fig. 1. This study begins with panoramic radiographic images, which are then processed using a segmentation model to detect three main components: CEJ, ABC, and tooth numbering. CEJ and ABC are used to determine the severity level, while tooth numbering is used to identify teeth according to the FDI standard. Each detected tooth is represented as an OBB, with its long axis serving as the vertical reference line. The results of CEJ, ABC, and tooth-numbering segmentation are then combined into a single overlay to obtain an integrated anatomical representation.

Furthermore, the intersection points between the CEJ line, the ABC line, and each tooth's long axis are calculated using the line intersection algorithm to determine the CEJ and ABC reference points precisely. Based on these points, bone loss is measured by calculating the distance between the CEJ and the ABC along the tooth's long axis. These distance values are then used to determine the severity of bone loss, which is classified into normal, mild, moderate, or severe categories based on the ratio of root length to bone loss [18]. The final result is a set of panoramic radiographic images that automatically include information on tooth numbering and the severity of bone loss.

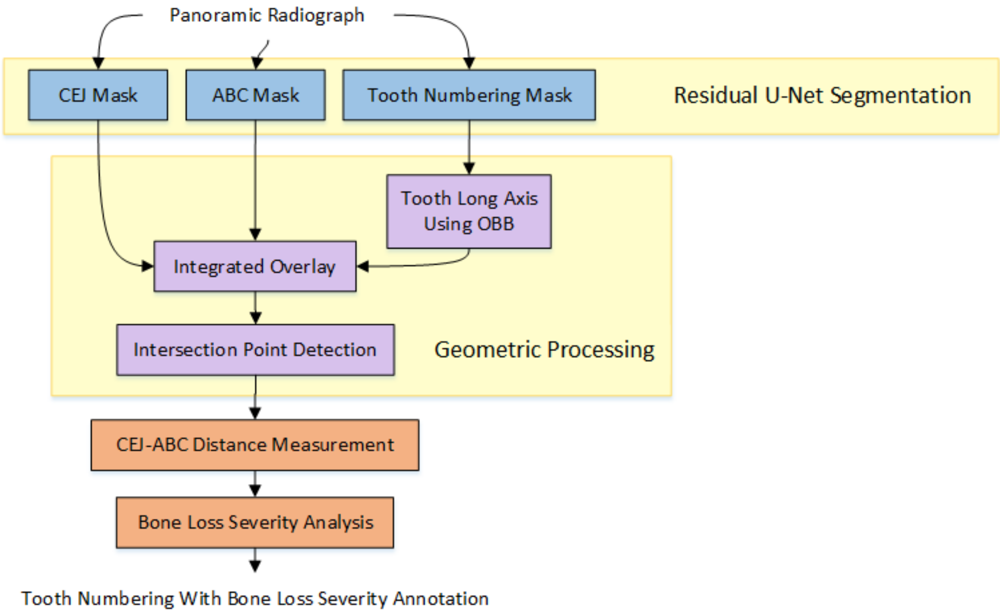


Fig. 1 Proposed methodology for tooth numbering and periodontal bone loss severity detection

2.1. Data collection

The datasets used in this study are collected from two sources of panoramic radiographs. First, the clinical dataset for CEJ and ABC segmentation is obtained from the Faculty of Dentistry, Universitas Airlangga, Indonesia, with ethical approval under certificate number 0254/HRECC.FODM/II/2025. Second, the dental numbering dataset is derived from a publicly available Kaggle dataset [19]. The Kaggle dataset provides tooth-numbering annotations according to the FDI two-digit numbering system, whereas the clinical dataset contains manually annotated CEJ and ABC boundaries. The first author prepares the initial annotations, which are subsequently reviewed and validated by an experienced dentomaxillofacial radiologist. The annotation process involves iterative review and verification to ensure anatomical consistency between the CEJ and ABC boundaries. However, formal inter-observer and intra-observer variability analyses are not performed because a single experienced dentomaxillofacial radiologist validated the final annotations. To the best of our knowledge, no publicly available panoramic radiograph dataset contains all three anatomical annotations (FDI tooth numbering, CEJ, and ABC) simultaneously. Therefore, these two complementary datasets are employed to support the proposed integrated framework for automated tooth numbering and periodontal bone loss severity assessment.

In this study, 800 labelled panoramic radiographic images serve for multi-class tooth-numbering tasks. Each panoramic radiograph corresponds to a unique patient. The datasets are randomly divided into training, validation, and testing subsets using a ratio of 80%, 10%, and 10%, respectively. Because each panoramic radiograph corresponds to a unique patient, the image-wise split is equivalent to a patient-wise split. Consequently, no patient appears simultaneously in more than one subset, thereby minimising the risk of data leakage between training and testing. Each image is annotated according to the FDI numbering system [20], with examples shown in Fig. 2(b). Because many panoramic radiographs contain missing teeth, extracted teeth, or incomplete dentition, the total number of annotated teeth is 2,309 rather than the theoretical maximum of 25,600 teeth (800 × 32). Therefore, Table 1 reports the distribution of annotated teeth instead of the number of radiographs. In addition, 200 panoramic radiographic images, each corresponding to a unique patient, are dedicated to CEJ and ABC segmentation, with example annotations shown in Fig. 2(c) and Fig. 2(d). The CEJ and ABC datasets are relatively small because each segmentation task consists of only two classes (background and object). In contrast, tooth numbering involves multiple classes and therefore demands a larger dataset. The tooth-numbering dataset trains the multi-class tooth segmentation model, whereas the clinical dataset trains the CEJ and ABC segmentation models.

