BMC Oral Health

RESEARCH

Open Access

# Cone-beam CT landmark detection for measuring basal bone width: a retrospective validation study

Juan Dai[1†], Xinge Guo[2,3†], Hongyuan Zhang[2,3†], Haoyu Xie[2,3], Jiahui Huang[2,3], Qiangtai Huang[2,3] and Bingsheng Huang[2,3*]

## Abstract

**Background** Accurate assessment of basal bone width is essential for distinguishing individuals with normal occlusion from patients with maxillary transverse deficiency who may require maxillary expansion. Herein, we evaluated the effectiveness of a deep learning (DL) model in measuring landmarks of basal bone width and assessed the consistency of automated measurements compared to manual measurements.

**Methods** Based on the U-Net algorithm, a coarse-to-fine DL model was developed and trained using 80 cone-beam computed tomography (CBCT) images. The model's prediction capabilities were validated on 10 CBCT scans and tested on an additional 34. To evaluate the performance of the DL model, its measurements were compared with those taken manually by one junior orthodontist using the concordance correlation coefficient (CCC).

**Results** It took approximately 1.5 s for the DL model to perform the measurement task in only CBCT images. This framework showed a mean radial error of 1.22 ± 1.93 mm and achieved successful detection rates of 71.34%, 81.37%, 86.77%, and 91.18% in the 2.0-, 2.5-, 3.0-, and 4.0-mm ranges, respectively. The CCCs (95% confidence interval) of the maxillary basal bone width and mandibular basal bone width distance between the DL model and manual measurement for the 34 cases were 0.96 (0.94–0.97) and 0.98 (0.97–0.99), respectively.

**Conclusion** The novel DL framework developed in this study improved the diagnostic accuracy of the individual assessment of maxillary width. These results emphasize the potential applicability of this framework as a computer-aided diagnostic tool in orthodontic practice.

**Keywords** Basal bone width, Maxillary transverse deficiency, CBCT, Deep learning

†Juan Dai, Xinge Guo and Hongyuan Zhang contributed equally to this work.

*Correspondence:
Bingsheng Huang
huangb@szu.edu.cn

[1]Department of Stomatology, Shenzhen University General Hospital, Shenzhen University Clinical Medical Academy, Xueyuan Avenue, Nanshan District, Shenzhen, Guangdong 518055, China
[2]Medical AI Lab, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen 518060, China
[3]Guangdong Key Laboratory of Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen 518060, China

## Background

Maxillary transverse deficiency (MTD) manifests as insufficient width in the maxilla, leading to disharmony between maxillary and mandibular widths [1, 2]. This dentofacial deformity is notably prevalent among adult orthodontic patients, with an approximate incidence rate of 30% among individuals seeking orthodontic treatment, impacting craniofacial development, function, aesthetics, and stability [2, 3]. Early assessment of the craniofacial skeleton in the transverse dimension is crucial to accurately determine the necessity for transverse maxillary expansion, thereby enhancing the efficiency and effectiveness of treatment. Consequently, there has been an ongoing evolution and development of diagnostic tools for evaluating the maxillary transverse dimension. Evaluation of maxillary width entails three primary measurements: maxillary dental arch width, alveolar arch width, and basal bone width [4]. Among these, basal bone width assessment is critical for distinguishing between Class I individuals with normal occlusion and Class III patients requiring surgical intervention [5, 6]. Thus, precise measurement of basal bone width is essential for delineating various malocclusions and formulating personalized treatment plans [5, 7].

Cone-beam computed tomography (CBCT) stands as the primary imaging modality in dental disease diagnosis and orthodontic treatment planning, enjoying widespread utilization in clinical settings [8, 9]. In accordance with the Yonsei analysis method [6], the tooth's center of resistance (CR) serves as a dependable indicator of its position. Hence, we chose the center point of the root bifurcation as the reference for measuring the width of the maxillary and mandibular basal bones (Fig. 1). This point also serves as the resistance center of the first molar, providing a critical reference for our analysis. However, precise localization of the center point of root bifurcation on upper and lower first permanent molars using CBCT images, which is crucial for accurate basal bone width measurement, poses a formidable challenge, even for experienced oral surgeons [10, 11]. This task involves meticulously identifying the position of the first molar, followed by accurately determining the center point at the root bifurcation. Such a process is time-consuming and demands a high level of radiographic interpretation proficiency and expertise.

In our study, we selected CBCT over MRI, traditional CT, and 2D digital radiography because of its particular advantages for orthodontic applications. CBCT provides high-resolution three-dimensional images that are essential for the precise identification of orthodontic landmarks, which is crucial for accurately measuring basal bone width. Unlike MRI, which although offers excellent soft tissue contrast, lacks detailed imaging of bone structures, CBCT excels in providing detailed visualizations of bone tissues. Additionally, compared to conventional CT scans, which subject patients to higher radiation levels, CBCT is favored because of its significantly lower radiation dose and shorter scanning times. These characteristics are particularly valuable in fields like dental implantology, orthodontics, and endodontics, where exact measurements of bone and tooth structures are imperative [12, 13]. The ability of CBCT to deliver high-quality images with minimal radiation exposure makes it ideal for complex dental and maxillofacial surgical planning. Furthermore, it is more cost-effective and accessible than both MRI and CT—advantages that render it highly suitable for frequent dental evaluations. Consequently, CBCT stands out as the most practical and appropriate choice for conducting detailed, efficient, and safe craniofacial assessments in our research.

