An Auxiliary Measurement Method for Marginal Bone Loss around Dental Implants Based on Periapical Radiographs

Tao GAO[1#], Rong TAN[2#], Su Kun TIAN[3], Hu CHEN[3], Hao Ran WANG[4], Jian Guo XU[2]

Objective: *To develop an efficient and accurate method for the precise quantification of peri-implant marginal bone loss, aiming to overcome the limitations of conventional assessments that rely heavily on expert subjective judgment and to enhance the monitoring of marginal bone level stability and facilitate the early detection of clinical risk factors.*

Methods: *This research proposes an efficient automatic quantification method for marginal bone loss based on the YOLOv8-pose model. By detecting and localising keypoints associated with peri-implant bone loss, this method enables automated identification of marginal bone loss characteristics. Model performance is evaluated using the marginal bone loss rate to assess its effectiveness in predicting the severity of resorption on both sides of the implant.*

Results: *The precision of keypoint detection for marginal bone loss features exceeds 98%, and the mean radial error of six keypoints in most images is less than 10 pixels, with the minimum error reaching only 2.89 pixels. In predicting marginal bone loss severity on both implant sides, the model achieves accuracies of 0.906 and 0.844, respectively, demonstrating robust predictive capability and application potential.*

Conclusion: *The YOLOv8-pose-based method effectively addresses the inefficiency and subjectivity of traditional assessments by enhancing the automation and objectivity of peri-implant marginal bone loss detection. It provides valuable technical support for accurate evaluation and clinical decision making in peri-implant marginal bone loss.*

Keywords: *convolutional neural network, dental implant, keypoint detection, marginal bone loss, periapical radiographs*

Chin J Dent Res 2025;28(4):261–276; doi: 10.3290/j.cjdr.b6745502

In recent years, implant restoration technology has evolved significantly, and due to its high success and retention rates, dental implantology has been recognised as a well-established and reliable approach for restoring missing teeth.[1] According to previous studies, the 10-year cumulative survival rate of implants at the implant level and patient level can reach 96.8% and 92.5%, respectively.[2] Although the success rate of dental implant surgery exceeds 90%, due to the large number of patients receiving implants, many patients remain at risk of implant failure. Thus, early and accurate identification of key lesion characteristics that lead to implant failure is essential, not only for enhancing implant success rates and safeguarding patient health but also for

1 Beijing Jishuitan Hospital, Capital Medical University, Beijing, P.R. China
2 College of Mechanical & Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, P.R. China
3 Center of Digital Dentistry, Peking University School and Hospital of Stomatology & National Engineering Research Center of Oral Biomaterials and Digital Medical Devices & NHC Key Laboratory of Digital Stomatology (Key Laboratory of Digital Stomatology, Chinese Academy of Medical Sciences), Beijing, P.R. China
4 Detroit Green Technology Institute, Hubei University of Technology, Wuhan, P.R. China

 # Equal contribution.

Corresponding authors: Dr Jian Guo XU, College of Mechanical & Electrical Engineering, Nanjing University of Aeronautics and Astronautics, 29 Yudao Road, Qinhuai District, Nanjing 210016, P.R. China. Email: xu_nuaa_edu@hotmail.com.

Dr Hu CHEN, Center of Digital Dentistry, Peking University School and Hospital of Stomatology, No.22 Zhongguancun South Street, Haidian District, Beijing 100081, P.R. China. Email: chenhu44@126.com.

This work was financially supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2023-PT320-09) and Beijing Natural Science Foundation and Haidian Original Innovation Joint Fund (no. L232145).

significantly improving patient satisfaction by meeting their expectations. In implant therapy, the stability of peri-implant marginal bone levels is critical for both the success of treatment and the long-term viability of the implant. Implant marginal bone loss (MBL) refers to the gradual reduction or degradation of bone tissue around the implant, and its progressive development has a significant adverse effect on the long-term survival of the implant. Specifically, MBL can lead to mucosal recession, loss of gingival papillae and exposure of the crown margins, which negatively affects the aesthetic outcomes of the restoration.[3] Furthermore, it leads to reduced bone support, compromises the mechanical stability of the implant and may ultimately result in implant loosening or failure.[4] Additionally, MBL fosters an environment conducive to bacterial colonisation, thereby increasing the risk of peri-implantitis. Studies have demonstrated that the extent of MBL around the implant correlates with the severity of peri-implantitis.[5,6] MBL around the implant is a key indicator for assessing the long-term success of the implant,[7,8] and dynamic monitoring of the MBL at various stages around the implant provides valuable clinical insights. On one hand, early detection of changes in MBL allows for accurate assessment of the stability of the surrounding bone tissue, enabling timely identification of potential clinical risks. On the other hand, the degree of MBL serves as essential evidence for preventing and intervening in peri-implantitis at an early stage. As such, monitoring the changes in MBL around the implant is of paramount importance for evaluating the success rate of implant treatment, ensuring long-term implant stability and optimising treatment outcomes, as it is one of the key determinants of implant restoration success.

Clinically, the stability of peri-implant bone tissue is usually monitored through periapical radiographs, which, due to their simplicity and widespread application, have become the preferred method for evaluating marginal bone levels.[9,10] Currently, the assessment of bone loss predominantly relies on the subjective judgment of experts. The evaluation process involves radiographic diagnostics, comparing the bone height around the implant with the immediate postoperative and post-restoration bone heights; however, this manual approach has notable limitations. Although radiographs can reveal bone loss, accurately quantifying its extent and progression still depends heavily on individual experience, and often results in inconsistent and imprecise measurements. With increasing case volumes, the repetitive nature of follow-up assessments places a considerable burden on clinicians, making manual evaluation both time-consuming and prone to diagnostic errors. Consequently, there is an urgent need for an intelligent method to monitor marginal bone loss around implants, thereby evaluating long-term implant stability.

In recent years, the rapid development of artificial intelligence (AI) technologies has provided efficient and convenient solutions for medical image analysis. By leveraging AI to assist in processing large volumes of medical images, it is not only possible to significantly alleviate the workload of clinicians, but also to generate standardised diagnostic results, effectively reducing human error in the diagnostic process. The successful application of convolutional neural networks (CNNs) in natural image classification has ushered in a new era of deep learning research in medical imaging. Currently, CNNs have been widely applied in various medical fields for image recognition and detection.[11,12] In dentistry, deep learning–based dental diagnosis and treatment technology has emerged, marking not only a profound integration of AI with oral medicine but also a pivotal milestone in the transition from digital dentistry to intelligent dentistry.[13,14] As this development progresses, CNNs are used extensively for the identification of periodontal damage,[15] caries,[16] alveolar bone loss[17,18] and prosthetic design.[19] In addition, CNNs have made notable progress in several key aspects of dental image processing, including automatic tooth detection,[20] precise 3D dental model reconstruction,[21] tooth segmentation[22,23] and intelligent removal of artifacts in medical images.[24] In the field of implantology, recent reviews[25-27] indicate that deep learning technologies have demonstrated significant application value in classification,[28,29] detection,[30,31] prognosis prediction[32] and prediction of implant position.[33] However, studies utilising deep learning techniques to quantify MBL around implants remain relatively limited, while some efforts have yielded preliminary results, yet significant challenges persist in their practical application. For example, Vera et al[30] employed deep learning methods to detect implant crowns and screws and measure bone loss; however, due to varying contrast between implants and surrounding tissues in different images, the robustness of this method in detecting bone loss keypoints was insufficient. Cha et al[34] used an improved R-CNN model to measure the MBL around implants in radiographs, and Liu et al[35] employed Faster R-CNN to detect MBL around dental implants. Although the models' performance was promising, the high training cost made it difficult to implement widely in resource-constrained settings. Lee et al[36] proposed a new perspective for quantitative research by staging bone loss based on bounding box areas, but the method did not achieve

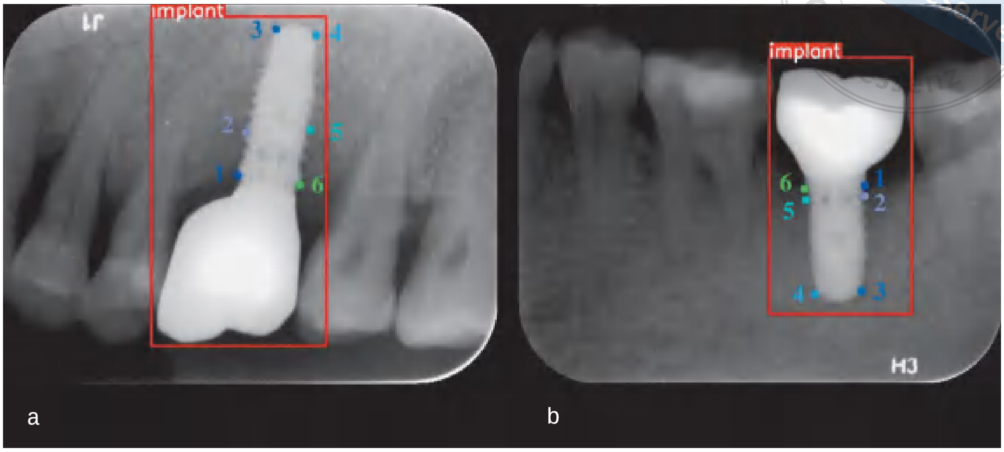

Fig 1a and b Keypoints for labelling peri-implant MBL: maxillary dental implant (a); mandibular dental implant (b).

accurate measurements of bone loss rate, limiting its ability to quantitatively assess long-term implant stability and prognosis in clinical practice.

