

Deep learning method to automatically diagnose periodontal bone loss and periodontitis stage in dental panoramic radiograph

Ting Xue^{a,*}, Lei Chen^a, Qinfeng Sun^b

^a Department of Stomatology, Qilu Hospital (Qingdao), Cheeloo College of Medicine, Shandong University, Qingdao, 266035, China

^b Department of Periodontology, School and Hospital of Stomatology, Cheeloo College of Medicine, Shandong University & Shandong Key Laboratory of Oral Tissue Regeneration, Jinan, 250012, China

ARTICLE INFO

Keywords:

Deep learning
Computer-aided diagnosis
Periodontal bone loss
Periodontitis stage

ABSTRACT

Objectives: Artificial intelligence (AI) could be used as an automatically diagnosis method for dental disease due to its accuracy and efficiency. This research proposed a novel convolutional neural network (CNN)-based deep learning (DL) ensemble model for tooth position detection, tooth outline segmentation, tooth tissue segmentation, periodontal bone loss and periodontitis stage prediction using dental panoramic radiographs.

Methods: The dental panoramic radiographs of 320 patients during the period January 2020 to December 2023 were collected in our dataset. All images were de-identified without private information. In total, 8462 teeth were included. The algorithms that DL ensemble model adopted include YOLOv8, Mask R-CNN, and TransUNet. The prediction results of DL method were compared with diagnosis of periodontists.

Results: The periodontal bone loss degree deviation between the DL method and ground truth drawn by the three periodontists was 5.28%. The overall PCC value of the DL method and the periodontists' diagnoses was 0.832 ($P < 0.001$). The ICC value was 0.806 ($P < 0.001$). The total diagnostic accuracy of the DL method was 89.45%.

Conclusions: The proposed DL ensemble model in this study shows high accuracy and efficiency in radiographic detection and a valuable adjunct to periodontal diagnosis. The method has strong potential to enhance clinical professional performance and build more efficient dental health services.

Clinical significance: The DL method not only could help dentists for rapid and accurate auxiliary diagnosis and prevent medical negligence, but also could be used as a useful learning resource for inexperienced dentists and dental students.

1. Introduction

Periodontal diseases like gingivitis and periodontitis are the most common human disease. As the 6th most prevalent disease worldwide, periodontitis could lead to tooth loss, alveolar bone loss, edentulism, and masticatory dysfunction, which indirectly influences nutrition and health [1,2]. These influences could impose huge socio-economic impacts and healthcare costs. Since there are large numbers of new scientific evidence emerging about periodontitis in the last 30 years, its classification standards have also been revised repeatedly [3]. In 2017, the new definition and classification framework for periodontitis based on a multidimensional staging and grading system were enacted by the American Academy of Periodontology and the European Federation of Periodontology [4]. Clinically, the diagnosis of periodontitis could be

conducted by measuring the clinical attachment loss (CAL) with probing pocket depths and gingival recession. However, the preciseness of this method is influenced by the differences of probing force, angulation, placement, and tip diameter of different dentists, and if under the situation of mild attachment loss or cemento-enamel junction (CEJ) localization is difficult to determine, the CAL detection is hard to be accurate [5-7].

Radiographic bone loss (RBL) detection method could be adopted when the CAL detection is limited [8]. As an objective and standard detection method, the radiographs could help access the condition of periodontal tissue, which includes the alveolar bone levels, bone defects, and furcation involvement [9-11]. The auxiliary radiographic detection information significantly reduces the impact of variability associated with manual clinical measurements and helps to more precise and

* Corresponding author at: Department of Stomatology, Qilu Hospital (Qingdao), Cheeloo College of Medicine, Shandong University, 758 Hefei Road, Qingdao, Shandong 266035, China.

E-mail address: xxtt.589@163.com (T. Xue).

reliable assessment of periodontal health. In addition to clinical examination, bidimensional (2D) imaging modalities (e.g. periapical, bite-wing, and panoramic radiographs) are the preferred diagnostic tools for both initial assessments and follow-up evaluations of clinical attachment loss and alveolar bone levels around dental roots, since the advantages of accessibility, affordability, and high-resolution capabilities [9,12-15]. The precision of this method also depends on the examiner, and sometimes this kind of inspection is subjective and individualized [16]. The dentist may be inexperienced or have too many patients to handle, and can not examine the radiographs carefully. Therefore, the automated assistance systems could be very helpful under this situation [17].

The computer-aided diagnosis (CAD) with AI could be adopted in many medical fields to help improve the quality of patient care [18]. The cases include identifying cavities and periodontitis lesions, as well as maxillary sinusitis, osteoporosis, and other pathologies in the oral and maxillofacial field. The deep learning (DL) assisted dental images detection method has been verified valuable and attracted more and more attention of dentists. The DL dental detection model combines with the convolutional neural network (CNN) algorithm and trained with dataset that includes images annotated with dentists [19,20]. Then, the trained model could automatically recognize the tooth diseases that learned with high accuracy. The low error rate and high diagnosis efficiency of this method is attractive for clinical dentists who are seeking better quality and efficiency diagnosis and treatment support.

There are many CNN-based DL models that could be used as the computer vision method to recognize the dental disease from the patient's radiographs [21]. The YOLO, R-CNN, and U-Net series are the most respective models [22-25]. YOLO series, such as YOLOv5 and YOLOv8, could be used for image classification detection. It has high detection accuracy and efficiency, and the novel YOLOv8 is employed for tooth recognition and classification detection in this study. The R-CNN series, such as Fast R-CNN and Mask R-CNN, shows strong performance in object detection and semantic segmentation tasks. The Mask R-CNN is used for the tooth outline segmentation in this study. Then is the U-Net series, its structure includes a special network architecture of downsampling and upsampling, which can realize efficient image segmentation. U-Net models are widely used in the field of medical image segmentation, especially in the processing of images with complex structure and details. The novel TransUNet, which combines the advantages of transform model and U-Net model, is employed for the tooth tissue segmentation in this study.

Although the DL assisted diagnosis has achieved very remarkable improvement in many complicated dental diseases detection, there is limited research on the DL detection method for comprehensive diagnosis and evaluation of interproximal bone level [22,26-29]. The low recognized accuracy of early version CNN-based models and training dataset with limited number of annotated images are the main reasons that influence the DL model performance. There are also staging standard differences in clinical practice, which influence the popularization and application of DL assisted method.

A novel CNN-based DL assisted periodontists detection method for panoramic radiograph is proposed in this research. The YOLOv8, Mask R-CNN and TransUNet are employed for tooth position, tooth shape detection, and alveolar bone loss degree (ABLD) detection, respectively. A total of 320 patients' panoramic radiographs that annotated by professionals were collected for training and test dataset. The DL detection results were compared with the professionals to show the accuracy of the method.

