

Article

An AI-Based Framework for Automated Radiographic Bone Loss Measurement Using Segmentation and Geometric Landmark Modeling

Mohammad Abdel-Majeed ^{1,2,*} , Iyad Jafar ¹, Omar AL-Karadsheh ³ , Shorouq Al-Awawdeh ⁴, Siraj Zabadi ⁵ and Mahdi Flefl ³

¹ Department of Computer Engineering, School of Engineering, The University of Jordan, Amman 11942, Jordan; iyad.jafar@ju.edu.jo

² Department of Electrical Engineering, Princess Sumaya University for Technology, Amman 11941, Jordan

³ Department of Oral and Maxillofacial Surgery, Oral Medicine, and Periodontology, School of Dentistry, The University of Jordan, Amman 11942, Jordan; o.karadsheh@ju.edu.jo (O.A.-K.); m.flefl@ju.edu.jo (M.F.)

⁴ Applied Science Department, Middle East University, Amman 11831, Jordan; s.awawdeh@meu.edu.jo

⁵ College of Dentistry, The Ohio State University, Columbus, OH 43210, USA; zabadi.1@osu.edu

* Correspondence: m.abdel-majeed@ju.edu.jo

Abstract

Accurate assessment of radiographic bone loss (RBL) is essential for periodontal diagnosis and staging; however, manual measurement from dental radiographs is labor-intensive, time-consuming and subject to inter- and intra-examiner variability. Existing AI-based methods primarily formulate bone loss assessment as classification, landmark prediction, or direct segmentation of thin anatomical structures, limiting measurement interpretability and robustness. This study proposes clinically interpretable two-phase framework for automated and clinically interpretable RBL estimation from periapical radiographs. The framework explicitly separates anatomical structure recognition from geometric measurement, improving transparency and reducing error propagation. In the first phase, deep learning models segment key anatomical structures, including the crown, root, third root and alveolar bone. In the second phase, a deterministic geometric algorithm extracts clinically relevant landmarks, including the cemento–enamel junction (CEJ), bone crest, and root apex, and computes root length, CEJ–bone crest distance, and radiographic bone loss following established periodontal measurement principles. The framework was evaluated on a curated dataset of annotated radiographs. DS-TransUNet achieved the best segmentation performance. Quantitative evaluation yielded mean absolute errors of 0.81 mm for CEJ–bone crest distance, 0.71 mm for root length, and 5.89% for RBL estimation. Bland–Altman analysis demonstrated minimal systematic bias (−1.03%) and good agreement with expert measurements across different disease severities, supporting the framework’s potential as an objective and clinically applicable tool for periodontal bone loss assessment.



Academic Editor: Francesc Pozo

Received: 10 May 2026

Revised: 27 June 2026

Accepted: 3 July 2026

Published: 8 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: radiographic bone loss (RBL); periodontal disease; periapical radiographs; medical image segmentation; geometric modeling; landmark extraction; deep learning; dental image analysis

1. Introduction

Periodontitis is a chronic inflammatory disease characterized by progressive destruction of the periodontal supporting tissues, including alveolar bone, periodontal ligament, and gingiva. It remains one of the leading causes of tooth loss in adults worldwide and has

been associated with systemic conditions such as cardiovascular disease [1,2]. Accurate assessment of alveolar bone loss is therefore central to diagnosis, staging, prognosis, and treatment planning.

Radiographic evaluation of periodontal bone level is routinely performed using periapical or panoramic radiographs [3–5]. Clinically, radiographic bone loss (RBL) is determined by measuring the distance between the cemento–enamel junction (CEJ) and the alveolar bone crest relative to total root length. This measurement-based approach forms the basis of contemporary periodontal staging systems [6]. However, manual radiographic measurement is inherently subject to inter- and intra-examiner variability [3–5]. Variability arises from difficulty in precisely identifying CEJ endpoints, bone crest contours, and root apices, particularly in images affected by contrast variation, anatomical superimposition, or multi-root morphology. Because periodontal staging thresholds depend on percentage-based measurements, small landmark localization discrepancies may result in clinically relevant differences in disease classification.

Recent advances in artificial intelligence (AI), particularly deep learning (DL), have enabled automated analysis of dental radiographs. Existing AI-based approaches for periodontal bone loss assessment can be broadly categorized into classification-based, segmentation-based, and hybrid geometric methods, each with distinct advantages and limitations. Early DL-based systems focused primarily on classification or detection of periodontal bone loss from panoramic images [7–10], demonstrating promising diagnostic performance but without explicitly modeling anatomical landmarks or reproducing the geometric measurement procedure used in clinical practice. These approaches function primarily as screening tools rather than quantitative measurement systems.

More recent approaches have adopted segmentation-based strategies using U-Net and related encoder-decoder architectures [11,12] to detect tooth structures, alveolar bone, and CEJ lines. Lee et al. proposed a deep learning framework integrating multiple segmentation networks to compute RBL percentages from radiographic images [11]. Although segmentation accuracy was reported to be high, modeling the CEJ as a thin segmentation output introduces sensitivity to pixel-level noise and discontinuities. Since RBL computation depends directly on CEJ–bone intersection points, segmentation inaccuracies may propagate into measurement variability.

Hybrid frameworks integrating object detection, segmentation, and geometric fitting have also been explored. Chen et al. proposed an ensemble deep learning approach combining YOLO-based tooth localization [13], U-Net segmentation [12], and ellipse fitting to estimate tooth long axes and compute bone loss percentages. While this approach integrates structural detection and geometric modeling, detailed specification of the axis-fitting procedure and robustness to anatomical variability, particularly in multi-root teeth, remains limited.

Despite the considerable progress achieved by recent AI-based periodontal assessment systems, several challenges remain. Many existing approaches formulate bone loss assessment as a classification or regression task, providing limited insight into the anatomical basis of the generated predictions. Segmentation-based methods improve interpretability but often rely on direct segmentation of thin anatomical structures such as the CEJ, making the resulting measurements sensitive to segmentation discontinuities and local noise. Furthermore, most existing frameworks tightly couple structural recognition with the final measurement process, limiting the ability to independently analyze segmentation and measurement errors. These limitations may become more pronounced in anatomically complex cases such as multi-root molars, where small landmark localization errors may propagate into clinically meaningful differences in radiographic bone loss estimation and periodontal staging.

These observations highlight a critical methodological gap. While deep learning has substantially improved anatomical structure recognition, existing methods either prioritize diagnostic prediction without explicit anatomical measurements or tightly couple segmentation with measurement generation, making the final RBL estimation highly dependent on segmentation errors. In addition, several geometric measurement approaches provide limited details regarding landmark derivation and geometric fitting procedures, making independent implementation and reproducibility more challenging. Furthermore, the handling of anatomically complex scenarios, such as multi-root molars and irregular root morphologies, is often not explicitly described or evaluated. Consequently, there remains a need for a framework that combines robust anatomical segmentation with transparent, reproducible, and clinically meaningful geometric measurement aligned with routine periodontal assessment.

To address this need, this paper proposes a two-phase framework that explicitly separates anatomical structure recognition from geometric measurement. In the first phase, deep learning-based segmentation is employed to identify crown, root, third root and alveolar bone structures. In the second phase, a deterministic geometric analysis algorithm extracts CEJ endpoints, root apex, and bone crest landmarks using morphological processing and shape-based constraints. Unlike approaches that directly segment the CEJ as a thin line, the proposed method derives CEJ landmarks from structural relationships between segmented anatomical regions, thereby reducing sensitivity to pixel-level segmentation noise. The framework explicitly handles multi-root teeth and enforces anatomical consistency in landmark localization.

By combining data-driven segmentation with anatomically grounded geometric modeling, the proposed system enables reproducible and clinically interpretable radiographic bone loss measurement. The proposed framework explicitly models clinically relevant anatomical landmarks and derives RBL using deterministic geometric analysis. This design enables objective quantitative assessment while remaining closely aligned with standard periodontal measurement protocols. The framework is evaluated on annotated periapical radiographs, with comprehensive analysis of segmentation performance and measurement accuracy.

The main contributions of this work can be summarized as follows:

- A two-phase framework that combines deep learning-based segmentation with deterministic geometric analysis for automated radiographic bone loss estimation.
- An anatomically constrained landmark extraction approach that enables clinically interpretable measurement of CEJ, bone crest, and root apex without requiring direct landmark prediction.
- A comprehensive evaluation protocol that integrates geometric measurement accuracy, Bland-Altman agreement analysis, and bone loss assessment.

2. Related Work

Artificial intelligence (AI), particularly deep learning (DL), has been increasingly applied to dental radiographic analysis, including tooth detection, segmentation, and periodontal bone loss assessment. Existing approaches can be broadly categorized into four groups: (1) classification-based detection methods, (2) segmentation-based measurement frameworks, (3) hybrid systems integrating multiple DL components, and (4) machine learning-based estimation approaches.

Early DL-based studies primarily focused on classification or detection of periodontal bone loss using panoramic radiographs. Krois et al. [7] developed a convolutional neural network (CNN) for detecting bone loss, while Kim et al. [8] proposed DeNTNet, a transfer learning-based approach for identifying periodontal disease. Similarly, Chang et al. [9]

introduced a hybrid DL framework for diagnosing and staging periodontitis. More recently, Kong et al. [10] proposed a two-phase detection framework for automated bone loss diagnosis in panoramic radiographs. While these approaches demonstrate strong diagnostic capability, they do not explicitly model anatomical landmarks or reproduce the geometric measurement procedure used in clinical periodontal assessment, limiting their interpretability and quantitative applicability.

More recent studies have shifted toward segmentation-based frameworks using intraoral radiographs to enable quantitative measurement of alveolar bone levels. Lee et al. [11] proposed a system integrating multiple U-Net-based models to segment tooth structures, alveolar bone, and the cemento–enamel junction (CEJ), enabling computation of radiographic bone loss (RBL). Similarly, Chen et al. [14] introduced a deep learning-based framework for automatic tooth recognition and bone loss measurement from digital radiographs. Although these approaches enable quantitative estimation, they often rely on direct segmentation of thin anatomical structures such as the CEJ. Such representations are inherently sensitive to pixel-level noise and discontinuities, which may propagate into measurement variability, particularly when CEJ–bone intersections are used directly for RBL computation.

Hybrid approaches combining detection, segmentation, and geometric modeling have also been explored. For example, Chen et al. [13] employed Faster R-CNN for tooth detection, while other systems integrate localization and segmentation with geometric fitting to estimate tooth axes and bone loss. These frameworks represent an important step toward combining structural recognition with measurement; however, the geometric modeling components are often insufficiently constrained or not fully specified, particularly in handling anatomical variability such as curved roots or multi-root teeth. In broader medical imaging domains, hybrid approaches combining segmentation with deterministic geometric modeling have demonstrated improved robustness. For example, automated glenoid bone loss assessment integrates segmentation with geometric reconstruction to compute defect percentages [15]. Although developed for orthopedic imaging, such frameworks highlight the benefits of separating structural recognition from measurement logic, improving interpretability and reproducibility. In periodontal radiography, however, explicit modeling of CEJ endpoints, root-specific landmarks, and multi-root handling remains limited. Most existing systems rely either on segmentation-derived intersections or implicit feature representations. Consequently, error propagation from segmentation to quantitative measurement remains an underexplored methodological challenge.

In addition to deep learning-based approaches, machine learning-based estimation methods have also been investigated. Tsoromokos et al. [16] proposed a framework in which bone loss is inferred from extracted features using traditional machine learning techniques. While such approaches provide an alternative to deep learning, they typically lack explicit anatomical modeling and do not replicate the geometric measurement procedures used in clinical practice.

