



OPEN Robust oriented object detection for posterior teeth in mixed dentition

Rui Hao^{1,4}, Yunhao Zheng^{2,4}, Yuxing Ma² & Xin Xiong^{2,3}✉

The mixed dentition stage presents significant diagnostic challenges in panoramic radiography due to the high structural density and spatial superposition of primary roots and developing permanent germs. Traditional horizontal bounding box (HBB) detection methods suffer from a “geometric mismatch,” failing to accurately localize tilted teeth in crowded dentitions. This study aims to develop and validate a high-precision oriented object detection (OBB) framework to resolve these localization ambiguities. A retrospective dual-center study was conducted utilizing 1,148 panoramic radiographs. A specialized dataset annotating primary molars, developing premolars, and first permanent molars was constructed. The YOLOv11-OBB architecture was trained on a primary cohort ($n=1,040$). Model performance was evaluated on an internal test set ($n=104$) and an independent external validation cohort ($n=108$) to assess generalizability. The YOLOv11-OBB-Large model achieved exceptional performance, yielding a Mean Average Precision (mAP@50–95) of 0.904 on the internal dataset and maintaining robust accuracy (mAP@50–95 = 0.896) on the external dataset. The Nano model demonstrated a superior efficiency-accuracy trade-off, achieving an inference latency of 141ms on a standard CPU (vs. 958ms for Large) with minimal performance loss (mAP@50–95 = 0.879). In the stratified analysis by tooth position, the model exhibited consistent reliability across both permanent and primary dentitions. Mandibular premolars and molars achieved the highest precision (mAP@50–95 > 0.94), while primary molars maintained robust detection rates (mAP@50–95 > 0.86) despite the complexities of root resorption and germ overlap. The proposed OBB framework effectively addresses the geometric limitations of traditional detectors, providing precise localization and rotation information essential for downstream tasks such as automated space analysis and eruption monitoring. The model demonstrates strong generalizability across different clinical centers. To foster reproducibility and facilitate future research, the source code and pre-trained weights are publicly available at <https://github.com/Zheng-Yunhao/Mixed-Dentition-OBB-Detection>.

The mixed dentition stage, defined as the developmental period commencing with the eruption of the first permanent tooth and concluding with the exfoliation of the last primary tooth, typically spanning ages 6 to 12, represents the most dynamic and diagnostically complex phase in stomatological development¹. This critical window is characterized by the concomitant exfoliation of primary teeth and the eruption of permanent successors, involving intricate physiological processes such as root resorption, alveolar bone remodeling, and vertical dental development^{2–4}.

Panoramic radiography serves as the gold standard for initial dental examinations—encompassing orthodontic assessment and caries screening—in pediatric dentistry, offering a comprehensive view of the maxillofacial complex⁵. However, the automated and accurate interpretation of the posterior region, comprising primary molars and developing premolars, remains a formidable challenge^{5,6}. The posterior mixed dentition exhibits high structural density where the roots of primary molars are often in close proximity to, or even spatially superimposed upon, the developing germs of permanent successors in two-dimensional projections⁷. This anatomical complexity frequently leads to substantial obstacles to the development of robust automated diagnostic algorithms^{8,9}.

Over the past decade, deep learning, particularly convolutional neural networks (CNNs), has revolutionized dental radiographic analysis, evolving from basic image classification to complex object detection and

¹Department of Pediatric Dentistry, Zhengzhou Stomatological Hospital, Zhengzhou, Henan, China. ²State Key Laboratory of Oral Diseases, National Centre for Stomatology, National Clinical Research Centre for Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, Sichuan, China. ³Department of Orthodontics West China School of Stomatology, Sichuan University, Chengdu 610041, Sichuan, China. ⁴ These authors contributed equally: Rui Hao and Yunhao Zheng. ✉email: drxiongxin@scu.edu.cn

segmentation¹⁰. Among these, the “You Only Look Once” (YOLO) family of detectors has gained prominent traction for panoramic radiograph interpretation¹¹. Unlike traditional two-stage detectors, YOLO’s single-stage architecture offers a superior balance of computational efficiency and global contextual understanding. This global perspective is crucial when analyzing the continuous, high-density structures of the dental arch, as it allows the model to leverage spatial relationships between adjacent teeth. Successive iterations of YOLO have continuously refined feature extraction and multi-scale fusion to address overlapping structures¹². The latest architecture, YOLOv11, introduces advanced structural designs and enhanced feature pyramids that significantly improve the detection accuracy of dense, occluded objects, making it highly suitable for the complex anatomical environment of the mixed dentition¹³.

Despite these rapid architectural advancements, the majority of existing tooth detection models—including earlier YOLO variants—still rely on the horizontal bounding box (HBB) paradigm^{14–16}. While HBB-based detectors perform adequately for isolated objects, they suffer from a fundamental “geometric mismatch” when applied to the mixed dentition. Posterior teeth often erupt at significant inclinations in panoramic radiography⁵. When an axis-aligned rectangular box is applied to a tilted tooth, it inevitably encompasses a large area of invalid background and adjacent structures. To address these geometric limitations, oriented bounding box (OBB) has emerged as a promising solution^{17,18}. By introducing an angular parameter to the prediction, OBB allows the bounding box to rotate and tightly adhere to the anatomical contour of the tooth. This rotational capability allows for effective decoupling of the tooth from the background, resolving the localization ambiguity caused by overlapping bounding boxes in mixed dentition.