Given the limitations of existing methods in terms of quantification accuracy for MBL around implants and their adaptability to resource-constrained settings, there is an urgent need to explore a more efficient, flexible and clinically applicable technical approach. The YOLOv8 model, released by the Ultralytics team (Frederick, MD, USA) in 2023, presents a promising solution to this challenge.[37] A powerful and versatile deep learning framework, YOLOv8 supports a wide range of computer vision tasks, including image classification, object detection, image segmentation and keypoint detection. As a branch dedicated to keypoint detection tasks, YOLOv8-pose is a lightweight, high-precision and real-time model that can directly predict the keypoint position and its visibility state through an end-to-end CNN. Initially used for human pose estimation, it has been widely applied in abnormal posture detection,[38] animal behaviour monitoring[39,40] and industrial tasks like instrument calibration[41] and equipment adjustment.[42] Despite these promising results, the application of YOLOv8-pose in the medical imaging field remains relatively scarce. Some preliminary studies have already demonstrated its potential. For example, Rashmi et al[43] used YOLOv8-pose to detect cranial keypoints, providing valuable practical insights for medical image analysis; however, due to the significant differences in greyscale distribution and texture features between medical and natural images, keypoint detection algorithms face new challenges and higher demands in this domain. Extending YOLOv8-pose to the field of medical imaging for the detection of lesion keypoints not only represents a significant innovation but also holds considerable potential for practical applications.

Inspired by the outstanding performance of the YOLOv8-pose model in natural images and its potential applicability to medical scenarios, the present study applies YOLOv8-pose to the quantification of MBL around implants. To overcome challenges in previous research, such as high training costs, limited robustness and inaccurate quantification, an innovative approach is developed that harnesses the model's efficiency and multitasking capabilities, offering an effective solution for precise quantification of MBL.

Materials and methods

Acquisition and preprocessing of data

This study was approved by the bioethics committee of Peking University School and Hospital of Stomatology (PKUSSIRB-2024106237). The study was conducted in accordance with institutional ethical guidelines. The data used in this study were collected from Beijing Jishuitan Hospital and Peking University School and Hospital of Stomatology and anonymised to ensure the protection of patient privacy. In this study, a total of 208 periapical radiographs of dental implants were collected, with image sizes ranging from approximately 300 × 300 to 500 × 400 pixels. The multicentre nature of the data provides a solid foundation for exploring the diagnosis and treatment of implant-related diseases. The annotations were performed by experienced oral surgeons from both hospitals. The study focuses on implants that have undergone restorative procedures, with six keypoints marked, three on each side of the implant. As shown in Fig 1, the points are labelled in a clockwise sequence, and include the highest points of the implant neck (points 1 and 6, located at the first thread closest to the crown),[44] the bone contact points (points 2 and 5, the

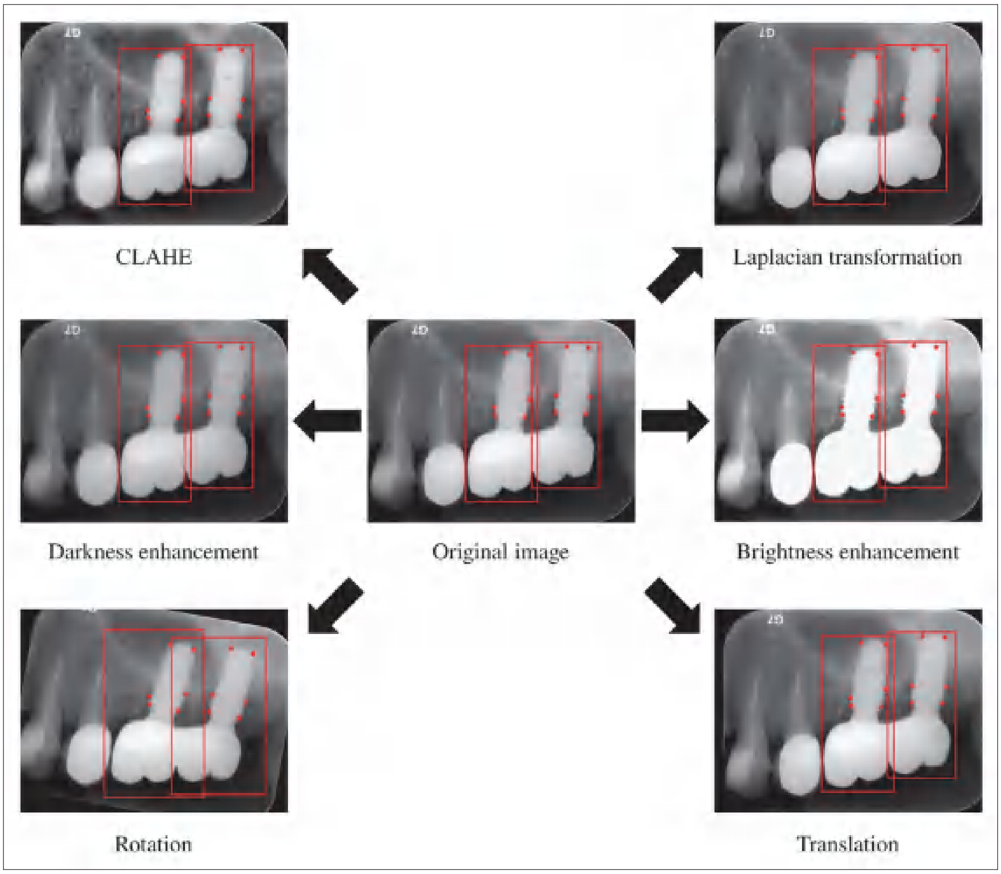

Fig 2 Data augmentation methods.

first contact points between the implant and the bone tissue) and the implant apex points (points 3 and 4). The distance between points 1 and 2, as well as that between points 5 and 6, represents the amount of bone loss.

The coordinates of the labelled keypoints were recorded by professional dental practitioners in an Excel sheet (Microsoft, Redmond, WA, USA), along with the patient ID and image ID. In this study, the original dataset was split into training, validation and test sets in a ratio of 8:1:1 to meet the requirements for model training, verification and evaluation. The allocation ratio ensured that the training set comprises the majority of the data to support the model's comprehensive learning, while reserving a sufficient portion for validation and test sets, enabling a thorough assessment of the model's generalisation ability and practical performance. Furthermore, to avoid deviations in model performance evaluation caused by data leakage, the principle of data independence was followed strictly, meaning that images from the same patient were assigned to only one set and could not appear in multiple sets simultaneously.

In this study, the labelled data annotated by dental professionals were converted into the format required for the pose task of the YOLOv8 model. The specific process is as follows: the target regions are labelled using the open-source software LabelImg and the bounding box information of the targets is recorded, then the coordinates of the keypoints annotated by professionals are normalised and sequentially documented in txt files corresponding to each image. Each line in a txt file represents one target. The values in [cls, x, y, w, h] specify the class, bounding box centre coordinates and dimensions, while the values in [xi, yi, visible] indicate the coordinates and visibility status of the keypoints. In this study, all keypoints are considered visible, with their visibility values consistently set to 2.