Prior to the widespread adoption of deep learning, classical image processing techniques were used for bone loss estimation. Lin et al. [17] applied edge detection and morphological operations to estimate alveolar bone levels. Although geometrically interpretable, these approaches were highly sensitive to image quality and required careful parameter tuning.

Accurate tooth and structure segmentation remains a fundamental component of most modern systems. Tuzoff et al. [18] demonstrated the use of CNNs for tooth detection and numbering in panoramic radiographs, while U-Net and its variants [12] have become standard architectures for biomedical image segmentation due to their ability to capture multi-scale contextual information. However, segmentation accuracy alone does not guar-

antee reliable clinical measurement, particularly when the target anatomical structures are narrow boundaries or require precise geometric interpretation.

Across the literature, several key limitations can be identified. First, CEJ localization is frequently modeled as a thin segmentation output, increasing sensitivity to noise and discontinuities. Second, bone loss measurement is often tightly coupled to segmentation quality, with limited geometric constraints to mitigate error propagation. Third, anatomical variability, particularly in multi-root teeth, is insufficiently addressed. Finally, many existing approaches lack explicit separation between structural recognition and measurement logic, limiting interpretability and reproducibility.

The proposed framework addresses these limitations by introducing a two-phase approach that separates deep learning-based segmentation from deterministic geometric landmark modeling. Instead of directly predicting thin anatomical structures, the method derives clinically relevant landmarks from structural relationships between segmented regions, enabling robust and interpretable measurement of CEJ bone crest distance and root length. Furthermore, the framework explicitly handles multi-root anatomical structures and provides a comprehensive evaluation of measurement accuracy bone loss measurement, supporting its clinical applicability.

3. Materials and Methods

In order to enable our system to measure bone loss levels, and after long discussions with the experienced dentists in the team from the School of Dentistry at the University of Jordan, we proposed implementing a two-phase computer tool to measure the bone loss. In the first phase, an AI model is trained to perform image segmentation on the radiographs to isolate the crowns, roots, third root and bone structures. In the second phase, we developed a low-level algorithm based on image processing and computer vision algorithms in order to mimic the process that dentists follow to measure the loss. Figure 1 shows the two phases of the developed procedure. In the following subsections, we will discuss the details of each phase.

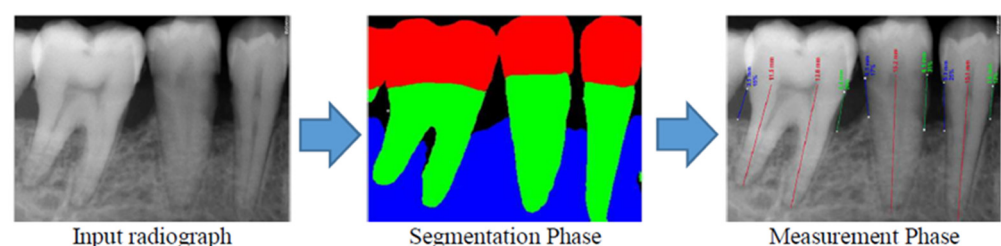


Figure 1. The two-phase procedure for bone loss measurement.

3.1. Phase 1: Image Segmentation

The first phase in this procedure is to segment the image into the four anatomical structures of interest: crown, root, third root and the alveolar bone. To achieve this, we have to develop a segmentation model that is capable of doing this accurately. Figure 2 shows the steps that we followed to obtain the segmentation model.



Figure 2. Steps for building and training a segmentation model for periapical images.

3.1.1. Dataset Collection and Labeling

In order to create the dataset, we collected 2127 radiographs from the Periodontology Dental Clinics at the University of Jordan, where the radiographs were taken under our supervision to make sure that they comply with the imaging standards. The labeling process was performed on the Supervisely platform (<https://supervise.ly/> (accessed on 25 June 2026)) by identifying the four different main regions in the radiographs. The labeled radiographs were reviewed by two researchers and independently verified by two board-certified periodontists, with discrepancies resolved by consensus. The total number of collected labels is listed in Table 1.

Table 1. Summary of the label count for the bone loss dataset prepared by the labeling team.

Label	Number of Labels
Bone	5048
Root	9783
Crown	9285
Third Root	371
Total	24,487

The selected labels were designed to facilitate the identification of key anatomical landmarks required for radiographic bone loss (RBL) measurement, particularly the cemento–enamel junction (CEJ) and the root structures. Accurate delineation of these components enables reliable estimation of the CEJ-to-bone crest distance and root length, which are essential for calculating RBL.

The Third Root label was introduced to explicitly represent the additional root branch commonly observed in multi-root molar teeth. During segmentation model training, it was treated as an independent class to improve the separation of anatomically distinct root structures and prevent merging adjacent roots into a single segmentation region. However, the measurement algorithm does not distinguish between primary and third-root structures; therefore, all root-related classes are merged into a unified root mask before performing geometric analysis. The inclusion of the Third Root label primarily improves segmentation accuracy in complex molar anatomies, enabling more reliable landmark extraction and root-length estimation. Figure 3 presents representative examples where the Third Root label facilitates the separation of molar roots. Without this label, multiple roots may be incorrectly segmented as a single structure, negatively affecting the subsequent measurement procedure.

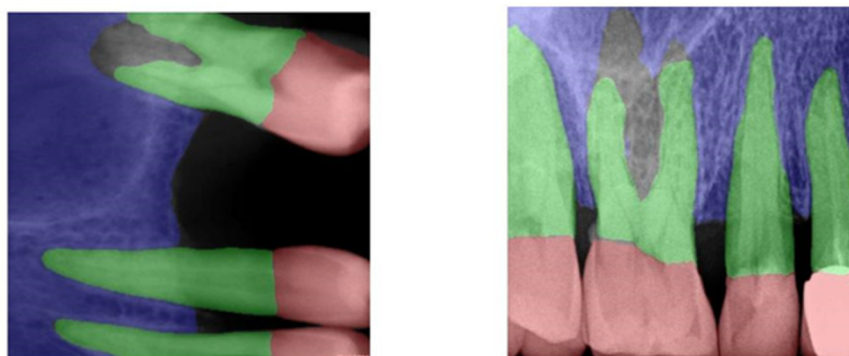


Figure 3. Examples demonstrating the use of the Third Root label to improve root separation in multi-root molars and facilitate accurate RBL measurement.

To evaluate the measurement phase of the proposed framework, a subset of 92 radiographs was independently analyzed by four board-certified periodontists. The experts evaluated CEJ endpoint locations, CEJ–bone crest distance, root length, and the resulting RBL values. To facilitate consistent measurements, the radiographs were displayed with an overlaid calibration grid consisting of 1 mm × 1 mm squares. Three board-certified periodontists performed independent measurements, while a fourth senior board-certified periodontist participated in a consensus review process to resolve discrepant cases and establish the final reference measurements used for algorithm evaluation.

3.1.2. Segmentation Dataset Preparation

After applying the data cleaning process and removing corrupted or unclear radiographs, a total of 2127 images were retained for segmentation model development and evaluation. The dataset was randomly divided at the image level into training, validation, and testing sets using an 80%, 10%, and 10% ratio, respectively. The training set was used for model optimization, the validation set was used for hyperparameter tuning and model selection, and the independent test set was reserved for final performance evaluation. Prior to training, all radiographs and corresponding segmentation masks were resized to 320 × 320 pixels. Data augmentation was applied only to the training set using random horizontal and vertical flipping to increase data variability and improve model generalization.

3.1.3. Segmentation Model Selection

There are several image segmentation models that can be used for radiograph segmentation. We studied several state-of-the-art models and selected a variety of them for training and testing. These models are Unet [12], TransUnet [19], DS-TransUnet [20], Segment Anything [21], and Yolo [22]. These models utilize state-of-the-art architectures, including encoder–decoder networks, Transformer-based models, and self-attention mechanisms.

3.1.4. Model Training and Hyperparameters Selection

We tried two different approaches to train the model. In the first approach, we trained the model for multi-class classification. In this case, the model is expected to predict the class for each pixel out of the five classes (root, bone, crown, third root and background). The background label denotes the region of the radiograph that does not contain root, third root, crown or bone. In the multi-class scenario, we have one model to predict the class of each pixel in the image.

In the second approach, we trained four binary classifiers, each of which was specialized in one class only; namely, the bone classifier, crown classifier, root classifier and third-root classifier. Each of these classifiers outputs a binary segmented image that labels pixels of the desired class by 1, while other pixels are labeled as 0.

For the multiclass classification scenario, the training labels are provided to the model as shown in Figure 4 where each class is colored/labeled with a certain color in a grayscale, where the color code for bone, root, third root and crown is 75, 150, 175 and 200 respectively. The background of the ground truth (unlabeled class) is colored black with color code 0.

For the binary classification scenario, the labels are specified such that for each classifier, the image has pixels colored as the designated class or background. Figure 5 shows an example of the binary masks for the labels in binary classification.

In the binary segmentation setting, each model independently predicts the mask of a specific anatomical structure. The resulting masks are subsequently merged to generate a unified segmentation map for the measurement phase. During this merging process, overlaps may occur in boundary regions where a pixel is predicted as belonging to multiple structures (e.g., Root and Bone). To resolve such conflicts, a priority-based strategy is

employed, where structures that are essential for accurate geometric measurement are assigned higher precedence. Specifically, the priority order is defined as Root > Third Root > Bone > Crown, ensuring preservation of root structures and boundaries, maintaining anatomical consistency, and preventing the loss of critical information required for subsequent geometric analysis and radiographic bone loss measurement.

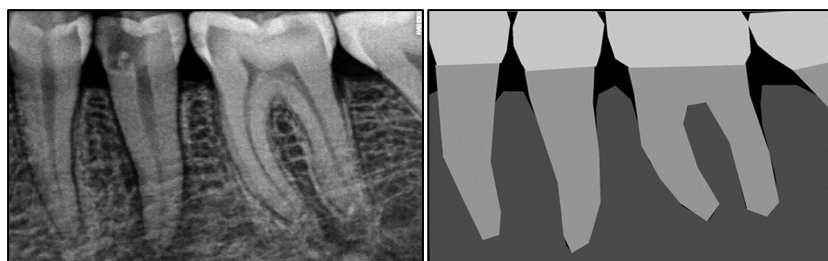


Figure 4. Original radiograph (left) and its labels (right) for multi-class segmentation.

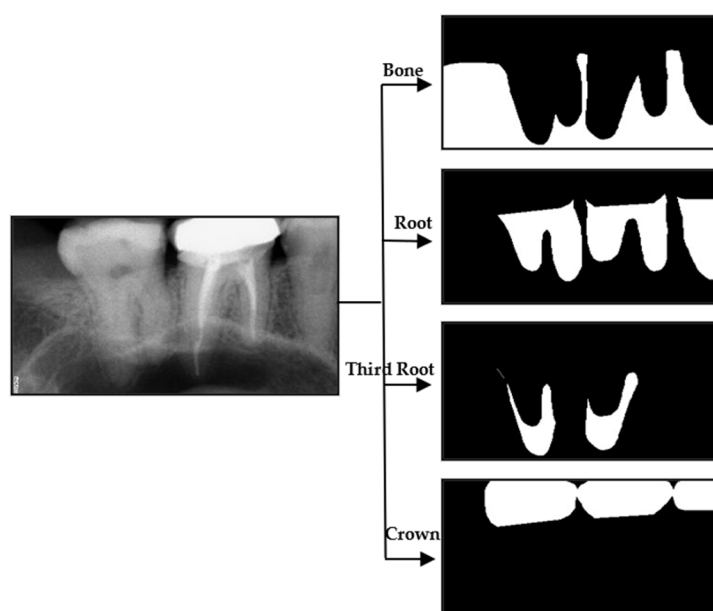


Figure 5. Original radiograph (left) and the binary labels for each class in binary classification.