Crucially, OBB-based detection serves as a vital foundational model for advanced downstream clinical tasks. The precise orientation of the OBB acts as a proxy for the biological long axis, which is essential for monitoring eruption trajectories and predicting impaction risks, like ectopic eruption of first molars^{19,20}. Furthermore, the rotatable box captures the true mesiodistal width of unerupted permanent premolars invariant to their inclination, thereby facilitating automated orthodontic space analysis. Additionally, by tightly outlining the vertical boundaries of the tooth crown, OBBs allow for the sensitive detection of vertical positioning discrepancies, aiding in the diagnosis of primary molars infraocclusion²¹.

In the present study, we constructed a specialized mixed dentition OBB dataset, meticulously annotating primary molars, developing premolars, and first permanent molars. Based on this dataset, we developed a high-precision detection framework using the YOLOv11-OBB architecture¹³. To bridge the gap between algorithmic theory and clinical utility, we conducted a rigorous external validation on an independent cohort, demonstrating the model’s robustness across different geographical centers. Furthermore, to address the scarcity of open resources in dental AI and foster reproducibility, we have made our complete source code and pre-trained model weights publicly available²². This open-science initiative aims to lower the technical barrier for clinical researchers and accelerate the development of fully automated diagnostic systems.

Methods

Data acquisition

This retrospective dual-center diagnostic study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. The study protocol was reviewed and approved by the Institutional Review Boards of both Zhengzhou Stomatological Hospital (Internal Center) and the West China School of Stomatology, Sichuan University (External Center). Informed consent was obtained from all subjects and/or their legal guardians. To protect patient privacy, all personal identification information was irreversibly anonymized and de-identified prior to analysis.

A total of 1148 digital panoramic radiographs were meticulously collected for this study. To ensure the clinical applicability of the deep learning model, specific inclusion criteria were established: (1) patients with mixed dentition stage; (2) radiographs exhibiting clearly visible posterior regions with both primary molars and erupting permanent successors; and (3) high-quality diagnostic images free from significant positioning errors.

The dataset was constructed from two geographically and technically distinct sources to ensure diversity and evaluate model generalizability. Specifically, a primary cohort of 1040 images was acquired from the radiology database of Zhengzhou Stomatological Hospital. To facilitate robust model training and unbiased evaluation, these images were randomly partitioned using a stratified sampling method into a training set of 832 images for parameter optimization, a validation set of 104 images for hyperparameter tuning and early stopping, and an internal testing set of 104 images for initial performance benchmarking. Beyond this internal data, an independent external validation cohort consisting of 108 images was collected from the West China School of Stomatology. This dataset was reserved strictly for the final inference phase and remained unseen by the model during training, serving as a gold standard to test the model’s robustness against domain shifts caused by scanner variations and patient demographics. The detailed demographic characteristics of the patients, including age and sex distributions across all subsets, are summarized in Table 1.

Panoramic radiographs at Zhengzhou Stomatological Hospital were acquired using an mDX-15Pro-K10 unit (Meyer, Hefei, China) operating at 100 kV and 8 mA. At West China Hospital of Stomatology, panoramic radiographs were obtained using either a Veraviewepocs unit (Morita, Kyoto, Japan) operating at 65 kV and 6 mA, or a PaX-i unit (VATECH, Hwaseong, Korea) operating at 73 kV and 10 mA.

Annotation protocol

To ensure high-quality ground truth, a rigorous two-stage annotation strategy was implemented. Initial annotations were performed by a senior dentist with 20 years of clinical experience, who defined the OBB for primary molars, developing premolars, and first permanent molars. The boxes were drawn to tightly encapsulate the tooth boundaries, with the rotation angle aligned with the tooth’s long axis (Fig. 1). Subsequently, a

Variable	Training set	Validation set	Internal test set	External test set
<i>Age</i>				
Range	5–12	5–12	5–12	5–12
Mean	7.76	7.86	7.95	7.84
SD	1.48	1.48	1.54	1.54
<i>Sex</i>				
Male	404 (48.56%)	46 (44.23%)	50 (48.08%)	46 (42.59%)
Female	428 (51.44%)	58 (55.77%)	54 (51.92%)	62 (57.41%)

Table 1. Demographic characteristics of the study population across the training, validation, and test sets.