2. Methods

2.1. Dataset gathering and preprocessing

The research was reviewed and approved by the academic committee of our hospital. A total of 320 patients and their panoramic radiographs were included, the X-ray images were anonymized and all

personal information was also concealed. All images were collected from our hospital during the period 2020 to 2023. **Table 1** shows the population characteristics. The radiographs of patients aged 12 years or younger, as well as images with severe noise or haziness or showing teeth that were partially present or severely distorted were excluded.

There are totally 8462 teeth analyzed for the following purpose: (1) Tooth position detection with YOLOv8, (2) Tooth shape detection & segmentation with Mask R-CNN, (3) Bone issue detection & segmentation with TransUNet, and (4) ABLD and cementoenamel junction (CEJ) level detection. This study employed three computer vision algorithms, and **Fig. 1** shows the detection progress.

To address the distortion problem of 2D radiograph, all the training and test samples used the panoramic radiographs. The panoramic radiograph filming was conducted by the experienced radiologist using unified process and parameters. Keep the position and distortion of the same number teeth are as consistent as possible in different images, to reduce the influence of distortion in model training and test process.

Before the images were used for training, the class type labeling and segmentation boundary annotation need to be processed with software named Labelme. The class type labeling is used for annotating the tooth position and tooth category in the images. The segmentation boundary is used for annotating the semantic border and bone level of the tooth. The panoramic images were randomly divided into training dataset (90%, 288 images) and test dataset (10%, 32 images), which were used for model training and evaluation, respectively.

To ensure the accuracy and consistency and reduce the inter-observer variability of the annotation process, the labeling work was conducted by three senior dentists (one chief physician, one attending physician, and one physician) with at least 8 years of working experience on periodontal and radiology backgrounds. Then, the annotation results were reviewed by two other senior dentists when the labeling work was completed.

2.2. Overview of processing methods

Fig. 1 shows the architecture of the whole detection system. There are three neural network models trained with the labeling images. The YOLOv8 was used to recognize the tooth types and label the tooth names, since YOLOv8 has the advantage in accuracy and speed on object detection. The YOLOv8 also could provide the outer rectangle of the whole tooth area. The subsequent segmentation operations will be performed in the rectangular area, which could save the computing resources. The Mask-RCNN [30] was used to separate the individual teeth, since it has advantage in semantic segmentation for objects that own clear boundaries. Among the three DL models, TransUNet needs the most computation resource. In this research, the TransUNet [31] was used to separate the teeth issues and key point localizations based on the results of Mask-RCNN, since it has advantage in processing medical images with complex structure and detail. The system proposed in this research integrated above three models, which could be used as a whole system, and also could be used separately for single function. The following are the details explanation for the component models.

Table 1
The characteristics statistics of the patients.

Characteristic	Statistics
Patients number	320
Female/male	170/150
Age	
30–40	57
40–50	131
50–60	96
60–70	36
History of periodontitis, n (%)	168(52.5%)

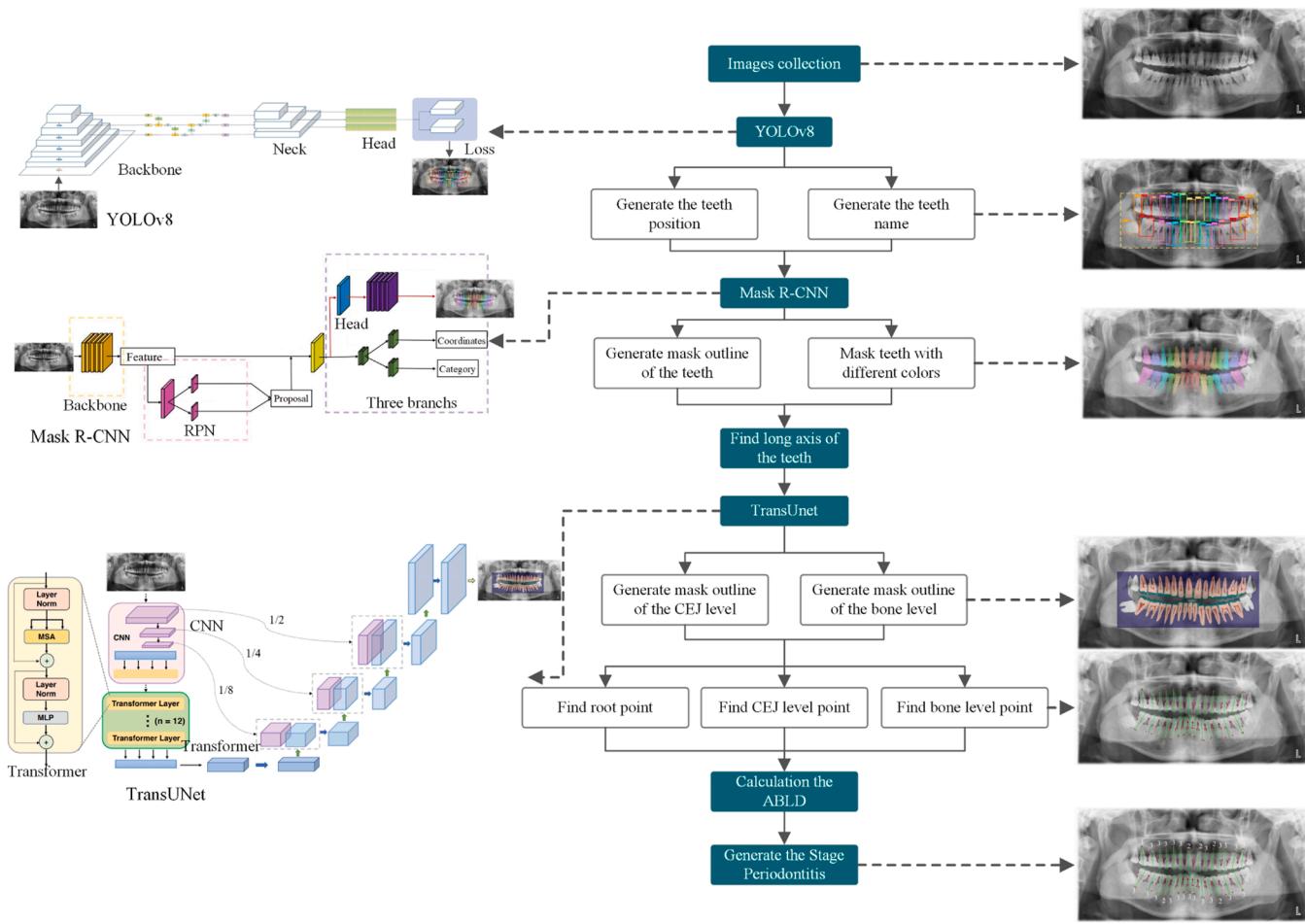


Fig. 1. Work flow of deep learning (DL) method for CAD that proposed in this research.