DL has been extensively utilized to analyze oral images [14], representing one of the most advanced machine learning techniques that enables precise localization of various anatomical landmarks [15–18]. However, previous studies have primarily focused on detecting orthodontic landmarks in radiographic images [15, 18] rather than specifically targeting precise basal bone width measurement. To the best of our knowledge, no study has reported on the utilization of DL specifically for this purpose.

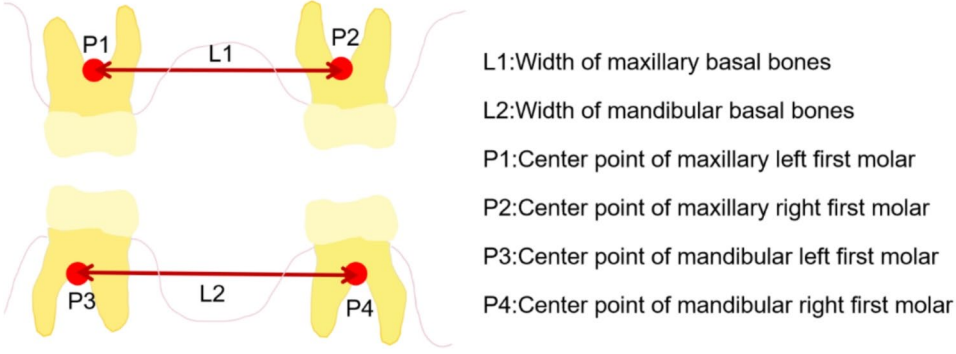


**Fig. 1** An example of landmarks and the widths of the maxilla and mandible

This study aimed to develop a DL model for detecting landmarks on CBCT images to facilitate basal bone width measurement. We hypothesized that there would be no difference between the automated measurements of the DL model and those conducted by a professional dental surgeon.

## Materials and methods

### Dataset description

In accordance with the Declaration of Helsinki [19], this retrospective study obtained approval from the Medical Ethics Committee of Shenzhen General Hospital (KYLL-20221219 A). A cohort of 124 patients (77 females and 47 males) with a mean age of 19.7 years (age range: 10–45 years) was included in the study, contributing to the comprehensive collection of 124 cone-beam computed tomography (CBCT) images. It is also noticeable that the involved patients underwent CBCT scans for oral health assessment and diagnosis of oral diseases between 2018 and 2023. No specific inclusion criteria were applied to force the established DL model to adapt as much as possible, as is needed in real-life scenarios. However, the following exclusion criteria were applied: (a) cases lacking first molars; (b) CBCT images with extensive metal artefacts due to the presence of orthodontic brackets or radioactive material restorations adjacent to dental implants which may make landmarks difficult to detect; and (c) motion artefacts causing blurred CBCT image quality. The patients underwent scanning using the Kavo 3D eXami CBCT unit (Imaging Sciences International, LLC, USA). The scanning protocol comprised a voxel size of 0.2 mm, a duration of 23 s, and a field of view (14×8 cm) with settings at 120 kV and 5 mA. All scans adhered to the manufacturer's recommended procedures, and all image data sets were acquired and stored in the Digital Imaging and Communications in Medicine (DICOM) format for subsequent analysis. The CBCT average image size was specified as '710×710×565', where '710×710' represents the dimensions of the transverse plane and '565' represents the corresponding depth size. Each voxel corresponded to a volume of 0.2 mm × 0.2 mm × 0.2 mm, and the image resolution within each slice had pixels of 0.1 mm in width and height.

### Model preparation

To ensure robust evaluation, the dataset was divided into three distinct subsets using random partition. The training set ($n$=80) was utilized to train the DL model, while the validation set ($n$=10) served for internal validation through hyperparameter optimization. Finally, the testing set ($n$=34) was employed to assess the performance of the DL model in comparison to manual annotation. These three datasets were mutually exclusive, with no data leakage.

### Landmark annotation

Manual landmark annotations were conducted using 3D Slicer software (version 5.0.2; 3D Slicer community and coordinated by Kitware, Inc., USA, and the Surgical Planning Laboratory at Harvard Medical School, USA; https://www.slicer.org/), which provides tools and features that facilitate precise landmark localization and measurement [20]. Two experienced doctors, each with >5 years of experience, performed the annotations. Both doctors underwent training on the specific techniques and guidelines for landmark identification and annotation using 3D Slicer software. This training aimed to standardize their approach and minimize interobserver variability. Moreover, a double-check approach was implemented. After the initial annotations were made by one doctor, the second doctor independently reviewed and verified the landmarks. Any discrepancies or disagreements were resolved through discussion and consensus between the two doctors. This iterative process helped to ensure the accuracy and consistency of the annotations. Final annotations were saved as location data in the JSON file format. Figure 2 shows an example of CBCT slices from different angles (axial, coronal, and sagittal) with the identified landmarks highlighted.