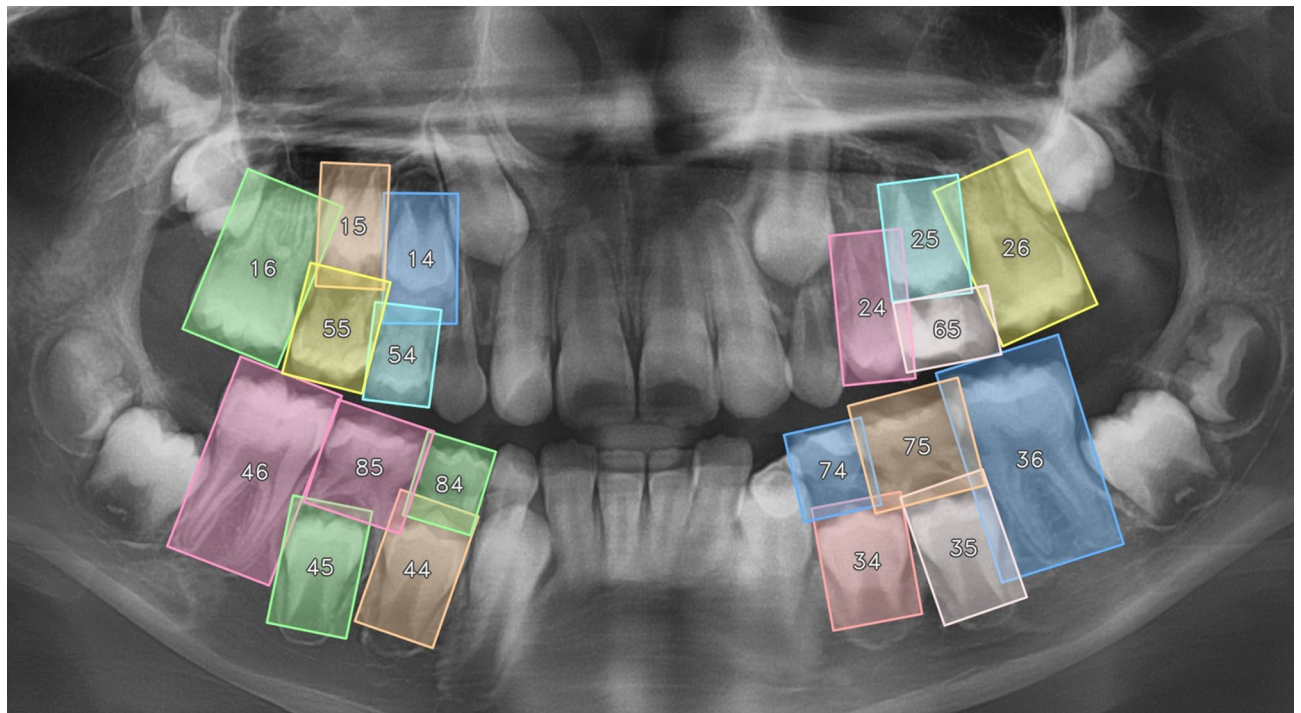


Fig. 1. Annotation of posterior teeth with oriented bounding boxes.

second dentist with 5 years of experience reviewed all annotations to correct any errors or inconsistencies. All modifications were made after discussion and consensus between the two dentists.

To quantitatively evaluate inter-observer reliability, 50 panoramic radiographs were randomly selected and independently annotated by both dentists. The results demonstrated a high level of agreement, with a mean intersection over union (IoU) of 0.882 (95% CI: 0.879–0.886) for the bounding boxes, and a Cohen's Kappa value of 1.0 for tooth classification.

All tooth positions were labeled according to the FDI World Dental Federation notation system. This two-digit system identifies each tooth based on its quadrant and its position relative to the midline. The first digit represents the quadrant, ranging from 1 to 4 for permanent teeth and 5 to 8 for primary teeth, numbered in a clockwise direction starting from the patient's upper right. The second digit denotes the specific tooth type, ranging from 1 to 8 for permanent teeth and 1 to 5 for primary teeth, proceeding from the central incisor toward the posterior region.

Model architecture

We utilized the YOLOv11-OBb architecture, which advances the single-stage detection paradigm¹³. The network incorporates the improved C3k2 feature extraction block in the backbone, enhancing the model's ability to capture multi-scale features—a critical capability for simultaneously detecting small premolar germs and large permanent molars. Additionally, the C2PSA module is employed to focus the model's attention on tooth textures while suppressing background noise from the alveolar bone. The prediction head is decoupled to independently regress the objectness, class probabilities, and the five-parameter tuple for oriented boxes. We trained four model scales—Nano, Small, Medium, and Large—to investigate the trade-off between computational efficiency and detection accuracy.

Training configuration and data augmentation

All experiments were conducted with an NVIDIA GeForce RTX 4070 SUPER GPU, utilizing the PyTorch 2.6.0 framework.

Input images were resized and padded to a standardized resolution of 1024×1024 pixels to preserve fine anatomical details. The models were trained for 300 epochs using the AdamW with a learning rate of 0.000417 and a momentum of 0.9. The batch size was set to 6. To prevent overfitting, early stopping was implemented with a patience of 50 epochs.

Data augmentation strategies included random rotation with $\pm 30^\circ$ and Mosaic augmentation, which was enabled to enhance the detection of small objects. Random brightness and contrast adjustments were also applied to simulate varying exposure conditions.

Model performance measurement

Model performance was quantitatively evaluated using standard metrics adapted for oriented detection. The primary metric was mean average precision (mAP), calculated as the area under the precision-recall curve. We reported both mAP@50 (IoU threshold = 0.5) and mAP@50–95 (averaged over IoU thresholds from 0.5 to 0.95), with the latter providing a more stringent assessment of localization accuracy. Additionally, precision and recall were calculated to evaluate the model's clinical safety profile in minimizing false positives and false negatives, respectively. Finally, 20 panoramic radiographs from the external validation set, encompassing a total of 160 premolars, were randomly sampled to compare the mean absolute error (MAE) between model-automated and manual measurements of the mesiodistal width to verify the feasibility for downstream tasks. The automated measurement of the mesiodistal width was defined as the Euclidean distance between the top-left vertex and the top-right vertex of the predicted oriented bounding box.

