Use of the deep learning approach to measure alveolar bone level

Chun-Teh Lee[1], **Tanjida Kabir**[2], **Jiman Nelson**[1], **Sally Sheng**[1], **Hsiu-Wan Meng**[1], **Thomas E. Van Dyke**[3,4], **Muhammad F. Walji**[5], **Xiaoqian Jiang**[2], **Shayan Shams**[2,6]

[1]Department of Periodontics and Dental Hygiene, The University of Texas Health Science Center at Houston School of Dentistry, Houston, Texas, USA

[2]School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, Texas, USA

[3]Center for Clinical and Translational Research, The Forsyth Institute, Cambridge, Massachusetts, USA

[4]Department of Oral Medicine, Infection, and Immunity, Faculty of Medicine, Harvard University, Boston, Massachusetts, USA

[5]Department of Diagnostic and Biomedical Sciences, The University of Texas Health Science Center at Houston School of Dentistry, Houston, Texas, USA

[6]Department of Applied Data Science, San Jose State University, San Jose, California, USA

Abstract

Aim: The goal was to use a deep convolutional neural network to measure the radiographic alveolar bone level to aid periodontal diagnosis.

Materials and Methods: A deep learning (DL) model was developed by integrating three segmentation networks (bone area, tooth, cemento-enamel junction) and image analysis to measure the radiographic bone level and assign radiographic bone loss (RBL) stages. The percentage of RBL was calculated to determine the stage of RBL for each tooth. A provisional periodontal diagnosis was assigned using the 2018 periodontitis classification. RBL percentage, staging, and presumptive diagnosis were compared with the measurements and diagnoses made by the independent examiners.

Results: The average Dice Similarity Coefficient (DSC) for segmentation was over 0.91. There was no significant difference in the RBL percentage measurements determined by DL and

Correspondence: Shayan Shams, One Washington Sq., San Jose, CA 95192, USA. shayan.shams@sjsu.edu.
Chun-Teh Lee and Tanjida Kabir contributed equally to this study.

examiners ($p = .65$). The area under the receiver operating characteristics curve of RBL stage assignment for stages I, II, and III was 0.89, 0.90, and 0.90, respectively. The accuracy of the case diagnosis was 0.85.

Conclusions: The proposed DL model provides reliable RBL measurements and image-based periodontal diagnosis using periapical radiographic images. However, this model has to be further optimized and validated by a larger number of images to facilitate its application.

1 | INTRODUCTION

Periodontitis is a biofilm-induced chronic inflammatory disease that is characterized by gingival inflammation and alveolar bone loss around teeth. According to the 2009–2014 National Health and Nutrition Examination Survey (NHANES), approximately 61 million adults (42.2%) over 30 years of age suffer from periodontitis with 7.8% having severe periodontitis (Eke et al., 2018). The latest 2018 periodontitis staging and grading classification was designed to allow clinicians to assess periodontitis severity, complexity, extent, as well as progression rate, and determine the patient's potential response to treatment (Papapanou et al., 2018). Disease severity can be quantified as clinical attachment loss, alveolar bone loss, or the number of teeth lost. The primary criteria for periodontitis grading are direct or indirect evidence of disease progression. Direct evidence consists of longitudinal documentation of progressive attachment loss and/or radiographic bone loss (RBL). When direct evidence is unavailable, the estimated rate of progression, as indirect evidence, can be quantified as RBL at the most severely affected site in relation to the patient's age (Tonetti et al., 2018; Kornman & Papapanou, 2020). Measuring RBL is important for making a proper periodontal diagnosis, especially when comprehensive and longitudinal periodontal charting is unavailable. Accurate interpretation of radiographs is important, but clinicians may have different interpretations depending on individual experience and knowledge. Developing a tool to assist clinicians in interpreting and measuring alveolar bone will aid an accurate and reliable periodontal diagnosis.

Deep learning (DL) models have been extensively utilized in different medical domains, including identifying anatomic structures and detecting pathological findings on radiographic images (Giger, 2018; Shams et al., 2018; Esteva et al., 2019). Recently, DL-based computer-aided diagnosis (CAD) in oral imaging was developed, but its adoption has been limited (Schwendicke et al., 2019). There are a few studies measuring alveolar bone level on panoramic radiographs using DL models (Lee et al., 2018; Kim et al., 2019; Krois et al., 2019; Chang et al., 2020). Panoramic images provide a quick overview of the dentition, but the unignorable distortion and a lack of detail prevent accurate and precise diagnosis of periodontitis and other oral diseases (Akesson et al., 1992; Pepelassi & Diamanti-Kipioti, 1997; Hellen-Halme et al., 2020). The current standard is a visual assessment of intra-oral radiographs, which is subject to error. State-of-the-art DL models can offer an objective method for a reliable periodontal diagnosis. The goal of this study was

to introduce a novel DL-based CAD model and compare it with clinicians' assessment based on periapical radiographs.

2 | MATERIALS AND METHODS

2.1 | Model development

The study was conducted in accordance with the guidelines of the World Medical Association's Declaration of Helsinki, a study checklist for artificial intelligence in dental research (Schwendicke et al., 2021), and approved by the University of Texas Health Science Center at Houston (UTHealth) Committee for the Protection of Human Subjects (HSC-DB-20–1340). A DL-based CAD model was developed for alveolar bone level assessment and periodontal diagnosis based on intra-oral radiographs. All digital intra-oral images were taken with a standard film positioning holder and approved by the radiology technicians or radiologists. Convolutional neural networks (CNNs), a neural network-based algorithm designed to process data that exhibit natural spatial invariance, were utilized (LeCun et al., 2015). The proposed model was designed specifically to provide high classification accuracy while maintaining interpretability. It integrated three segmentation networks (bone area, tooth, and cemento-enamel junction, CEJ) and image analysis (Figure 1).

2.2 | Annotation

An open-source tool, Computer Vision Annotation Tool (CVAT), was used to annotate regions of interest (ROIs) from each intra-oral radiograph. The patients' images were selected from the electronic health record (EHR) system by the periodontal diagnosis history, and the images were extracted from the image database. The extracted Digital Imaging and Communications in Medicine (DICOM) files were converted to Portable Network Graphics (PNG) format with additional metadata such as region and image laterality. Each full-mouth series (FMS) radiograph was uploaded to the CVAT under a unique ID for annotation. Multiple ROIs including bone area, tooth, and others (e.g., restorations, vertical defects) were annotated using polygon, and CEJ was annotated using polyline on each image via the coordinates and names. When the CEJ was fully covered by the crown and could not be recognized, the margin of the crown was annotated as CEJ. Each clinical examiner received a unique account. They could log in and perform the annotation without observing other examiners' annotations.