(a) Original image (b) Tooth numbering annotation (c) CEJ segmentation dataset (d) ABC segmentation dataset

Fig. 2 Example of panoramic radiograph images

To ensure compatibility between the two datasets, identical preprocessing procedures are applied before model training. All panoramic radiographs are converted into PNG format, resized to 512×1024 pixels, and processed using the same image preparation pipeline. The tooth-numbering dataset is exclusively used for training the multi-class tooth segmentation model, whereas the clinical dataset serves for CEJ and ABC segmentation. During inference, the outputs of these independently trained models are integrated within a unified framework, where tooth numbering, CEJ segmentation, and ABC segmentation are combined to estimate the tooth long axis using an OBB and subsequently perform periodontal bone loss severity classification.

Table 1 Dataset tooth numbering distribution

Label	Tooth number	Count	Label	Tooth number	Count
1	11	80	17	31	93
2	12	95	18	32	93
3	13	95	19	33	93
4	14	78	20	34	81
5	15	70	21	35	63
6	16	56	22	36	42
7	17	62	23	37	47
8	18	27	24	38	55
9	21	84	25	41	84
10	22	94	26	42	92
11	23	93	27	43	95
12	24	80	28	44	91
13	25	63	29	45	70
14	26	52	30	46	54
15	27	70	31	47	62
16	28	37	32	48	58

Table 1 presents the distribution of tooth numbering in the dataset according to the FDI two-digit system. Each label corresponds to a specific tooth number and the total number of samples available for that tooth. Overall, the dataset shows a relatively balanced distribution across most tooth classes, though some variation is observed. Teeth 12, 13, and 43 have the highest representation (up to 95 samples), while 18 and 36 have comparatively fewer (27 and 42, respectively). This variation reflects the natural differences in data availability across tooth types. Despite these disparities, the dataset remains sufficiently diverse to support robust analysis and model development. Although slight variations exist among tooth classes, no additional class-balancing strategy is applied during training because the overall class distribution remains sufficiently representative for model development.

2.2. Residual U-Net architecture for segmentation of CEJ, ABC, and tooth numbering

This study employs a residual U-Net architecture for three segmentation tasks: (1) CEJ segmentation, (2) ABC segmentation, and (3) tooth numbering segmentation. To evaluate the influence of feature extraction capability, several encoder backbones are investigated, including ResNet18, ResNet34, ResNet50, and ResNet101 [14,21-22]. The network follows the standard U-Net architecture consisting of an encoder, a decoder, and skip connections, as illustrated in Fig. 3. The encoder extracts hierarchical image features. In contrast, the decoder reconstructs the segmentation map at the original image resolution. Skip connections preserve spatial information by combining low-level and high-level features. The integration of residual learning enables more effective feature extraction and facilitates the segmentation of complex dental anatomical structures.

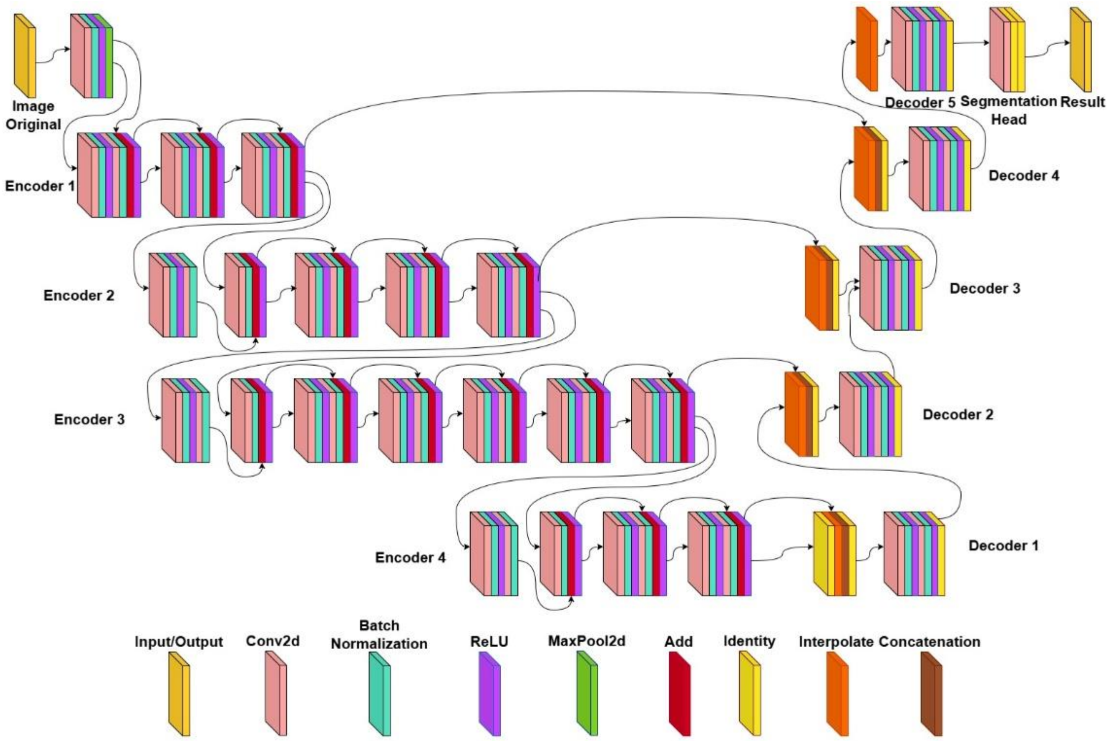


Fig. 3 ResNet34-U-Net architecture

2.3. OBB for tooth axis

The OBB is used in this study to represent the shape and orientation of each tooth based on the segmentation results, in contrast to the AABB, which is only aligned with the coordinate axes. The OBB is formed by segmenting each tooth, calculating its contour, and determining the extreme points to construct a bounding box aligned with the main orientation of the teeth. From the OBB, the centerline is obtained, which represents the tooth's long axis, which is the main axis of the tooth from the ABC to the root. These lines serve as geometric references in the analysis, particularly for determining the point of intersection between the CEJ and ABC lines.