2.3. Tooth recognize and position detection

The tooth position and type recognize was conducted with the one-step object detection method YOLOv8 (<https://github.com/ultralytics/ultralytics>), and the pre-training weight trained with COCO dataset were adopted. The images size feed in the YOLOv8 model were 640-pixel × 640-pixel. The Fédération Dentaire Internationale (FDI) tooth numbering system was employed during the tooth types labeling process. The system named the upper teeth as 18 to 11, 21 to 28 from right to left, and lower teeth as 48 to 41, 31 to 38 from right to left. There are 7590 labeling teeth in the panoramic images used for training and 872 labeling teeth in the images for test. Because the position of wisdom teeth is much lower than that of the others, it will be excluded and not participate in the following detection process, to prevent the data imbalance in ABLD evaluation. With YOLOv8, the teeth that need to be recognized and their location areas in the images could be located quickly. Then, the following detection process is only conducted in the bounding rectangle (yellow dotted) of all teeth, which could reduce the computing resource of the following process. Fig. 2 row b) shows the teeth position recognition results. As shown in the figure, the yellow dotted box on the outside represents the tooth detection area, and the solid rectangular boxes of different colors represent different teeth in the detection results.

2.4. Tooth mask segmentation

The tooth shape segmentation could be realized with the semantic boundaries of Mask R-CNN [32]. The algorithm could add an output branch that predicts the detection object's mask. To get better detection

results with limited size of training dataset, the test-time augmentation algorithm was adopted. The images fed into the model were made into five copies and rotated with 15° steps. The Hungarian algorithm was employed for excluding the mask that with insufficient pairing overlap, the IoU threshold was set to 0.75, and the mask with highest confidence score will be retained. The method could effectively improve the accuracy of mask segmentation. Fig. 2 row c) shows the teeth segmentation results. As shown in the figure, the different colored areas represent different teeth and their boundaries.

2.5. Tissue type classification

The tooth tissue segmentation is realized with the TransUNet algorithm, which merits both Transformers and U-Net [33], as a strong model for medical image segmentation. In TransUNet [31], the transformer encodes tokenized image patches from a convolution neural network (CNN) feature map as the input sequence for extracting global contexts, and the decoder upsamples the encoded features which are then combined with the high-resolution CNN feature maps to enable precise localization. The TransUNet could realize the pixel-level semantic segmentation, and achieves superior performances to various competing methods on different medical applications including multi-organ segmentation and teeth segmentation.

The data augmentation methods including random horizontal flipping, random cropping, and random intensity shift were adopted to improve the training dataset. The output results of TransUNet in this research includes six channels, which represent the six teeth tissue types: bone, enamel, dentin, pulp, artificial substances and background. In order to reduce computing resources, the teeth tissue segmentation was

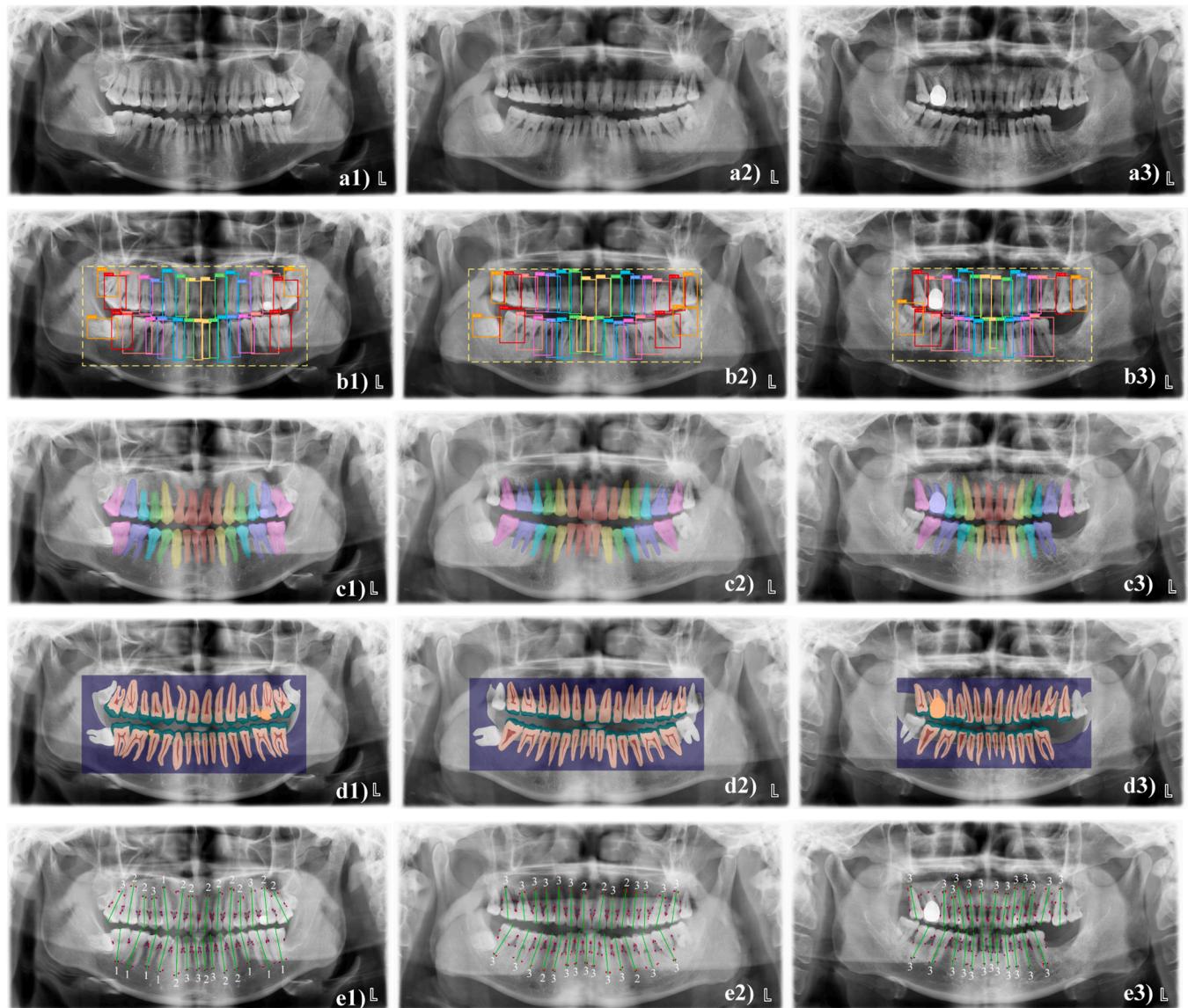


Fig. 2. Results of the detection process. a) original panoramic radiographs, b) position recognition results, c) whole teeth segmentation results, d) tissue segmentation results, e) periodontitis stage classification. For 1, 2, 3 mean periodontitis stage.

only conducted in the teeth distribution area of outer bounding rectangle that get from the YOLOv8 detection results before. Sometimes, there are small gaps between segmentation boundaries of Mask R-CNN and TransUNet, the morphological reconstruction method is employed to make them more consistent as the post-processing step. Fig. 2 row d) shows the teeth tissue segmentation results. As shown in the figure, the different colored areas represent different tissue segmentation, the blue areas represent the bone, the green areas represent the enamel, the pale yellow areas represent the dentin, the red areas represent the pulp, the dark yellow areas represent the artificial substances.