Since the evaluated segmentation models are based on different architectures and training frameworks, a unified training configuration was not suitable for all models. Therefore, model-specific hyperparameter adjustments were performed to determine appropriate training settings, including the optimizer, learning rate, loss function, and architecture-specific parameters, when applicable. The final configuration of each model was selected based on validation performance, and the corresponding training settings are summarized in Table 2.

Table 2. Training configurations of the evaluated segmentation models.

Model	Input Size	Optimizer	Learning Rate	Loss Function	Model Specific Tuning
Unet	320×320	Adam	0.0001	BCE + Dice loss (binary); Cross-Entropy + Dice loss (multi-class)	Encoder: 4 layers, Decoder: 4 layers, Middle layer: 1, Dropout: 0.1

Table 2. Cont.

Model	Input Size	Optimizer	Learning Rate	Loss Function	Model Specific Tuning
TransUnet	320×320	SGD	0.01	Binary_Crossentropy	Filters [32, 64, 128, 256, 512, 1024]; ViT embedding = 768; 12 attention heads; 12 transformer layers
DS-TransUnet	320×320	AdamW	0.0003	Weighted CE + Dice + IoU (0.4, 0.4, 0.2)	Base channels = 32; ViT embedding = 256; 4 attention heads; 4 transformer layers
SAM	320×320	Adam	0.0001	Focal Dice loss + IoU loss	Default SAM-Med2D architecture
YOLOv5	320×320	SGD	0.01	YOLOv5 segmentation loss	Pre-trained: Yolo5x-seg model

3.2. Phase 2: Bone Loss Measurement

Radiographic bone loss (RBL) is commonly assessed by identifying anatomical landmarks on a dental radiograph, most notably the cemento–enamel junction (CEJ) and the alveolar bone crest. Figure 6 shows the main structures, landmarks and distances used in the measurement process when projected on the original radiograph and its corresponding segmented version. Clinically, the CEJ is located at the cervical region of the tooth, where the enamel covering the crown meets the cementum covering the root. The bone crest corresponds to the most coronal point of the supporting alveolar bone adjacent to the tooth. Once these landmarks are identified, the distance from the CEJ to the bone crest (d_1) is measured and normalized by the root length (d_2), which is the distance from the midpoint of the CEJ line to the root tip.

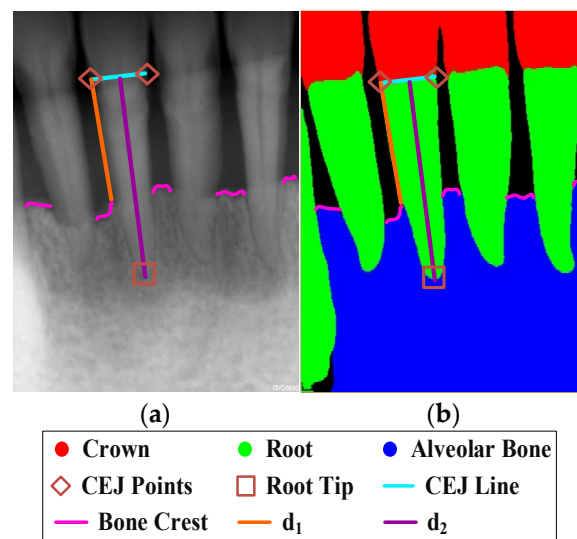


Figure 6. Landmarks used to measure bone loss. (a) Original radiograph. (b) Segmented radiograph.

Given that the distances are denoted by d_1 and d_2 , the radiographic bone loss (RBL) percentage is computed as

$$RBL = \frac{d_1 - 2 \text{ mm}}{d_2 - 2 \text{ mm}} \% \quad (1)$$

where the 2 mm correction accounts for the fact that, under healthy conditions, the alveolar crest is normally located approximately 1–2 mm apical to the CEJ [23].

All geometric measurements produced by the proposed framework were initially calculated in image pixel coordinates and subsequently converted to physical distances by multiplying the pixel distances by the radiograph spatial resolution provided by the imaging system (0.019 mm/pixel). This calibration procedure is applied to CEJ localization error, CEJ–bone crest distance, and root length measurements. Then, RBL is calculated using these calibrated measurements.

To automate radiographic bone loss measurement, we developed a deterministic geometric landmark extraction algorithm that emulates the clinical measurement procedure performed by periodontists. The complete workflow is summarized in Algorithm 1 and illustrated in Figure 7. Starting from the segmented crown, root, and bone structures, the algorithm identifies the CEJ endpoints, bone crest points, and root apex, and subsequently computes root length, CEJ–bone crest distance, and radiographic bone loss (RBL). The individual stages are described below.

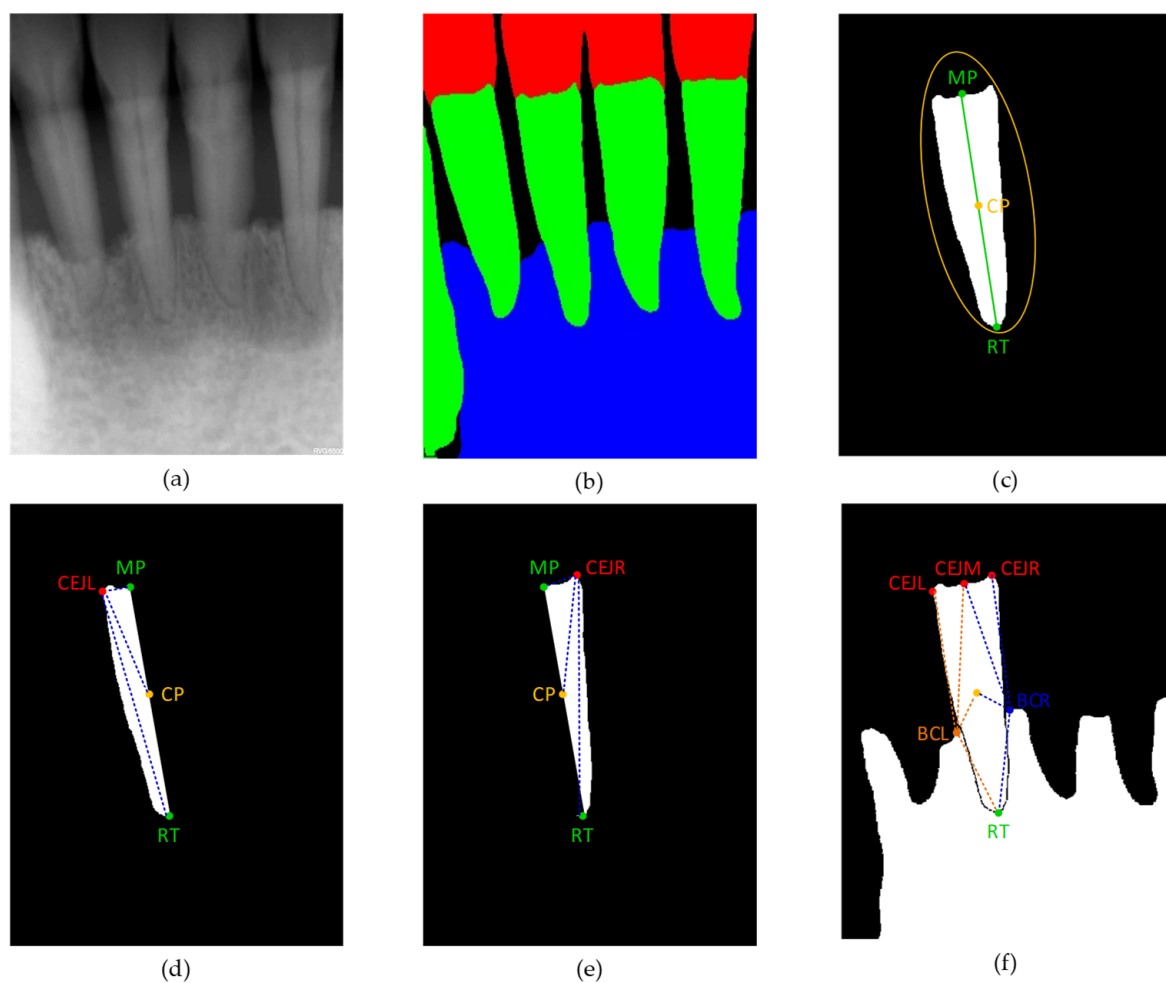


Figure 7. (a) Original radiograph; (b) segmented image; (c) root major axis and RT and MP points; (d) locating CEJL point; (e) locating CEJR point; (f) locating BCL and BCR points.

Algorithm 1. Geometric Landmark Extraction and Radiographic Bone Loss Measurement

Input: Segmented radiograph (I_s) containing crown, root, third-root, and bone masks

Output: CEJ endpoints ($(CEJL, CEJR)$), bone crest points ((BCL, BCR)), root apex (RT), root length (d_2), CEJ–bone distance (d_1), and RBL percentage

- 1: Extract the crown, root, third-root, and bone masks from (I_s)
 - 2: Merge root and third-root masks into a unified root mask
 - 3: Apply pixel-level conflict resolution using the priority order Root > Third Root > Bone > Crown
 - 4: Apply morphological filtering to remove small artifacts and separate adjacent structures
 - 5: Remove components smaller than 5% of the largest corresponding root or bone Component
 - 6: Extract the bone boundary from the filtered bone mask
 - 7: Identify connected components in the root mask
 - 8: **for** each root component (Ri) **do**
 - 9: Compute the centroid point (CPI) and principal orientation of (Ri)
 - 10: Estimate the root apex (RTi) and coronal midpoint (MPI) using the major axis
 - 11: **If** (Ri) touches the image border **then** exclude it from measurement
 - 12: **else**
 - 13: Perform scan-line discontinuity analysis to determine whether (Ri) is a molar root
 - 14: **if** (Ri) is classified as molar **then**
 - 15: Split (Ri) into left and right root branches using the main axis
 - 16: Compute branch-specific axes for the left and right branches
 - 17: Estimate ($CEJL$) from the upper boundary of the left branch
 - 18: Estimate ($CEJR$) from the upper boundary of the right branch
 - 19: **else**
 - 20: Split (Ri) into left and right regions using the main axis
 - 21: Estimate ($CEJL$) from the upper-left root boundary
 - 22: Estimate ($CEJR$) from the upper-right root boundary
 - 23: **end if**
 - 24: Compute the CEJ midpoint ($CEJM$) as the midpoint between ($CEJL$) and ($CEJR$)
 - 25: Refine ($CEJM$) by intersecting the line connecting ($CEJM$) and (RTi) with the root boundary
 - 26: Locate (BCL) and (BCR) on the bone boundary using ($CEJL$), ($CEJR$), ($CEJM$), (CPI), and (RTi) as anatomical anchors
 - 27: Compute the CEJ–bone crest distance (d_1)
 - 28: Compute the root length (d_2)
 - 29: Convert pixel distances to millimeters
 - 30: Compute RBL
 - 31: **Return** the extracted landmarks and measurements
 - 32: **end for**
-

As shown in Algorithm 1, the first phase of the measurement process focuses on preprocessing the segmented image and extracting the anatomical structures required for landmark localization. The segmented RGB image is first decomposed into three single-channel images corresponding to the crown, root, and bone regions. Morphological erosion is then applied to separate adjacent structures and suppress minor and small segmentation artifacts. Connected-component labeling is applied to the root mask to identify and label

individual roots. Small false-positive root components are removed if their area is less than the T_{root} threshold, which is set to 5% of the largest root component. A similar filtering step is applied to the bone mask. After filtering, the bone boundary is extracted by subtracting an eroded version of the bone mask from the original mask.