To estimate the orientation of each tooth, an OBB is computed from the segmented tooth region. Let $S = (x_i, y_i)_{i=1}^{N}$ denote the set of pixel coordinates belonging to a segmented tooth, where N is the total number of pixels in the segmented tooth region, and (x_i, y_i) represents the coordinates of the i-th pixel. The centroid of the tooth is defined as:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{2}$$

where $\bar{x}$ and $\bar{y}$ denote the centroid coordinates of the segmented tooth along the horizontal and vertical image axes, respectively.

The covariance matrix C of the point set is computed as:

$$C = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{bmatrix} \tag{3}$$

which describes the spatial distribution of the segmented tooth pixels around the centroid. The elements σ_{xx}, σ_{yy}, and σ_{xy} represent the variance along the x-axis, the variance along the y-axis, and the covariance between the x- and y-coordinates, respectively. These terms are defined as:

$$\sigma_{xx} = \frac{1}{N}\sum (x_i - \bar{x})^2 \tag{4}$$

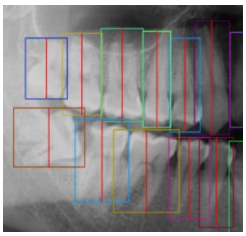
$$\sigma_{yy} = \frac{1}{N}\sum (y_i - \bar{y})^2 \tag{5}$$

$$\sigma_{xy} = \frac{1}{N}\sum (x_i - \bar{x})(y_i - \bar{y}) \tag{6}$$

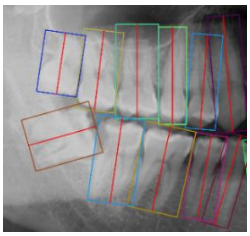
where x_i and y_i are the coordinates of the i-th pixel belonging to the segmented tooth region, while $\bar{x}$ and $\bar{y}$ are the centroid coordinates computed in Eqs. (1)-(2). The covariance terms quantify the spread of the tooth pixels relative to the centroid and are subsequently used to estimate the principal orientation of the tooth.

The principal direction of the tooth is obtained from the eigenvector corresponding to the largest eigenvalue of the covariance matrix. This direction represents the tooth's long axis, which is used as a reference for subsequent bone loss measurement. Compared to axis-aligned bounding boxes, this approach provides a more anatomically consistent estimation of tooth orientation.

The use of OBB provides an advantage in maintaining dental orientation information more accurately than the AABB method, as shown in Fig. 4. In Fig. 4(a), the estimate of the tooth axis using AABB is still not representative of the actual direction because of its limitations that are only parallel to the axis of image coordinates. On the other hand, in Fig. 4(b), the OBB-based method can adjust the orientation to the tooth shape, thereby producing a more stable and anatomically consistent estimate of the tooth axis. Thus, the OBB approach enables more precise tooth-axis estimation.



(a) Conventional AABB-based axis estimation



(b) OBB-based tooth axis estimation

Fig. 4 Comparison of tooth long-axis estimation using AABB and OBB methods

2.4. Bone loss severity classification

The classification of periodontal bone loss severity is based on the contemporary framework proposed by Caton et al., which was a follow-up to the study by Tonetti et al. [18] in the 2018 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions (2018 World Workshop) [23]. In this framework, the severity of periodontitis is determined by the percentage of alveolar bone loss relative to root length, providing a standardized, clinically interpretable measure for disease staging. This approach enables objective quantification of bone loss and facilitates consistent classification across different cases.

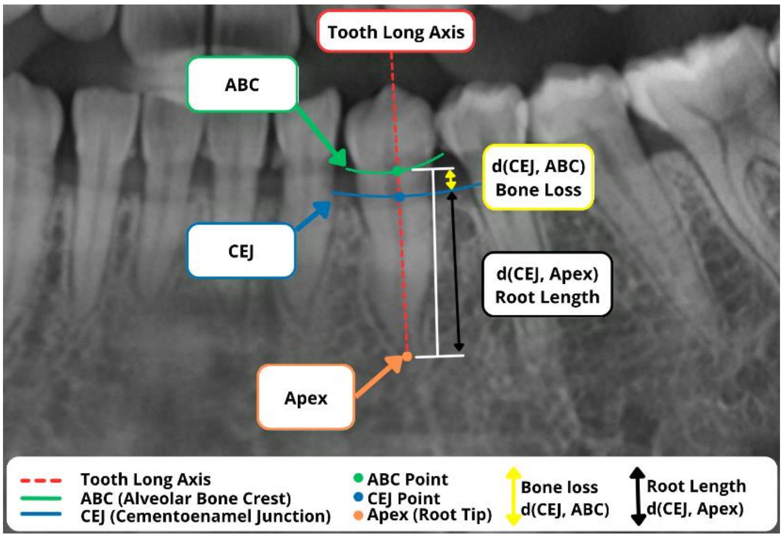


Fig. 5 Illustration of bone loss severity classification

Fig. 5 shows the geometric process for calculating the percentage of periodontal bone loss along the tooth's length axis. The percentage of bone loss is then calculated using a comparison between the distance of bone loss and the length of the root, namely:

$$\text{Bone Loss} = \text{d (CEJ, ABC)} \tag{7}$$

$$\text{Root Length} = \text{d (CEJ, Apex)} \tag{8}$$

$$Bone\ Loss\ Ratio\ (\%) = \frac{Bone\ Loss}{Root\ Length} \times 100\% \tag{9}$$

According to the staging framework proposed by Tonetti et al. [18], the calculated bone loss ratio is categorized into four severity levels: normal (<15%), mild (15%–33%), moderate (>33%–66%), and severe (>66%). These percentage ranges provide a quantitative implementation of the radiographic bone loss stages described in the 2018 World Workshop, allowing continuous bone loss measurements to be converted into clinically interpretable severity categories.

3. Result and Discussion

This section presents the quantitative evaluation and performance analysis of the proposed residual U-Net models for CEJ segmentation, ABC segmentation, and tooth numbering. Segmentation performance is evaluated using the intersection over union (IoU) and dice similarity coefficient (DSC), which are widely adopted metrics in medical image segmentation to measure the overlap between predicted masks and ground-truth annotations [24-25].