2.6. The principle axis and key point localization

The principal axis of the tooth or the implant was determined by applying the principal axes of inertia to their boundary images [34–36]. The long-axis orientation is defined as the minor axis of the minimum moment of inertia that calculated by the geometrical moments. The green lines in Fig. 2 row e) represent the principal axis of the tooth. Then, the three key points that used to calculate the bone loss degree could be localized with the Mask R-CNN and TransUNet models. The

specific calculation process is as follows:

- (1) For key point A (apex of the tooth root (APEX)), the first key point is defined as the lowest point of the mask. Then, two lines from this point were drawn on each side at 45° upward. The other key point A is defined as the pixel position in the mask that has the largest vertical distance below the lines.
- (2) For key point B (the alveolar bone crest (ALC)), which is defined as the intersect position that dentin, bone, and background tissue segmentation boundaries. If the two teeth are so close and there is no pixel of background between them to localize the key point B, the point on the tooth boundary, which is nearest to the mirror position of key point B at the other side of the tooth, will be taken as the key point B.
- (3) For key point C (CEJ), which is defined as the lowest pixel of the enamel tissue boundary along the tooth mask boundary.

Fig. 3 shows the key point localization annotation in single-tooth magnification images of panoramic radiographs. The crosses were annotated by professionals, and the circles were annotated by the DL

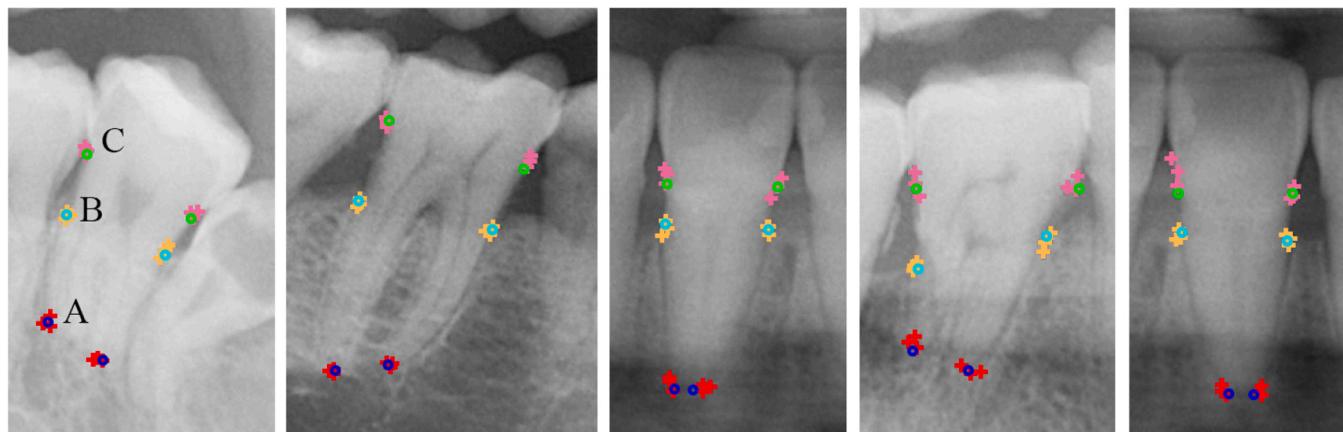


Fig. 3. Key point localization annotation (A: APEX, B: ALC, C: CEJ) in magnification images of panoramic radiographic.

models. The two left images show the similar annotation by professionals and DL model. The right three images show the significant deviations between annotation by professionals and DL model.

2.7. Length-based alveolar bone loss degree measurement

The normal alveolar bone crest is located at a distance of 1–2 mm from the CEJ towards the APEX. If there is bone loss, the alveolar bone crest is located over 2 mm apical to the CEJ. The bone crest level is defined as the point along the root where an intact lamina is found.

The ABLD calculation method in this research is based on the new criteria that proposed in 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. The value of ABLD is decided by two metrics: BL and TR. The BL denotes the length between the positions of 2 mm below CEJ and ALC. The TR denotes the distance between CEJ and APEX. Following is the calculation equation [37]:

$$ABLD = (BL / TR) \times 100\% = \frac{|CEJ - ALC - 2mm|}{|CEJ - APEX - 2mm|} \times 100\%$$

Based on the value of the ABLD, the periodontitis could be divided into three stages. Following are the classification criteria of different stages [37]:

- Stage I: value of ABLD $\leq 15\%$ (in the coronal third of the root).
- Stage II: $15\% < \text{value of ABLD} \leq 33.3\%$ (in the coronal third of the root).
- Stage III: value of ABLD $> 33.3\%$ (extending to the middle third of the root and beyond).

The sample periodontitis stage results that calculated by the ABLD should be validated by the periodontists' clinical diagnosis. When the results were different, the radiograph required re-labeled by the periodontists, and the periodontitis stage required recalculated. Meanwhile, in clinical application of the DL detection method, the stage prediction results that got by the DL model are also recommended to be validated by clinical examination, when the periodontists in doubt about the results.

Fig. 2 row e) shows the teeth periodontitis stage classification results.

2.8. Statistical analysis

The data were summarized as mean-standard deviation, and the between-group differences were tested by independent *t*-test. Percentages were summarized as categorical variables. Differences between groups were compared using the chi-square test for categorical variables.

The whole dataset includes 320 panoramic radiographs and was divided into training dataset and test dataset as 9:1 (288 and 32). Table 2 shows the periodontitis stage classification of all sample teeth. The training dataset was used for tooth classification recognition, whole tooth and tissue segmentation learning and create optimal DL models. The test dataset was used to verify the performance of the optimal models. The parameters including accuracy, precision, recall, F1 score and confusion matrix were chosen as the evaluation criteria. The Pearson correlation coefficients (PCC) and Intraclass correlation coefficients (ICC) were calculated between the DL and dentist's detection results using Python.

3. Results

3.1. Model training information

Fig. 4a) shows the PRC of the YOLOv8 model. It could be seen that the mean mAP score of YOLOv8 for all teeth classes is 94.0%. Fig. 4b-d) are the loss curves for YOLOv8, Mask R-CNN, and TransUNet. It could be seen that the loss values decrease rapidly until convergence and no overfitting is present.

3.2. Tooth recognition

The tooth position and number recognition results are shown in Fig. 2 and Table 3. It could be seen that the overall tooth accuracy is 90.51%, the highest individual tooth accuracy recorded for tooth 21 at 97.96% and the lowest for tooth 28 at 76.09%. Overall, the recognition accuracy of single-rooted teeth was higher than that of multi-rooted teeth. The recognition accuracy for premolars is 90.26%, and for molars is 86.35%.