2.3 | Segmentation model

Different variations of U-Net (Ronneberger et al., 2015) were trained and evaluated to find the best architecture and hyperparameter settings for each segmentation model. Specifically, U-Net with CNN, ResNet-34, and ResNet-50 encoder (He et al., 2016) were evaluated for the bone area, tooth, and CEJ line segmentations. Additionally, the hyperparameters, such as kernel size and the number of CNN layers, were adjusted to optimize the performance of the model.

The U-Net architecture consists of three paths: the contraction path is composed of a contraction block (CNN) to shrink the input's width and height and double the number

of feature maps; the expansion path consists of CNN and Upsampling layers. After each expansion block, the number of feature maps halves while their width and height double to maintain the symmetry; the cross-connection path appends the feature maps of the expansion block to feature maps of corresponding contraction blocks to ensure that features learned during contraction are used in the reconstruction.

U-Net with ResNet-34 encoder (Table S1) outperformed other variations for the bone area and tooth segmentations, and U-Net with CNN blocks (Table S2) provided the best result for the CEJ line segmentation (Table 1).

The model was trained for 100 epochs with a batch size of eight using binary-cross entropy loss and the Adam optimizer. Equation (1) illustrates the formula to calculate binary-cross entropy loss:

$$\text{Loss} = -\left(\frac{1}{|B|}\right) \sum_{m=1}^{b} \sum_{i=1}^{n} \left(y_i^m \hat{y}_i^m log + (1 - y_i^m) log(1 - \hat{y}_i^m)\right), \tag{1}$$

where $\hat{y}_i^m$ denotes the predicted value for pixel i for sample m and $\hat{y}_i^m$ is the target pixel value in the mask, and B is the batch size. The segmentation models' outputs were integrated and utilized for further image analysis to measure alveolar bone level. The model was developed using TensorFlow version 2.0 (Abadi et al., 2016), trained and evaluated using the NVIDIA Tesla V100 GPU. The average inference time for a radiographic image is 53.4 ms. The source code and data for the project can be obtained by contacting the corresponding author.

2.4 | Alveolar bone level measurement

A sequence of post-processing steps, such as noise removal (Gaussian filtering) and precise contour detection, were performed to improve the quality of the predicted mask (Figure 2). The parameters in Equation (2), used to calculate RBL percentage, were obtained by the following steps:

a.   Intersecting points of the bone area and each tooth were obtained by performing logical AND operation between the bone area contour and tooth masks (Figure 3a).

b.   Intersecting points of the CEJ line and each tooth were obtained by performing logical AND operation between the predicted CEJ line and tooth masks (Figure 3b).

c.   The minimum points on the left and right sides of each tooth axis were identified to get the apices of each root (Figure 3c).

d.   Line 1L and line 1R were drawn from the intersecting points of the tooth and CEJ line (Figure 3d).

e.   Line 2L and line 2R were drawn from each root apex to the CEJ line parallel to the tooth axis (Figure 3e).

f. The lines 1L, 2L, 1R, and 2R were transformed to millimetres by multiplying this distance with the imager pixel spacing from the DICOM header. The RBL percentage could be obtained (Figure 3f) as follows:

$$\text{RBL}\% = \max\left(\frac{1L}{2L}, \frac{1R}{2R}\right) \times 100. \tag{2}$$

Following the 2018 periodontitis classification, stage I is RBL < 15% (in the coronal third of the root); stage II is 15% ≤ RBL ≤ 33% (in the coronal third of the root); and stage III is extending to the middle third of root and beyond (RBL > 33%). No bone loss (stage 0) was assigned if the distance between the CEJ and alveolar bone level is less than 1.5 mm (Hausmann et al., 1991; Xie, 1991) disregarding the RBL percentage. The RBL stage for each tooth was assigned based on the higher RBL stage of the mesial site or the distal site when the two sites have different stages.

2.5 | Model training and validation

In total, 693 periapical radiographic images from randomly selected 37 periodontitis patients were included in the original dataset. All images were annotated and examined by three independent examiners, including two periodontists and one periodontal resident (S.S., H.M., J.N.) familiar with 2018 periodontitis classification. The image dataset was randomly divided to 70%, 10%, and 20% for training, validation, and testing, respectively. The model was further evaluated on 644 additional periapical images ("additional dataset") from randomly selected 46 cases to assess the accuracy of RBL percentage measurement, RBL stage assignment, and whole-case periodontal diagnosis. These patients included in the "additional dataset" and the original dataset were completely different to avoid any data snooping bias.

All examiners were calibrated for annotation and RBL percentage measurements using three sets of FMS intra-oral radiographs. Before starting the annotations, the Dice similarity coefficient (DSC) for annotations between examiners reached at least 0.84, and the RBL percentage results between examiners were not significantly different ($p > .05$ calculated by Student's t-test). The examiners used the software MiPACS (Medicor Imaging, Charlotte, NC) to measure RBL percentage, and the time was documented. The segmentation models were trained using the annotated bone area, teeth, and the CEJ line. The stages of RBL were assigned by the examiner before measuring the RBL percentage to avoid potential biases. If there was a conflict among the examiners' RBL stage assignment, the stage was decided by the measured RBL percentage to get the ground truth for the final stage assignment.

Forty-six cases' FMS periapical radiographs ("additional dataset") were analysed by the DL model to assign a periodontal diagnosis for individual cases following the 2018 periodontitis classification. A periodontitis case should have inter-dental RBL detectable at ≥ 2 non-adjacent teeth. In each case, the RBL percentage or stage of each tooth was assigned based on the highest when a tooth was present in more than one image. The model-assigned periodontal diagnosis (extent, stage, and grade) was based on the proportion of RBL stages, RBL percentages of the teeth, and the patient's age. These cases were also diagnosed by the

three examiners following the same principle. If there were inconsistent results among the examiners, a diagnosis would be decided by the majority rule.