The IoU metric is defined as:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{10}$$

where A and B denote the predicted and ground-truth segmentation masks, respectively.

The DSC is calculated as:

$$DSC = \frac{2|A \cap B|}{|A|+|B|} \tag{11}$$

Higher IoU and DSC values indicate better agreement between the predicted segmentation and the reference annotation.

3.1. Training evaluation of segmentation results

The model is implemented using PyTorch 2.8.0 and the segmentation models pytorch (SMP) library on a workstation equipped with an NVIDIA GeForce RTX 3060 GPU (12 GB VRAM). The CEJ, ABC, and tooth-numbering datasets are randomly divided into training, validation, and testing subsets using an 80:10:10 ratio. The validation set is used for model selection and early stopping, whereas the testing set is reserved exclusively for final performance evaluation and remains completely isolated during model development.

The training process is evaluated to assess the model's stability and effectiveness. The analysis includes observations of training and validation metrics—including loss values, accuracy, IoU, and F1-score—to ensure that the model converges well without overfitting. The following table summarizes the hyperparameter configurations used in all experiments.

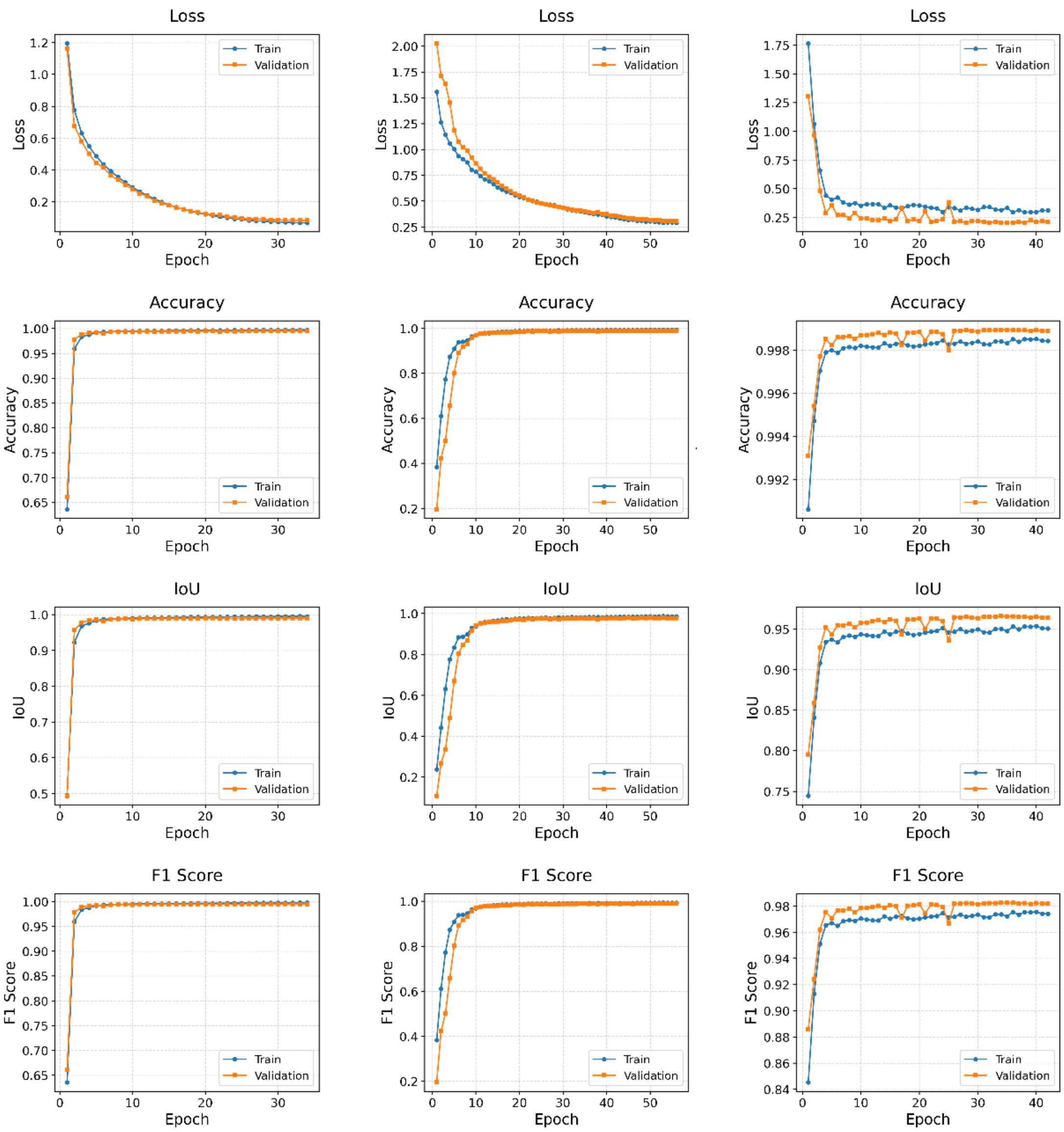
Table 2 presents the hyperparameter configurations used for CEJ, ABC, and tooth numbering segmentation. A learning rate of 1e-4 is used for binary segmentation to ensure stable boundary refinement. In contrast, a slightly higher rate of 3e-4 is used for multi-class segmentation to accelerate convergence in more complex classification settings. The batch size is set to 4 for binary segmentation and 2 for multi-class segmentation, taking into account GPU memory limitations and model stability.