Results

Model training

The training process demonstrated stable convergence across all four model scales. As shown in Fig. 2, for YOLOv11-OB- Nano both the bounding box regression loss and classification loss decreased steadily and stabilized after approximately 250 epochs.

Model performance

Table 2 summarizes the quantitative performance on both the internal (Zhengzhou Stomatological Hospital) and external (West China Hospital of Stomatology) datasets. On the internal testing set, the large scale model achieved the highest performance, yielding a mAP@50 of 0.9914 and a mAP@50–95 of 0.9035. The Nano model also demonstrated high accuracy with a mAP@50–95 of 0.8882, showing only a minor performance gap compared to the L model. When evaluated on the independent external dataset, the models maintained robust performance. The Large model achieved a mAP@50–95 of 0.8958, representing a minimal performance drop of less than 1% compared to the internal results. Similarly, the Nano model retained a mAP@50–95 of 0.8790. These results indicate that the proposed method generalizes well across different scanners and patient populations.

Table 3 presents the performance comparison between the OBB model and the HBB model. The experimental results demonstrate that on both the internal and external test sets, the OBB model outperforms the HBB model, particularly in fine-grained metrics such as mAP@50–95.

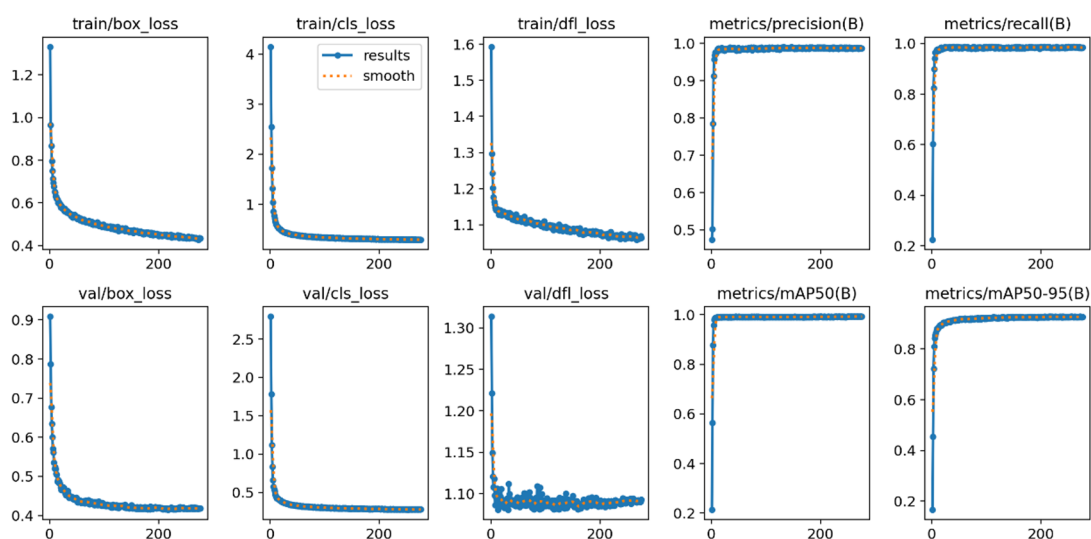


Fig. 2. (A) Visualization of the training and validation process for the YOLOv11-OB- Nano model. box_loss: Bounding box regression loss; cls_loss: Classification loss; df_l_loss: Distribution Focal Loss; (B) Metrics calculated for Bounding boxes; mAP50: mean Average Precision at an IoU threshold of 0.5; mAP50-95: mean Average Precision averaged over IoU thresholds from 0.5 to 0.95.

Model scale	Dataset	mAP@50	mAP@50-95	Precision	Recall
Nano	Internal	0.9911	0.8882	0.9854	0.9839
	External	0.9904	0.8790	0.9775	0.9820
Small	Internal	0.9925	0.8917	0.9852	0.9875
	External	0.9907	0.8840	0.9856	0.9806
Medium	Internal	0.9917	0.8974	0.9864	0.9926
	External	0.9913	0.8928	0.9867	0.9829
Large	Internal	0.9914	0.9035	0.9912	0.9875
	External	0.9919	0.8958	0.9899	0.9849

Table 2. Evaluation results of the proposed models on internal and external datasets.

Model type	Dataset	mAP@50	mAP@50-95	Precision	Recall
Oriented Bounding Boxes	Internal	0.9911	0.8882	0.9854	0.9839
	External	0.9904	0.8790	0.9775	0.9820
Horizontal Bounding Boxes	Internal	0.9910	0.6459	0.7141	0.9965
	External	0.9814	0.6512	0.7193	0.9856

Table 3. Performance evaluation of YOLOv11 Nano with oriented and horizontal bounding boxes.