The threshold T_{root} was empirically determined through a sensitivity analysis by evaluating threshold values ranging from 0% to 50%. For each value, the number of remaining detected root structures after filtering was compared with the ground-truth number of roots identified by the clinical experts. As shown in Figure 8a, thresholds below approximately 3% retained small false-positive structures, while values above approximately 5% removed true root regions. Therefore, $T_{\text{root}} = 5\%$ was selected as a compromise between eliminating false detections and preserving valid anatomical structures.

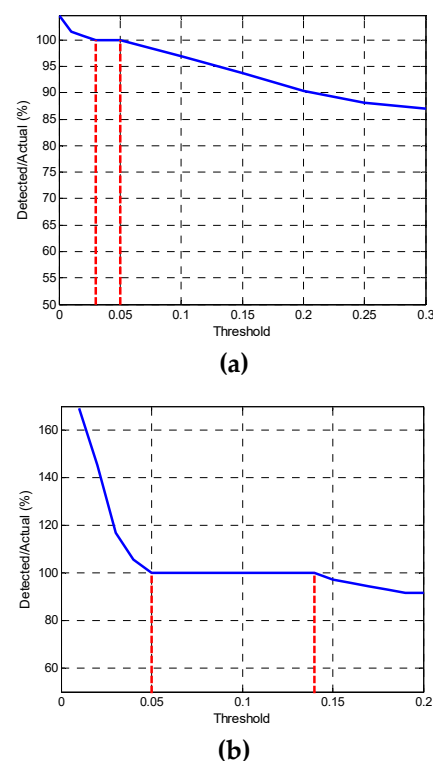


Figure 8. Sensitivity analysis of the selected geometric thresholds: (a) root filtering threshold T_{root} and (b) molar identification threshold T_{molar} . The blue curves represent the ratio of detected to actual roots in (a) and detected to actual molars in (b) as the threshold varies. The red dashed vertical lines indicate the range of threshold values for which the detected-to-actual ratio is 100%.

Because changing T_{root} directly affects the number of retained root components, the resulting measurements are not always directly comparable across different threshold values. Increasing the threshold may remove clinically valid small root structures, whereas decreasing it may retain false-positive fragments that do not correspond to true anatomical roots. Accordingly, threshold selection was based primarily on preserving clinically valid root structures while eliminating false-positive detections, rather than solely minimizing downstream measurement errors. Consequently, $T_{\text{root}} = 5\%$ was adopted as a stable operating point that balances the removal of false-positive components while preserving clinically valid root structures.

Each connected root component is then processed independently, as shown in Figure 7c. The root centroid (CP) and principal orientation of the root are computed by fitting an ellipse around the root. The resulting major axis is used to locate the root tip (RT) and the middle point on the top of the root (MP), which provides an initial estimate of the

middle point on the CEJ line. Roots on the image borders are flagged and excluded because they may correspond to incomplete teeth and may lead to unreliable measurements.

Because molar teeth typically contain two root branches, a dedicated procedure is used to detect molar roots. The algorithm performs a horizontal scan-line analysis on each root component. Starting from the apical end of the root and moving upward toward the coronal region, each image row intersecting the root is examined. For a given row, the leftmost and rightmost root pixels are identified to determine the horizontal span of the root. The actual number of root pixels in that row is then compared with this span. For a compact single root, the row is nearly continuous and the number of foreground pixels is close to the span length. In contrast, a molar root with two branches often contains an internal gap between the branches, causing the number of foreground pixels to be smaller than the horizontal span. In our implementation, if this discontinuity is observed in more than 10% of the root rows, the root is classified as a molar. We will refer to this threshold as T_{molar} .

The threshold T_{molar} was selected based on a sensitivity analysis by evaluating values between 0% and 20%. The number of detected molars was compared with the ground-truth molar annotations provided by the experts. Figure 8b shows that threshold values below 5% resulted in false identification of non-molar teeth as molars, while values above approximately 14% caused some true molars to be missed. Therefore, $T_{\text{molar}} = 10\%$ was selected because it lies within the stable operating range and provides a robust balance between false and missed molar detections.

Unlike continuous measurement parameters, T_{molar} determines the processing pathway by classifying teeth as molars or non-molars. Consequently, changing this threshold alters the subsequent geometric measurement procedure rather than simply perturbing the final measurements. Accordingly, T_{molar} was selected based on agreement with expert tooth-type annotations and the stability of the resulting classification, rather than solely on downstream measurement errors. The adopted value of $T_{\text{molar}} = 10\%$ provided the best balance between missed and incorrectly identified molars.

For non-molar roots, the major axis is used to split the root into left and right halves that are processed separately. To identify the left CEJ point (CEJL), the root boundary is cropped to retain only the upper-left region of the root (the region above the CP point). Among the remaining boundary points, the point maximizing the sum of its Euclidean distances from the RT, CP, and MP points is selected as CEJL, as shown in Figure 7d. The right CEJ point (CEJR) is determined the same way, except that the search is restricted to the upper-right boundary, as shown in Figure 7e. The midpoint of the CEJ line (CEJM) is then computed as the average of the coordinates of CEJL and CEJR. This midpoint is further refined by intersecting the line joining CEJM and RT with the root boundary.

For molar roots, the original root component is divided into left and right halves using a mask generated from the main axis passing through the centroid. Each half is then processed separately as an independent root branch. A secondary root axis is computed for each branch, and branch-specific coronal and apical points are obtained. In this case, the left branch is used to estimate CEJL, and the right branch is used to estimate CEJR.

Once the CEJ points have been determined, the algorithm identifies the left and right bone crest points on the bone boundary, as shown in Figure 7f. Instead of selecting only the nearest bone point to the CEJ, several anchor points are used to guide the search. Specifically, the CEJL, CEJR, CEJM, CP, and RT points are used to locate the bone crest points on the left and right sides of the root. For the left side of the root, the left bone crest point (BCL) is defined as the point on the bone boundary around the root that is closest to the CEJL, CEJM, CP, and RT points. The right bone crest point (BCR) is identified in a similar manner, but with respect to the CEJR point. Given the identified CEJ and bone crest

landmarks, the algorithm calculates the distances d_1 and d_2 required for the RBL calculation given in Equation (1).

4. Results

This section presents the experimental evaluation of the proposed framework for automated radiographic bone loss measurement. The evaluation is structured to reflect the two-phase design of the system. First, the performance of the deep learning-based segmentation models is analyzed to assess the accuracy of anatomical structure extraction. Next, the reliability of the geometric landmark localization and the resulting measurement accuracy are evaluated by comparing automated outputs with expert annotations. Both quantitative metrics and qualitative visualizations are provided to comprehensively demonstrate the accuracy, robustness, and clinical applicability of the proposed approach.

4.1. Segmentation Models Evaluation

All evaluated segmentation models were trained and tested on the collected dataset for bone loss assessment. Performance was evaluated using the Dice Similarity Coefficient (Dice) and Intersection over Union (IoU), which measure the overlap between the predicted segmentation and the ground truth. Dice emphasizes overall agreement between regions, while IoU provides a stricter measure of spatial overlap.

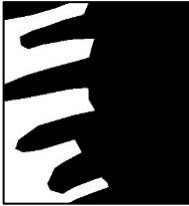
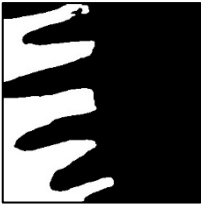
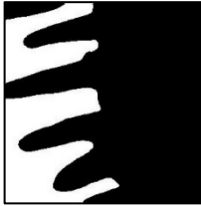
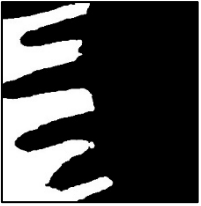
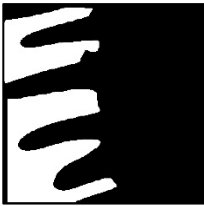
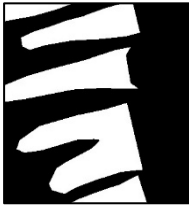
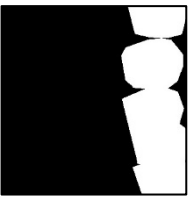
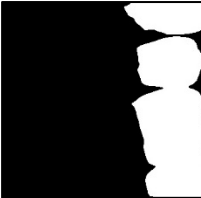
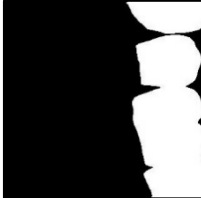
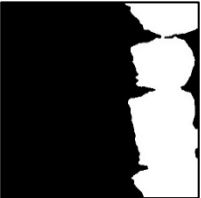
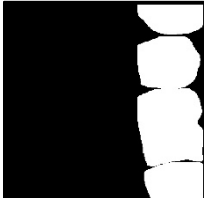
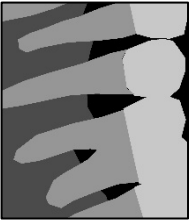
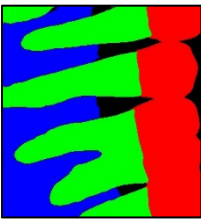
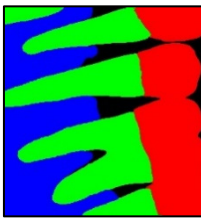
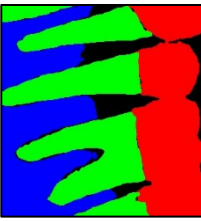
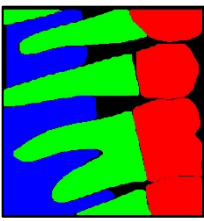
Quantitative performance results are summarized in Table 3, while qualitative comparisons are presented in Table 4. First, we explored multiclass vs. binary segmentation, and the results indicate that binary-class segmentation consistently outperforms multi-class segmentation, particularly for anatomically complex structures such as roots and alveolar bone. These results are listed in the first two rows of Table 3. Hence, we adopted binary classification approach for the remaining segmentation algorithms. This improvement can be attributed to the reduced class competition during training, allowing each model to specialize in detecting a single anatomical structure with higher precision.

Table 3. Evaluation results for segmentation models.

Model	Bone		Crown		Root	
	Dice	IoU	Dice	IoU	Dice	IoU
Unet-Multiclass	0.913	0.84	0.93	0.88	0.85	0.75
Unet-Binary	0.931	0.872	0.947	0.899	0.897	0.814
TransUnet-Binary	0.9112	0.8395	0.9162	0.8468	0.8581	0.7554
DS-TransUnet-Binary	0.9394	0.8877	0.9426	0.8927	0.9098	0.8244
SAM-Binary	0.9243	0.8620	0.9302	0.8713	0.8887	0.8063
YOLOv5-Binary	0.9313	0.8738	0.9223	0.8587	0.9008	0.8256

Among the evaluated models, DS-TransUNet achieved the best overall performance, demonstrating superior Dice and IoU scores across all anatomical classes. This performance gain can be attributed to its hybrid architecture, which combines convolutional feature extraction with transformer-based global context modeling. The ability to capture both local texture details and long-range spatial dependencies is particularly beneficial for dental radiographs, where anatomical boundaries are often subtle and affected by noise, contrast variations, and structural overlap.