2.6 | Statistical analysis

The segmentation models were validated by DSC, Jaccard index (JI), and pixel accuracy (PA). DSC (Equation 3) compares the similarity of model's output against reference masks. JI (Equation 4) measures the similarity and diversity of the model's prediction and reference masks. Pixel accuracy (Equation 5) illustrates the percentage of images that are correctly classified by the model with respect to the reference images.

$$DSC = \frac{(2 \times \text{area of overlap})}{(\text{Total number of pixels in both images})}, \tag{3}$$

$$JI = \frac{(\text{Area of overlap})}{(\text{Area of union})}, \tag{4}$$

$$PA = \frac{(\text{Number of correctly classified pixels})}{(\text{Total number of pixels})}. \tag{5}$$

The results of RBL stage assignment by the model were evaluated by the area under the receiving operating characteristics curve (AUROC), sensitivity, specificity, and accuracy. The RBL percentage measurement and required time were compared with examiners' results by Student's t-test. The agreement of RBL stage assignment between the CAD and examiners was calculated by Cohen's Kappa coefficient. The case diagnosis accuracy was presented as the number of cases where the CAD-assigned diagnosis is consistent with the examiners' divided by the total number of cases. When validating the model assigned RBL stage, an RBL percentage variation (±3%) was considered in all stage assignments.

Assuming a standard deviation of 0.15, power analysis was conducted for a two-sided t-test with a null hypothesis that the difference between means of the RBL percentages (between different examiners) is no larger than 0.03 at a significance level of .05 and a power of .8. The analysis showed that at least 394 samples were required.

3 | RESULTS

3.1 | Segmentation model validation

For the bone area, tooth shape, and CEJ line segmentations, DSC was 0.96, 0.95, and 0.91; JI was 0.93, 0.91, and 0.88, and PA was 0.96, 0.89, and 0.99, respectively.

3.2 | Staging validation

The proposed CAD achieved high AUROC, sensitivity, specificity, and accuracy for RBL stage assignment. While analysing the "additional dataset", the AUROC values of RBL stage assignment for no bone loss, stage I, stage II, and stage III were 0.98, 0.89, 0.90, and 0.90, respectively. Sensitivity, specificity, and accuracy for different stages are all over 0.8 (Table 2).

The RBL stage assignment between the proposed CAD and the ground truth showed higher agreement ($\mathscr{K} = 0.81$) compared to the assignment between individual examiners and the ground truth (examiner 1: $\mathscr{K} = 0.72$; examiner 2: $\mathscr{K} = 0.66$; examiner 3: $\mathscr{K} = 0.52$).

3.3 | RBL percentage and level measurements

There was no significant difference in RBL percentage measurements between the DL model and examiners ($p = .65$, Figure 4) in analyzing the "additional dataset". However, the time required to complete RBL percentage measurements and data entry for each FMS radiograph by examiners (mean ± standard deviation: 137 ± 62 min) was significantly longer than the time required by the CAD (9 ± 0.51 min).

The mean RBL percentages for stage I, II, and III measured by the model were $12.11 \pm 2.84\%$, $21.04 \pm 2.67\%$, and $36.12 \pm 7.84\%$, respectively. The mean distances between CEJ and alveolar bone level for sites with RBL stage I, II, III estimated using CAD were 1.73 ± 0.48, 2.93 ± 0.50, and 4.71 ± 1.37 mm, respectively. In the sites assigned with no bone loss, the mean distance was 1.46 ± 0.43 mm.

3.4 | Periodontitis diagnosis assignment

The 46 cases from the "additional dataset" for diagnosis testing included 16 cases with localized, stage III, grade B diagnoses; 10 cases with localized, stage III, grade C diagnoses; 10 cases with generalized, stage III, grade C diagnoses; and cases with other diagnoses (Table S3). The accuracy of periodontitis case diagnosis was 0.85 in the 46 cases. The accuracy for extent (0.96), stage (0.87), and grade (0.96) was generally high.

4 | DISCUSSION

The proposed DL-based CAD model demonstrates high accuracy of alveolar bone level measurements and provisional periodontitis diagnosis based on periapical radiographs. Although, ideally, periodontal diagnosis should be made based on periodontal examinations, radiographic images, and clinical judgement together, the DL model using radiographic findings can provide a rapid and reliable preliminary periodontal diagnosis. Notably, the CAD analysis was considerably faster than clinicians (9 vs. 137 min). The proposed DL model is helpful when a comprehensive periodontal examination is missing, an inexperienced clinician cannot make a proper periodontal diagnosis, or a large number of patients' periodontal diagnosis and radiographic findings have to be re-assessed for diagnosis quality control and research purposes.

Few studies have investigated the use of DL in assessing the alveolar bone level and periodontitis staging. Lee et al. developed a DL model to classify periodontally compromised teeth from periapical radiographs (Lee et al., 2018). Kim et al. and Krois et al. utilized DL to detect RBL or calculate RBL percentage from panoramic radiographs (Kim et al., 2019; Krois et al., 2019). However, these models did not assign the RBL stage to each tooth or determine the case-level periodontitis stage and grade. Chang et al. developed a hybrid model using image processing and DL to assign periodontitis stage from panoramic radiographs (Chang et al., 2020). However, the AUC for RBL stage assignment and the whole-case diagnosis accuracy were not provided. Comparing the current results

with the results in the literature, the segmentation accuracy of the proposed model was comparable or superior to the other study whose DSC, JI, and PA ranged from 0.83 to 0.93 (Chang et al., 2020). The current model also achieved a comparable or higher accuracy of detecting different bone loss levels as compared to the other studies (Kim et al., 2019; Krois et al., 2019). It should be noted that those published studies analysed panoramic radiographs. The proposed model is the only model that measured RBL and assigned the periodontitis extent, stage, and grade compatible with the 2018 periodontitis classification based on periapical radiographs, the standard radiographic images for periodontal diagnosis (American Academy of Periodontology, 2003; Do et al., 2018; Tetradis et al., 2018). Although the accuracy and reliability of the other models appear to be good, assessing the bone level on panoramic radiographs to make periodontitis diagnosis is generally not recommended because of distorted images, overlapping objects, and low resolution.