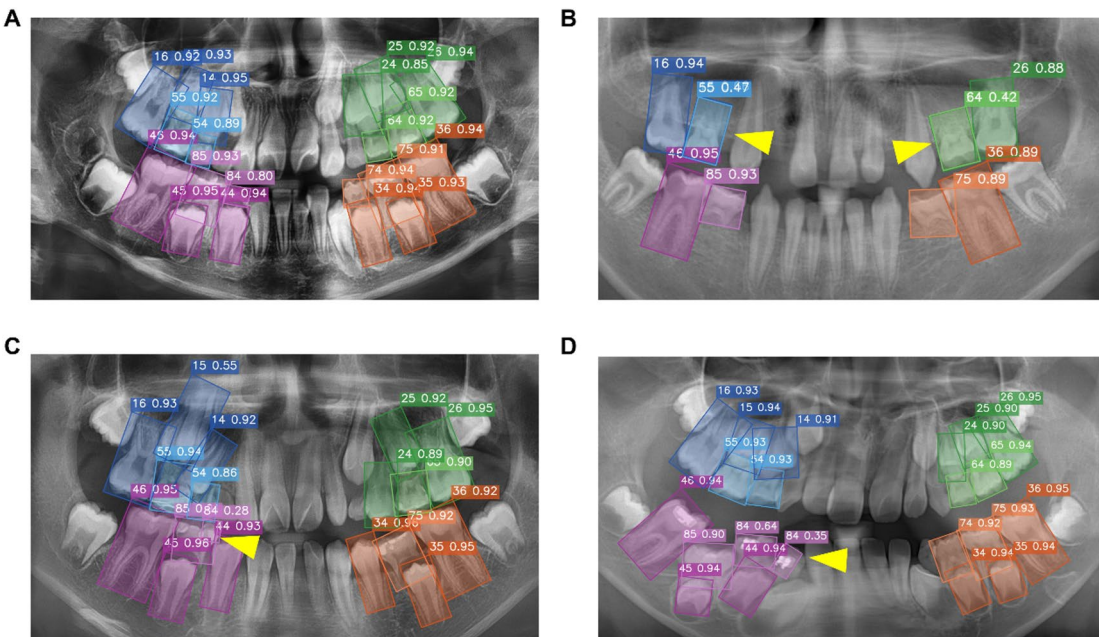


Fig. 3. Visualization of model prediction results. (A) A sample with correct predictions; (B) Prediction errors caused by missing permanent teeth, with yellow arrows pointing to the incorrectly predicted teeth; (C) Prediction errors caused by dental restorations, with yellow arrows pointing to the incorrectly predicted teeth; (D) Prediction errors caused by orthodontic brackets, with yellow arrows pointing to the incorrectly predicted teeth.

Figure 3 Presents the visualization results of the model. Although the model demonstrates favorable visual performance (Fig. 3A), prediction errors still occur when encountering rare pathological conditions (Fig. 3B–D). Across the 20 randomly sampled panoramic radiographs, the mean absolute error between the model-automated and manual measurements of the premolar mesiodistal width was 0.345 mm (95% CI: 0.318–0.376). The detailed results concerning the mesiodistal width errors for each specific tooth position are presented in Table S1.

Stratified performance by tooth category

To assess clinical reliability, we analyzed the detection accuracy across specific tooth types using the YOLOv11-OBB-Large model (Table 4). The model demonstrated exceptional performance across all categories, with

Position	Internal		External	
	mAP@50	mAP@50-95	mAP@50	mAP@50-95
#14	0.993	0.901	0.983	0.874
#15	0.984	0.875	0.985	0.873
#16	0.995	0.927	0.995	0.930
#24	0.994	0.860	0.990	0.853
#25	0.983	0.878	0.989	0.890
#26	0.995	0.933	0.989	0.927
#34	0.995	0.946	0.995	0.956
#35	0.995	0.950	0.995	0.945
#36	0.987	0.946	0.991	0.946
#44	0.995	0.949	0.995	0.925
#45	0.988	0.952	0.995	0.954
#46	0.982	0.946	0.979	0.939
#54	0.981	0.867	0.987	0.839
#55	0.995	0.898	0.995	0.885
#64	0.994	0.873	0.994	0.875
#65	0.990	0.885	0.995	0.903
#74	0.995	0.904	0.995	0.882
#75	0.995	0.911	0.995	0.899
#84	0.992	0.865	0.995	0.864
#85	0.995	0.910	0.995	0.910

Table 4. Performance of the YOLOv11-OBb-large model stratified by tooth position on internal and external test sets.

mAP@50 consistently exceeding 0.98. Notably, mandibular premolars (#34, #35, #44, #45) exhibited the highest localization precision, achieving a mAP@50–95 of approximately 0.95, indicating precise oriented bounding box regression. First Permanent Molars also showed high consistency with a mAP@50–95 ranging from 0.92 to 0.94, likely due to their larger crown size and distinct calcification patterns. The detection of maxillary premolars (e.g., #15, #24) and first primary molars (e.g., #54, #84) was slightly more challenging, yielding mAP@50–95 values between 0.86 and 0.88. Despite these variations, the model maintained robust performance on the external dataset. Figure S1 shows the precision-recall curves for individual tooth positions across the internal and external test sets.

Impact of model scaling and efficiency

To evaluate the trade-off between detection performance and computational cost, we conducted a comprehensive benchmark across four model scales (Nano, Small, Medium, Large) on the external test set. Figure 4 illustrates the efficiency frontier, mapping the relationship between mean average precision and inference latency. As anticipated, increasing model depth and width yielded consistent performance gains. The Large model achieved the highest accuracy (mAP@50–95 = 0.8958), outperforming the Nano model (mAP@50–95 = 0.8790) by approximately 1.7%. However, this marginal performance improvement incurred a disproportionate computational cost. The Large model contains 26.17 M parameters, nearly a tenfold increase compared to the Nano model (2.67 M). On a standard CPU environment, the Nano model demonstrated superior efficiency with an average latency of 141ms, whereas the large model required 958ms, approaching the limit of real-time interaction.