Table 4. Sample segmentation result for different segmentation models.

Original Image	Ground Truth	Model (Predicted Images for Each Class)			
		TransUnet	DS_TransUnet	SAM	YOLOv5
	 Bone				
	 Root (Tooth)				
	 Crown				
Merged Result					
	Dice Score	90.47%	95.5%	94.96%	93.83%

A closer examination of class-wise performance reveals that bone and crown regions are segmented more accurately than root structures. This is expected, as root boundaries are often less distinct due to overlapping anatomical features and lower contrast in periapical radiographs. Despite this challenge, DS-TransUNet maintains strong performance on root segmentation, which is critical for accurate root length estimation and subsequent bone loss computation. The qualitative results in Table 4 further confirm these observations, showing that DS-TransUNet produces more consistent and anatomically coherent segmentation outputs compared to other models. In particular, it demonstrates improved continuity in bone boundaries and more precise delineation of root structures, both of which are essential for reliable landmark extraction.

It is important to point out that segmentation accuracy directly impacts the subsequent geometric measurement phase. Errors in bone or root segmentation can propagate inaccuracies in CEJ localization, bone crest detection, and root axis estimation. Therefore, the improved segmentation performance achieved by DS-TransUNet provides a strong

foundation for the geometric analysis phase and contributes to the overall accuracy and robustness of the proposed bone loss measurement framework.

To complement the quantitative segmentation results, Figure 9 presents representative segmentation outputs produced by the best-performing model. The selected examples cover varying levels of anatomical complexity, including single-root teeth, multi-root molars, third-root cases, complex root morphologies, and severe radiographic bone loss. These examples demonstrate the ability of the proposed segmentation framework to accurately delineate crown, root, third-root, and alveolar bone structures across diverse clinical scenarios.

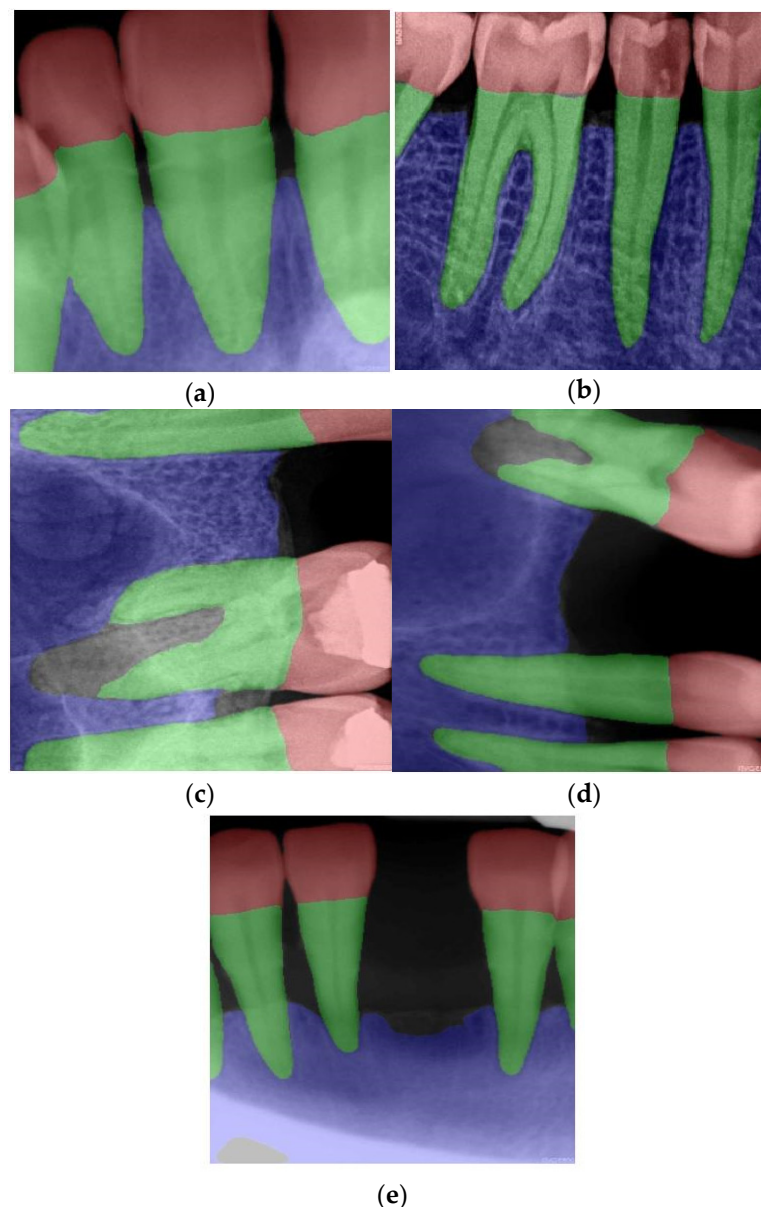


Figure 9. Representative segmentation results for cases of varying complexity: (a) single-root tooth, (b) multi-root molar, (c) third-root case, (d) complex root morphology, and (e) severe radiographic bone loss.

It is worth mentioning segmentation in the proposed framework is not an end goal but a foundational step for subsequent geometric analysis. Rather than optimizing segmentation solely for pixel-level accuracy, the objective is to achieve anatomically consistent delineations that support reliable landmark extraction and measurement. The reported segmentation performance demonstrates that the model provides sufficiently accurate

structural representations that are expected to enable robust downstream computation of clinically relevant metrics.

4.2. Geometric Landmark Localization and Measurement Accuracy

Accurate radiographic bone loss (RBL) estimation requires reliable localization of key anatomical landmarks, including the cemento–enamel junction (CEJ), the alveolar bone crest, and the root apex. Rather than predicting these landmarks using a dedicated landmark detection network, the proposed framework derives them deterministically from segmented anatomical structures through geometric analysis, as described in Section 3. This design enables explicit landmark localization while shifting the primary evaluation focus from isolated coordinate prediction toward clinically meaningful downstream geometric measurements derived from these landmarks.

Three key geometric variables were analyzed: CEJ endpoint localization (CEJL and CEJR), CEJ–bone crest distance, and root length. CEJ endpoint localization directly depends on segmentation quality, while CEJ–bone crest distance depends on accurate identification of CEJ endpoints (CEJL and CEJR) and bone crest points (BCL and BCR). Root length further depends on accurate localization of the CEJ midpoint and root apex. Therefore, differences between automated and expert-derived measurements provide both landmark-level and clinically meaningful assessments of anatomical localization accuracy.

Since the proposed framework is evaluated against expert-derived measurements, the reliability of the reference standard was first assessed. Independent measurements of root length and CEJ–bone crest distance were obtained for a subset of 92 radiographs, and inter-examiner agreement was quantified using the Intraclass Correlation Coefficient (ICC). Prior to consensus, excellent agreement was observed across all evaluated measurements, with ICC values of 0.944, 0.873, 0.932, and 0.945 for left root length, right root length, left CEJ–bone crest distance, and right CEJ–bone crest distance, respectively. Following expert review and consensus, agreement increased substantially, yielding ICC values of 0.988, 0.989, 0.995, and 0.996 for the corresponding measurements. These results indicate that the final expert measurements provide a highly reliable reference standard for evaluating the proposed automated measurement framework. Table 5 summarizes the inter-examiner agreement results.

Table 5. Inter-examiner agreement for expert measurements before and after consensus.

Measurement	ICC Before Consensus	ICC After Consensus
Left Root Length	0.944	0.988
Right Root Length	0.873	0.989
Left CEJ–Bone Distance	0.932	0.995
Right CEJ–Bone Distance	0.945	0.996

Having established the reliability of the expert reference measurements, a detailed comparison of automated and expert-derived radiographic bone loss measurements in terms of mean absolute error (MAE) and root mean squared error (RMSE) is presented in Table 6. The table summarizes the quantitative agreement between automated and four board-certified expert periodontists' annotations for key geometric variables, including CEJ Endpoint Localization (CEJL/CEJR), CEJ–bone crest distance, root length, and the resulting radiographic bone loss (RBL). CEJ localization error was computed as the average Euclidean discrepancy between automated and expert CEJ endpoints after compensating for marker visualization radius, yielding an MAE of 0.45 mm and an RMSE of 0.68 mm. These measurements directly reflect the accuracy of the extracted anatomical landmarks

and provide a clinically meaningful assessment of the system's ability to reproduce the manual measurement procedure used in periodontal diagnosis.

Table 6. Landmark localization and measurement error statistics between automated and expert annotations for CEJ localization, CEJ–bone crest distance, root length, and radiographic bone loss (RBL).

Measurement	MAE (mm)	RMSE (mm)
CEJR and CEJL Localization	0.45	0.68
CEJ–Bone distance	0.8109	1.1639
Root length	0.7128	1.1216
RBL (%)	5.8895	9.5697

The results indicate that the proposed framework achieves high landmark and geometric measurement accuracy, with CEJ localization MAE of 0.45 mm, sub-millimeter CEJ–bone and root-length errors, and low RBL percentage error. These results demonstrate that segmentation-derived geometric landmark estimation can reliably reproduce clinically relevant periodontal measurements. Since RBL percentage is directly derived from CEJ–bone and root-length measurements, the low downstream error further validates the robustness of the landmark derivation process.

To further illustrate the variability and distribution of these measurement differences, Figure 10 presents the distribution of Euclidean CEJ endpoint localization error between automated and expert annotations. The median CEJ localization error was notably low (0.065 mm), indicating that a substantial proportion of CEJ endpoints were localized with near-exact agreement. Although the interquartile range and upper whisker indicate moderate variability in some cases, the majority of localization errors remained within clinically acceptable limits, as further supported by an MAE of 0.45 mm and an RMSE of 0.68 mm. The observed upper spread is primarily attributable to a limited subset of challenging cases, whereas most samples remained tightly clustered near the expert-defined landmarks. These findings demonstrate that the proposed segmentation-derived geometric landmark framework can reliably localize CEJ endpoints while maintaining robustness across diverse radiographic conditions.

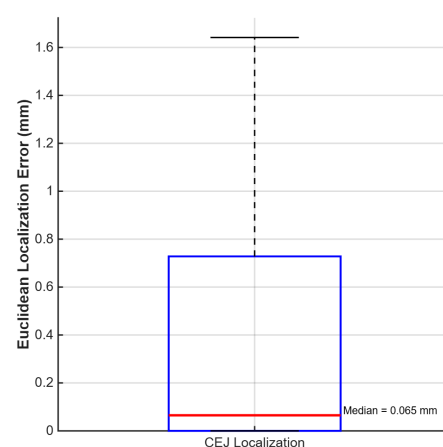


Figure 10. Distribution of CEJ endpoint localization error (mm) between automated and expert annotations. The blue box represents the interquartile range (IQR), the red horizontal line indicates the median localization error (0.065 mm), and the black dashed whiskers with horizontal caps represent the minimum and maximum non-outlier values.

Also, Figure 11 presents boxplots for CEJ–bone crest distance and root length. As shown, most errors are concentrated within a narrow range, with limited outliers, supporting the stability and robustness of the proposed geometric measurement approach. Unlike prior approaches that rely on implicit prediction of disease severity, the proposed framework provides explicit, verifiable measurements that can be directly interpreted in a clinical context.