### Overall algorithm of the landmark detection framework

The landmark detection framework in this study comprised two stages, which involved the regression of 4-channel heatmaps for landmarks in a coarse-to-fine manner, as illustrated in Fig. 3. Both stages utilized the same U-Net structure (Fig. 3c) but were assigned different learning scopes. Notably, our approach exclusively relied on 3D views for capturing landmarks. This method allowed for accurate spatial relationship capture, essential for developing a deep learning model that predicts landmark positions from all angles. Unlike standardized slices, the 3D perspective utilizes the full spatial context, reducing the risk of missing or inaccurately locating landmarks [21]. This enhances the model's robustness and reliability, providing a detailed and holistic representation of anatomical structures, crucial for precise landmark detection in orthodontic applications.

In Stage 1, referred to as the "global stage," the U-Net was trained to focus on the entire image, resulting in the generation of a coarse segmentation mask for the teeth in CBCT images. This mask defined the initial region of interest (ROI), crucial for ensuring that landmark detection was centered on the relevant dental structures. The segmentation did not rely on greyscale thresholding; instead, it was based on the model's learned features, ensuring consistent processing across all cases.

Stage 2, known as the "local stage," employed a patch-based U-Net model. Building upon the initial dental region mask, the ROI was precisely excised and dilated

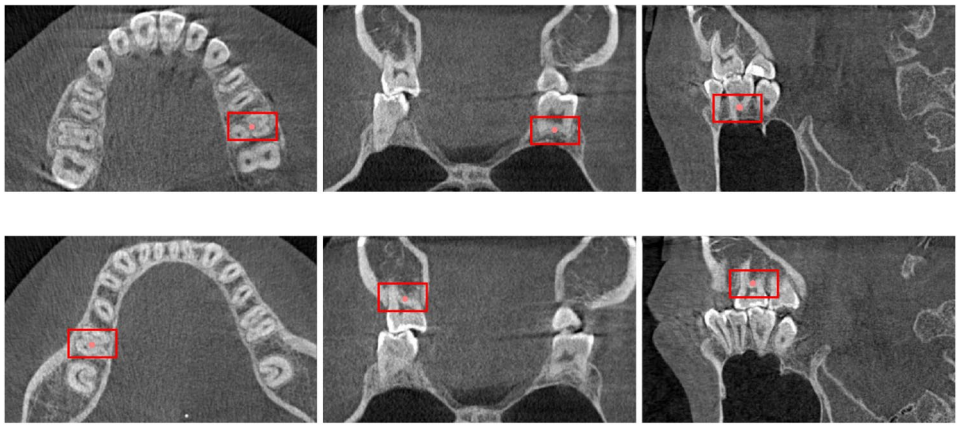

**Fig. 2** Example of CBCT slices from different angles (axial, coronal, and sagittal) showing the identified landmarks. Each landmark's location is visualized in the 3D Slicer software, using a JSON file containing the physical coordinates. Only one landmark from the maxilla and one from the mandible are shown across three slices to illustrate the method of landmark assessment and visualization

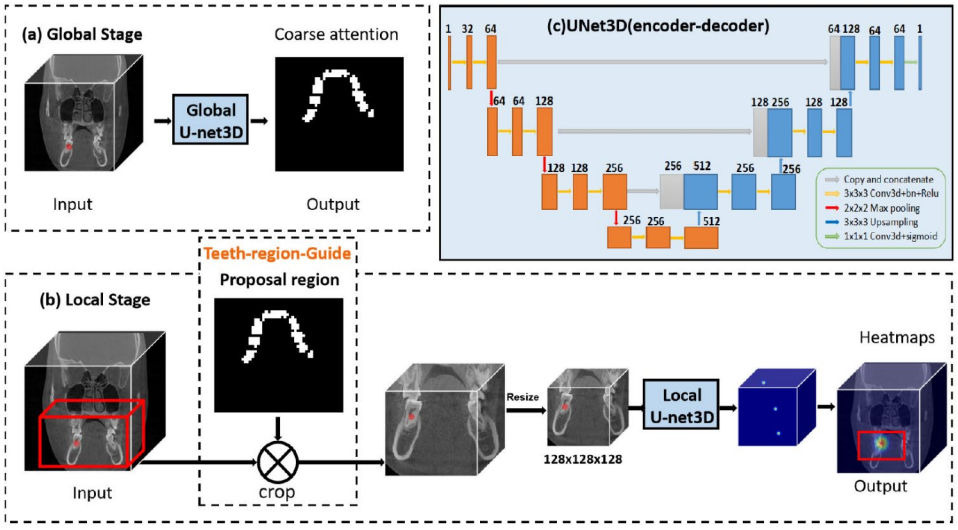


**Fig. 3** Overall framework of the attention-guided deep regression model. (**a**) The global stage is shown at the top left. (**b**) The local stage, embedded with the teeth-region guide, is shown at the bottom. For clarity, only one landmark is displayed here, though the model was trained with four landmarks. (**c**) U-Net is illustrated as an encoder-decoder for global U-Net and patch-based local U-Net

by 10 pixels to ensure comprehensive coverage of the dental area. This segmentation is fundamental to our method as it allows for focused and detailed analysis within the specifically relevant dental area. Subsequently, the region containing the dental structures was excised from the CBCT image. This extracted segment was resized to a volumetric resolution of 128×128×128 pixels before being introduced into a U-Net architecture for further processing. The model was tasked with the prediction of four pivotal landmarks within the dental area. Postprediction, the coordinates of these landmarks were recalculated to align with the co-ordinate system of the original CBCT image. A heatmap was then constructed to visually represent the location of these critical points, ensuring precise and patient-specific landmark detection.