The current results showed that clinicians do not assign bone loss stage accurately without manually calculating the RBL percentage for each tooth, which is a very time-consuming process, suggesting that a CAD tool designed to assess RBL will be helpful in clinical decision making. However, the RBL percentage can be misleading in some clinical situations, such as teeth with short roots or wide supra-crustal tissue attachment. Our DL model also provides information on the distance between the CEJ and alveolar bone level, which was not reported in any published DL model, for the clinician's reference. The average CEJ–bone distance for no-bone-loss sites in this study was similar to those in the literature (Hausmann et al., 1991; Xie, 1991) and close to the distance for sites with stage I bone loss. These findings support the concept that initial radiographic bone loss should be diagnosed based on multiple factors, including the shape of the alveolar bone crest, discontinuity of the crestal lamina dura, bone level relative to other teeth, and longitudinal radiographic bone level change (Zaki et al., 2015).

The accuracy of the whole dentition periodontal diagnosis suggested by the model was generally high. As compared to the extent and grade, the accuracy of stage was lower. In this exercise, stage is decided by the teeth with more severe RBL than others (Kornman & Papapanou, 2020). The current DL model tends to underestimate RBL at vertical defects, endo-periodontal lesions, and at highly overlapped teeth, where defining the bone level is more challenging. This could explain the lower accuracy of stage assignment in the diagnosis.

The change in clinical attachment level (CAL) is considered the primary parameter used to define the severity of periodontitis, because the periodontal destruction has to be progressive to a certain degree before it can be visualized in the radiograph (Goodson et al., 1984). Use of RBL for periodontal diagnosis may result in under-detection of incipient periodontitis and underestimate of disease severity, although CAL, which is very sensitive to the degree of inflammation present, is moderately to highly associated with the RBL (Zhang et al., 2018; Farook et al., 2020). However, CAL is not commonly documented in clinical practice (Patel et al., 2020) because it is less relevant to clinical treatment decisions compared to probing depth and level of bone loss. Additionally, the accuracy of CAL is questioned for inexperienced clinicians (Vandana & Gupta, 2009). Although RBL cannot fully represent

periodontal destruction, assessing RBL might be a more practical approach to determine periodontal disease severity than measuring CAL in many clinical settings.

In addition to assisting clinical interpretation of RBL, the model can be useful when a large number of radiographic images have to be reviewed. In a teaching clinic, the diagnosis accuracy and reliability of images are primary components of quality assurance. The proposed DL model can efficiently review bone level interpretation results of many students. To study periodontal disease progression and treatment outcomes in a large cohort, a retrospective review of periodontal charting is frequently performed. However, inconsistency of periodontal charting between different care providers may affect data reliability (Buduneli et al., 2004; Lafzi et al., 2007). Bone level changes in periapical images reviewed by the proposed DL model can be used to validate the results of periodontal charting (Machtei et al., 1997; Zaki et al., 2015).

Although the proposed model demonstrated high accuracy and reliability in measuring RBL and assigning a periodontitis diagnosis, it has some limitations. This model is not able to precisely identify vertical defect depth as well as angulation, which can be important for periodontal diagnosis in some cases. Images with poor quality, such as overlapping teeth and distorted tooth length which may mislead the diagnosis, cannot be automatically excluded from the analysis. When multiple teeth are missing, the model is not able to accurately identify the tooth number (position). To assess tooth-level or case-level RBL of a large cohort, the current model has to be trained and validated by more images to further optimize its performance and efficiency. Additionally, the model is not designed to replace periodontal charting and other clinical data. An accurate periodontal diagnosis should always be made based on the results of periodontal charting, radiographic findings, and patient history.

With the limitations of this study, the DL model provides an accurate and reliable alveolar bone level measurement, RBL stage assignment, and preliminary periodontal diagnosis based on periapical radiographs. DL can be utilized as a tool to assist clinicians in diagnosing periodontitis in the clinic and further making the proper treatment plan.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

DATA AVAILABILITY STATEMENT

The source code and data for the project can be obtained by contacting the corresponding author.

REFERENCES

Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving J, Isard M & Kudlur M (2016). Tensorflow: A system for large-scale machine learning. In 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16) (pp. 265–283).

Akesson L, Hakansson J, & Rohlin M (1992). Comparison of panoramic and intraoral radiography and pocket probing for the measurement of the marginal bone level. Journal of Clinical Periodontology, 19(5), 326–332. 10.1111/j.1600-051x.1992.tb00654.x [PubMed: 1517478]

American Academy of Periodontology (2003). Position Paper: Diagnosis of periodontal diseases. Journal of Periodontology, 74(8), 1237–1247. 10.1902/jop.2003.74.8.1237

Buduneli E, Aksoy O, Kose T, & Atilla G (2004). Accuracy and reproducibility of two manual periodontal probes. An in vitro study. Journal of Clinical Periodontology, 31(10), 815–819. 10.1111/j.1600-051x.2004.00560.x [PubMed: 15367182]

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, Kim T-Il., Yi W-J (2020). Deep learning hybrid method to automatically diagnose periodontal bone loss and stage periodontitis. Scientific Reports, 10(1), 7531. 10.1038/s41598-020-64509-z [PubMed: 32372049]

Do J, Takei H, & Carranza F (2018). Chapter 32. Periodontal examination and diagnosis. In Newman M, Takei H, Klokkevold P, & Carranza F (Eds.), Newman and Carranza's clinical periodontology (13th ed.). Elsevier.

Eke PI, Thornton-Evans GO, Wei L, Borgnakke WS, Dye BA, & Genco RJ (2018). Periodontitis in US adults: National Health and Nutrition Examination Survey 2009–2014. Journal of the American Dental Association, 149(7), 576–588 e576. 10.1016/j.adaj.2018.04.023 [PubMed: 29957185]

Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, Cui C, Corrado G, Thrun S, & Dean J (2019). A guide to deep learning in healthcare. Nature Medicine, 25(1), 24–29. 10.1038/s41591-018-0316-z

Farook FF, Alodwene H, Alharbi R, Alyami M, Alshahrani A, Almohammadi D, Alnasyan B, & Aboelmaaty W (2020). Reliability assessment between clinical attachment loss and alveolar bone level in dental radiographs. Clinical and Experimental Dental Research, 6(6), 596–601. 10.1002/cre2.324 [PubMed: 32918518]

Giger ML (2018). Machine learning in medical imaging. Journal of the American College of Radiology, 15(3 Pt B), 512–520. 10.1016/j.jacr.2017.12.028 [PubMed: 29398494]

Goodson JM, Haffajee AD, & Socransky SS (1984). The relationship between attachment level loss and alveolar bone loss. Journal of Clinical Periodontology, 11(5), 348–359. 10.1111/j.1600-051x.1984.tb01331.x [PubMed: 6585374]

Hausmann E, Allen K, & Clerehugh V (1991). What alveolar crest level on a bite-wing radiograph represents bone loss? Journal of Periodontology, 62(9), 570–572. 10.1902/jop.1991.62.9.570 [PubMed: 1941497]

He K, Zhang X, Ren S & Sun J (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770–778).