Table 2 Hyperparameter Configuration for CEJ, ABC, and Tooth Numbering Segmentation

Components	CEJ & ABC	Tooth numbering
Segmentation type	Binary	Multi-class
Number of classes	2	33
Optimizer	Adam	Adam
Learning rate	1e-4	3e-4
Batch size	4	2
Early stop	Yes	Yes
Image size	512 × 1024	512 × 1024
Loss function	Binary cross	Categorical cross

Fig. 6 illustrates the training and validation performance curves for all segmentation tasks. Overall, the models exhibit stable convergence, characterized by decreasing loss values and improving accuracy, IoU, and F1-score throughout training. The CEJ and ABC models converge rapidly because they involve binary segmentation tasks. Conversely, the tooth-numbering model exhibits slightly larger fluctuations due to the increased complexity of multi-class segmentation. Nevertheless, the close agreement between training and validation curves suggests good generalization performance with limited overfitting.

It should be noted that the metrics presented in Fig. 6 and Table 3 are computed using different evaluation strategies. Fig. 6 reports micro-averaged metrics obtained during the training and validation stages, where all pixels from all classes are aggregated before metric calculation. In comparison, Table 3 reports mean IoU and mean dice scores evaluated on the independent test set by averaging the performance across individual classes. Consequently, the values reported in Fig. 6 are generally higher, particularly for the tooth-numbering task, because micro-averaging is influenced by the dominant background pixels. In contrast, the class-wise evaluation in Table 3 provides a more challenging and clinically relevant assessment of segmentation performance.

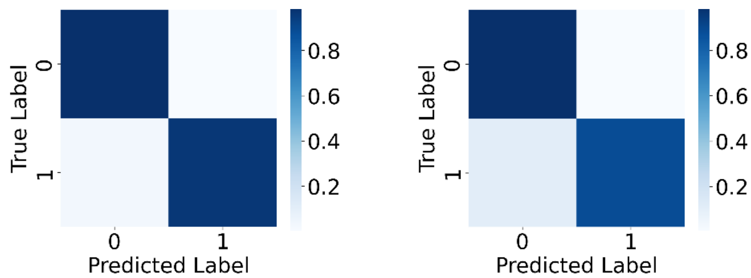


(a) CEJ (b) ABC (c) Tooth numbering

Fig. 6 Training and validation performance of the CEJ, ABC, and tooth numbering models

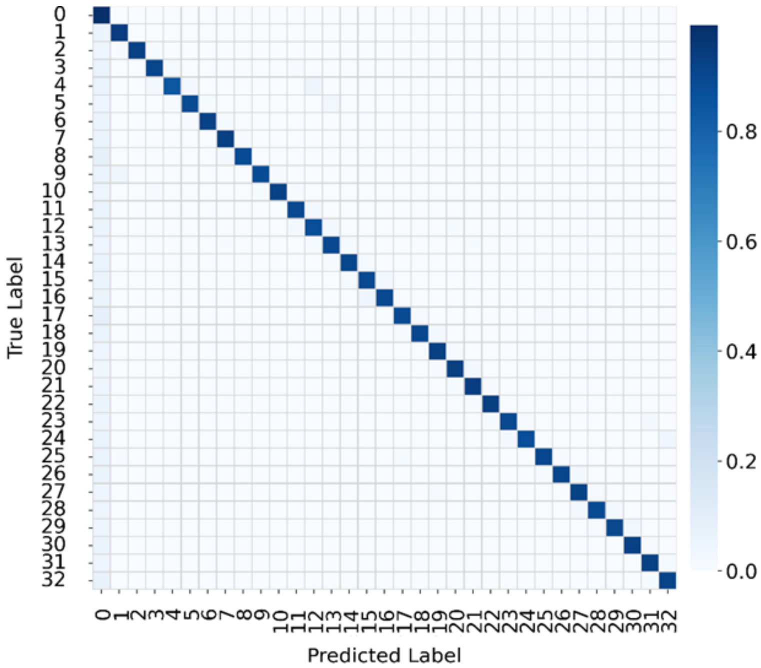
3.2. Testing evaluation of segmentation results

The segmentation performance of the CEJ, ABC, and tooth-numbering models is evaluated using accuracy, precision, recall, dice coefficient (F1-score), and IoU metrics derived from the confusion matrix. Fig. 7 shows that most predictions are concentrated along the main diagonal of the confusion matrices, indicating high agreement between predicted and ground-truth labels with relatively few misclassifications.



(a) CEJ training confusion matrix (b) ABC training confusion matrix

Fig. 7 Confusion matrix



(c) Tooth numbering training confusion matrix

Fig. 7 Confusion matrix (continued)

Table 3 Performance comparison of U-Net with different backbones across three segmentation tasks

Object	Backbone	IoU	Dice coefficient
CEJ	U-Net (vanilla)	0.936	0.966
	ResNet18	0.936	0.966
	ResNet34	0.937	0.968
	ResNet50	0.936	0.966
	ResNet101	0.936	0.966
ABC	U-Net (vanilla)	0.923	0.958
	ResNet18	0.917	0.955
	ResNet34	0.925	0.960
	ResNet50	0.923	0.958
	ResNet101	0.918	0.955
Tooth numbering	U-Net (vanilla)	0.744	0.851
	ResNet18	0.584	0.732
	ResNet34	0.743	0.852
	ResNet50	0.484	0.641
	ResNet101	0.665	0.796

Table 3 shows that ResNet34 provides a favourable balance between segmentation performance and model complexity across the evaluated segmentation tasks. Although the vanilla U-Net achieves a marginally higher IoU for tooth numbering, ResNet34 obtains the highest overall dice performance and is therefore selected as the backbone of the proposed framework.

These findings indicate that the proposed residual U-Net with a ResNet34 backbone provides a favourable balance between segmentation accuracy and computational complexity. Furthermore, the proposed framework integrates anatomical segmentation, automatic tooth numbering, and geometry-based periodontal bone loss assessment within a unified pipeline, improving the consistency and efficiency of panoramic radiograph analysis.