#### Implementation details and training setup

All experiments were based on the Python API implementation of PyTorch (version 2.0.1; Facebook AI Research lab (FAIR), USA; https://pytorch.org/), which provides developers with a flexible and powerful environment to create and train their models effectively [22]. The optimizer used in the experiments was Adam [23], which employed an initial learning rate of 0.001. A learning decay rate of 0.95 per epoch was applied to adjust the learning rate throughout the training process. In addition, the input image size was set to 128×128×128 voxels. We utilized an early stopping mechanism on the validation loss with a patience of 10 epochs to avoid overfitting. For data augmentation, we adopted random mirroring, rotation and contrast adjustment strategies for all training data. All experiments were run on a single A100

GPU (Nvidia Corp., Santa Clara, CA, USA) and Xeon(R) Gold 6326 CPU (Intel Corp., Santa Clara, CA, USA) running at 2.90 GHz with 16 cores.

### Model evaluation metrics

To assess the effectiveness of our developed DL model, we considered two aspects: (i) the precision of landmark detection and (ii) the measurement error of widths of maxillary and mandibular basal bone.

To gauge the precision of our model, we employed two metrics, namely the mean radial error (MRE) and successful detection rate (SDR), to examine the effectiveness of the four landmarks that were measured [15]. MRE quantifies the discrepancy between the actual landmarks and the detected landmarks obtained from our DL algorithm. This offers a reliable assessment of the discrepancy between the actual data and the outcomes produced by the DL model. SDR, on the other hand, measures the proportion of accurately identified landmarks within various error thresholds, including 2.0, 2.5, 3.0, and 4.0 mm. SDR measures the accuracy between the ground truth and the automatic results. This metric has demonstrated its efficacy in assessing the quality of localization and is extensively employed in applications involving landmark detection [15, 18, 24].

To evaluate the measurement error of maxillary and mandibular basal bone widths, we computed the mean square error (MSE) between the measurements derived from manually marked landmarks and the measurements predicted by our DL model [25]. The MSE serves as the evaluation criterion for examining the model's predictions in quantifying the basal bone width of the maxilla and mandible.

### Statistical analysis

In this study, all statistical analyses were conducted using MedCalc statistical software (version 20.0.9.0; MedCalc Software Ltd., Ostend, Belgium; https://www.medcalc.org/) and Python (version 3.10.9; Python Software Foundation, Wilmington, Delaware, USA; https://www.python.org/). All tests conducted were two-sided, and a p value of less than 0.05 was deemed statistically significant.

1) Sample size: According to a previous report [26], the sample size was estimated using a power analysis tailored for a one-sample T-test [27]. The purpose of this t-test was to verify that there is no statistically significant difference between the length measurements predicted by the deep learning model and those calculated by dentists based on ground truth landmarks. In specific, we compared the error values between the AI-generated measurements and those provided by the dentists against zero using a t-test. Based on the following inputs and assumptions: 80% power, a two-sided significance level of 0.05, and an effect size (Cohen's d) of 0.5, a minimum of 27 CBCT scans was required for both the training and testing cohorts. Therefore, the sample sizes of 80 scans in the training cohort and 34 scans in the test cohort were deemed sufficient to detect a difference in length measurements.
2) Location performance: We used the concordance correlation coefficient (CCC) [28] to assess the consistency or agreement of landmark locations and reported the corresponding 95% confidence interval. The one-sample T test was used to compare basal bone width measurement error.

**Table 1** Quantitative evaluation results of landmark detection. The MRE of each landmark and SDR within the 2.0-, 2.5-, 3.0-, 4.0-mm range criteria are listed

| Landmarks | SDR2-mm(%) | SDR2.5-mm(%) | SDR3-mm(%) | SDR4-mm(%) | MRE(mm) |
|---|---|---|---|---|---|
| L1 | 70.59 | 82.35 | 85.29 | 88.24 | 2.09 ± 2.68 |
| L2 | 76.47 | 82.35 | 88.24 | 88.24 | 2.30 ± 3.17 |
| L3 | 67.65 | 85.29 | 91.18 | 97.06 | 1.60 ± 0.89 |
| L4 | 70.59 | 75.47 | 82.35 | 91.18 | 2.02 ± 2.43 |
| Average | 71.34 | 81.37 | 86.77 | 91.18 | 2.00 ± 2.46 |

MRE: mean radial error; SDR: successful detection rate

## Results

### Landmark detection success rate

Table 1 displays the overall landmark detection performance. The SDR within the 2.0-, 2.5-, 3.0-, and 4.0-mm error ranged across different landmarks, with L1 achieving high SDR values of 70.59%, 82.35%, 85.29%, and 88.24%, respectively. L1 and L2, which are located on the maxillary basal bone, denoting the left and right points respectively, are crucial for measuring the maxillary width which is essential for assessing the transverse dimension of the upper jaw. Similarly, L3 and L4 exhibited varying SDR values within these range criteria. L3 and L4, which are set on the mandibular basal bone, representing the left and right points respectively, are used for mandibular width measurement which is critical for evaluating the lower jaw's transverse dimension. The average SDR across all landmarks was 71.34%, 81.37%, 86.77%, and 91.18% for the 2-, 2.5-, 3-, and 4-mm range criteria, respectively. Measuring the distances between L1 and L2, as well as between L3 and L4, provides key diagnostic information for identifying maxillary transverse deficiency and planning appropriate treatments. Furthermore, MRE was measured for each landmark; it ranged from 1.60 to 0.30 mm. The average MRE across all landmarks was 2.00 mm. All four landmarks exhibited

CCC values greater than 0.9 on the x-, y-, and z-axes (Table 2). Figure 4 provides a visual representation of the outcome.