Discussion

In this dual-center retrospective study, we developed and validated a foundational model for the oriented detection of posterior teeth using panoramic radiographs of mixed dentition. Leveraging the YOLOv11-OBb architecture to navigate the anatomical complexities of the mixed dentition stage, our model achieved near-perfect performance, with a mAP@50 exceeding 99% on both internal and external test sets. Specifically, the mAP@50–95—a metric reflecting high localization tightness—reached 0.896 for the large size model in the external test set. These results demonstrate that the proposed model possesses the robustness and generalizability required for real-world clinical deployment. It can serve as a robust cornerstone for downstream tasks, such as the analysis of permanent tooth eruption pathways, impaction prediction, the diagnosis of deciduous infraocclusion, and orthodontic space analysis.

The mixed dentition has long been a challenging domain for AI-based object detection due to the severe overlap between primary roots and permanent germs. Recent studies have predominantly relied on traditional HBB detectors^{14,16,23}. While HBB-based approaches have achieved high detection rates for tooth counting and tooth numbering, they suffer from significant limitations in geometric precision. Although these models achieved a respectable mAP@50, their mAP@50–95 dropped drastically. For instance, Peker et al.¹⁶ reported a mAP@50–95 of only 0.696. Similarly, Incerti et al.²³ found that mAP@50–95 values for first permanent molars,

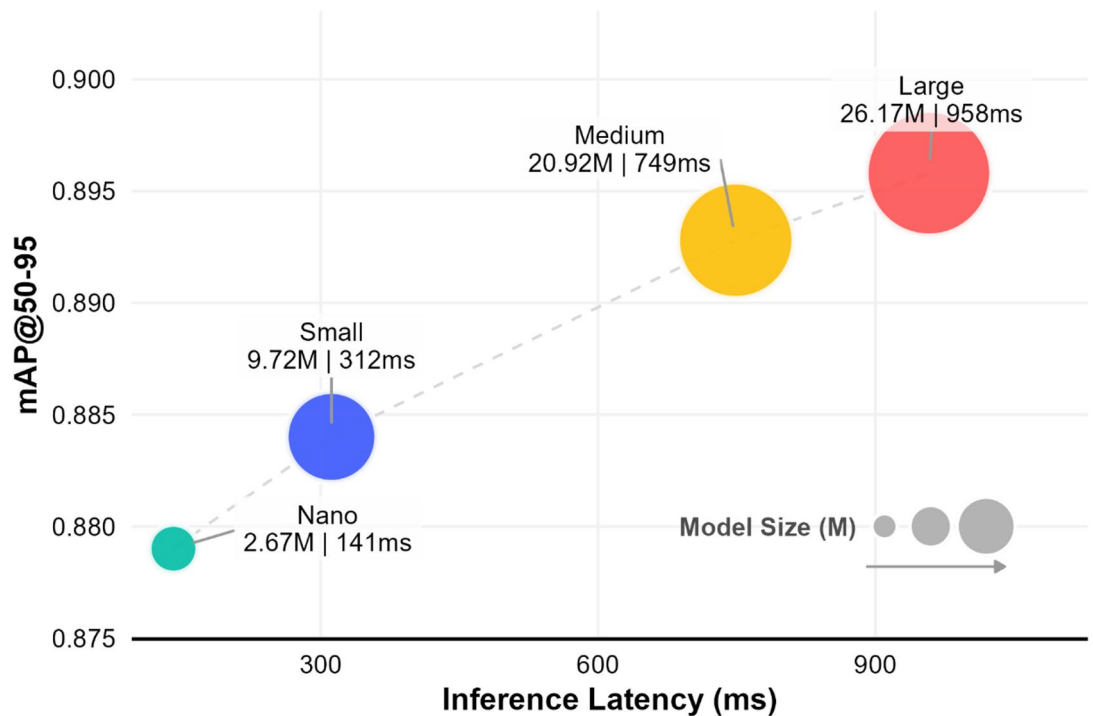


Fig. 4. Efficiency-accuracy trade-off analysis on the external test set.

premolars, and deciduous molars ranged from 0.127 to 0.563, with particularly low accuracy for deciduous molars. In stark contrast, our YOLOv11-OBB model maintained a mAP@50–95 of 0.896 in the external test set. This massive gap empirically validates the ‘Geometric Mismatch’ hypothesis: HBBs fail to tightly encapsulate inclined teeth, leading to loose bounding boxes that include excessive background noise, whereas OBBs maintain pixel-level tightness.

A critical mechanism underlying our model’s success is its ability to overcome the “HBB Paradox” during non-maximum suppression²⁴. In the posterior region, premolars often erupt at oblique angles beneath primary molars. In HBB paradigms, the bounding box of a tilted premolar frequently overlaps significantly with that of the primary molar. This often triggers non-maximum suppression algorithms to erroneously suppress the “less confident” box—typically the developing germ—resulting in false negatives. By utilizing rotatable boxes, our approach minimizes the physical overlap between distinct biological targets, ensuring that even when a premolar is spatially superimposed “behind” a primary root, the model recognizes them as separate, oriented instances²⁵. This geometric decoupling is likely the primary driver behind the high recall observed across all test sets.