Given the limited sample size of the original dataset, this study expands the training and validation sets, as shown in Fig 2. The augmentation process simultaneously handled images and their corresponding labels, implementing various data enhancement strategies, such as rotation, translation, brightness adjustment and contrast enhancement. Rotation is applied at angles ranging from 1 to 10 degrees, in both clockwise and counterclockwise directions. Translation involves horizontal shifts of 4, 8, 12, 16 and 20 pixels. Brightness adjustment is performed by scaling

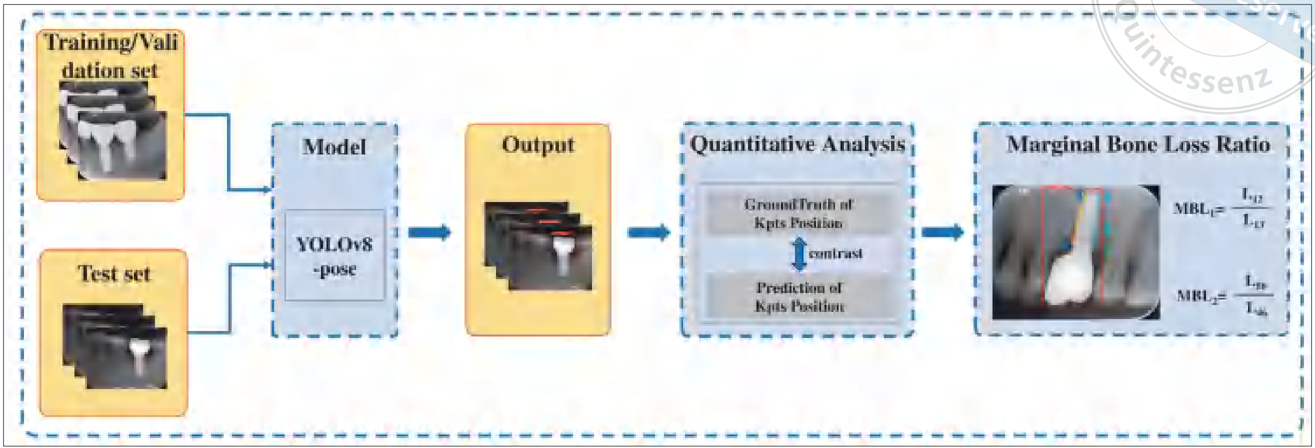

Fig 3 Flow diagram of the proposed methodology.

pixel intensities using factors of 0.9 for darkening and 1.4 for brightening. Contrast enhancement is carried out using two approaches: Laplacian transformation, which sharpens the image by emphasising edge features through convolution-based operations, and contrast-limited adaptive histogram equalisation (CLAHE), with a clip limit of 1.5 and a tile grid size of 8 × 8, to enhance local contrast while limiting noise amplification. Among these, brightness and contrast adjustments do not alter the original labels, whereas translation and rotation, being affine transformations, require corresponding modifications to the labels. The rotation adjustments were performed according to Formula 1, where (x,y) represents the coordinates of a point in the image plane, (c_x,c_y) denotes the image centre and the rotation pivot, θ is the rotation angle and (x',y') are the transformed coordinates after rotation. Similarly, translation adjustments followed Formula 2, where (tx,ty) represents the translation vector and (x",y") are the coordinates after the translation transformation. These transformations effectively increased the data sample size and enhanced the model's robustness to variations in positional information, lighting conditions and image quality, ultimately improving its generalisation capability. To ensure data integrity, the augmented samples were subjected to further filtering, excluding those where any keypoints extended beyond the image boundaries. After filtering, the final dataset comprised 3,940 images in the training set, 530 images in the validation set and 32 images in the test set.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & c_x(1-\cos\theta)+c_y\sin\theta \\ \sin\theta & \cos\theta & c_y(1-\cos\theta)-c_x\sin\theta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \textbf{(1)}$$

$$\begin{bmatrix} x'' \\ y'' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \textbf{(2)}$$

Overview of the workflow

This study aims to explore an auxiliary method for measuring peri-implant MBL on periapical radiographs within a deep learning framework, investigating the YOLOv8-Pose model and conducting training and testing on a private dataset to evaluate its performance and quantify indicators of MBL. The workflow of the proposed method is illustrated in Fig 3.

As shown in Fig 3, first, the YOLOv8-pose model was trained on a constructed dataset in this study, yielding a parameter-optimised version of the model. Subsequently, the best-trained weights were loaded to evaluate the model's performance on the test set, and the predicted coordinates of keypoints were extracted. Then, a quantitative analysis of MBL around implants was performed by comparing the ground truth coordinates of keypoints with their predicted values. Finally, the precision of the model in localising keypoints associated with MBL was assessed based on the results of the quantitative analysis, and the accuracy of the model in quantifying the severity of MBL was examined. This study integrates deep learning models with medical image analysis, enabling the efficient and accurate localisation of keypoints associated with peri-implant MBL. Additionally, it facilitates the quantitative analysis of bone loss based on the identified keypoints, providing critical data support and diagnostic insights for clinical applications. Furthermore, the methodology proposed in this research establishes a new paradigm for studying peri-implant marginal bone loss, offering an innovative approach to address this specific medical challenge. Beyond this, the successful implementation of the method presents a novel perspective and technical framework for research in other medical fields involving keypoint localisation and quantitative analysis.

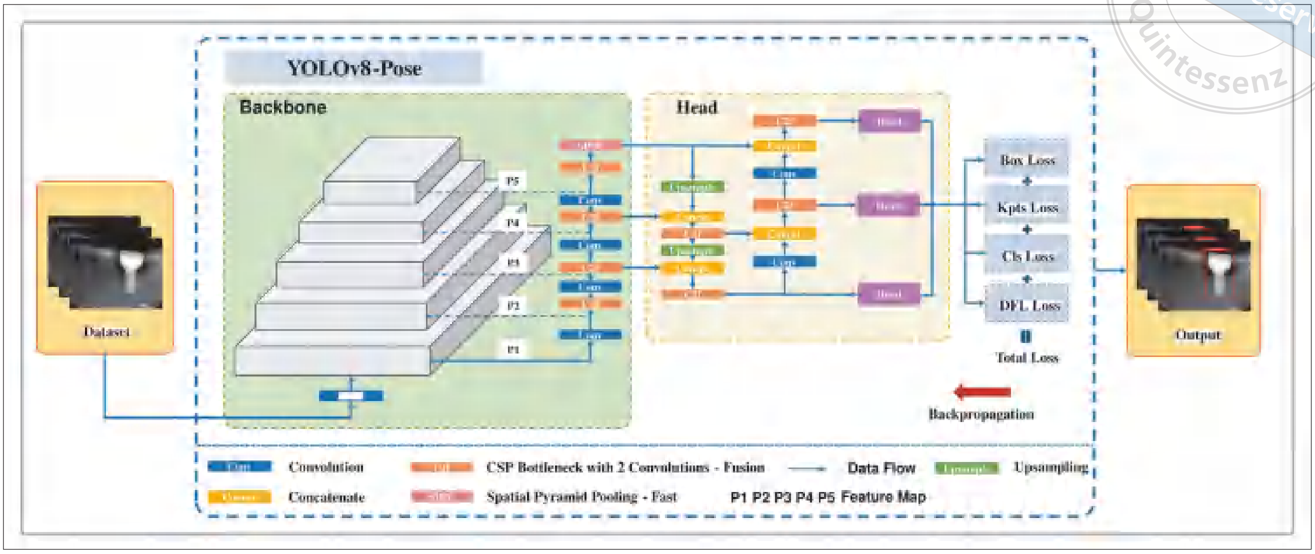

Fig 4 Input-output workflow of the YOLOv8-pose model.

Model architecture and details

YOLOv8 builds upon the core principles of the YOLO (You Only Look Once) framework, performing object detection through a single forward pass of a neural network. As illustrated in Fig 4, the YOLOv8 model consists of two main components: the backbone and the detection head. The backbone is responsible for extracting features from the input image, comprising multiple convolutional and CSP Bottleneck with 2 Convolutions-Fusion (C2f) units, and concludes with a Spatial Pyramid Pooling Fast (SPPF) unit. By progressively down-sampling, the backbone captures features ranging from low-level spatial details to high-level semantic information. The detection head is composed of a feature fusion layer and the final detection heads. The feature fusion layer integrates multi-scale features for detecting objects at different scales, employing Concat units to sum the channel dimensions while preserving the resolution of the feature map. Upsampling units double the feature map resolution to align with the C2f unit outputs. The detection head generates the final predictions, including bounding box coordinates, class confidences and, in this study, keypoint locations.