### Basal bone width measurement

Table 3 reveals the analogy between the consequence of manual measurements and the suggested method. The maxillary basal bone width demonstrated a CCC of 0.96 (95% confidence interval [CI]: 0.94–0.97) and MSE of 1.22 mm. Similarly, the mandibular basal bone width exhibited a CCC of 0.98 (95% CI: 0.97–0.99), with an MSE error of 0.68 mm. Notably, the p values of the T test for these comparisons were 0.37 and 0.67, respectively.

**Table 2** Quantitative evaluation results of CCC for each landmark on the x-, y-, and z-axes

| | L1's CCC [95% CI] | L2's CCC [95% CI] | L3's CCC [95% CI] | L4's CCC [95% CI] |
|---|---|---|---|---|
| x-axis | 0.72 [0.51–0.85] | 0.95 [0.91–0.98] | 0.97 [0.93–0.98] | 0.62 [0.40–0.78] |
| y-axis | 0.97 [0.94–0.98] | 0.90 [0.81–0.95] | 0.99 [0.98–1.00] | 0.91 [0.83–0.96] |
| z-axis | 0.99 [0.97–0.99] | 0.98 [0.96–0.99] | 0.95 [0.91–0.98] | 0.99 [0.99–1.00] |

CCC: concordance correlation coefficient; CI: confidence interval

## Discussion

In this study, a DL model for measuring basal bone width was developed and validated, and its performance was