Beyond robust detection, the OBB output serves as a high-fidelity “geometric primitive” that unlocks advanced downstream clinical tasks. Unlike standard detection boxes, the oriented box acts as a proxy for the biological long axis. This enables the automatic monitoring of eruption trajectories, allowing clinicians to screen for impaction or ectopic eruption of permanent molars without manual measurement²⁶. Furthermore, the OBB architecture is particularly advantageous for automated space analysis. Traditional HBB width measurements are inherently distorted by tooth inclination. Conversely, our OBB captures the true mesiodistal width invariant to rotation, facilitating precise, automated crowding analysis²⁷. Additionally, by providing a tighter region of interest that excludes the roots of adjacent teeth, this model offers a superior pre-processing step for developmental staging classifiers compared to loose horizontal crops.

Our efficiency benchmark reveals a crucial insight for clinical deployment. While the YOLOv11-OBB-Large scale model achieved the highest accuracy, the Nano model demonstrated a compelling trade-off, retaining 98.1% of the large model’s accuracy (mAP@50–95 of 0.879 vs. 0.896) while reducing inference latency by approximately 85% (140ms vs. 958ms). This finding suggests that the Nano model is highly suitable for deployment on edge devices or standard clinic workstations where computational resources may be limited, whereas the large model is ideal for server-side batch processing where maximum precision is paramount.

A recurring limitation in current dental AI literature, including large-scale studies on panoramic radiographs, is the lack of open-source code and model weights. This opacity hinders reproducibility and prevents clinical researchers from building upon existing work. To address this scarcity, we have made our source code and pre-trained YOLOv11-OBB weights publicly available. We believe this initiative will lower the technical barrier for other researchers to validate our findings or adapt the model for specific local populations, fostering a more collaborative ecosystem in digital dentistry.

Despite the promising results, this study has limitations. First, our model is currently restricted to deciduous molars, premolars, and first permanent molars. Although the model is directly applicable to downstream tasks

such as diagnosing deciduous infraocclusion, predicting premolar impaction, and orthodontic space analysis, future research should incorporate anterior teeth, second molars, and third molars to provide comprehensive OBB results for the complete dentition. Second, our model was trained on common mixed dentition classes; rare anomalies such as supernumerary teeth were not explicitly categorized and may lead to misclassification. Future work will focus on integrating longitudinal data to predict development trends and expanding the class categories to include a wider spectrum of dental anomalies.

Conclusion

This study presents a robust, high-precision oriented object detection framework for mixed dentition analysis using YOLOv11-OBB. By addressing the geometric limitations of traditional bounding boxes, our model effectively resolves the challenges of tooth overlap and inclination, achieving state-of-the-art performance in both internal and external validations. The OBB paradigm not only improves detection accuracy but also provides clinically relevant geometric parameters—such as angulation and true mesiodistal width—that are essential for automated space analysis and eruption monitoring. Furthermore, the demonstrated efficiency of the Nano architecture suggests that high-performance AI can be effectively deployed in resource-constrained clinical environments. We anticipate that this foundational model, along with our open-source contribution, will serve as a cornerstone for the next generation of automated dental diagnostics.

Data availability

The source code and pre-trained models developed for this study are publicly available at <https://github.com/Zheng-Yunhao/Mixed-Dentition-OBB-Detection>. The datasets are available from the corresponding author on reasonable request.