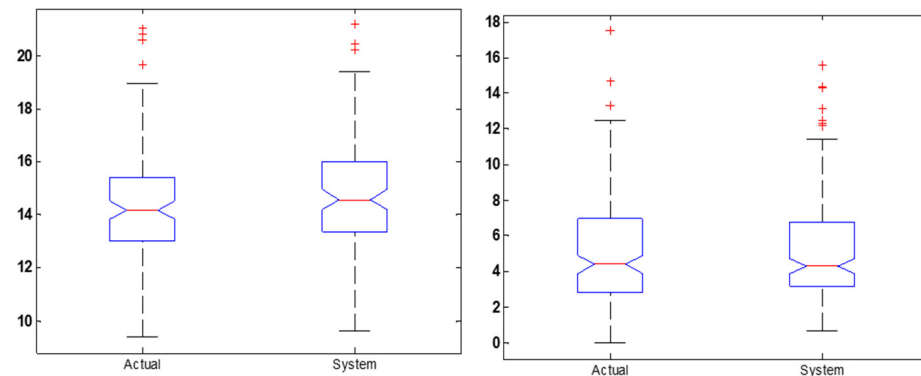


Figure 11. Boxplots showing the distribution of differences between automated and expert measurements for CEJ–bone crest distance (left) and root length (right). The blue boxes represent the interquartile range (IQR), the red horizontal lines indicate the median values, the black dashed whiskers extend to the minimum and maximum non-outlier values, and the red ‘+’ symbols denote outliers.

Since RBL is computed directly from these geometric measurements, the accuracy of the final RBL (%) estimation was also evaluated. As shown in Table 6, the automated system achieved an MAE of 5.89% for RBL estimation with an RMSE of 9.57%, demonstrating good agreement with expert measurements. The distribution of the RBL measurement differences is shown in Figure 12, where the majority of the errors remain within a relatively small range, indicating consistent bone loss estimation across the evaluated cases.

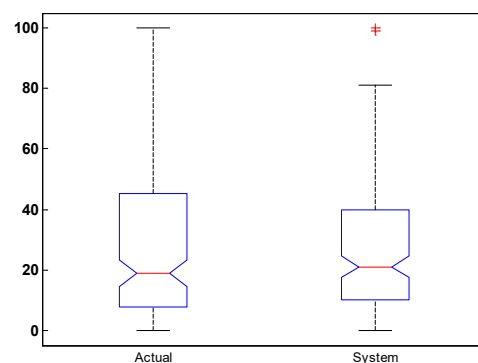


Figure 12. Boxplot showing the distribution of differences between automated and expert measurements of radiographic bone loss (RBL%). The blue boxes represent the interquartile range (IQR), the red horizontal lines indicate the median values, the black dashed whiskers extend to the minimum and maximum non-outlier values, and the red ‘+’ symbol indicates an outlier.

Importantly, unlike many prior deep learning approaches that attempt to directly segment the CEJ as a thin boundary structure [11], the proposed framework derives CEJ endpoints from geometric relationships between segmented crown and root regions. Thin-boundary segmentation is inherently vulnerable to pixel-level noise, discontinuities, and annotation uncertainty, which can propagate substantial measurement variability. By separating structural recognition from geometric landmark derivation, the proposed method

introduces anatomical consistency constraints that improve robustness, reproducibility, and interpretability.

To evaluate the agreement between automated and expert-derived radiographic bone loss (RBL) measurements, a Bland–Altman analysis was conducted. Figure 13 presents the Bland–Altman plot for RBL measurements. The mean difference (bias) between the two methods was -1.03% , indicating a slight tendency of the automated system to underestimate RBL values. The 95% limits of agreement ranged from -19.94% to $+17.89\%$, within which the majority of the measurements were distributed. As shown, most data points cluster around the zero-difference line, indicating strong agreement between the automated and expert measurements.

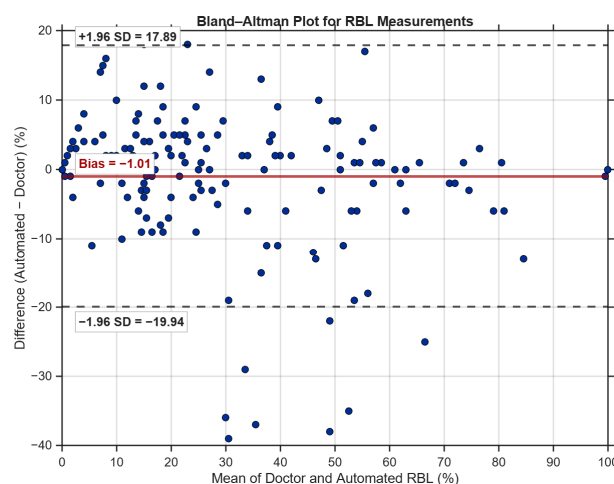


Figure 13. Bland–Altman plot showing the agreement between automated and expert-derived radiographic bone loss (RBL) measurements. The solid red line represents the mean difference (bias), while the dashed lines indicate the 95% limits of agreement.

To further investigate the effect of disease severity on measurement agreement, the Bland–Altman analysis was stratified according to the ground-truth radiographic bone loss (RBL) percentage. Cases were categorized as healthy (0% RBL), mild ($>0\%$ and $<15\%$ RBL), moderate ($15\text{--}33\%$ RBL), and severe ($>33\%$ RBL). The resulting bias and limits of agreement are summarized in Table 7.

Table 7. Stratified Bland–Altman analysis of automated and expert-derived radiographic bone loss (RBL) measurements.

Group	Number of Samples	Bias%	SD (%)	Lower LOA (%)	Upper LOA (%)
Healthy	32	2.69	4.70	-6.53	11.90
Mild	41	3.37	5.68	-7.76	14.49
Moderate	49	-0.73	6.41	-13.29	11.82
Severe	55	-4.24	10.71	-25.22	16.75

The stratified analysis reveals that agreement was strongest in healthy and mild cases, which exhibited narrow limits of agreement and relatively small bias. Moderate cases showed the smallest systematic bias (-0.73%), indicating excellent agreement between automated and expert measurements. In contrast, severe cases demonstrated larger variability and wider limits of agreement, with a tendency toward underestimation of bone loss. This increase in variability is expected given the greater anatomical complexity and reduced landmark visibility associated with advanced periodontal destruction. Nevertheless, the observed bias remained substantially smaller than the overall measurement range,

indicating that the proposed framework maintains clinically meaningful agreement across all severity levels.

Although both the overall and severity-stratified analyses indicate larger deviations in some cases, particularly for severe bone loss, the low average error and small overall bias indicate good agreement between automated and expert measurements. The observed variability is mainly influenced by a limited number of challenging cases, whereas most measurements remain closely distributed, as reflected by the low MAE and boxplot distributions. While point-level CEJ localization errors may contribute to measurement variability, the deterministic geometric framework reduces the propagation of these errors, resulting in improved consistency in CEJ–bone distance, root length, and final RBL estimation. Overall, the results suggest that the proposed framework can reproduce the clinical measurement procedure with reasonable accuracy and has the potential to support objective radiographic bone loss assessment and disease staging.

4.3. Tooth Type-Aware Analysis and Its Impact on Measurement Accuracy

Correct identification of molar teeth is important for the proposed measurement framework because molars often contain multiple root branches that require special handling during landmark localization. To evaluate the performance of the molar identification step, the predicted tooth type was compared with expert annotations.

Table 8 presents the confusion matrix for molar classification. The proposed method achieved an accuracy of 90.97%, with a precision of 0.93, recall of 0.86, and an F1-score of 0.88. Most teeth were correctly identified, with only a small number of molars misclassified as non-molars. These results indicate that the algorithm reliably distinguishes between molar and non-molar teeth, enabling appropriate handling of multi-root structures during the bone loss measurement process.

Table 8. Confusion matrix for molar identification.

		Predicted	
		Not Molar	Molar
Actual	Not Molar	99	1
	Molar	12	32

Table 9 presents a stratified analysis of measurement performance across molar and non-molar teeth to evaluate the impact of anatomical complexity on the proposed framework. The results indicate that the CEJ–bone crest distance estimation remains highly consistent across both groups, with similar MAE values for molars (0.8178 mm) and non-molars (0.8066 mm), suggesting robust CEJ and bone crest localization regardless of tooth type. However, root length estimation shows higher error in molars (0.8707 mm) compared with non-molars (0.5338 mm), while the resulting RBL estimation error was also slightly higher in molars (6.7463%) than in non-molars (5.3860%). These differences are likely due to the increased structural complexity of molars, including multiple roots, overlapping anatomical structures, and greater variation in root morphology. However, statistical significance tests for molar and non-molar differences using Mann-Whitney U tests indicated that the differences in CEJ–Bone crest distance, root length and RBL measurement errors between molar and non-molar teeth were not statistically significant ($p = 0.550$, 0.100 , and 0.814 , respectively). These findings highlight that while the proposed segmentation–geometry framework generalizes well across tooth categories despite the increased anatomical complexity of molars. Nevertheless, the consistently higher errors observed for molars indicate that multi-root anatomy remains more challenging for geometric

measurement, supporting the inclusion of dedicated third-root labeling and root-separation strategies in the segmentation.

Table 9. MAE comparison of CEJ–bone distance, root length, and RBL (%) between molar and non-molar teeth.

Measurement	Molar		Non-Molar		<i>p</i> -Value
	MAE	RMSE	MAE	RMSE	
CEJ–Bone distance	0.8178	1.1914	0.8066	1.1466	0.550
Root length	0.8707	1.3434	0.5338	0.7989	0.100
RBL (%)	6.7463	10.8724	5.3860	8.7138	0.814

4.4. Sample Measurement Results

Figure 14 presents representative measurement examples from eight radiographic images to qualitatively illustrate the performance of the proposed framework across diverse clinical scenarios. The selected cases include both molar and non-molar teeth from different radiographs, providing visual and numerical examples of the system’s ability to generalize across varying tooth anatomies and image characteristics. By including multiple tooth types and imaging conditions, the presented results demonstrate how the framework performs in practical cases, ranging from simpler single-root structures to more anatomically complex multi-root molars. These examples complement the aggregate quantitative results by offering case-level insight into measurement consistency, landmark localization quality, and RBL estimation across heterogeneous dental structures.

Overall, the experimental results demonstrate that the proposed framework achieves accurate anatomical localization, reliable quantitative measurements, and robust RBL percentages. These findings highlight the effectiveness of the proposed approach for automated radiographic bone loss assessment.