3.3. Alveolar bone loss degree (ABLD) measurement

Table 4 shows the ABLD deviation between diagnosis results and ground truth. The ground truth is the average value of the three periodontists' diagnosis results. It could be seen that the average deviation of DL detection results to ground truth is 5.28%. The average deviation

Table 2
Periodontitis stage classification of all sample teeth.

	Training dataset	Test dataset	Total
Stage I	2115	228	2343
Stage II	2259	261	2520
Stage III	3216	383	3599
Total	7590	872	8462

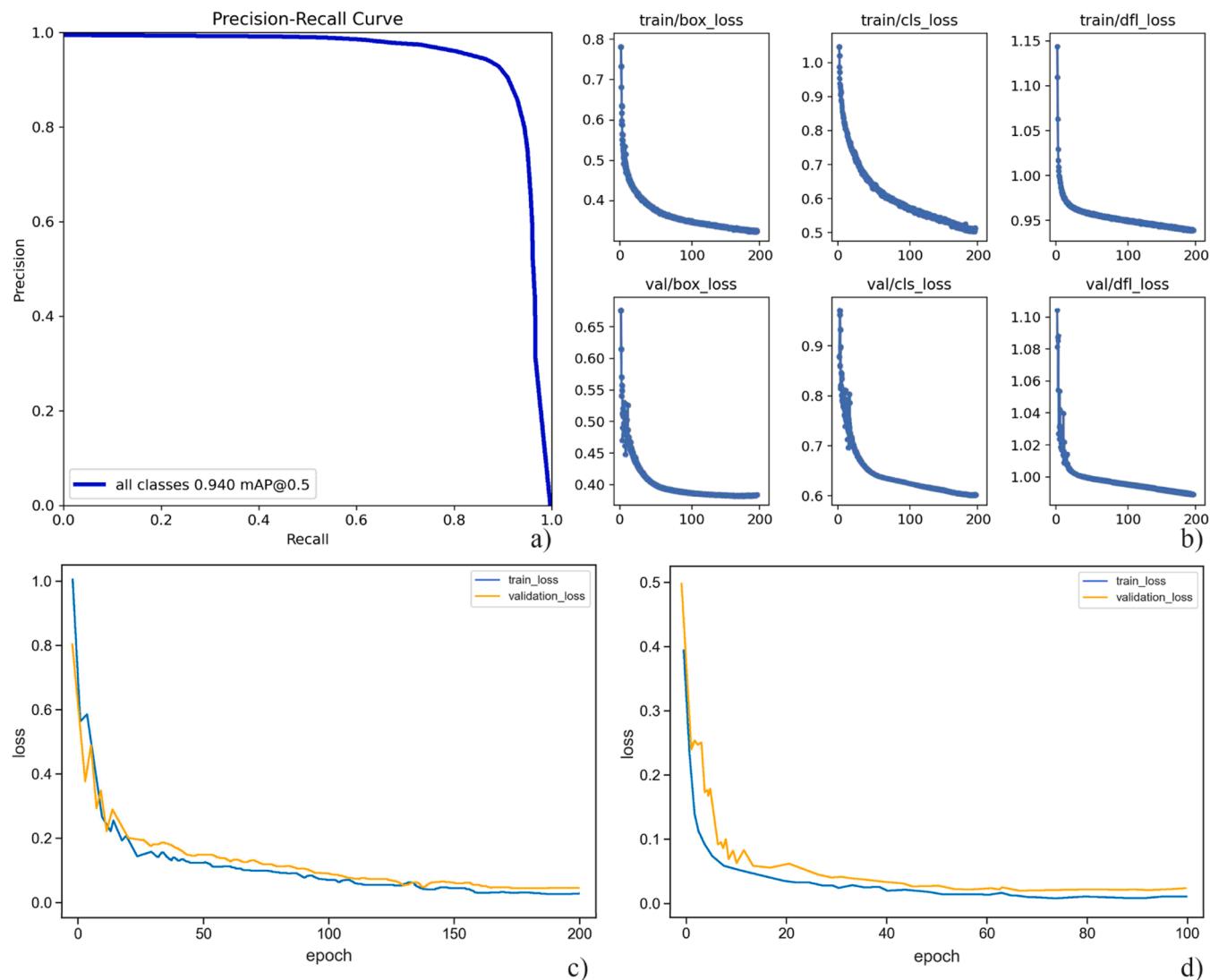


Fig. 4. PRC and loss curves in the training process. a) PRC for YOLOv8, b) loss curve for YOLOv8, c) loss curve for Mask R-CNN, d) loss curve for TransUNet.

Table 3
Tooth recognition accuracy.

Tooth	Accuracy (%)						
11	96.43	21	97.96	31	96.14	41	94.88
12	94.75	22	94.14	32	96.35	42	95.10
13	95.17	23	94.93	33	95.25	43	92.94
14	88.76	24	90.51	34	89.23	44	89.37
15	89.82	25	90.24	35	87.68	45	89.23
16	88.45	26	86.76	36	89.04	46	91.44
17	89.13	27	90.58	37	86.72	47	87.65
18	81.32	28	76.09	38	87.96	48	82.44

Table 4
The ABLD deviation between detection results and ground truth.

Detection results	The ABLD deviation
Chief physician	4.83%
Attending physician	5.65%
Physician	7.37%
DL output	5.28%

of other three periodontists are 4.83%, 5.65% and 7.37%, respectively.

3.4. Correlations of stage classification performance between the DL method and the three periodontists

As shown in Table 5, the PCC between DL output and the periodontist's diagnosis are 0.863 ($P < 0.001$), 0.827 ($P < 0.001$), 0.805 ($P < 0.001$), respectively. The overall correlation between them is 0.832. The PCC results show high correlation between DL output and periodontist's diagnosis. Except that, the diagnosis results of the three periodontists also show strong correlations, as the PCC value is 0.816.

Table 5

The PCC between stages were obtained using the DL method and those diagnosed by the periodontists.

	DL output	Chief Physician	Attending physician	Physician
DL output	1	0.863	0.827	0.805
Chief Physician	0.863	1	0.836	0.795
Attending physician	0.827	0.836	1	0.779
Physician	0.805	0.795	0.779	1

Table 6

The ICC between stages were obtained using the DL method and those diagnosed by the periodontists.

	DL output	Chief Physician	Attending physician	Physician
DL output	1	0.841	0.804	0.772
Chief Physician	0.841	1	0.791	0.746
Attending physician	0.804	0.791	1	0.733
Physician	0.772	0.746	0.733	1

Table 6 shows the ICC values between DL output and the periodontist's diagnosis, which are 0.841 ($P < 0.001$), 0.804 ($P < 0.001$), 0.772 ($P < 0.001$), respectively. The overall ICC value is 0.806. Results show high correlation between them. Both PCC and ICC results show that the DL output is excellent reliable on ABLD and periodontitis stage diagnosis.