YOLOv8 supports multi-task outputs, including object detection, image segmentation and keypoint detection. Its Pose branch specialises in detecting and localising keypoints within an image, enabling human pose estimation. With its efficient end-to-end architecture, real-time performance, robustness and multi-task capabilities, YOLOv8-pose has demonstrated exceptional performance in human pose estimation. In the context of medical imaging, the application of keypoint detection remains relatively underexplored. However, as the technology matures, an increasing number of studies are investigating its potential in this domain. In this study, the YOLOv8-pose model is employed to detect keypoints related to peri-implant marginal bone loss in periapical radiographs. The performance of different YOLOv8-pose model variants is evaluated on a custom dataset, providing valuable baseline data for future research.

Loss functions

In the YOLOv8-pose model, the total loss (as shown in Formula (3)) is composed of several components: bounding box loss (L_{box}), keypoint loss (L_{kpts}), classification loss (L_{cls}) and distributive focal loss (L_{dfl}). Here, the λ values represent the loss gains, which balance the contributions of each loss component to the total loss. L_{box} measures the difference between the predicted and ground truth bounding boxes, optimising the position accuracy of the predicted boxes using the Intersection over Union (IoU) loss. L_{kpts} represents the keypoint loss, which is derived by combining L_{pose} and L_{kobj}, each weighted by their respective loss gains (as shown in Formula (4)). Specifically, L_{pose} quantifies the discrepancy between the predicted keypoint coordinates and the ground truth keypoints, with the loss computed based on Euclidean distance. On the other hand, L_{kobj} evaluates the difference between the model's prediction of keypoint visibility and the ground truth label, calculated using binary cross-entropy loss. L_{cls} measures the difference between the predicted and true class distributions, similarly calculated using binary cross-entropy loss, optimising the model's ability to predict the correct object of each predicted box as a probability distribution rather than a specific value. This approach accounts for the multiple potential errors between the predicted and ground truth boxes. The incorporation of L_{dfl} enhances the model's accuracy and robustness by better capturing the complexities of bounding box regression tasks.

$$L_{total}=\lambda_{box}\cdot L_{box}+\lambda_{cls}\cdot L_{cls}+\lambda_{dfl}\cdot L_{dfl}+L_{kpt} \quad \textbf{(3)}$$

$$L_{kpts}=\lambda_{pose}\cdot L_{pose}+\lambda_{kobj}\cdot L_{kobj} \quad \textbf{(4)}$$

Experimental settings

The model in this study is implemented using the PyTorch deep learning framework and executed on an NVIDIA 4080 GPU computing platform to accelerate the training process. During training, the number of epochs is set to 300, with a batch size of 16. Both the initial and final learning rates are configured at 0.01. The keypoint shape configuration parameter is defined as (6,3). The patience parameter is set to 100, meaning the training would terminate early if no significant improvement was observed within the last 100 epochs, thereby avoiding unnecessary resource consumption. The values of each loss gain are set as follows: the bounding box loss gain (λ_{box}), classification loss gain (λ_{cls}), distribution focal loss gain (λ_{dfl}), pose loss gain (λ_{pose}) and keypoint objectness loss gain (λ_{kobj}) are set to 7.5, 0.5, 1.5, 12 and 1, respectively. Additionally, no keypoint flipping or connection rules are applied during the training phase.

Evaluation indicators

This study conducts statistical analyses on the discrepancies between predicted values and ground truth to quantify the error distribution characteristics and evaluate the model's performance in the marginal bone loss quantification task. Metrics such as Euclidean distance (ED), mean Euclidean distance (MED) and mean radial error (MRE) are utilised to assess the model's predictive accuracy and consistency. Furthermore, to validate the model's robustness in clinical applications, ratio of MBL is introduced as the core quantification metric.

ED is a critical metric for quantifying the pixel-level discrepancy between predicted and ground truth keypoint coordinates. It is defined as the ED between the predicted and actual coordinates of a keypoint. A smaller ED indicates that the predicted keypoint position is closer to the ground truth, reflecting higher localisation accuracy. In this study, the index i ranges from 1 to 6, corresponding to the six keypoints in each image. As shown in Formula (5), P_i represents the predicted coordinates of the i-th keypoint, while G_i denotes the ground truth coordinates. The Euclidean distance between P_i and G_i is denoted as ED_i.

MED is computed as the mean ED across all keypoints in the test set, providing a representation of the model's localisation accuracy for different keypoints. As shown in Formula (6), N_T denotes the total number of samples. A smaller MED_i value signifies higher precision in the localisation of keypoints.

Additionally, MRE serves as a more comprehensive evaluation metric, measuring the mean radial error between predicted and ground truth keypoints localisation for all keypoints in a single image. MRE is defined as the mean of pixel-level errors across all keypoints, offering a holistic assessment of the model's localisation performance at the image level. As indicated in Formula (7), P_i and G_i hold the same meanings as in Formula (5) and Formula (6), while N denotes the total number of keypoints within the image. A lower MRE indicates that the predicted keypoints are closer to the true positions, demonstrating higher overall detection accuracy.

$$ED_i=\|\mathbf{P}_i-\mathbf{G}_i\|_2 \quad \textbf{(5)}$$

$$MED_i=\frac{\|\mathbf{P}_i-\mathbf{G}_i\|_2}{N_T} \quad (6)$$

$$\mathrm{MRE}=\frac{\sum_{i=1}^{N}\|\mathbf{P}_i-\mathbf{G}_i\|_2}{N} \quad (7)$$

Generally, 1 mm MBL during the first year followed by mean annual loss of 0.2 mm is considered acceptable.[45] However, MBL exceeding 2 mm or continuous loss thereafter may indicate pathological changes, such as peri-implantitis. The extent of MBL serves as a critical indicator for evaluating the severity of peri-implantitis,[46] highlighting the importance of quantitative analysis in this context. Although MBL can often be qualitatively observed by experts on radiographs, accurately quantifying the extent of MBL and assessing the progression of pathological changes remains a challenging and time-intensive task. The ratio of MBL offers a quantitative measure of MBL around the implant, calculated as the ratio of bone loss length to the total implant length.[34] According to Monje et al,[46] the severity of implant defects can be graded according to the MBL as slight (MBL < 25% of the implant length), moderate (MBL ≥ 25% to 50% of the implant length) or advanced (MBL > 50% of the implant length).

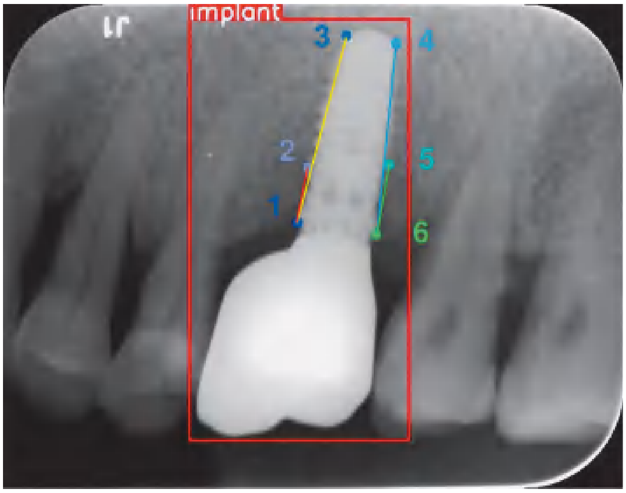

Fig 5 Schematic diagram of MBL length and the total implant length.

As illustrated in Fig 5 and Formula (8), L_{ij} represents the ED between two points. In Formulas (9) and (10), L_{12} and L_{56} denote the length of MBL, while L_{13} and L_{46} represent the total length of the implant. This study calculates the ratio of MBL on both sides of the implant using Formulas (9) and (10). Accuracy is employed as the evaluation metric, comparing the predicted severity of MBL with the ground truth values to assess the model's performance in quantifying bone loss severity and its potential clinical applicability. Accuracy is a commonly used indicator to measure the performance of classification, which reflects the proportion of correct predictions made by the model in all samples. As defined in Formula (11), N_T denotes the total number of samples and N_C denotes the number of correct predictions. A higher accuracy signifies the model's ability to effectively identify and quantify the severity of MBL from the data, providing reliable support for clinical diagnoses.

$$L_{ij}=\sqrt{(x_i\text{-}x_j)^2+(y_i\text{-}y_j)^2} \quad (8)$$

$$MBL_1=\frac{L_{12}}{L_{13}} \quad (9)$$

$$MBL_2=\frac{L_{56}}{L_{46}} \quad \mathbf{(10)}$$

$$Accuracy=\frac{N_c}{N_T} \quad \mathbf{(11)}$$

Through a comprehensive analysis of the evaluation metrics, this study not only offers a scientific foundation for optimising model performance but also introduces an efficient and reliable method for the precise quantification of MBL severity. This approach addresses the limitations of traditional qualitative analysis, laying the groundwork for automated and standardised assessment of MBL around dental implants.