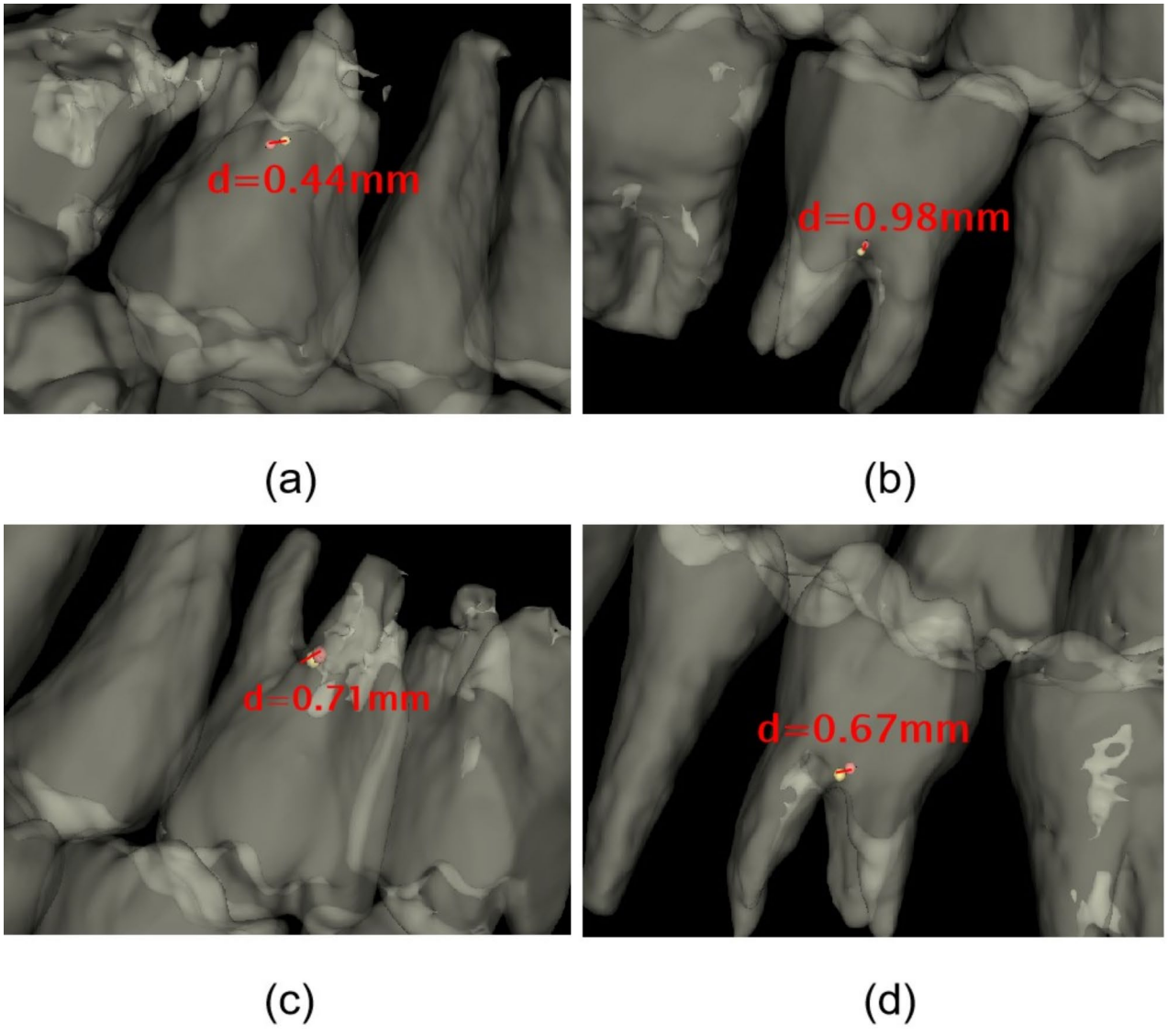


**Fig. 4** Qualitative evaluation results of landmark detection. The red dot represents the ground truth, the yellow dot represents the prediction result, and the red line segment represents the error distance. **(a)** Example of landmark detection on the maxillary left first molar. **(b)** Example of landmark detection on the maxillary right first molar. **(c)** Example of landmark detection on the mandibular left first molar. **(d)** Example of landmark detection on the mandibular right first molar.

**Table 3** Results of the comparative analysis of maxillary and mandibular basal bone width measurements between the DL model and doctor manual assessments. Figure 5 provides a visual representation of the outcome

| | CCC [95% CI] | MSE (mm) | T test *p* value (< 0.05) |
|---|---|---|---|
| L1-L2(maxillary basal bone width) | 0.96 [0.94–0.97] | 1.22 ± 1.93 | 0.37 |
| L3-L4(mandibular basal bone width) | 0.98 [0.97–0.99] | 0.68 ± 0.82 | 0.67 |

DL: deep learning; CCC: concordance correlation co-efficient; CI: confidence interval

compared with that of trained doctors. The average SDR of the DL model for landmarks within a 2-mm range was 71.34%. The error in measuring the maxillary basal bone width between the DL model and the doctor measurements was 1.22 mm ($p$=0.37). CCC between the two measurements was 0.96 (95% CI: 0.94–0.97). For the mandibular basal bone width, the error between the DL model and doctor measurements was 0.68 mm, which is equivalent to the level of error observed in doctors (CCC: 0.98, 95% CI: 0.97–0.99; $p$=0.67).

Manual measurement of dental parameters from CBCT images is a time-consuming and intricate process, highlighting the necessity for a precise and efficient technique to evaluate basal bone width. Our investigation demonstrated that our DL model obtained an excellent average inference speed of 1 s per CBCT image, which represents a substantial enhancement compared to the conventional 2-minute manual measurements. This DL-based technology has the capacity to increase the speed of the measuring process by 120 times, making it easier to perform manual measurements, enhancing the efficiency of doctors' work, and accelerating the early clinical evaluation of occlusal connections.

In the last ten years, DL approaches have been exploited to build automatic measurement methods [15, 29], significantly enhancing the precision and efficiency of orthodontic diagnoses. Our study extends these advancements by applying machine learning techniques specifically to maxillary transverse deficiency (MTD), developing a reliable model for measuring basal bone widths. This model not only minimizes the variability between different operators but also reduces the overall workload for doctors, facilitating the development of more intelligent medical systems.

Furthermore, the integration of DL in dental practices goes beyond mere measurement. For example, a systematic review on the prevalence of bruxism among alcohol abusers illustrates the potential of DL to unravel the complex interdependencies of dental conditions and systemic health factors [30]. By leveraging such technologies, clinicians are equipped to make more informed decisions regarding occlusion and treatment planning, which in turn minimizes the inefficient use of social resources. This comprehensive approach promises to revolutionize dental diagnostics, making it possible to not just accelerate but also enhance the accuracy of clinical evaluations, thereby significantly improving patient outcomes.

This study had some limitations that should be mentioned. First, this study only focused on basal bone width measurement using Yonsei transverse analysis. Among the CBCT-based diagnostic methods, three main approaches were identified: (1) the Yonsei transverse analysis method, which utilizes the CR of teeth as a reliable indicator of their position, measured by assessing the distances between the bilateral centers of root furcation points of the first permanent molars as basal bone widths. (2) The University of Pennsylvania (Penn) CBCT analysis method [31], which selects the bilateral jugal points as maxillary measuring points and the buccal cortex at the level of the CR of the mandibular first molars as mandibular measuring points. (3) Miner et al. at the Boston University (BU) Goldman School of Dental Medicine introduced a transverse analysis method [32], measuring the transverse distance between bilateral maxillary palatal and mandibular lingual cortex. Zhang et

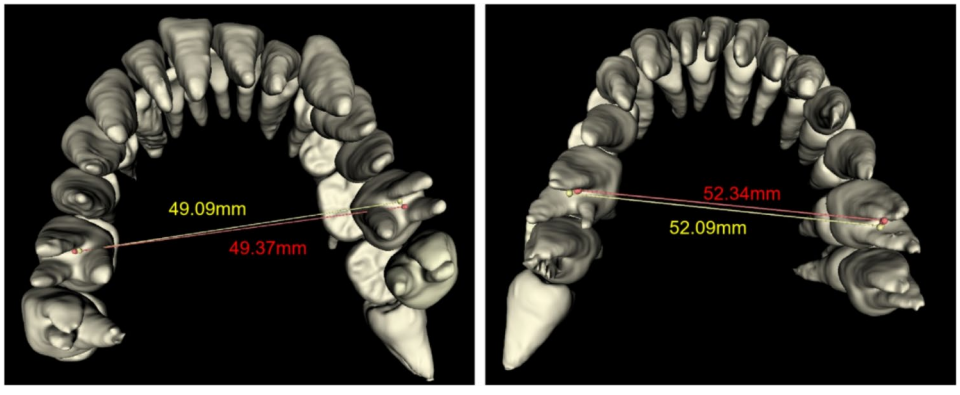


**Fig. 5** Qualitative evaluation results of measurement. The red line represents the ground truth, and the yellow line represents the prediction result