Hellen-Halme K, Lith A, & Shi XQ (2020). Reliability of marginal bone level measurements on digital panoramic and digital intraoral radiographs. Oral Radiology, 36(2), 135–140. 10.1007/s11282-019-00387-0 [PubMed: 31004276]

Kim J, Lee H-S, Song I-S, & Jung K-H (2019). DeNTNet: Deep Neural Transfer Network for the detection of periodontal bone loss using panoramic dental radiographs. Scientific Reports, 9(1), 17615. 10.1038/s41598-019-53758-2 [PubMed: 31772195]

Kornman KS, & Papapanou PN (2020). Clinical application of the new classification of periodontal diseases: Ground rules, clarifications and "gray zones". Journal of Periodontology, 91(3), 352–360. 10.1002/JPER.19-0557 [PubMed: 31891194]

Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, Dörfer C, & Schwendicke F (2019). Deep learning for the radiographic detection of periodontal bone loss. Scientific Reports, 9(1), 8495. 10.1038/s41598-019-44839-3 [PubMed: 31186466]

Lafzi A, Mohammadi AS, Eskandari A, & Pourkhamneh S (2007). Assessment of intra- and inter-examiner reproducibility of probing depth measurements with a manual periodontal probe. Journal of Dental Research, Dental Clinics, Dent Prospects, 1(1), 19–25. 10.5681/joddd.2007.004

LeCun Y, Bengio Y, & Hinton G (2015). Deep learning. Nature, 521(7553), 436–444. 10.1038/nature14539 [PubMed: 26017442]

Lee JH, Kim DH, Jeong SN, & Choi SH (2018). Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. Journal of Periodontal & Implant Science, 48(2), 114–123. 10.5051/jpis.2018.48.2.114 [PubMed: 29770240]

Machtei EE, Hausmann E, Grossi SG, Dunford R, & Genco RJ (1997). The relationship between radiographic and clinical changes in the periodontium. Journal of Periodontal Research, 32(8), 661–666. 10.1111/j.1600-0765.1997.tb00576.x [PubMed: 9409461]

Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH, Flemmig TF, Garcia R, Giannobile WV, Graziani F, Greenwell H, Herrera D, Kao RT, Kebschull M, Kinane DF, Kirkwood KL, Kocher T, Kornman KS, Kumar PS, … Tonetti MS (2018). Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. Journal of Periodontology, 89, S173–S182. 10.1002/jper.17-0721 [PubMed: 29926951]

Patel T, Vayon D, Ayilavarapu S, Zhu L, Jensen S, & Lee CT (2020). A comparison study: Periodontal practice approach of dentists and dental hygienists. International Journal of Dental Hygiene, 18(3), 314–321. 10.1111/idh.12441 [PubMed: 32384212]

Pepelassi EA, & Diamanti-Kipioti A (1997). Selection of the most accurate method of conventional radiography for the assessment of periodontal osseous destruction. Journal of Clinical Periodontology, 24(8), 557–567. 10.1111/j.1600-051x.1997.tb00229.x [PubMed: 9266343]

Ronneberger O, Fischer P & Brox T (2015). U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention (pp. 234–241). Springer, Cham.

Schwendicke Falk., Golla T, Dreher M, Krois J (2019). Convolutional neural networks for dental image diagnostics: A scoping review. Journal of Dentistry, 91, 103226. 10.1016/j.jdent.2019.103226 [PubMed: 31704386]

Schwendicke F, Singh T, Lee JH, Gaudin R, Chaurasia A, Wiegand T, Uribe S, & Krois J, & IADR e-oral health network and the ITU WHO focus group AI for Health. (2021). Artificial intelligence in dental research: Checklist for authors, reviewers, readers. Journal of Dentistry, 107, 103610. 10.1016/j.jdent.2021.103610 [PubMed: 33631303]

Shams S, Platania R, Zhang J, Kim J, Lee K & Park S-J (2008). Deep generative breast cancer screening and diagnosis. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 859–867). Springer, Cham.

Tetradis S, Mallya S, & Takei H (2018). Chapter 33. Radiographic aids in the diagnosis of periodontal disease. In Newman M, Takei H, Klokkevold P, & Carranza F (Eds.), Newman and Carranza's clinical periodontology (13th ed.). Elsevier.

Tonetti MS, Greenwell H, & Kornman KS (2018). Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. Journal of Clinical Periodontology, 45(Suppl 20), S149–S161. 10.1111/jcpe.12945 [PubMed: 29926495]

Vandana KL, & Gupta I (2009). The location of cemento enamel junction for CAL measurement: A clinical crisis. Journal of Indian Society Periodontology, 13(1), 12–15. 10.4103/0972-124X.51888

Xie JX (1991). Radiographic analysis of normal periodontal tissues. Zhonghua Kou Qiang Yi Xue Za Zhi, 26(6), 339–341 388–389. [PubMed: 1820240]

Zaki HA, Hoffmann KR, Hausmann E, & Scannapieco FA (2015). Is radiologic assessment of alveolar crest height useful to monitor periodontal disease activity? Dental Clinics of North America, 59(4), 859–872. 10.1016/j.cden.2015.06.009 [PubMed: 26427571]

Zhang W, Rajani S, & Wang BY (2018). Comparison of periodontal evaluation by cone-beam computed tomography, and clinical and intraoral radiographic examinations. Oral Radiology, 34(3), 208–218. 10.1007/s11282-017-0298-4 [PubMed: 30484028]

Clinical Relevance

Scientific rationale for study:

Assessing radiographic bone level (RBL) is important for periodontal diagnosis. The interpretation of radiographic images is subjective, and accuracy depends on a clinician's experience and knowledge. Artificial intelligence and image analysis can improve reliability.

Principal findings:

The proposed deep learning-based aided diagnosis can reliably assess the RBL and assign bone loss stage.