Received: 28 January 2026; Accepted: 20 April 2026

Published online: 24 April 2026

References

- Nelson, S. Development and eruption of the teeth. In: *Wheeler's Dental Anatomy, Physiology, and Occlusion*. 9th edn. pp. 28–29. (2010).
- Garcete Delvalle, C. S., Bruna Del Cojo, M., Mourelle Martínez, M. R. & De Nova García, M. J. Root resorption pattern and root length of mandibular primary molars in children: a cross-sectional radiographic study. *Eur. Arch. Paediatr. Dent.* **26** (5), 977–986. <https://doi.org/10.1007/s40368-025-01052-3> (2025).
- Orilisi, G. et al. Physiological root resorption of deciduous teeth: an ATR-FTIR approach. *J. Clin. Med.* **14** (1), 48. <https://doi.org/10.3390/jcm14010048> (2024).
- Fonseca-Souza, G. et al. Association between dental age and malocclusions: a systematic review. *BMC Oral Health.* **24** (1), 383. <https://doi.org/10.1186/s12903-024-04143-7> (2024).
- Sams, C. M., Dietsche, E. W., Swenson, D. W., DuPont, G. J. & Ayyala, R. S. Pediatric panoramic radiography: techniques, artifacts, and interpretation. *Radiographics* **41** (2), 595–608. <https://doi.org/10.1148/rq.2021200112> (2021).
- Bumann, E. E. et al. A novel collaborative learning model for mixed dentition and fillings segmentation in panoramic radiographs. *J. Dent.* **140**, 104779. <https://doi.org/10.1016/j.jdent.2023.104779> (2024).
- Tunç, H., Akkaya, N., Aykanat, B. & Ünsal, G. U-net-based deep learning for simultaneous segmentation and agenesis detection of primary and permanent teeth in panoramic radiographs. *Diagnostics (Basel)*. **15** (20), 2577. <https://doi.org/10.3390/diagnostics15202577> (2025).
- Mine, Y. et al. Detecting the presence of supernumerary teeth during the early mixed dentition stage using deep learning algorithms: a pilot study. *Int. J. Paediatr. Dent.* **32** (5), 678–685. <https://doi.org/10.1111/ipd.12946> (2022).
- Mine, Y. et al. Challenges and limitations of multimodal large language models in interpreting pediatric panoramic radiographs. *Int J. Paediatr. Dent.* **16** <https://doi.org/10.1111/ipd.70029> Published online September (2025).
- Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun. ACM*. **60** (8), 84–90. <https://doi.org/10.1145/3065386> (2017).
- Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: unified, real-time object detection. <https://doi.org/10.48550/arXiv.1506.02640>. Preprint posted online May 9, 2016
- Ali, M. L. & Zhang, Z. The YOLO framework: a comprehensive review of evolution, applications, and benchmarks in object detection. *Computers*. <https://doi.org/10.3390/computers13120336> (2024).
- Khanam, R. & Hussain, M. YOLOv11: an overview of the key architectural enhancements. <https://doi.org/10.48550/arXiv.2410.17725>. Preprint posted online October 23, 2024.
- Beser, B. et al. YOLO-V5 based deep learning approach for tooth detection and segmentation on pediatric panoramic radiographs in mixed dentition. *BMC Med. Imaging*. **24** (1), 172. <https://doi.org/10.1186/s12880-024-01338-w> (2024).
- Şevik, U. & Mutlu, O. Detection of dental anomalies in digital panoramic images using YOLO: a next generation approach based on single stage detection models. *Diagnostics (Basel)*. **15** (15), 1961. <https://doi.org/10.3390/diagnostics15151961> (2025).
- Peker, R. B. & Kurtoglu, C. O. Evaluation of the performance of a YOLOv10-based deep learning model for tooth detection and numbering on panoramic radiographs of patients in the mixed dentition period. *Diagnostics (Basel)*. **15** (4), 405. <https://doi.org/10.3390/diagnostics15040405> (2025).
- Huang, Z., Li, W., Xia, X. G., Wang, H. & Tao, R. Task-wise sampling convolutions for arbitrary-oriented object detection in aerial images. *IEEE Trans. Neural Netw. Learn. Syst.* **36** (3), 5204–5218. <https://doi.org/10.1109/TNNLS.2024.3367331> (2025).
- Nie, G. & Huang, H. Multi-oriented object detection in aerial images with double horizontal rectangles. *IEEE Trans. Pattern Anal. Mach. Intell.* **45** (4), 4932–4944. <https://doi.org/10.1109/TPAMI.2022.3191753> (2023).
- Garrocho-Rangel, A., Benavidez-Valadez, P., Rosales-Berber, M. A. & Pozos-Guillén, A. Treatment of ectopic eruption of the maxillary first permanent molar in children and adolescents: a scoping review. *Eur. J. Paediatr. Dent.* **23** (2), 94–100. <https://doi.org/10.23804/ejpd.2022.23.02.02> (2022).
- Kalavritinos, M. et al. Incidence of incisor root resorption associated with the position of the impacted maxillary canines: a cone-beam computed tomographic study. *Am. J. Orthod. Dentofac. Orthop.* **157** (1), 73–79. <https://doi.org/10.1016/j.ajodo.2019.02.016> (2020).
- Akgöl, B. B., Üstün, N. & Bayram, M. Characterizing infraocclusion in primary molars: prevalence, accompanying findings, and infraocclusion severity and treatment implications. *BMC Oral Health.* **24** (1), 661. <https://doi.org/10.1186/s12903-024-04428-x> (2024).

22. Schwendicke, F. & Krois, J. Data dentistry: how data are changing clinical care and research. *J. Dent. Res.* **101** (1), 21–29. <https://doi.org/10.1177/00220345211020265> (2022).
23. Incerti Parenti, S. et al. Artificial intelligence-aided tooth detection and segmentation on pediatric panoramic radiographs in mixed dentition using a transfer learning approach. *Diagnostics (Basel)*. **15** (20), 2615. <https://doi.org/10.3390/diagnostics15202615> (2025).
24. Girshick, R., Donahue, J., Darrell, T. & Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. *arXiv Preprint posted online Oct. 22*, 13112524. <https://doi.org/10.48550/arXiv.1311.2524> (2014).
25. Xie, X., Cheng, G., Wang, J., Yao, X. & Han, J. Oriented R-CNN for object detection. <https://doi.org/10.48550/arXiv.2108.05699>. Preprint posted online August 12, 2021.
26. Liu, J. et al. Artificial intelligence-aided detection of ectopic eruption of maxillary first molars based on panoramic radiographs. *J. Dent.* **125**, 104239. <https://doi.org/10.1016/j.jdent.2022.104239> (2022).
27. Camcı, H. & Salmanpour, F. Estimating the size of unerupted teeth: Moyers vs deep learning. *Am. J. Orthod. Dentofac. Orthop.* **161** (3), 451–456. <https://doi.org/10.1016/j.ajodo.2021.03.015> (2022).

Author contributions

Conceptualization: X.X. and Y.Z.; Methodology: Y.Z.; Software: Y.Z.; Validation: Y.Z. and R.H.; Formal analysis: Y.Z. and R.H.; Investigation: R.H.; Resources: X.X. and R.H.; Data Curation: R.H. and Y.M.; Writing—original draft preparation: Y.Z. and R.H.; Writing—review & editing: X.X.; Visualization: Y.Z.; Supervision: X.X.; Project administration: X.X. R.H. and Y.Z. contributed equally to this work.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-026-50297-5>.

Correspondence and requests for materials should be addressed to X.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026