al. found that among the three diagnostic methods [33], the Yonsei analysis method demonstrated the highest interexaminer reliability, followed by those of Penn and BU [6, 34]. In the future, we intend to incorporate other transverse width measures to expand the applicability of our model to real-life clinical scenarios, and thus, provide a reference for the variety of diagnostic methods for MTD patients. Second, the study occurred at a specific research institution, limiting our study's focus to a single center, and future research will involve data collection from additional hospitals or institutions to further validate the generalizability of this developed model. Third, the study's retrospective approach introduces constraints, necessitating additional prospective investigations to authenticate the precision and efficacy of the DL model. Despite the inherent limitations of this study, the obtained results remained unaffected and demonstrated positive outcomes.

In summary, this study introduces a DL model for basal bone measurement based on landmark detection. This model achieved an impressive accuracy rate of 71.34% in detecting landmarks with a precision of 2 mm, meeting the clinical requirements. Additionally, the DL model achieved a measurement error of < 1.22 mm in basal bone width, suggesting promising results in its application.

## Conclusion

In this study, we successfully developed and validated a DL measurement model for detecting orthodontic target landmarks, demonstrating its remarkable consistency with manual measurements and affirming its reliability. The innovative method not only holds substantial promise for extensive utilization in the realm of medical technology and diagnosis but also opens avenues for future research on its application in more complex diagnostic tasks and its integration into clinical workflows.

#### Abbreviations

| | |
|---|---|
| MTD | Maxillary Transverse Deficiency |
| CBCT | Cone-Beam Computed Tomography |
| DL | Deep Learning |
| SDR | Successful Detection Rate |
| MRE | Mean Radial Error |
| CCC | Concordance Correlation Coefficient |
| MRI | Magnetic Resonance Imaging |
| CT | Computed Tomography |
| ROI | Region of Interest |
| CR | Center of Resistance |
| MSE | Mean Square Error |
| DICOM | Digital Imaging and Communications in Medicine |
| CI | Confidence Interval |
| API | Application Programming Interface |

#### Acknowledgements

Not applicable.

#### Author contributions

JD was involved in data annotation, original draft writing, investigation, methodology, resource allocation, and supervision. XEG contributed to original draft writing, data curation, validation, model training, and formal analysis. HYZ participated in review and editing, validation, and formal analysis. HYX was responsible for investigation, and methodology. JHH contributed to investigation and methodology. QTH played a role in investigation and methodology. BSH contributed to conceptualization, review and editing, and project administration. All authors have reviewed and approved the final manuscript.

#### Funding

The present work was supported by the Science Technology and Innovation Committee of Shenzhen (No.20220808170713001) , Medicine Plus Program of Shenzhen University (No. 2024YG020) , the Beijing Natural Science Foundation(Grant No. L232112) and the National Natural Science Foundation of China (No. U2013205 and No. 62371303).

#### Data availability

The datasets generated and analyzed during the current study are not publicly available due to the fact that the datasets may be part of an ongoing research project or collaboration, but they are available from the corresponding author on reasonable request. While the datasets will not be publicly available, we will make our source code accessible on our GitHub account to ensure the reproducibility of our work. The code repository is available at https://github.com/Guo777777/Landmark-detection.

## Declarations

#### Ethics approval and consent to participate

This retrospective study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Research Ethics Committee of Shenzhen University General Hospital, Shenzhen University Medical School (KYLL-20221219 A). All the methods in the study were carried out in accordance with the relevant guidelines and regulations. An informed consent agreement was signed by the adult patients and a guardian for minors.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

Received: 13 November 2023 / Accepted: 23 August 2024
Published online: 14 September 2024

#### References

1. McNamaraa JA. Maxillary transverse deficiency. Am J Orthod Dentofac Orthop. 2000;117:567–70.
2. Betts N, Vanarsdall R, Barber H, Higgins-Barber K, Fonseca R. Diagnosis and treatment of transverse maxillary deficiency. Int J Adult Orthodon Orthognath Surg. 1995;10:75–96.
3. Proffit W, White R Jr. Who needs surgical-orthodontic treatment? Int J Adult Orthodon Orthognath Surg. 1990;5:81–9.
4. Shu R, Han X, Wang Y, Xu H, Ai D, Wang L, et al. Comparison of arch width, alveolar width and buccolingual inclination of teeth between class II division 1 malocclusion and class I occlusion. Angle Orthodont. 2013;83:246–52.
5. Lee KJ, Choi SH, Choi TH, Shi KK, Keum BT. Maxillary transverse expansion in adults: Rationale, appliance design, and treatment outcomes. Sem Orthodont. 2018;24:52–65.
6. Koo YJ, Choi SH, Keum BT, Yu HS, Hwang CJ, Melsen B, et al. Maxillomandibular arch width differences at estimated centers of resistance: comparison between normal occlusion and skeletal class III malocclusion. Korean J Orthodont. 2017;47:167–75.
7. Zou W, Jiang J, Xu T, Wu J. Relationship between mandibular dental and basal bone arch forms for severe skeletal class III patients. Am J Orthod Dentofac Orthop. 2015;147:37–44.
8. De Vos W, Casselman J, Swennen GR. Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: a systematic review of the literature. Int J Oral Maxillofac Surg. 2009;38:609–25.
9. Bayome M, Park JH, Kim YJ, Kook. YA. 3D analysis and clinical applications of CBCT images. Sem Orthodont. 2015;21:254–62.