Results

In this study, the present authors trained and evaluated the performance of five different YOLOv8-pose model variants on the task of predicting keypoints for MBL around dental implants. Specifically, Table 1 presents the performance on validation set and evaluation metrics for the five model variants: YOLOv8n-pose, YOLOv8s-pose, YOLOv8m-pose, YOLOv8l-pose and YOLOv8x-pose.

Furthermore, Fig 6 illustrates the training performance of the five model variants across various metrics. Considering the model's effectiveness and performance, the present authors proceed to evaluate the YOLOv8l-pose model's performance on the test set. In this study, the YOLOv8l-pose model demonstrates exceptional performance in predicting keypoints associated with MBL around dental implants on the test set. From a quantitative analysis perspective, for each image, the ED between the predicted and ground truth coordinates of the six keypoints, as well as the MRE across all keypoints within the same images, are summarised in Table 2, which shows that remarkable performance was exhibited by the trained model in the keypoint localisation task. To provide a more intuitive representation of the results, Fig 7 visualises the ED between predicted values and ground truth for each of the six keypoints in individual images. Meanwhile, Fig 8 illustrates the MRE of all keypoints within single images. As shown in Fig 9, the MED between predicted values and ground truth for each keypoint on the test set is presented.

As illustrated in Fig 10a, we compare the ratio of predicted and ground truth values of MBL on both sides of the implants. Based on the staging criteria mentioned in previous studies,[46] the model's predictions for the severity of MBL on the test set are analysed, as shown in Fig 10b.

From a qualitative analysis perspective, the performance of the YOLOv8l-pose model in predicting keypoints for MBL around dental implants on the test set is extensively investigated in this study. As shown in Fig 11, the "Ground Truth" represents the gold-standard keypoint locations annotated by professional clinicians, whereas the "Prediction" denotes the model's predicted keypoint locations for the test set.

Table 1 Performance evaluation of YOLOv8-pose model variants

Model	P^{pose}	mAP50-95(pose)	Parameter count	Training duration/h
YOLOv8n-pose	0.991	0.990	3.1	1.053
YOLOv8s-pose	0.984	0.994	11.4	1.187
YOLOv8m-pose	0.982	0.994	26.4	1.698
YOLOv8l-pose	0.999	0.994	44.4	2.103
YOLOv8x-pose	1	0.993	69.4	3.065

Table 2 ED and MRE between prediction and ground truth (pixels)

Tooth ID	ED_1	ED_2	ED_3	ED_4	ED_5	ED_6	MRE
3008567	6.32	17.12	14.14	16.16	34.56	10.00	16.38
3008568	2.00	10.05	8.60	11.40	33.24	7.07	12.06
3008577	3.16	22.00	6.32	7.07	2.24	7.62	8.07
3008578	4.47	7.81	9.49	5.66	12.04	4.12	7.27
3008585	2.83	3.00	3.61	2.83	2.00	4.12	3.06
3008933	6.32	2.24	3.16	7.28	12.04	9.49	6.76
3008935	5.39	1.00	1.41	1.00	10.20	5.83	4.14
3008938	3.16	11.70	4.00	3.16	4.47	5.39	5.31
3008940	3.61	5.10	2.24	1.00	1.41	4.00	2.89
3008954	3.16	11.18	10.30	2.24	3.00	5.66	5.92
3008957	17.00	4.47	5.66	6.32	5.10	15.52	9.01
3008959	8.06	9.22	7.62	7.81	1.00	7.28	6.83
3008974	8.49	5.00	6.71	7.62	8.06	13.04	8.15
3008986	2.00	3.16	2.83	6.71	9.49	3.16	4.56
3008990	9.22	4.00	5.39	9.49	8.54	1.00	6.27
3009006	1.41	5.00	2.00	6.32	4.47	2.00	3.54
3009007	3.16	10.77	3.61	3.61	10.20	11.05	7.06
3009202	5.00	11.40	3.16	2.83	1.00	6.00	4.90
3009210	7.28	11.31	14.42	10.77	13.15	1.41	9.73
3009211	6.40	7.62	17.69	12.81	9.49	2.83	9.47
3009212	3.61	3.16	19.10	15.56	11.70	7.28	10.07
3009215	11.40	9.85	6.71	10.30	9.00	5.00	8.71
3009219	8.94	8.06	5.00	4.12	2.24	6.40	5.79
3009220	5.39	5.10	5.39	7.07	11.18	6.08	6.70
3009225	5.10	5.39	4.47	4.47	12.17	6.08	6.28
3009242	6.08	5.00	6.08	7.07	9.90	8.60	7.12
3009248	9.49	7.81	4.24	6.08	5.00	12.17	7.46
3009250	13.15	4.24	5.00	3.16	4.47	11.40	6.91
3009255	3.61	6.00	1.00	7.62	7.62	4.47	5.05
3009256	2.83	10.20	6.32	3.16	12.00	9.06	7.26
3009257	4.12	2.24	9.22	6.08	14.87	7.07	7.27
3009258	11.18	14.56	11.00	11.40	5.10	14.32	11.26

Discussion

This study comprehensively evaluates the performance of the YOLOv8l-pose model in localising keypoints and quantifying MBL around dental implants.

As shown in Table 1, Precision (P) measures the proportion of correctly identified positive samples out of all predicted positives. Importantly, higher precision indicates fewer false positives, signifying better alignment between predicted keypoints and ground truth. Meanwhile, mAP50-95 represents the mean of average precision (AP) across IoU thresholds ranging from 0.50 to 0.95. This metric is widely regarded as a robust standard for assessing object detection performance. A higher mAP50-95 score reflects stronger robustness of the model under varying IoU thresholds. Given the focus on keypoints associated with MBL, P^{pose} was employed to evaluate the model's prediction accuracy in keypoint detection. Additionally, mAP50-95(pose) was used to assess the precision of the predicted keypoint loca-

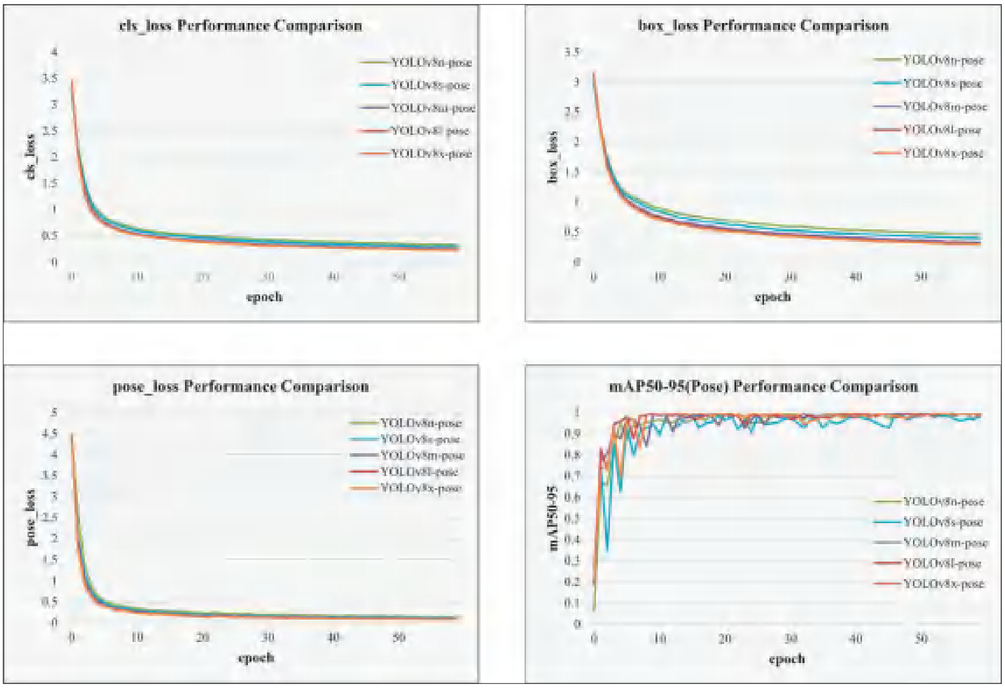

Fig 6 Comparison of loss functions and evaluation metrics.

tions. As shown in Fig 6, all models demonstrated rapid convergence, achieving high mAP50-95(pose) scores. Among these, the larger models exhibited faster loss convergence, with the YOLOv8l-pose model achieving a mAP50-95(pose) score exceeding 0.95 at the fastest rate. Additionally, considering the Ppose metric and training time shown in Table 1, the YOLOv8l-pose model is selected for the quantitative analysis of MBL around dental implants.