3.3. Visualization of segmentation results

This section presents visualization results to complement the quantitative evaluation, illustrating the performance of the proposed framework in CEJ segmentation, ABC segmentation, tooth numbering, and periodontal bone loss assessment. As shown in Fig. 8, the blue regions indicate the CEJ boundaries successfully identified by the model. Furthermore, Fig. 9 illustrates the predicted ABC segmentation results, where the green regions represent the detected alveolar bone crest boundaries. Finally, Fig. 10 shows the tooth segmentation and numbering results. The model successfully separates individual teeth and assigns FDI tooth numbers consistently across both dental arches.

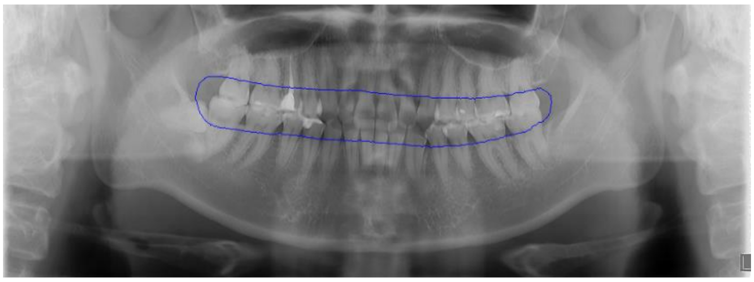


Fig. 8 Prediction of CEJ segmentation

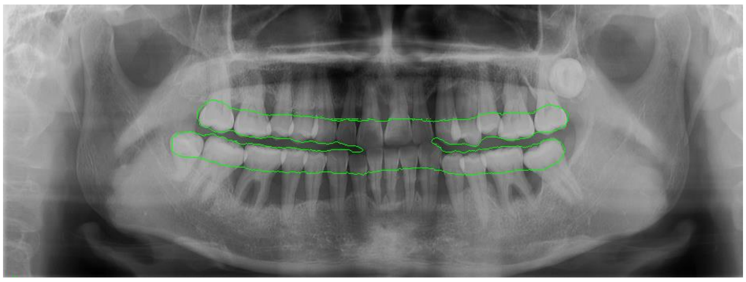


Fig. 9 Prediction of ABC segmentation

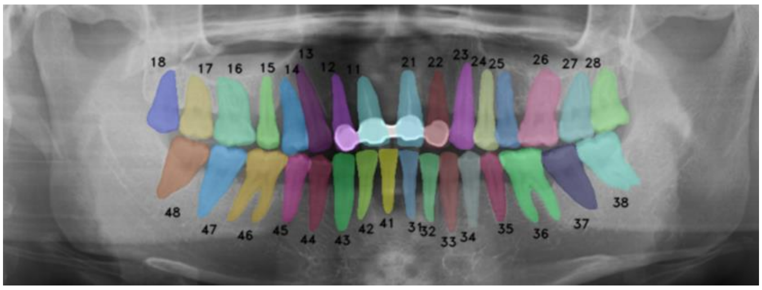


Fig. 10 Prediction of tooth numbering segmentation

Fig. 11 shows the final results of the proposed framework on panoramic radiographs. The predicted CEJ and ABC boundaries are shown in blue and green, respectively. The red line indicates the tooth long axis for each detected tooth. Roman numerals (I, II, III, IV) indicate the predicted bone loss severity, whereas the black numbers represent the assigned tooth numbers. By integrating these components into a single visualization, the proposed framework provides anatomically consistent assessment of periodontal bone loss and automatic tooth numbering.

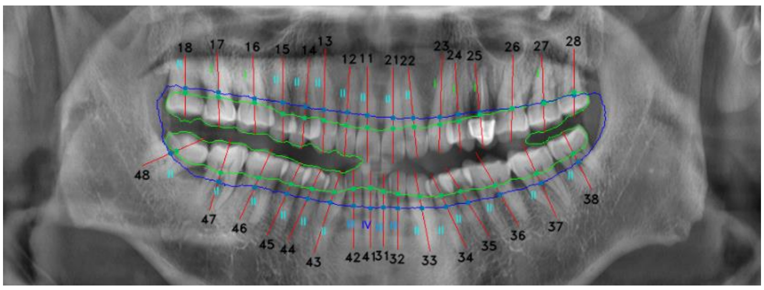


Fig. 11 Visualization of bone loss classification and tooth numbering predictions

To evaluate the contribution of the proposed OBB-based tooth axis estimation, bone loss severity classification using OBB is compared with the conventional AABB approach. An experienced dento-maxillofacial radiologist provided ground-truth labels. The evaluation involves 265 teeth from 10 panoramic radiographs and is assessed using accuracy, precision, recall, macro-F1 score, and balanced accuracy.

Table 4 compares the performance of AABB and OBB. Although OBB achieves slightly lower precision than AABB (0.54 vs 0.57), it yields higher accuracy, recall, macro-F1, and balanced accuracy. However, McNemar's test shows no statistically significant difference (p = 0.263). These results suggest that anatomical tooth orientation may offer a modest advantage in assessing the severity of periodontal bone loss.

Table 4 Comparison of bone loss severity classification using AABB and OBB

Method	Accuracy	Precision	Recall	Macro-F1	Balanced accuracy
AABB	0.61	0.57	0.44	0.45	0.44
OBB (Proposed)	0.63	0.54	0.46	0.47	0.46

Table 5 presents the class-wise performance of the proposed OBB method. The mild category achieves the highest F1-score (0.73), whereas the severe category has the lowest (0.12). The limited number of severe cases (n = 13) likely causes class imbalance and reduces classification reliability. Fig. 12 shows the confusion matrix of the proposed OBB-based bone loss severity classification. Most predictions lie on the main diagonal, indicating good agreement with the reference labels. Misclassifications mainly occur between adjacent severity categories, whereas the Severe class exhibits the lowest performance due to the limited number of samples.