3.5. Assessment of the DL diagnostic output accuracy

Fig. 5 shows the confusion matrix between DL prediction results and ground truth about the periodontitis stage diagnosis. The diagonal elements represent the DL predicted the periodontitis stage correctly, elements above the diagonal represent the DL method gives the less severe prediction, and elements below the diagonal represent the DL method gives the more severe prediction. The ground truth of periodontitis stage diagnosis results depends on the majority voting among the three periodontists. Shown as **Table 7**, the total accuracy of DL diagnosis results is 89.45%. Specifically, stage III has the highest recall value with 92.95%, and for stage I is 88.16%, for stage II is 85.44%. Stage III also has the highest F1 score with 92.71%.

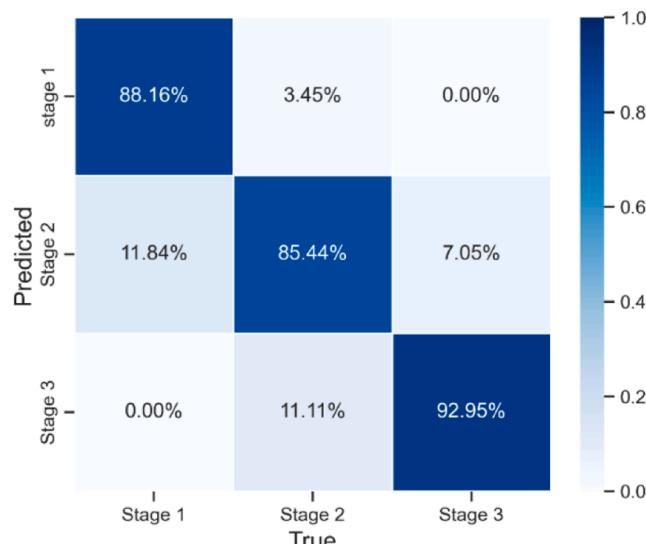


Fig. 5. Confusion matrix of DL output for test dataset.

Table 7

Performance of the DL algorithm on the test dataset.

Performance	stage 1	stage 2	stage 3
Total accuracy	89.45% (for total of three stages)		
Precision	95.71%	80.51%	92.47%
Recall	88.16%	85.44%	92.95%
F1	91.78%	82.90%	92.71%

3.6. Evaluation of stage classification performance

Fig. 6 and **Table 8** show the periodontitis stage classification accuracy comparison of DL method and professionals with ground truth. It could be seen that the chief physician (90.48%) has the highest diagnosis accuracy, and 83 teeth were misdiagnoses compared with the ground truth. Then is the DL method (89.45%), 92 teeth were misdiagnosed. Specifically, 56 teeth were more severe and 36 teeth were less severe. **Table 9** shows the pairwise comparison of DL diagnosis results with the periodontists. It could be seen that, compared with the diagnosis results of three periodontists, there are 68, 79, 112 teeth are classified into the more severe classification, and 37, 58, 44 teeth are classified into the less severe classification, respectively. Therefore, the diagnosis of the DL method usually is more severe than that of periodontists.

4. Discussion

Comparison results show that the ABLD deviation between the DL method and ground truth is 5.28%. In previous research, the deviation values are 6.50% [38], 9.50% [39] and 10.69% [40], respectively. Therefore, the DL method proposed in this research provides the better performance. The diagnosis results of the DL model could be improved if more information is considered, like additional imagery data (e.g., cone-beam computerized tomography). Further assessment of the periodontal bone loss morphology (horizontal or vertical bone loss), rate of periodontal loss (% bone loss/age), and other radiographically assessable items (root morphology and furcation involvement) may be helpful. Integrated this auxiliary information, the DL method could provide more accurate information to dentists in clinical diagnoses. The patient's personal information (systematic diseases [e.g., diabetes mellitus], clinical records [e.g., clinical attachment level], smoking habits) could help the periodontists to make corresponding treatment plans.

The PCC between DL output and periodontists is 0.832, and the ICC is 0.816. Results show that the automatic ABLD calculation and periodontists stage classification with DL method has high accuracy and excellent diagnostic reliability. Research shows that the training dataset size and annotation accuracy by the periodontists are key factors of the model prediction performance. If the number of the images in training dataset is insufficient or annotation quality of the periodontists is unsatisfactory, the model may overfitting or has poor generalization ability, and shows poor performance on new dataset [41]. Thus, the adequate dataset with annotation by specialists in this research could be helpful for improving the detection performance of the DL model.

The prediction accuracy of this method is 90.26% for premolars and 86.35% for molars, which is better than 82.8% for premolars and 73.4% for molars in another research before [22]. The total prediction accuracy for all periodontitis stages is 89.45% in this research, and for stage I is 88.16%, for stage II is 85.44%, for stage III is 92.95%, respectively. In previous research, the overall stage diagnostic accuracy is 72.8%, and for Stage I is 64.2%, for Stage II is 74.3%, for Stage III is 94.0% [38]. Compared with the previous research, the DL method proposed in this research has large improvement, especially for the diagnosis of Stage I and Stage II. This is because new tissue segmentation algorithm TransUNet was employed, which could provide more accurate tooth tissue segmentation. Except that, much more radiographs were included into the training dataset, which could improve the accuracy of DL models.

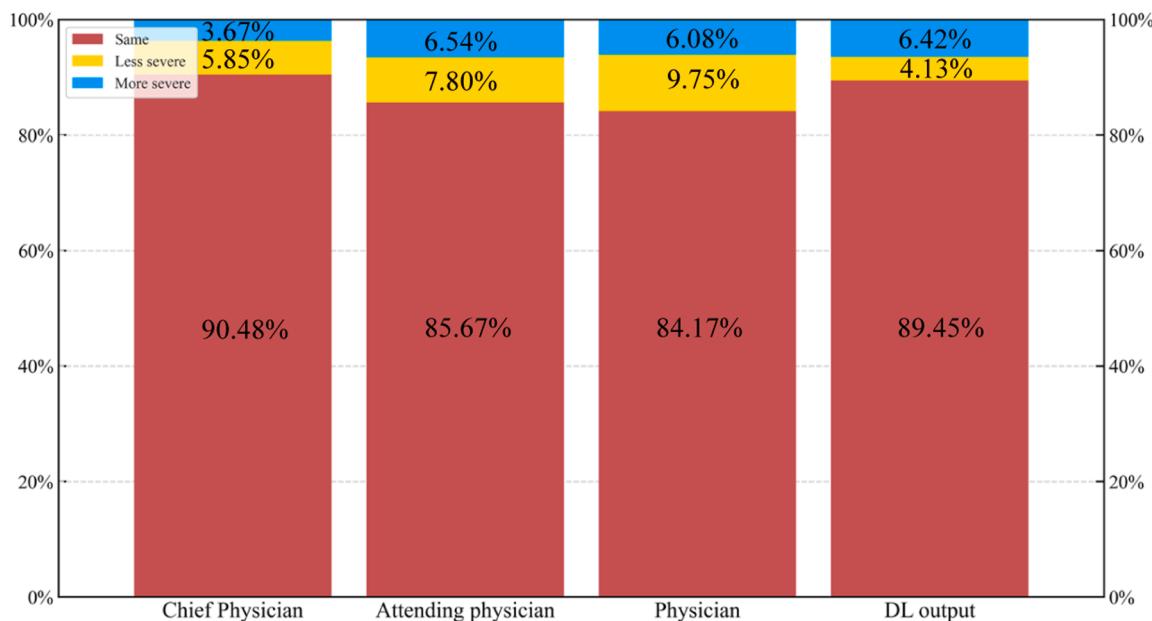


Fig. 6. Periodontitis stage classification accuracy comparison with ground truth.