From a quantitative analysis perspective, as shown in Table 2, the results demonstrate that YOLOv8l-pose achieves high precision and robustness in localising peri-implant keypoints and quantifying MBL. Specifically, the ED for the majority of keypoints is observed to be less than 10 pixels, with the minimum localisation error reaching an impressive accuracy level of 1.00 pixel. This result strongly validates the model's high accuracy and stability in the keypoint localisation task for peri-implant marginal bone loss. Notably, the model exhibits excellent performance in predicting the MRE for six keypoints in individual images, with the MRE for most images also being less than 10 pixels. These findings confirm the robustness and generalisation capability of the model. Figure 7 visually demonstrates the precise alignment between predicted values and ground truth for each of the six keypoints across individual images, highlighting the model's exceptional localisation accuracy. The model exhibits remarkably high prediction accuracy, as evidenced by the results in Fig 8, where the minimum MRE across all keypoints for a single image is only 2.89 pixels, and the maximum MRE reaches just 16.38 pixels. These findings further validate the model's reliability and practical applicability.

Nevertheless, as shown in Fig 9, slightly higher prediction errors are observed at bone contact points (kpt2 and kpt5). The following factors may be attributed to this phenomenon: first, bone contact points are typically located at the interface between the implant edge and bone tissue, where low image contrast and relatively vague feature information are present; second, due to the influence of shooting angles and imaging quality, substantial variation in the morphological characteristics of bone contact points is exhibited across different samples, thereby increasing the learning difficulty for the model. Insights regarding the improvement of implant edge bone resorption quantification are provided by these results.

In addition, the study demonstrates high accuracy in predicting the severity of MBL. As illustrated in Fig 10a, for this critical clinical indicator, the model demonstrates excellent performance in quantifying the extent of MBL. On the test set, the model's predicted MBL values (Pred_MBL) align closely with the ground truth MBL values (GT_MBL) for most of the samples. As illustrated in Fig 10b, its strong performance in terms of MBL indicates that a high level of predictive accuracy in assessing the severity of MBL is exhibited by the model,

Fig 7 Performance of ED on the test set.

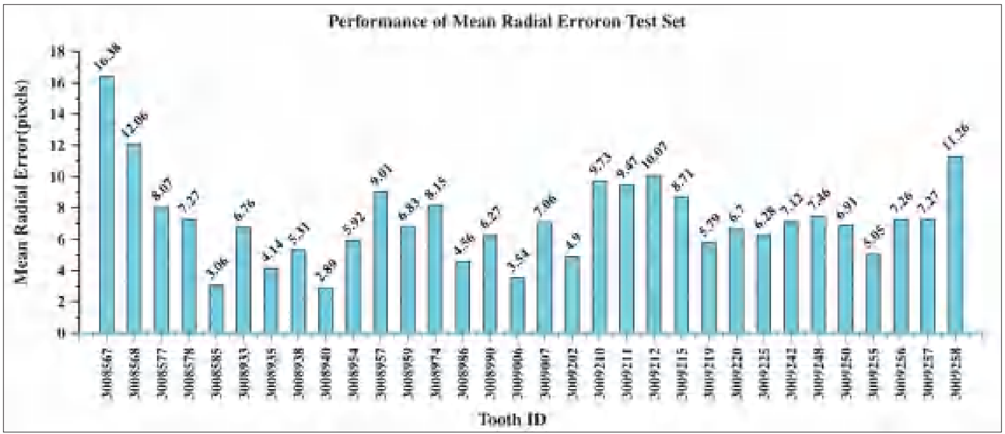

Fig 8 Performance of MRE on the test set.

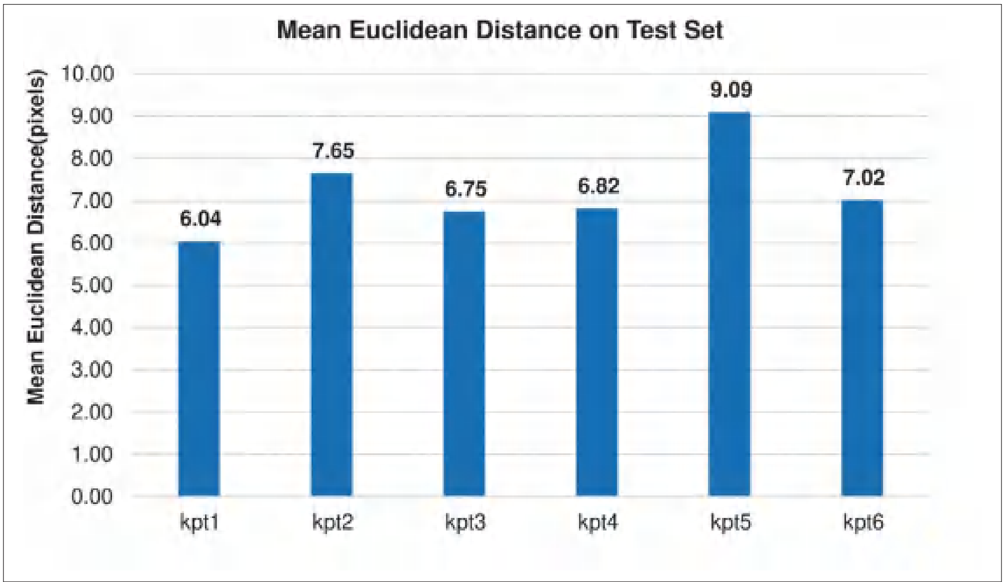

Fig 9 Performance of MED on the test set.

further supporting its effectiveness in clinical applications. Valuable insights are provided by this study into the use of AI for regularly assessing the progression of implant MBL, evaluating risks, and maintaining implant health. From a qualitative analysis perspective, as shown in Fig 11, a high degree of alignment between the predicted keypoint positions and the ground truth positions is revealed by a comparison, demonstrating excellent localisation accuracy and stability. Very close alignment is observed between all predicted keypoint positions and the ground truth, especially for the positions of the highest points of the implant neck and the implant apex. The model's outstanding performance on the test set highlights its high potential for accurately locating keypoints associated with marginal bone loss around implants, providing reliable technical support for medical diagnosis and research.

Limitations

Despite the excellent performance of the YOLOv8-pose model in predicting MBL keypoints around implants on the test set, several limitations remain. First, constrained by a relatively small training dataset that was insufficient to cover the wide variety of complex scenarios associated with keypoints of MBL around implants, the study is subject to limitations. Consequently, the model's generalisation capability on the test set may be limited, particularly when extreme or atypical image samples are encountered, where all keypoints may struggle to be consistently predicted by the model.

Second, the significant variation in greyscale values within the gingival tissue around the implant margin poses challenges for the model's robustness in detecting MBL keypoints. In certain images, large deviations in the prediction of bone-level points reflect the model's