Table 5 Class-wise performance of the proposed OBB method

Severity	Precision	Recall	F1-score	Samples
Normal	0.71	0.50	0.58	94
Mild	0.64	0.87	0.73	120
Moderate	0.50	0.42	0.45	38
Severe	0.33	0.07	0.12	13
Macro average	0.54	0.46	0.47	265

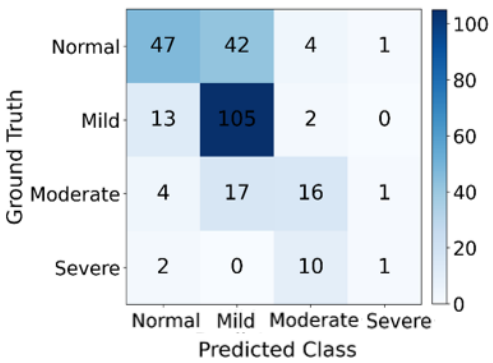


Fig. 12 Confusion matrix of four-class periodontal bone loss severity classification

3.4. Discussion

The experimental results demonstrate the effectiveness of the proposed framework for automatic tooth numbering and assessment of periodontal bone loss severity. Nevertheless, several limitations should be acknowledged. First, the CEJ and ABC segmentation models are trained on a smaller clinical dataset than the tooth-numbering model. Second, two independent datasets are used because no publicly available panoramic radiograph dataset provides FDI tooth numbering, CEJ, and ABC annotations simultaneously. Although identical preprocessing is applied, differences in image characteristics and annotation protocols may affect model generalization. Furthermore, although instance segmentation may provide explicit separation of individual teeth, semantic tooth-numbering segmentation is adopted because it simultaneously provides pixel-level anatomical delineation and tooth identity required for the proposed integrated framework.

The tooth-numbering dataset shows moderate class imbalance, with class frequencies ranging from 27 to 95 samples. No balancing strategy is applied because the distribution remains representative and the model maintains strong performance across tooth classes. All backbone architectures are trained under identical settings to ensure a fair comparison rather than to optimize each architecture individually. Formal inter-observer and intra-observer variability analyses are beyond the scope of this study. Although the final annotations are reviewed and validated by an experienced dento-maxillofacial radiologist, future work should investigate the reliability of multi-expert annotation and backbone-specific optimization.

The proposed framework may encounter challenges in panoramic radiographs with overlapping teeth, extensive restorations, severe artefacts, or missing teeth. In such cases, segmentation errors may propagate to tooth long-axis estimation and affect the classification of periodontal bone loss severity. Despite these limitations, the proposed framework shows potential as a clinical decision-support tool by providing automatic tooth numbering and anatomically consistent assessment of bone loss. The system is intended to assist rather than replace clinicians.

4. Conclusions

This study proposed an integrated deep learning framework for automatic tooth numbering and periodontal bone loss severity assessment on panoramic radiographs using a Residual U-Net architecture with a ResNet34 backbone. The proposed framework combined CEJ segmentation, ABC segmentation, FDI-based tooth numbering, and OBB-based tooth long-axis estimation within a unified periodontal assessment pipeline. Based on the experimental results, the following conclusions can be drawn:

(1) The Residual U-Net with a ResNet34 backbone achieved strong segmentation performance, obtaining dice scores of 0.968 and 0.960 for CEJ and ABC segmentation, respectively.

(2) The proposed framework successfully performed automatic tooth numbering according to the FDI system, achieving a dice score of 0.852 and demonstrating reliable tooth identification performance on panoramic radiographs.

(3) The OBB-based tooth long-axis estimation enabled anatomically consistent periodontal bone loss measurement and improved bone loss severity classification compared with conventional AABB-based estimation.

(4) The integration of anatomical segmentation, automatic tooth numbering, and geometry-based bone loss assessment within a single framework provided a comprehensive approach for periodontal evaluation and may support improved diagnostic consistency and clinical efficiency.

Future work will focus on validating the proposed framework using larger multi-center datasets and investigating transformer-based and foundation-model architectures, including transformers make strong encoders for medical image segmentation (TransUNet), unet-like pure transformer for medical image segmentation (Swin-UNet), not-another transformer for volumetric medical image segmentation (nnFormer), medical segment anything model (MedSAM), and SAM, to further improve robustness and generalization.

Acknowledgments

Ethical approval for this study was obtained from the Health Research Ethical Clearance Commission, Faculty of Dental Medicine, Universitas Airlangga, Indonesia, under approval number 0254/HRECC.FODM/II/2025.

Conflicts of Interest

The authors declare no conflict of interest.

Statement of Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Statement of informed consent

Informed consent was obtained from all individual participants included in the study.

References

[1] F. Chairunisa, A. Ramadhani, S. Takehara, K. M. Thwin, T. Z. Tun, H. Okubo, L. Hanindriyo, et al., "Oral Health Status and Oral Healthcare System in Indonesia: A Narrative Review," Journal of International Society of Preventive and Community Dentistry, vol. 14, no. 5, pp. 352-361, 2024.

[2] N. T. Hashim, R. Babiker, V. Padmanabhan, A. T. Ahmed, N. C. S. K. Chaitanya, R. Mohammed, et al., "The Global Burden of Periodontal Disease: A Narrative Review on Unveiling Socioeconomic and Health Challenges," International Journal of Environmental Research and Public Health, vol. 22, no. 4, article no. 624, 2025.

[3] M. Zhang, Y. Liu, H. Afzali, and D. T. Graves, "An Update on Periodontal Inflammation and Bone Loss," Frontiers in Immunology, vol. 15, article no. 1385436, 2024.

[4] E. Tian, J. Hong, Z. Tang, R. Ren, S. Li, A. A. Abdulqader, et al., "Development and Validation of a Polyfit Approach for Assessing Alveolar Bone Loss Using Panoramic Radiography," BMC Oral Health, vol. 25, no. 1, article no. 417, 2025.

[5] S. Kurt-Bayrakdar, I. S. Bayrakdar, M. B. Yavuz, N. Sali, Ö. Çelik, O. Köse, et al., "Detection of Periodontal Bone Loss Patterns and Furcation Defects from Panoramic Radiographs Using Deep Learning Algorithm: A Retrospective Study," BMC Oral Health, vol. 24, no. 1, article no. 155, 2024.