Practical implications:

The proposed model can assist the clinician in making an accurate periodontal diagnosis. Furthermore, it is useful for the review of a large number of intra-oral radiographic images for quality control of clinical diagnosis and research purposes related to periodontal diseases.
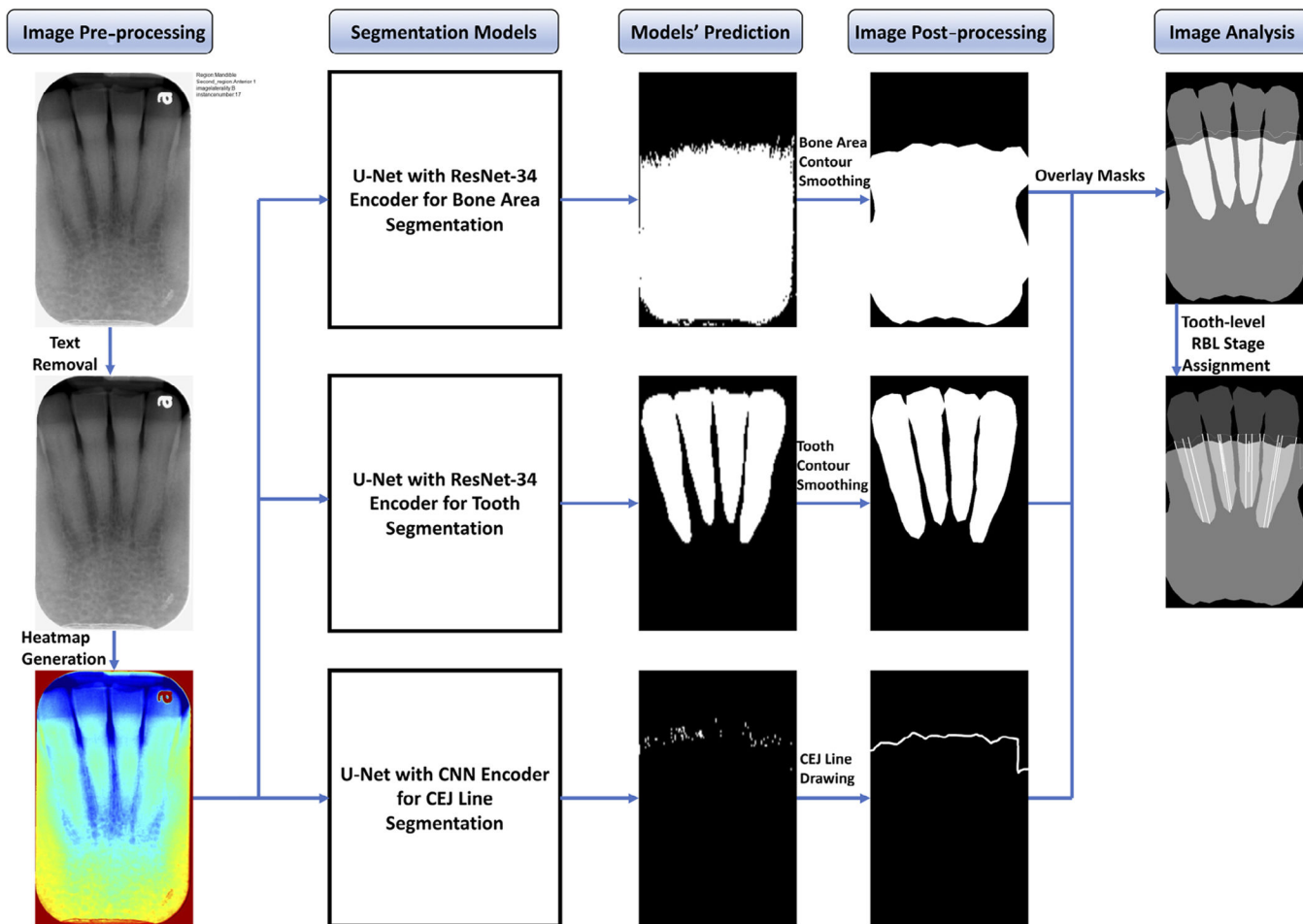
FIGURE 1.