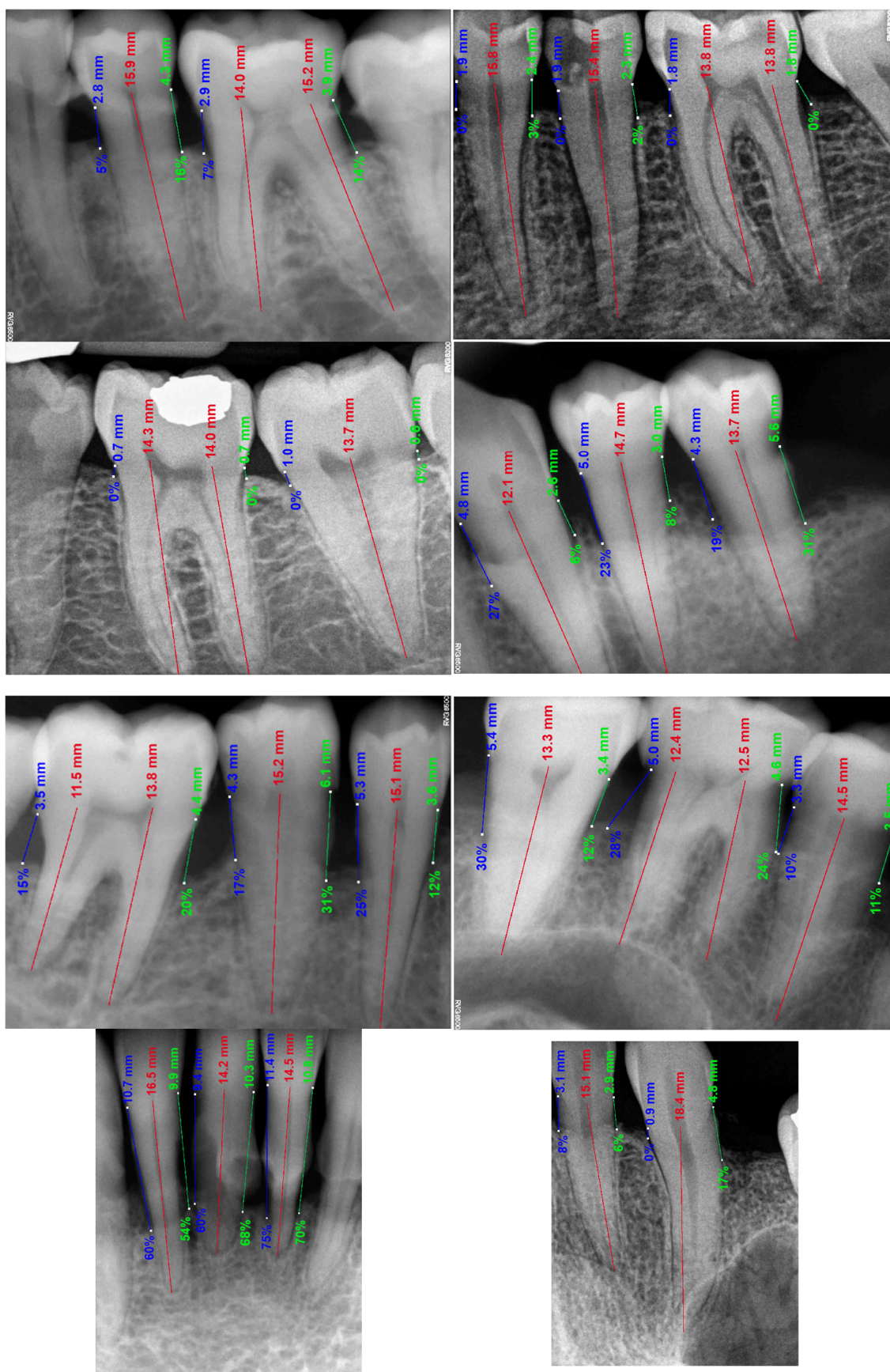


Figure 14. Representative examples of automated radiographic bone loss (RBL) measurements on periapical radiographs. The figure illustrates cases for both single-root teeth and multi-root molars. Detected anatomical landmarks (CEJ, bone crest, and root apex), measurement lines and computed RBL percentages overlaid on the radiographs.

5. Discussion

The findings of this study demonstrate that accurate automated radiographic bone loss assessment depends not only on robust anatomical segmentation, but also on the reliability of the downstream measurement strategy used to translate segmented structures into clinically meaningful values. While deep learning models are highly effective for structural recognition, direct end-to-end prediction of periodontal measurements may obscure error sources and reduce interpretability. The proposed two-phase framework addresses this limitation by explicitly separating anatomical segmentation from deterministic geometric analysis, thereby aligning the measurement process more closely with established clinical principles while enabling independent evaluation of structural and measurement reliability.

A central advantage of this design is its reliance on anatomical relationships rather than direct prediction of isolated landmarks or thin anatomical boundaries. Dental radiographs frequently contain noise, low contrast, overlapping roots, and discontinuous CEJ boundaries, all of which can degrade direct landmark prediction accuracy. By deriving CEJ points, bone crest, and root apex relationships through structural segmentation and geometric constraints, the proposed framework reduces sensitivity to local segmentation imperfections and improves measurement consistency in anatomically challenging cases. This is particularly relevant in molars, where multi-root complexity and structural overlap can substantially increase measurement variability.

From a clinical perspective, explicit geometric decomposition also improves transparency. Rather than functioning solely as a diagnostic classifier, the framework provides intermediate anatomical structures and measurement steps that can be visually inspected and verified by clinicians. This may improve trust, facilitate workflow integration, and support practical adoption in clinical environments where explainability is essential. Furthermore, separate reporting of CEJ localization, CEJ–bone distance, root length, and final RBL errors enables a clearer understanding of where measurement inaccuracies originate, which may be difficult to achieve in black-box architectures.

Earlier deep learning systems for periodontal assessment, including classification-oriented approaches such as Krois et al. [7] and Kim et al. [8], primarily focused on disease detection without explicit anatomical measurement modeling. While these systems demonstrated diagnostic utility, they functioned primarily as screening tools rather than quantitative radiographic measurement systems. More advanced approaches, such as Chang et al. [9], incorporated bone-loss-oriented segmentation but continued to rely largely on implicit feature learning. Segmentation-focused frameworks such as Lee et al. [11] improved quantitative RBL estimation, yet direct segmentation of thin CEJ structures remains sensitive to discontinuities and local noise. Hybrid detection-segmentation approaches, including Chen et al. [13,14], further advanced the field but often lacked anatomically constrained measurement logic, potentially limiting reproducibility in structurally complex teeth.

Despite the promising measurement accuracy and low systematic bias observed in the Bland-Altman analysis, the observed limits of agreement indicate that individual cases may still exhibit differences between automated and expert measurements. These variations may be associated with challenging anatomical conditions, severe bone loss patterns, or cases with less distinct anatomical boundaries. Furthermore, the current validation was performed using radiographs acquired from a single clinical center under a standardized imaging protocol. Future studies involving external multi-center datasets with broader variations in patient populations and imaging conditions are required to further evaluate the generalizability of the proposed framework.

5.1. Comparison with Related Work

To better contextualize the proposed framework within the current literature, Table 10 summarizes representative automated periodontal bone loss systems with an emphasis on anatomical modeling strategy, physical measurement outputs, clinical interpretability and quantitative performance. Because prior studies report heterogeneous metrics—including segmentation overlap, pixel-space landmark error, percentage agreement, and ICC—direct numerical comparison should be interpreted cautiously. Therefore, the comparison extends beyond quantitative performance to examine methodological differences in anatomical modeling, the availability of clinically meaningful millimeter-scale measurements, and the interpretability of the resulting measurement process. The following discussion further contrasts the proposed framework with representative studies from these complementary perspectives.

Table 10. Quantitative comparison of representative automated periodontal bone loss measurement studies.

Study	Anatomical Modeling	Physical Measurement Output	Clinical Interpretability	Best Reported Result
Danks et al. (2021) [24]	Landmark localization on pre-separated individual tooth image	PBL (%)	Moderate; reports final PBL but limited decomposition of intermediate anatomical measurements	Mean PBL error: 6.82%
Lee et al. (2022) [11]	Segmentation (tooth, bone, and thin-line CEJ segmentation with geometric analysis)	CEJ-Bone distance, root length, RBL (%)	Good; segmentation-based measurements, but no independent evaluation of CEJ localization or explicit molar-specific processing	CEJ-line DSC: 0.91
Tsoromokos et al. (2022) [16]	ML regression	ABL (%)	Moderate; estimates bone loss without explicit anatomical landmark modeling	ICC: 0.601–0.763
Ameli et al. (2024) [25]	Bone-loss segmentation + YOLO apex detection with CEJ center approximation	Root length, BL (%)	Good; evaluates final measurements but does not separately assess intermediate anatomical localization	ICC: 0.959
Proposed Framework	Multi-class anatomical segmentation (crown, root, bone, third root) + deterministic geometric landmark extraction	CEJ localization (mm), CEJ–bone distance (mm), root length (mm), RBL (%)	High; anatomically interpretable measurements with explicit evaluation of each measurement component and dedicated handling of multi-root teeth	RBL MAE: 5.89%; CEJ MAE: 0.45 mm; Root MAE: 0.71 mm

Danks et al. [24] provide one of the strongest methodological comparators due to their landmark-based PBL estimation framework. Their method operates on pre-separated and individually cropped tooth images, requiring each tooth to be isolated before landmark localization and periodontal bone loss estimation. Bone loss is then computed using pixel-based geometric measurements derived from the predicted landmarks. In contrast, the proposed framework operates directly on the segmented anatomical structures extracted

from the original radiograph, without requiring prior tooth cropping. Furthermore, rather than relying solely on landmark prediction, the proposed method performs multi-class anatomical segmentation of the crown, root, bone, and third root, followed by deterministic geometric landmark extraction to identify CEJ endpoints, bone crest, and root apices. Unlike Danks et al., who primarily report the final bone loss percentage, our framework provides clinically meaningful measurements in physical millimeter units, including CEJ localization accuracy, CEJ–bone crest distance, and root length, enabling detailed analysis of how anatomical localization errors propagate to the final radiographic bone loss estimation. This anatomical decomposition improves the clinical interpretability of the framework by allowing clinicians to identify the source of measurement errors rather than relying solely on the final percentage prediction.

Lee et al. [11] similarly demonstrated that segmentation-based frameworks can achieve clinician-level bone loss estimation using periapical radiographs. Their method segments the CEJ line and bone-loss region, then estimates root length using the center of the segmented CEJ together with the detected root apex. Although this produces accurate bone loss measurements, the framework does not explicitly model individual anatomical structures such as crowns, roots, or third roots, nor does it incorporate dedicated processing for multi-root molars. In contrast, the proposed framework performs explicit anatomical segmentation followed by geometry-based landmark extraction using CEJ endpoints, bone crest, and root apices. This enables direct physical measurements in millimeters for CEJ localization, CEJ–bone crest distance, and root length, while also supporting dedicated handling of complex molar morphology. Consequently, the proposed approach provides greater clinical interpretability by decomposing the overall measurement error into anatomically meaningful components rather than relying solely on the final radiographic bone loss estimate.

Ameli et al. [12] also proposed an automated framework for radiographic bone loss assessment on periapical radiographs by combining U-Net-based bone-loss segmentation with YOLO-based apex detection. Their anatomical modeling is primarily centered on the segmented bone-loss region, with the CEJ approximated as a single center point extracted from the segmented mask and root length measured between this point and the detected apex. Unlike the proposed method, no explicit modeling of crowns, roots, third roots, or dedicated molar analysis is performed. In contrast, the proposed framework explicitly segments multiple anatomical structures and extracts clinically relevant landmarks, including CEJ endpoints, bone crest, and root apices, to perform deterministic geometric measurements. In addition to estimating radiographic bone loss, our framework quantitatively evaluates CEJ localization, CEJ–bone crest distance, and root length in physical millimeter units, providing a comprehensive assessment of measurement reliability. This anatomical and geometric decomposition enhances clinical interpretability by allowing clinicians to identify the origin of measurement discrepancies and better understand framework performance across different tooth morphologies.

Overall, while previous automated periodontal assessment systems have significantly advanced RBL estimation, they differ substantially in their anatomical modeling strategies, measurement outputs, and levels of clinical interpretability. Existing approaches often rely on direct landmark prediction, object detection, or thin-line CEJ segmentation to estimate bone loss. In contrast, the proposed framework adopts an anatomy-driven strategy by first segmenting complete dental structures and subsequently applying deterministic geometric analysis to derive CEJ, alveolar bone crest, and root apex relationships. This decomposition of the measurement process improves methodological transparency and enables the reporting of clinically meaningful physical measurements in millimeters, in addition to RBL percentages. To the best of our knowledge, few prior studies provide

a comprehensive measurement-oriented evaluation that jointly reports CEJ localization accuracy, CEJ–bone crest distance, root-length estimation, and final radiographic bone loss error within a unified and clinically interpretable framework.