Table 8

Periodontitis stage classification accuracy comparison with ground truth.

Periodontitis stage diagnosis	Diagnosis accuracy	Number of misdiagnosis	Less severe classification	Same classification	More severe classification
Chief Physician	90.48%	83	51	789	32
Attending physician	85.67%	125	68	747	57
Physician	84.17%	138	85	734	53
DL output	89.45%	92	36	780	56

Table 9

Pairwise comparison between the DL method and professional's diagnosis results.

Comparison object	Less severe classification	Same classification	More severe classification
Chief Physician	37	767	68
Attending physician	58	735	79
Physician	44	706	122

As the high prediction accuracy, this method could serve as a very helpful tool for periodontitis diagnosis, especially for the teeth with stage III. The DL diagnosis method could help dentists for rapid and accurate auxiliary diagnosis and prevent medical negligence. The method also could help to assess the status of alveolar bone following various types of nonsurgical and surgical therapies. The method also could be used as a useful learning resource for inexperienced dentists and dental students.

Although there is not such a large training dataset request for models like TransUnet or Mask R-CNN, a larger training dataset may obtain model with better prediction performance. The DL algorithm with quality architecture is important, and the high-quality annotated training dataset also is critical. Therefore, the dataset for model training will continue to increase in the future. There is still large space to improve for periodontitis diagnosis like stage I and stage II.

Another limitation is that the 2-dimensional periapical radiographs only could provide limited information, and it is difficult for the DL model to get the complete diagnosis information. So, the work that constructs dataset and DL models trained with 3D images (such as CBCT) will be conducted in the further.

5. Conclusion

In conclusion, the DL method proposed in this research could help dentists to diagnose and monitor periodontitis systematically and precisely on panoramic radiographs. Therefore, not only the dental professionals, but also the inexperienced dentist could benefit from the method. Unlike the professional dental clinics in many economically developed areas, there are no professional periodontists in many town hospitals and clinics in the vast less developed areas of China. The diagnosis of dental diseases in these clinics is done by general dentists. For these dentists who may be inexperienced on periodontitis, this AI auxiliary method is significant.

CRediT authorship contribution statement

Ting Xue: Writing – original draft, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lei Chen:** Writing – review & editing, Validation. **Qinfeng Sun:** Supervision, Methodology, Conceptualization.

Declaration of competing interest

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

Acknowledgments

The authors would like to thank PhD Naihua Yue for the help with software and programming.