Flow diagram of the proposed computer-aided diagnosis (CAD) model. Segmentation models predicted the bone area, teeth, and cemento-enamel junction (CEJ) masks. Masks are processed to remove noises then overlaid to extract bone area, teeth, and CEJ line for radiographic bone loss (RBL) measurement and stage assignment for each tooth
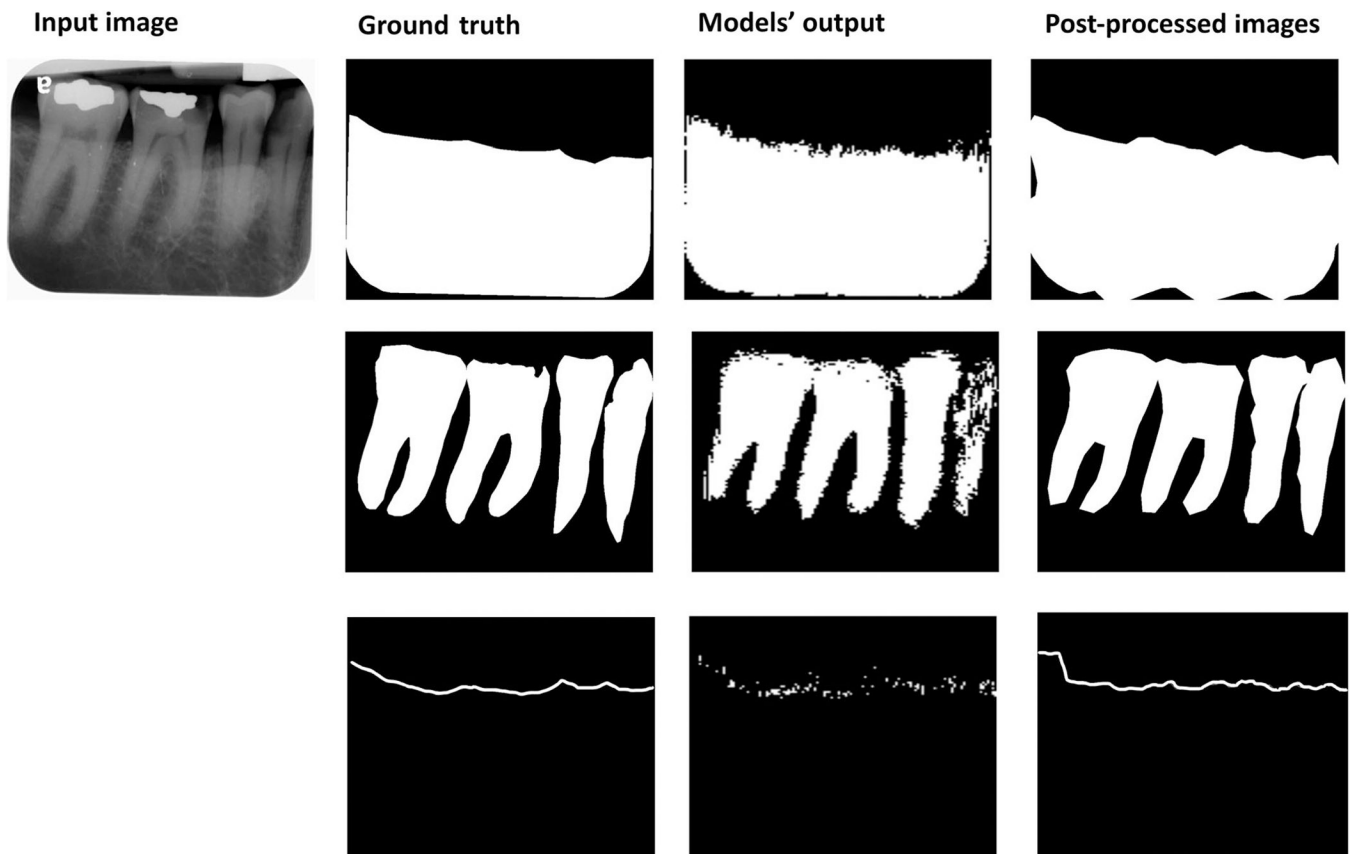
FIGURE 2.
Input image, ground truth (bone area, tooth, and cemento-enamel junction [CEJ] line),
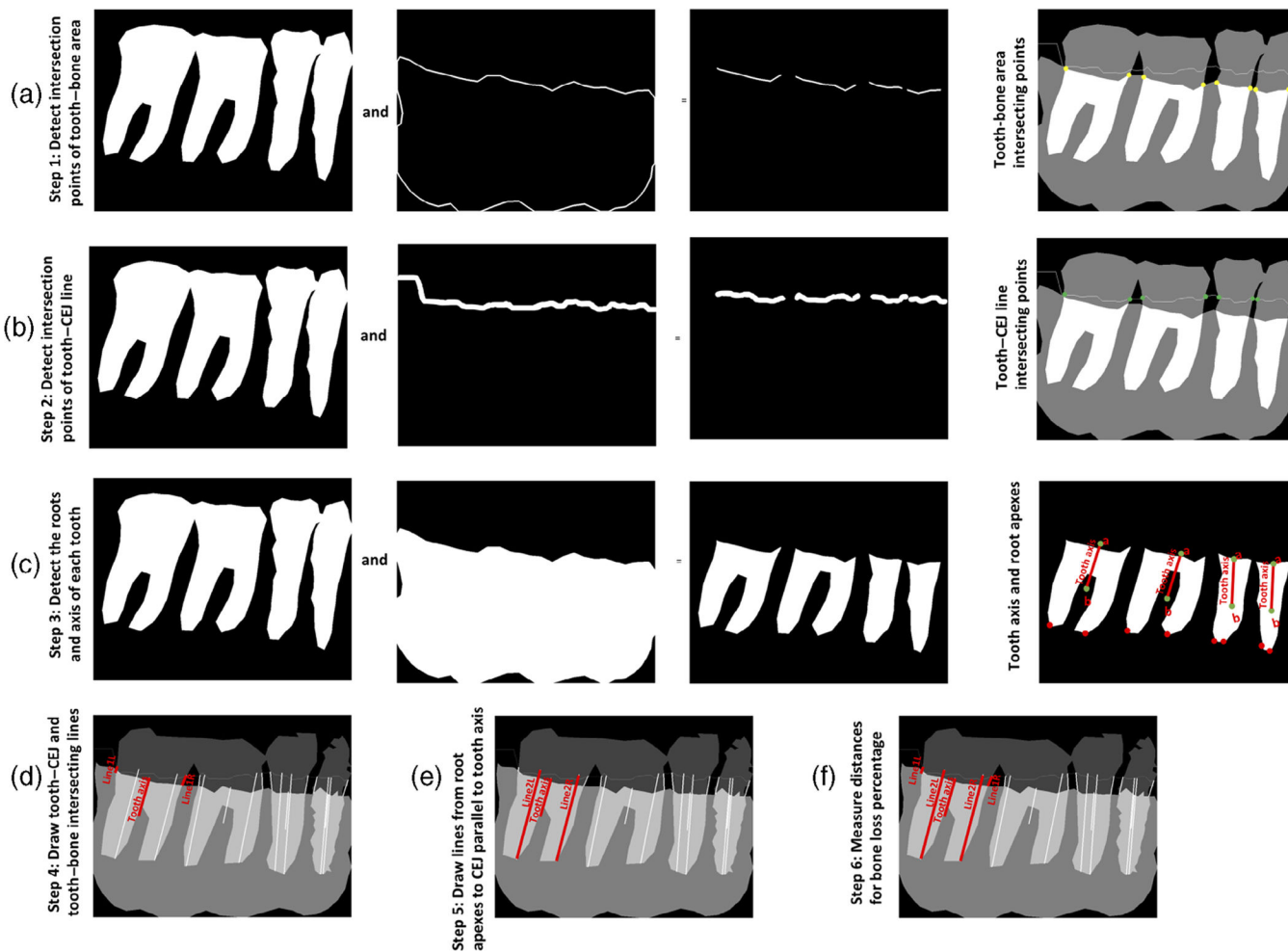segmentation models' output, and images after post-processing

FIGURE 3.