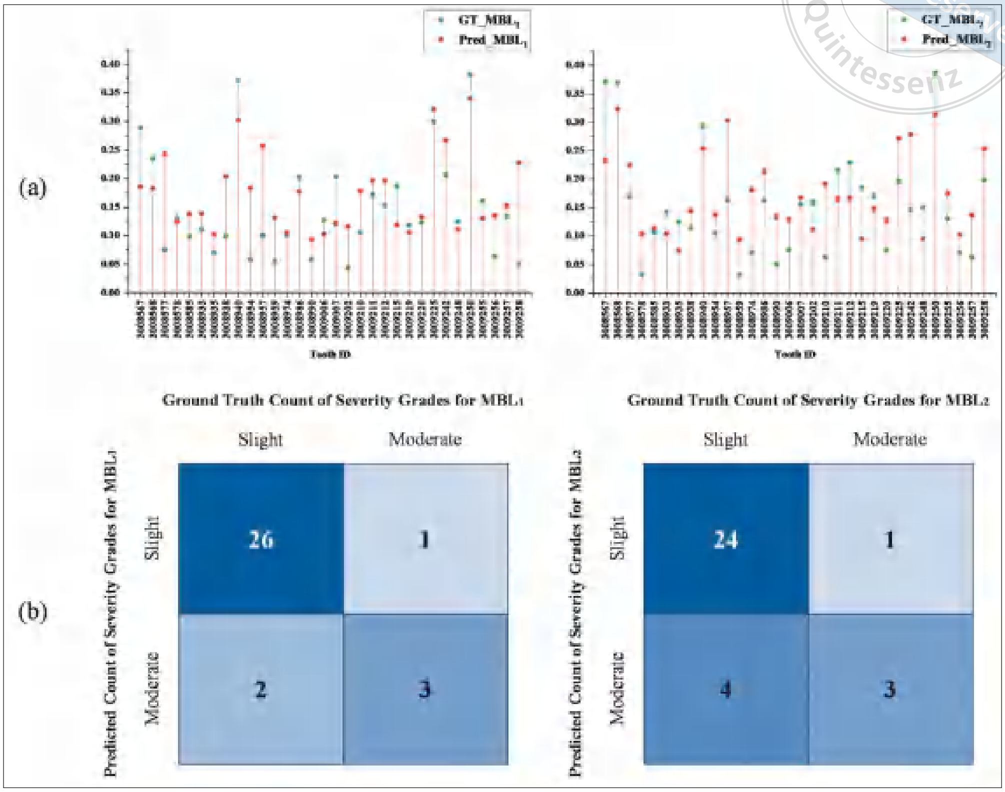

Fig 10a and b Performance of MBL on the test set: comparison of the predicted MBL values and ground truth MBL values on the test set (a); comparison of predicted MBL severity and ground truth values on the test set (b).

limited adaptability to regions with variable greyscale distribution. Furthermore, artifacts from the implant or occlusions caused by gingival tissue can restrict feature extraction, affecting the model's ability to accurately predict keypoints of MBL around the implant. This interference highlights the need for the model to demonstrate greater resilience to such complexities.

Conclusion

In this study, an innovative intelligent technological solution is proposed to address the clinical challenge of quantifying peri-implant MBL, providing a new approach for achieving automated and precise measurement of bone loss. The issues associated with traditional manual measurement methods, such as low efficiency, subjectivity and poor repeatability, are not only resolved in this research, but new pathways are also opened for the early diagnosis and precise treatment of peri-implant diseases. Specifically, the first multicentre keypoint database for measuring peri-implant MBL is established, laying a solid foundation for further exploration in the field of implant bone loss. The proposed method demonstrated strong robustness, accurately detecting bone loss keypoints even under varying greyscale conditions, and achieved precise quantification with high consistency. Furthermore, by combining deep learning algorithms with medical analysis of implant apical radiographs, the research successfully achieved full-process automation from feature detection to quantitative analysis of peri-implant MBL, significantly improving measurement efficiency. On the clinical application level, the efficient detection and precise quantification of peri-implant MBL provide reliable data support for the scientific evaluation of implant long-term stability, offering significant guidance for clinical diagnosis and treatment planning, and contributing to the improvement of long-term implant success rates.

Conflicts of interest

The authors declare no conflicts of interest related to this study.

Author contribution

Dr Tao GAO contributed to the conceptualisation, data collection, and label annotation, and participated in the scheme design and manuscript revision; Dr Rong TAN developed the detailed methodology, performed experimental analysis and data processing and drafted the manuscript; Dr Su Kun TIAN provided guidance on

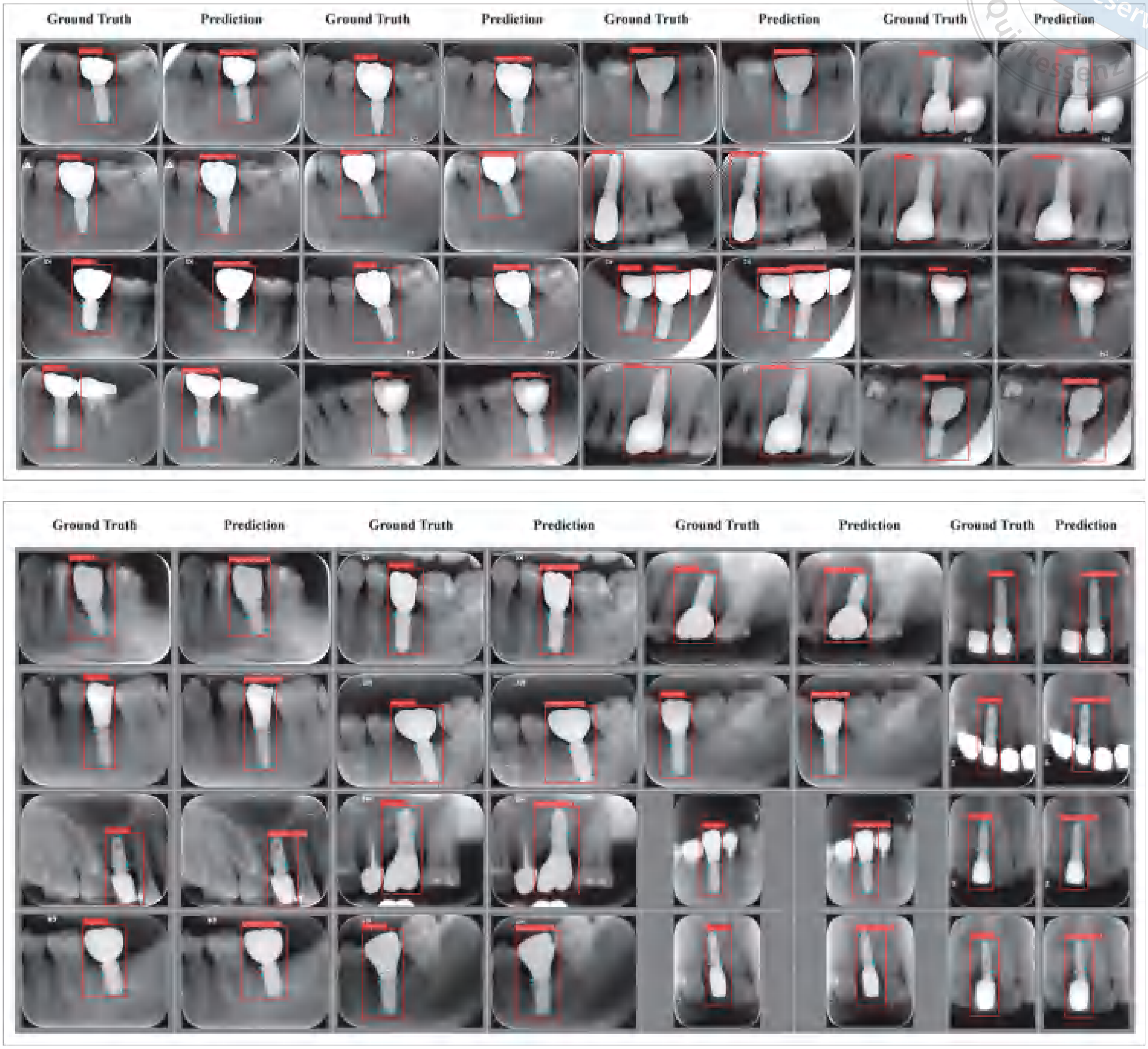

Fig 11 Comparison of predicted keypoints and ground truth for MBL around implants on the test set.

the scheme design and manuscript review; Dr Hu CHEN was involved in guiding the data collection and labelling, and assisted in the scheme design; Dr Hao Ran WANG assisted in designing the scheme; Dr Jian Guo XU guided the scheme design, implementation, experimental analysis, and manuscript revision.

(Received Mar 07, 2025; accepted Jul 10, 2025)

References

1. Albrektsson T, Donos N, Working Group 1. Implant survival and complications. The Third EAO consensus conference 2012. Clin Oral Implants Res 2012;23(suppl 6):63–65.
2. French D, Ofec R, Levin L. Long term clinical performance of 10871 dental implants with up to 22 years of follow-up: A cohort study in 4247 patients. Clin Implant Dent Relat Res 2021;23:289–297.