10. Mayhall J, Alvesalo L. Sexual dimorphism in the three-dimensional determinations of the maxillary first molar: cusp height, area, volume and position. Structure, function and evolution of Teeth. Tel Aviv: Freund; 1992.
11. Ronsivalle V, Marrapodi MM, Siurkel Y, Cicciù M, Minervini G. Prevalence of Bruxism in alcohol abusers: a systematic review conducted according to PRISMA guidelines and the cochrane handbook for systematic reviews of interventions. BMC Oral Health. 2024;24:108.
12. Pauwels R, Beinsberger J, Collaert B, Theodorakou C, Rogers J, Walker A, et al. Effective dose range for dental cone beam computed tomography scanners. Eur J Radiol. 2012;81:267–71.
13. Ludlow JB, Ivanovic M. Comparative dosimetry of dental CBCT devices and 64-slice CT for oral and maxillofacial radiology. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2008;106:106–14.
14. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44.
15. Wang CW, Huang CT, Hsieh MC, Li CH, Chang SW, Li WC, et al. Evaluation and comparison of anatomical landmark detection methods for cephalometric x-ray images: a grand challenge. IEEE Trans Med Imaging. 2015;34:1890–900.
16. Qian J, Cheng M, Tao Y, Lin J, Lin H, CephaNet. An improved faster R-CNN for cephalometric landmark detection. 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019). IEEE. 2019:868 – 71.
17. Chen R, Ma Y, Chen N, Liu L, Cui Z, Lin Y, et al. Structure-aware long short-term memory network for 3d cephalometric landmark detection. IEEE Trans Med Imaging. 2022;41:1791–801.
18. Lee JH, Yu HJ, Kim MJ, Kim JW, Choi J. Automated cephalometric landmark detection with confidence regions using bayesian convolutional neural networks. BMC Oral Health. 2020;20:1–10.
19. Williams JR. The declaration of Helsinki and public health. Bull World Health Organ. 2008;86:650–2.
20. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging. 2012;30:1323–41.
21. Zeng L, Chen L, Bao W et al. 3d-aware facial landmark detection via multi-view consistent training on synthetic data[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 12747–12758.
22. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G et al. Pytorch: an imperative style, high-performance deep learning library. Adv Neural Inform at Process Syst. 2019; 32.
23. Kingma DP, Ba J, Adam. A method for stochastic optimization. Arxiv Database. https://arxiv.org/ftp/arxiv/papers/1906/1906.07549.pdf
24. Zhong Z, Li J, Zhang Z, Jiao Z, Gao X. An attention-guided deep regression model for landmark detection in cephalograms. Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22. Springer International Publishing, 2019: 540–548.
25. Cramon-Taubadel N, Frazier BC, Lahr MM. The problem of assessing landmark error in geometric morphometrics: theory, methods, and modifications. Am J Phys Anthropol. 2007;134:24–35.
26. Park S, Kim YH, Bang HI, Park Y. Sample size calculation in clinical trial using R. J Minim Invasive Surg. 2023;26(1):9.
27. Bland M. An introduction to medical statistics. Oxford, UK: Oxford University Press; 2015.
28. Lawrence I, Lin K. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 1989;44:255–68.
29. Lisboa Cde O, Masterson D, da Motta AF, Motta AT. Reliability and reproducibility of three-dimensional cephalometric landmarks using CBCT: a systematic review. J Appl Oral Sci. 2015;23:112–9.
30. Ronsivalle V, Casella F, Fichera G, Bennici O, Conforte C, Lo Giudice A. Root resorption of maxillary posterior teeth after rapid maxillary expansion: a comprehensive review of the current evidence from in-vitro and in-vivo studies. Open Dent J. 2021;15:97–101.
31. Tamburrino R, Boucher N, Vanarsdall R, Secchi A. The transverse dimension: diagnosis and relevance to functional occlusion. RWISO J. 2010;2:13–22.
32. Miner RM, Al Qabandi S, Rigali PH, Will LA. Cone-beam computed tomography transverse analysis. Part I: normative data. Am J Orthod Dentofac Orthop. 2012;142:300–7.
33. Zhang CX, Guo QY, Liu W, Tang Y, Yuan R. Maxillary transverse deficiency diagnosed by 3 methods and its relationship with molar angulation in patients with skeletal class III malocclusion. Am J Orthod Dentofac Orthop. 2023;164:5–13.
34. Zhang CX, Tan XM, Wu W, Liu H, Liu Y, Qu XR, et al. Reliability of 2 methods in maxillary transverse deficiency diagnosis. Am J Orthod Dentofac Orthop. 2021;159:758–65.


## Publisher's note