Image analysis steps for calculating bone loss percentage for each tooth. Steps to calculate RBL percentage are as follows: (a) identify the intersecting points (yellow dots) between the tooth and bone area; (b) identify the intersecting points (green dots) between the tooth and CEJ line; (c) identify the mesial and distal root axes parallel to the tooth axis to locate the roots' apexes (red dots); (d) calculate the distance between the CEJ and alveolar bone level at both mesial and distal sites (the line connecting the green dots and the yellow dots); (e) calculate the root length by identifying a line from each root apex to the CEJ line parallel to the tooth axis; and (f) divide the distance between the CEJ and alveolar bone level by the distance between the CEJ and root apex
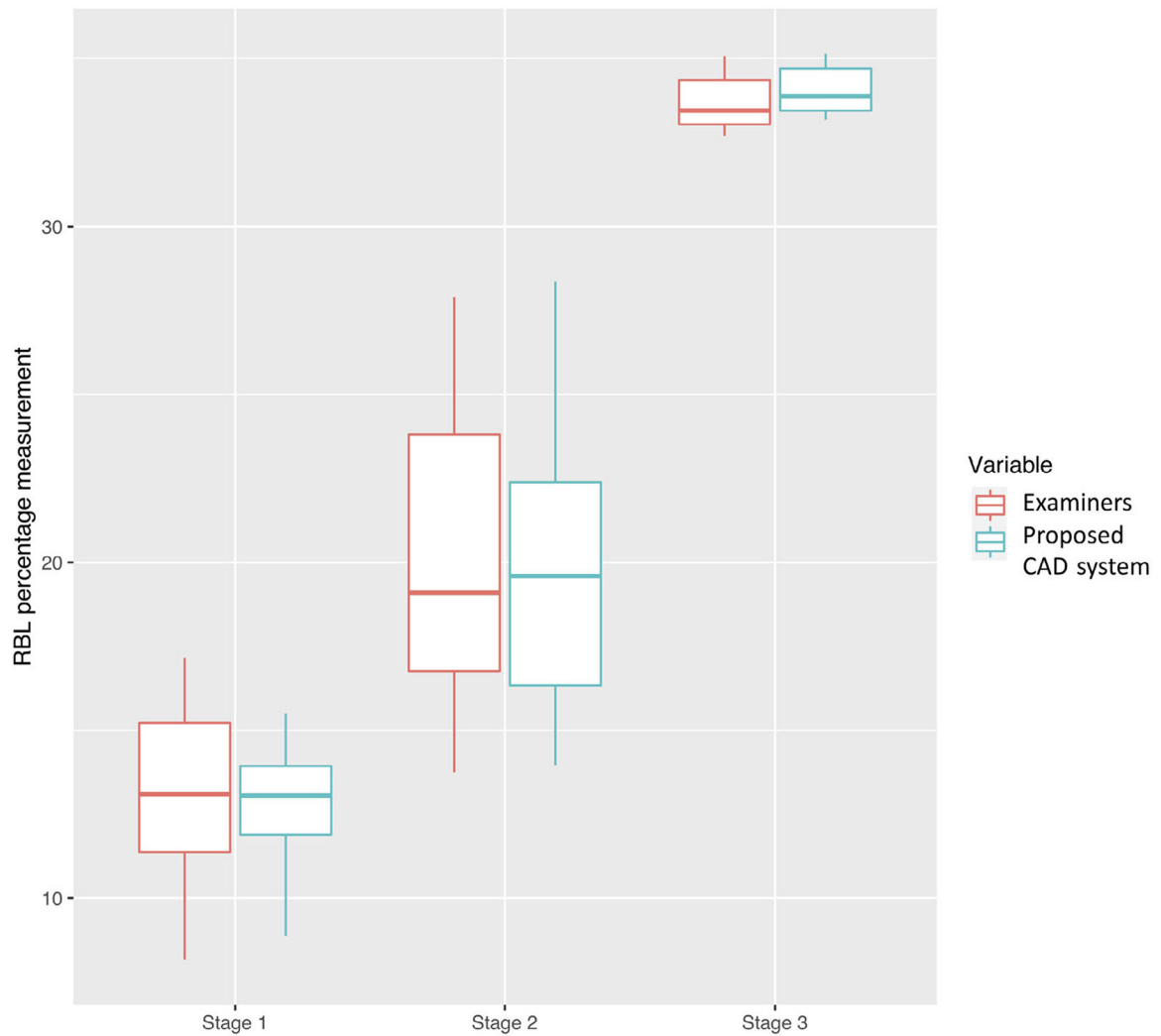
FIGURE 4.

Radiographic bone loss (RBL) percentage measurement distribution of the proposed computer-aided diagnosis (CAD) model and examiners. There was no significant difference in RBL percentage measurements between the CAD and examiners (p-value for all cases, stage I cases, stage II cases, and stage III cases = .65, .32, .27, .96). The bar inside the box represents the median. The upper end of the box represents the third quartile and the lower end of the box represents the first quartile. The ends of the whiskers represent maximum and minimum

TABLE 1

Performance evaluation matrices of different models for segmenting bone area, tooth and cementoenamel junction (CEJ) line

Task name	Model name	Pixel accuracy	Dice similarity coefficient	Jaccard index
Segmenting bone area	U-Net	0.9570	0.9621	0.9319
	U-Net with ResNet-34 Encoder	0.9603	0.9635	0.9343
	U-Net with ResNet-50 Encoder	0.9610	0.9229	0.9333
Segmenting tooth	U-Net	0.8413	0.8994	0.8225
	U-Net with ResNet-34 Encoder	0.8660	0.9470	0.9143
	U-Net with ResNet-50 Encoder	0.8898	0.9533	0.9026
Segmenting CEJ line	U-Net with 3X3 kernel	0.9456	0.4287	0.2753
	U-Net with 5X5 kernel	0.9966	0.6719	0.5143
	U-Net with 7X7 kernel	0.9850	0.9129	0.8776

TABLE 2

Radiographic bone loss (RBL) stage assignment performance

Performance	AUROC	Sensitivity	Specificity	Accuracy
Stage 1 RBL	0.89	0.82	0.97	0.91
Stage 2 RBL	0.90	0.93	0.86	0.88
Stage 3 RBL	0.90	0.80	0.99	0.99
No bone loss	0.98	0.96	1.00	0.99

Note: Sensitivity = true positive/(true positive + false negative). Specificity = true negative/(true negative + false positive). Accuracy = (true positive + true negative)/(true positive + true negative + false positive + false negative).

Abbreviation: AUROC, area under the receiving operating characteristics curve.