3. Lai HC, Zhang ZY, Wang F, Zhuang LF, Liu X, Pu YP. Evaluation of soft-tissue alteration around implant-supported single-tooth restoration in the anterior maxilla: The pink esthetic score. Clin Oral Implants Res 2008;19:560–564.
4. Kim JJ, Lee JH, Kim JC, Lee JB, Yeo IL. Biological responses to the transitional area of dental implants: Material-and structure-dependent responses of peri-implant tissue to abutments. Materials(Basel) 2019;13:72.
5. Pope J, Harrel S. Advanced therapeutics for peri-implant problems. Clin Dent Rev 2020;4:7.
6. Maghsoudi P, Slot DE, van der Weijden FGA. Bone remodeling around dental implants after 1-1.5 years of functional loading: A retrospective analysis of two-stage implants. Clin Exp Dent Res 2022;8:680–689.
7. Schwartz-Arad D, Herzberg R, Levin L. Evaluation of long-term implant success. J Periodontol 2005;76:1623-1628.
8. Galindo-Moreno P, Catena A, Pérez-Sayáns M, Fernández-Barbero JE, O'Valle F, Padial-Molina M. Early marginal bone loss around dental implants to define success in implant dentistry: A retrospective study. Clin Implant Dent Relat Res 2022;24:630-642.
9. Hermann JS, Cochran DL, Nummikoski PV, Buser D. Crestal bone changes around titanium implants. A radiographic evaluation of unloaded nonsubmerged and submerged implants in the canine mandible. J Periodontol 1997;68:1117–1130.
10. El Hage M, Nurdin N, Abi Najm S, Bischof M, Nedir R. Osteotome sinus floor elevation without grafting: A 10-year study of cone beam computerized tomography vs periapical radiography. Int J Periodontics Restorative Dent 2019;39:e89–e97.
11. Zou R, Zhang J, Wu Y. Skin lesion segmentation through generative adversarial networks with global and local semantic feature awareness. Electronics 2024;13:3853.
12. Huang P, He P, Tian S, et al. A ViT-AMC network with adaptive model fusion and multiobjective optimization for interpretable laryngeal tumor grading from histopathological images. IEEE Trans Med Imaging 2023;42:15–28.
13. Tian S, Wang M, Yuan F, et al. Efficient computer-aided design of dental inlay restoration: A deep adversarial framework. IEEE Trans Med Imaging 2021;40:2415–2427.
14. Tian S, Wang M, Dai N, et al. DCPR-GAN: Dental crown prosthesis restoration using two-stage generative adversarial networks. IEEE J Biomed Health Inform 2022;26:151–160.
15. Lee JH, Kim DH, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci 2018;48:114–123.
16. Casalegno F, Newton T, Daher R, et al. Caries detection with near-infrared transillumination using deep learning. J Dent Res 2019;98:1227–1233.
17. Erturk M, Öziç MÜ, Tassoker M. Deep convolutional neural network for automated staging of periodontal bone loss deverity on bite-wing radiographs: An Eigen-CAM explainability mapping approach. J Imaging Inform Med 2025;38:556–575.
18. Cerda Mardini D, Cerda Mardini P, Vicuña Iturriaga DP, Ortuño Borroto DR. Determining the efficacy of a machine learning model for measuring periodontal bone loss. BMC Oral Health 2024;24:100.
19. Shen X, Zhang C, Jia X, et al. TranSDFNet: Transformer-based truncated signed distance fields for the shape design of removable partial denture clasps. IEEE J Biomed Health Inform 2023;27:4950–4960.
20. Chen G, Qin J, Amor BB, et al. Automatic detection of tooth-gingiva trim lines on dental surfaces. IEEE Trans Med Imaging 2023;42:3194-3204.
21. Mei L, Fang Y, Zhao Y, et al. DTR-Net: Dual-space 3D tooth model reconstruction from panoramic X-Ray images. IEEE Trans Med Imaging 2024;43:517–528.
22. Liu Z, He X, Wang H, et al. Hierarchical self-supervised learning for 3D tooth segmentation in intra-oral mesh scans. IEEE Trans Med Imaging 2023;42:467-480.
23. Li P, Gao C, Lian C, Meng D. Spatial prior-guided bi-directional cross-attention transformers for tooth instance segmentation. IEEE Trans Med Imaging 2024;43;3936–3948.
24. Du M, Liang K, Zhang L, Gao H, Liu Y, Xing Y. Deep-learning-based metal artefact reduction with unsupervised domain adaptation regularization for practical CT images. IEEE Trans Med Imaging 2023;42:2133–2145.
25. Alharbi SS, Alhasson HF. Exploring the applications of artificial intelligence in dental image detection: A systematic review. Diagnostics (Basel) 2024;14:2442.
26. Şalgău CA, Morar A, Zgarta AD, et al. Applications of machine learning in periodontology and implantology: A comprehensive review. Ann Biomed Eng 2024;52:2348–2371.
27. Mohammad-Rahimi H, Motamedian SR, Pirayesh Z, et al. Deep learning in periodontology and oral implantology: A scoping review. J Periodontal Res 2022;57:942–951.
28. Sukegawa S, Yoshii K, Hara T, et al. Deep neural networks for dental implant system classification. Biomolecules 2020;10:984.
29. Hadj Saïd M, Le Roux MK, Catherine JH, Lan R. Development of an artificial intelligence model to identify a dental implant from a radiograph. Int J Oral Maxillofac Implants 2020;35:1077–1082.
30. Vera M, Gómez-Silva MJ, Vera V, et al. Artificial intelligence techniques for automatic detection of peri-implant marginal bone remodeling in intraoral radiographs. J Digit Imaging 2023;36:2259–2277.
31. Lee DW, Kim SY, Jeong SN, Lee JH. Artificial intelligence in fractured dental implant detection and classification: Evaluation using dataset from two dental hospitals. Diagnostics (Basel) 2021;11:233.
32. Zhang C, Fan L, Zhang S, Zhao J, Gu Y. Deep learning based dental implant failure prediction from periapical and panoramic films. Quant Imaging Med Surg 2023;13:935–945.
33. Yang X, Li X, Li X, et al. Two-stream regression network for dental implant position prediction. Expert Syst Appl 2024;235:121135.1–121135.10.
34. Cha JY, Yoon HI, Yeo IS, Huh KH, Han JS. Peri-implant bone loss measurement using a region-based convolutional neural network on dental periapical radiographs. J Clin Med 2021;10:1009.
35. Liu M, Wang S, Chen H, Liu Y. A pilot study of a deep learning approach to detect marginal bone loss around implants. BMC Oral Health 2022;22:11.
36. Lee WF, Day MY, Fang CY, et al. Establishing a novel deep learning model for detecting peri-implantitis. J Dent Sci 2024;19:1165–1173.
37. Jocher G, Chaurasia A, Qiu J. Ultralytics YOLOv8. 2023. https://github.com/ultralytics/ultralytics. Accessed 1 July 2025.
38. Liu S, Zhang T, Ji H, et al. A novel YOLOv8-pose-based algorithm for abnormal behavior analysis of quadruped robots. In 2024 IEEE 13th Data Driven Control and Learning Systems Conference 2024:1538–1543.
39. Sledevič T, Serackis A, Matuzevičius D, Plonis D, Andriukaitis D. Keypoint-based bee orientation estimation and ramp detection at the hive entrance for bee behavior identification system. Agriculture 2024;14:1890.

40. Li J, Liu Y, Zheng W, Chen X, Ma Y, Guo L. Monitoring cattle ruminating behavior based on an improved keypoint detection model. Animals (Basel) 2024;14:1791.
41. Janisch L, Schulz D, Schmidt A, Reisch R, Kamps T, Franke J. Comparative study of keypoint detection and ArUco marker methods for optical 6D pose estimation in electronics packaging. In 2024 IEEE 20th International Conference on Automation Science and Engineering 2024:3969–3974.
42. Ren B, Zheng X, Guan T, Wang J. Vibrator rack pose estimation for monitoring the vibration quality of concrete using improved YOLOv8-pose and vanishing points. Buildings 2024;14:3174.
43. Rashmi S, Srinath S, Deshmukh S, Prashanth S, Patil K. Cephalometric landmark annotation using transfer learning: Detectron2 and YOLOv8 baselines on a diverse cephalometric image dataset. Comput Biol Med 2024;183:109318.
44. Dalago HR, Schuldt Filho G, Rodrigues MA, Renvert S, Bianchini MA. Risk indicators for peri-implantitis. A cross-sectional study with 916 implants. Clin Oral Implants Res 2017;28:144–150.
45. Albrektsson T, Zarb G, Worthington P, Eriksson AR. The long-term efficacy of currently used dental implants: A review and proposed criteria of success. Int J Oral Maxillofac Implants 1986;1:11–25.
46. Monje A, Pons R, Insua A, Nart J, Wang HL, Schwarz F. Morphology and severity of peri-implantitis bone defects. Clin Implant Dent Relat Res 2019;21:635–643.