5.2. Limitations and Future Directions

Despite the promising performance of the proposed framework, several limitations should be acknowledged. Although CEJ–bone distance measurements showed comparable accuracy between molar and non-molar teeth, higher errors were observed for molars in root length and final RBL percentage estimation. This is primarily due to the increased anatomical complexity of molars, including multiple root branches, overlapping structures, and variations in root morphology, which complicate accurate root apex localization and subsequent geometric measurement. Furthermore, a limited number of outlier cases were observed in the Bland–Altman analysis, particularly in severe bone loss conditions where the CEJ, alveolar bone crest, or root boundaries become less distinguishable. Future work will focus on improving robustness in these challenging scenarios by incorporating larger and more diverse datasets, enhancing segmentation of complex anatomical structures, and integrating uncertainty estimation methods to identify measurements that may require additional clinical verification. In addition, the proposed framework will be extended to detect additional periodontal findings, such as furcation involvement, vertical bone defects, and other clinically relevant periodontal abnormalities. Integrating these findings with automated radiographic bone loss measurement could provide clinicians with a more comprehensive assessment of periodontal status and support more informed diagnosis and treatment planning.

6. Conclusions

This study presents a measurement-oriented framework for automated radiographic bone loss assessment on periapical radiographs by explicitly separating anatomical segmentation from deterministic geometric analysis. Unlike conventional end-to-end or segmentation-based approaches, the proposed method derives clinically relevant landmarks—including CEJ endpoints, bone crest, and root apex—through structured anatomical relationships, enabling reproducible and interpretable measurement of CEJ–bone distance, root length, and radiographic bone loss (RBL). The results demonstrate that reliable bone loss estimation can be achieved without direct landmark prediction while maintaining robustness in anatomically complex cases such as multi-root molars. By reporting measurement errors in physical units (mm) alongside percentage-based RBL, the framework provides clinically meaningful evaluation beyond pixel-space metrics or aggregate agreement measures. Overall, the proposed approach highlights the importance of anatomically constrained measurement modeling in improving transparency, robustness, and clinical interpretability, supporting its potential for integration into AI-assisted periodontal assessment systems.

Author Contributions: Conceptualization, M.A.-M., I.J. and O.A.-K.; methodology, M.A.-M., I.J., O.A.-K. and S.Z.; software, M.A.-M., I.J. and S.A.-A.; validation, M.A.-M., I.J., O.A.-K., S.Z., M.F. and S.A.-A.; formal analysis, M.A.-M., I.J., O.A.-K., S.Z. and S.A.-A.; investigation, M.A.-M., I.J., O.A.-K., S.Z., M.F. and S.A.-A.; resources, M.A.-M., I.J. and O.A.-K.; data curation, M.A.-M., O.A.-K., S.Z., M.F. and S.A.-A.; writing—original draft preparation, M.A.-M., I.J. and O.A.-K.; writing—review and editing, M.A.-M., I.J. and O.A.-K.; visualization, M.A.-M., I.J., O.A.-K., S.Z. and S.A.-A.; supervision, M.A.-M., I.J. and O.A.-K.; project administration, M.A.-M., I.J. and O.A.-K.; funding acquisition, M.A.-M. and O.A.-K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Jordanian Scientific Research and Innovation Support Fund [Project Number IE/1/39/2019] and the Deanship of Scientific Research at the University of Jordan.

Institutional Review Board Statement: This prospective clinical–computational study was conducted at Jordan University Hospital between 2022 and 2024. Ethical approval was obtained from the Institutional Review Board of Jordan University Hospital (Approval No. 2021/252). The study was conducted in accordance with the Declaration of Helsinki.

Informed Consent Statement: Written informed consent was obtained from all participants included in the study. The study was reported in compliance with the STARD guidelines and recommended standards for artificial intelligence research in dental imaging.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available because they contain clinical data that are subject to institutional ethical and privacy restrictions.

Acknowledgments: During the preparation of this work, the authors used ChatGPT (GPT-5.5, OpenAI) to help rephrase some paragraphs for clarity. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
DL	Deep Learning
CNN	Convolutional Neural Network
RBL	Radiographic Bone Loss
PBL	Periodontal Bone Loss
ABL	Alveolar Bone Loss
CEJ	Cemento–Enamel Junction
CEJL	Cemento–Enamel Junction (Left)
CEJR	Cemento–Enamel Junction (Right)
CEJM	Cemento–Enamel Junction Midpoint
BCL	Bone Crest (Left)
BCR	Bone Crest (Right)
RT	Root Tip
CP	Centroid Point
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
IoU	Intersection over Union
DSC	Dice Similarity Coefficient

References

1. Pihlstrom, B.L.; Michalowicz, B.S.; Johnson, N.W. Periodontal diseases. *Lancet* **2005**, *366*, 1809–1820. [[CrossRef](#)] [[PubMed](#)]
2. Sanz, M.; Marco del Castillo, A.; Jepsen, S.; Gonzalez-Juanatey, J.R.; D’Aiuto, F.; Bouchard, P.; Chapple, I.; Dietrich, T.; Gotsman, I.; Graziani, F.; et al. Periodontitis and systemic diseases: Consensus report. *J. Clin. Periodontol.* **2020**, *47*, 268–288. [[PubMed](#)]
3. Akesson, L.; Hakansson, J.; Rohlin, M. Comparison of panoramic and intraoral radiography for marginal bone level measurement. *J. Clin. Periodontol.* **1992**, *19*, 326–332. [[PubMed](#)]
4. Pepelassi, E.A.; Diamanti-Kipioti, A. Selection of the most accurate radiographic method for assessment of periodontal osseous destruction. *J. Clin. Periodontol.* **1997**, *24*, 557–567. [[PubMed](#)]
5. Hellen-Halme, K.; Lith, A.; Shi, X.Q. Reliability of marginal bone level measurements on digital panoramic and intraoral radiographs. *Oral Radiol.* **2020**, *36*, 135–140. [[PubMed](#)]
6. Tonetti, M.S.; Greenwell, H.; Kornman, K.S. Staging and grading of periodontitis: Framework and proposal of a new classification. *J. Clin. Periodontol.* **2018**, *45*, S149–S161. [[CrossRef](#)] [[PubMed](#)]
7. Krois, J.; Ekert, T.; Meinhold, L.; Golla, T.; Kharbot, B.; Wittemeier, A.; Dörfer, C.; Schwendicke, F. Deep learning for the radiographic detection of periodontal bone loss. *Sci. Rep.* **2019**, *9*, 8495. [[CrossRef](#)] [[PubMed](#)]

8. Kim, J.; Lee, H.S.; Song, I.S.; Jung, K.H. DeNTNet: Deep neural transfer network for the detection of periodontal bone loss using panoramic dental radiographs. *Sci. Rep.* **2019**, *9*, 17615. [[CrossRef](#)] [[PubMed](#)]
9. Chang, H.-J.; Lee, S.-J.; Yong, T.-H.; Shin, N.-Y.; Jang, B.-G.; Kim, J.-E.; Huh, K.-H.; Lee, S.-S.; Heo, M.-S.; Choi, S.-C.; et al. Deep learning hybrid method to automatically diagnose periodontal bone loss and stage periodontitis. *Sci. Rep.* **2020**, *10*, 7531. [[CrossRef](#)] [[PubMed](#)]
10. Kong, Z.; Ouyang, H.; Cao, Y.; Huang, T.; Ahn, E.; Zhang, M.; Liu, H. Automated periodontitis bone loss diagnosis in panoramic radiographs using a bespoke two-stage detector. *Comput. Biol. Med.* **2023**, *152*, 106374. [[PubMed](#)]
11. Lee, C.; Kabir, T.; Nelson, J.; Sheng, S.; Meng, H.; Van Dyke, T.E.; Walji, M.F.; Jiang, X.; Shams, S. Use of the deep learning approach to measure alveolar bone level. *J. Clin. Periodontol.* **2022**, *49*, 260–269. [[PubMed](#)]
12. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), Munich, Germany, 5–9 October 2015; pp. 234–241.
13. Chen, H.; Zhang, K.; Lyu, P.; Li, H.; Zhang, L.; Wu, J.; Lee, C.-H. A deep learning approach to automatic teeth detection and numbering based on object detection. *Sci. Rep.* **2019**, *9*, 3840. [[CrossRef](#)] [[PubMed](#)]
14. Chen, C.C.; Wu, Y.F.; Aung, L.M.; Lin, J.C.; Ngo, S.T.; Su, J.N.; Lin, Y.M.; Chang, W.J. Automatic recognition of teeth and periodontal bone loss measurement in digital radiographs using deep-learning artificial intelligence. *J. Dent. Sci.* **2023**, *18*, 1301–1309. [[CrossRef](#)] [[PubMed](#)]
15. Liu, Z.; Gu, H.; Li, Q.; Fox, M.; Levin, J.M.; Mazurowski, M.A.; Lau, B.C. Automated Glenoid Bone Loss Measurement and Segmentation in CT Scans for Pre-Operative Planning in Shoulder Instability. *arXiv* **2025**, arXiv:2511.14083. [[CrossRef](#)]
16. Tsoromokos, N.; Parinussa, S.; Claessen, F.; Moin, D.A.; Loos, B.G. Estimation of alveolar bone loss in periodontitis using machine learning. *Int. Dent. J.* **2022**, *72*, 621–627. [[CrossRef](#)] [[PubMed](#)]
17. Lin, P.; Huang, P.; Huang, P. Automatic methods for alveolar bone loss degree measurement in periapical radiographs. *Comput. Methods Programs Biomed.* **2017**, *148*, 1–11. [[CrossRef](#)] [[PubMed](#)]
18. Tuzoff, D.V.; Tuzova, L.N.; Bornstein, M.M.; Krasnov, A.S.; Kharchenko, M.A.; Nikolenko, S.I.; Sveshnikov, M.M.; Bednenko, G.B. Tooth detection and numbering in panoramic radiographs using CNNs. *Dentomaxillofac. Radiol.* **2019**, *48*, 20180051. [[PubMed](#)]
19. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. TransUNet: Transformers make strong encoders for medical image segmentation. *arXiv* **2021**, arXiv:2102.04306.
20. Lin, A.; Chen, B.; Xu, J.; Zhang, Z.; Lu, G.; Zhang, D. DS-TransUNet: Dual Swin Transformer U-Net for medical image segmentation. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 4005615. [[CrossRef](#)]
21. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. Segment Anything. *arXiv* **2023**, arXiv:2304.02643.
22. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
23. Armitage, G.C. Development of a classification system for periodontal diseases and conditions. *Ann. Periodontol.* **1999**, *4*, 1–6. [[CrossRef](#)] [[PubMed](#)]
24. Danks, R.P.; Bister, D.; Kumar, A.; Arnett, C.; Owen, J.; Rees, J.; Petrie, A.; Singh, P.; Ashley, P.; Kalsi, H.; et al. Automated measurement of periodontal bone loss on periapical radiographs using deep learning. *J. Dent.* **2021**, *114*, 103790.
25. Ameli, N.; Gibson, M.P.; Kornerup, I.; Lagravere, M.; Gierl, M.; Lai, H. Automating bone loss measurement on periapical radiographs for predicting the periodontitis stage and grade. *Front. Dent. Med.* **2024**, *5*, 1479. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.