[6] B. G. Zancan, J. A. Carneiro, C. U. Martins, C. Tirapelli, C. P. Capel, E. D. Costa, et al., "AI-Powered Precision in Dental Radiographic Analysis Using Tailored CNNs for Tooth Numbering and Cavity Detection," PLOS Digital Health, vol. 4, no. 11, article no. e0001074, 2025.

[7] A. AlGhaihab, A. J. Moretti, J. Reside, L. Tuzova, Y.-S. Huang, and D. A. Tyndall, "Automatic Detection of Radiographic Alveolar Bone Loss in Bitewing and Periapical Intraoral Radiographs Using Deep Learning Technology: A Preliminary Evaluation," Diagnostics, vol. 15, no. 5, article no. 576, 2025.

[8] M. B. Yavuz, N. Sali, S. K. Bayrakdar, C. Ekşi, B. S. İmamoğlu, İ. Ş. Bayrakdar, et al., "Classification of Periapical and Bitewing Radiographs as Periodontally Healthy or Diseased by Deep Learning Algorithms," Cureus, vol. 16, no. 5, article no. e60550, 2024.

[9] R. H. Putra, E. R. Astuti, I. K. E. Wijaksana, A. Fariza, R. M. Sitalaksmi, and N. Yoda, "Automated Periodontal Bone Loss Detection on Panoramic Radiographs Using You Only Look Once v8 (YOLOv8): A Retrospective AI Approach," Journal of International Oral Health, vol. 17, no. 3, pp. 203-211, 2025.

[10] K. C. Li, Y. C. Mao, M. F. Lin, Y. Q. Li, C. A. Chen, T. Y. Chen, et al., "Detection of Tooth Position by YOLOv4 and Various Dental Problems Based on CNN with Bitewing Radiograph," IEEE Access, vol. 12, pp. 11822-11835, 2024.

[11] H. Tunç, N. Akkaya, B. Aykanat, and G. Ünsal, "U-Net-Based Deep Learning for Simultaneous Segmentation and Agenesis Detection of Primary and Permanent Teeth in Panoramic Radiographs," Diagnostics, vol. 15, no. 20, article no. 2577, 2025.

[12] J. Ryu, D.-M. Lee, Y.-H. Jung, O. Kwon, S. Park, J. Hwang, et al., "Automated Detection of Periodontal Bone Loss Using Deep Learning and Panoramic Radiographs: A Convolutional Neural Network Approach," Applied Sciences, vol. 13, no. 9, article no. 5261, 2023.

[13] T. Xue, L. Chen, and Q. Sun, "Deep Learning Method to Automatically Diagnose Periodontal Bone Loss and Periodontitis Stage in Dental Panoramic Radiograph," Journal of Dentistry, vol. 150, article no. 105373, 2024.

[14] S. S. Alharbi, A. A. AlRugaibah, H. F. Alhasson, and R. U. Khan, "Detection of Cavities from Dental Panoramic X-Ray Images Using Nested U-Net Models," Applied Sciences, vol. 13, no. 23, article no. 12771, 2023.

[15] M. A. Makrygiannakis, A. Ntokos, and E. G. Kaklamanos, "Development and Evaluation of an Attention-Gated U-Net Model for Binary Segmentation of Teeth versus Background in Panoramic Radiographs for Orthodontic Applications," European Journal of Orthodontics, vol. 48, no. 1, article no. cjaf114, 2026.
[16] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment Anything in Medical Images," Nat Commun, vol. 15, no. 1, article no. 654, 2024.
[17] X. Sun, J. Liu, H. Shen, X. Zhu, and P. Hu, "On Efficient Variants of Segment Anything Model: A Survey," International Journal of Computer Vision, vol. 133, no. 10, pp. 7406-7436, 2025.
[18] M. S. Tonetti, H. Greenwell, and K. S. Kornman, "Staging and Grading of Periodontitis: Framework and Proposal of a New Classification and Case Definition," Journal of Periodontology, vol. 89, no. S1, pp. S159-S172, 2018.
[19] W. Zhou, X. Lu, D. Zhao, M. Jiang, L. Fan, W. Zhang, et al., "A Dual-Labeled Dataset and Fusion Model for Automatic Teeth Segmentation, Numbering, and State Assessment on Panoramic Radiographs," BMC Oral Health, vol. 24, no. 1, article no. 1201, 2024.
[20] J. C. Turp and K. W. Alt, "Designating Teeth: The Advantages of the FDI's Two-Digit System," Quintessence International, vol. 26, no. 7, pp. 501-504, 1995.
[21] Y. Pamungkas, E. Triandini, W. Yunanto, and Y. Thwe, "Impact of hyperparameter tuning on ResNet-UNet models for enhanced brain tumor segmentation in MRI scans," International Journal of Robotics and Control Systems, vol. 5, no. 2, pp. 917-936, 2025.
[22] Z. Huang, B. Li, Y. Cheng, and J. Kim, "Odontogenic Cystic Lesion Segmentation on Cone-Beam CT Using an Auto-Adapting Multi-Scaled UNet," Frontiers in Oncology, vol. 14, article no. 1379624, 2024.
[23] J. G. Caton, G. Armitage, T. Berglundh, I. L. C. Chapple, S. Jepsen, K. S. Kornman, et al., "A New Classification Scheme for Periodontal and Peri-Implant Diseases and Conditions – Introduction and Key Changes from the 1999 Classification," Journal of Periodontology, vol. 89, no. S1, pp. S1-S8, 2018.
[24] D. Müller, I. Soto-Rey, and F. Kramer, "Towards a Guideline for Evaluation Metrics in Medical Image Segmentation," BMC Research Notes, vol. 15, no. 1, article no. 210, 2022.
[25] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, et al., "Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index: Scientific Reports," Academic Radiology, vol. 11, no. 2, pp. 178-189, 2004.