References

- [1] M.S. Tonetti, S. Jepsen, L.J. Jin, J. Otomo-Corgel, Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: A call for global action, *J. Clin. Periodontol.* 44 (2017) 456–462, <https://doi.org/10.1111/jcpe.12732>.
- [2] C. Lee, K. Zhang, W. Li, K. Tang, Y. Ling, M.F. Walji, X. Jiang, Identifying predictors of the tooth loss phenotype in a large periodontitis patient cohort using a machine learning approach, *J. Dent.* 144 (2024) 104921, <https://doi.org/10.1016/j.jdent.2024.104921>.
- [3] J.G. Caton, A new classification scheme for periodontal and peri-implant diseases and conditions - Introduction and key changes from the 1999 classification, *J. Periodontol.* 89 (2018) 1–8, <https://doi.org/10.1111/jcpe.12935>.
- [4] M.S. Tonetti, H. Greenwell, K.S. Kornman, Staging and grading of periodontitis: Framework and proposal of a new classification and case definition, *J. Periodontol.* 89 (2018) 159–172, <https://doi.org/10.1002/JPER.18-0006>.
- [5] J.J. Garnick, L. Silverstein, Periodontal probing: Probe tip diameter, *J. Periodontol.* 71 (2000) 96–103, <https://doi.org/10.1902/jop.2000.71.1.96>.
- [6] R. Xi, M. Ali, Y. Zhou, M. Tizzano, A reliable deep-learning-based method for alveolar bone quantification using a murine model of periodontitis and micro-computed tomography imaging, *J. Dent.* (2024) 105057, <https://doi.org/10.1016/j.jdent.2024.105057>.
- [7] G. Pelekos, M. Fok, A. Kwok, M. Lam, E. Tsang, M.S. Tonetti, A pilot study on the association between soft tissue volumetric changes and non-surgical periodontal treatment in stage III periodontitis patients. A case series study, *J. Dent.* 134 (2023) 104536, <https://doi.org/10.1016/j.jdent.2023.104536>.
- [8] L. Akesson, J.K. Ha, M. Rohlin, Comparison of panoramic and intraoral radiography and pocket probing for the measurement of the marginal bone level, *J. Clin. Periodontol.* 19 (1992) 326–329, <https://doi.org/10.1111/j.1600-051x.1992.tb00654.x>.
- [9] R. Jacobs, R.C. Fontenelle, P. Lahoud, S. Shujaat, M. Bornstein, Radiographic diagnosis of periodontal diseases – Current evidence versus innovations, *Periodontology 95* (2024) 51–69, <https://doi.org/10.1111/prd.12580>, 2000.
- [10] W. Scarfe, B. Azevedo, L. Pinheiro, M. Priaminiarti, M. Sales, The emerging role of maxillofacial radiology in the diagnosis and management of patients with complex periodontitis, *Periodontol 74* (2017) 116–139, <https://doi.org/10.1111/prd.12193>, 2000.
- [11] E. Corbet, D. Ho, S. Lai, Radiographs in periodontal disease diagnosis and management, *Aust. Dent. J.* 54 (Suppl 1) (2009) S27–S43, <https://doi.org/10.1111/j.1834-7819.2009.01141.x>.
- [12] T. Dietrich, P. Ower, M. Tank, Periodontal diagnosis in the context of the 2017 classification system of periodontal diseases and conditions-implementation in clinical practice, *Br. Dent. J.* 226 (2019) 16–22, <https://doi.org/10.1038/sj.bdj.2019.3>.
- [13] M. Reddy, Radiographic methods in the evaluation of periodontal therapy, *J. Periodontol.* 63 (Suppl 12S) (1992) 1078–1084, <https://doi.org/10.1902/jop.1992.63.12s.1078>, 11.
- [14] M. Jeffcoat, Current concepts in periodontal disease testing, *J. Am. Dent. Assoc.* 125 (1994) 1071–1078, <https://doi.org/10.14219/jada.archive.1994.0136> 12.
- [15] V. Faria, K. Evangelista, C. Rodrigues, C. Estrela, T. Sousa, M. Silva, Detection of periodontal bone loss using cone beam CT and intraoperative radiography, *Dentomaxillofac. Radiol.* 41 (2012) 64–69, <https://doi.org/10.1259/dmfr/13676777>.
- [16] I. Choi, A. Cortes, E.S. Arita, M. Georgetti, Comparison of conventional imaging techniques and CBCT for periodontal evaluation: a systematic review, *Imaging Sci. Dent.* 48 (2018) 79–86, <https://doi.org/10.5624/isd.2018.48.2.79>.
- [17] P. Lin, P. Huang, P. Huang, Automatic methods for alveolar bone loss degree measurement in periodontitis periapical radiographs, *Comput Methods Prog Biomed* 148 (2017) 1–11, <https://doi.org/10.1016/j.cmpb.2017.06.012>.
- [18] N. Ahmed, M. Abbasi, F. Zuberi, Artificial intelligence techniques: analysis, application, and outcome in dentistry-a systematic review, *Biomed. Res. Int.* (2021) 9751564, <https://doi.org/10.1155/2021/9751564>.
- [19] M. Sun, Y. Chai, G. Chai, Fully automatic robot-assisted surgery for mandibular angle split osteotomy, *J. Craniofac. Surg.* 31 (2020) 336–339, <https://doi.org/10.1097/SCS.00000000000005587>.
- [20] S. Li, J. Liu, Z. Zhou, Z. Zhou, X. Wu, Y. Li, S. Wang, W. Liao, S. Ying, Z. Zhao, Artificial intelligence for caries and periapical periodontitis detection, *J. Dent.* 122 (2022) 104107, <https://doi.org/10.1016/j.jdent.2022.104107>.
- [21] A. Barr, E.A. Feigenbaum, *The Handbook of Artificial Intelligence*, 1, William Kaufmann, 1981.
- [22] J. Lee, D. Kim, S. Jeong, Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm, *J. Periodontal. Impl. Plant Sci.* 48 (2018) 114–123, <https://doi.org/10.5051/jpis.2018.48.2.114>.
- [23] X. Meng, F. Mao, Z. Mao, Q. Xue, J. Jia, M. Hu, Multi-stage Unet segmentation and automatic measurement of pharyngeal airway based on lateral cephalograms, *J. Dent.* 136 (2023) 104637, <https://doi.org/10.1016/j.jdent.2023.104637>.
- [24] H. Chang, S. Lee, T. Yong, Deep learning hybrid method to automatically diagnose periodontal bone loss and stage periodontitis, *Sci. Rep.* 10 (2020) 1–8, <https://doi.org/10.1038/s41598-020-64509-z>.
- [25] M. Büttner, L. Schneider, A. Krasowski, V. Pitchika, J. Krois, H. Meyer-Lueckel, F. Schwendicke, Conquering Class Imbalances in Deep Learning-based Segmentation of Dental Radiographs with Different Loss Functions, *J. Dent.* (2024) 105063, <https://doi.org/10.1016/j.jdent.2024.105063>.
- [26] J. Garnick, L. Silverstein, Periodontal probing: probe tip diameter, *J. Periodontol.* 71 (2000) 96–103, <https://doi.org/10.1902/jop.2000.71.1.96>.
- [27] L. Trombelli, R. Farina, C. Silva, Plaque-induced gingivitis: case definition and diagnostic considerations, *J. Clin. Periodontol.* 45 (2018) 44–67, <https://doi.org/10.1111/jcpe.12939>.
- [28] H. Mohammad-Rahimi, R. Rokhshad, S. Bencharit, J. Krois, F. Schwendicke, Deep learning: A primer for dentists and dental researchers, *J. Dent.* 130 (2023) 104430, <https://doi.org/10.1016/j.jdent.2023.104430>.
- [29] M. Xu, Y. Wu, Z. Xu, P. Ding, H. Bai, X. Deng, Robust automated teeth identification from dental radiographs using deep learning, *J. Dent.* 136 (2023) 104607, <https://doi.org/10.1016/j.jdent.2023.104607>.
- [30] T. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, S. Belongie, Feature Pyramid Networks for Object Detection. 2017.
- [31] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A.L. Yuille, Y. Zhou, TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. (2021).
- [32] K. He, G. Gkioxari, P. Dollar, R. Girshick, R-CNN. Mask, *IEEE Trans. Pattern. Anal. Mach. Intell.* 42 (2020) 386–397.
- [33] R.P. Olaf, T. Brox, Dental X-ray image segmentation using a U-shaped deep convolutional network, *International Symposium on Biomedical Imaging* 1 (2015) 1–13.
- [34] D.S. Kim, Principal direction of inertia for 3D trajectories from patient-specific TMJ movement, *Comput. Biol. Med.* 43 (2013) 169–175, <https://doi.org/10.1016/j.combiomed.2012.12.007>.
- [35] X. Huang, L. Zou, R. Yao, S. Wu, Y. Li, Effect of preparation design on the fracture behavior of ceramic occlusal veneers in maxillary premolars, *J. Dent.* 97 (2020) 103346, <https://doi.org/10.1016/j.jdent.2020.103346>.
- [36] W.J. Yi, Direct measurement of trabecular bone anisotropy using directional fractal dimension and principal axes of inertia. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* 104 (2007) 110–116, <https://doi.org/10.1016/j.tripleo.2006.11.005>.
- [37] M. Tonetti, H. Greenwell, K. Kornman, Staging and grading of periodontitis: framework and proposal of a new classification and case definition, *J. Periodontol.* 89 (2018) 159–172, <https://doi.org/10.1002/JPER.18-0006>.
- [38] I. Chen, C. Lin, M. Lee, T. Chen, T. Lan, C. Chang, T. Tseng, T. Wang, J. Du, Convolutional-neural-network-based radiographs evaluation assisting in early diagnosis of the periodontal bone loss via periapical radiograph, *J. Dent. Sci.* 19 (2024) 550–559, <https://doi.org/10.1016/j.jds.2023.09.032>.
- [39] P. Lin, P. Huang, P. Huang, Automatic methods for alveolar bone loss degree measurement in periodontitis periapical radiographs, *Comput Methods Prog Biomed* 148 (2017) 1–11, <https://doi.org/10.1016/j.cmpb.2017.06.012>.
- [40] R. Danks, S. Bano, A. Orishko, Automating periodontal bone loss measurement via dental landmark localization, *Int. J. Comput. Assist. Radiol. Surg.* 16 (2021) 1189–1199, <https://doi.org/10.1007/s11548-021-02431-z>.
- [41] H. Chang, S. Lee, T. Yong, Deep learning hybrid method to automatically diagnose periodontal bone loss and stage periodontitis, *Sci. Rep.* 10 (2020) 7531–7538, <https://doi.org/10.1038/s41598-020-64509-z>.