BMC Pediatrics

https://doi.org/10.1186/s12887-026-07354-9

Article in Press

# Automated detection of periapical lesions in pediatric panoramic radiographs using YOLOv7: a retrospective internal validation study

Received: 29 October 2025

Accepted: 14 July 2026

Published online: 20 July 2026

Cite this article as: Shahmaleki P., Gezginci P.A., Kırelli Y. *et al.* Automated detection of periapical lesions in pediatric panoramic radiographs using YOLOv7: a retrospective internal validation study. *BMC Pediatr* (2026). https://doi.org/10.1186/s12887-026-07354-9

Parmis Shahmaleki, Pelin Alcan Gezginci, Yasin Kırelli, Fatma Yuce & Cansu Buyuk

# Automated Detection of Periapical Lesions in Pediatric Panoramic Radiographs Using YOLOv7: A Retrospective Internal Validation Study

Parmis Shahmaleki[1], Pelin Alcan Gezginci[1], Yasin Kırelli[2], Fatma Yuce[3], Cansu Buyuk[4]

1Istanbul Okan University, Faculty of Engineering and Natural Sciences, Department of Industrial Engineering, Istanbul, Turkey.

2 Kütahya Dumlupınar University, Tavşanlı Faculty of Applied Sciences, Department of Management Information Systems, Kütahya, Turkey.

3 Istanbul Kent University, Faculty of Dentistry, Department of Dentomaxillofacial Radiology, Istanbul, Turkey.

4 Istanbul Okan University, Faculty of Dentistry, Department of Dentomaxillofacial Radiology, Istanbul, Turkey.

**Parmis Shahmaleki,** Assist Prof., Istanbul Okan University, Faculty of Engineering and Natural Sciences, Department of Industrial Engineering, Istanbul, Turkey, 0000-0002-3863-9323.

parmis.shahmaleki@okan.edu.tr,

**Pelin Alcan Gezginci,** Assist Prof., Istanbul Okan University, Faculty of Engineering and Natural Sciences, Department of Industrial Engineering, Istanbul, Turkey, 0000-0001-7155-7621.

pelin.alcan@okan.edu.tr,

**Yasin Kırelli**, Assist Prof., Kütahya Dumlupınar University, Tavşanlı Faculty of Applied Sciences, Department of Management Information Systems, Kütahya, Turkey, 0000-0002-3605-8621.

yasinkirelli@gmail.com

**Fatma Yuce**, Assist Prof., Istanbul Kent University, Faculty of Dentistry, Department of Dentomaxillofacial Radiology, Istanbul, Turkey, 0000-0002-9328-4895.

dtfatmayuce@gmail.com

**Cansu Buyuk,** Assoc Prof., Istanbul Okan University, Faculty of Dentistry, Department of Dentomaxillofacial Radiology, Istanbul, Turkey, 0000-0001-8126-0928.

cansu.buyuk@okan.edu.tr

**Corresponding Author:** Assist. Prof. Fatma Yuce
E-mail: dtfatmayuce@gmail.com



**Abstract**

**Aim:** This study aimed to assess the diagnostic capability of a YOLOv7 deep learning algorithm for the computerized detection of periapical lesions from pediatric panoramic radiographs. Its potential utility as a supportive diagnostic tool and accurate diagnosis in mixed dentition cases was further assessed by comparing the algorithm's performance with the diagnoses made by dental students.

**Material and Methods:** In this study, a total of 408 panoramic radiographs were used, consisting of 333 original images and 75 images generated through feature-based preprocessing expansion. The YOLOv7 model was trained on 302 images, which included 227 original radiographs and 75 images specifically enhanced via grayscale conversion, noise reduction, and edge detection filters to emphasize structural pathological features. A relatively larger set was allocated for testing in order to enhance robustness despite the small sample size. The diagnostic capability of the algorithm and trainees was compared using accuracy, sensitivity, specificity, precision, F1 score, and error rate.

**Results:** YOLOv7 achieved higher diagnostic performance compared with the student group. Its sensitivity (76.1%) was also higher than that of the students (55.2%). The algorithm further demonstrated superior specificity (99.8% vs. 97.3%), precision (99.8% vs. 95.3%), and F1 score (86.4% vs. 69.9%).

**Conclusion:** The findings suggest promising potential in the YOLOv7 algorithm's performance for detecting periapical lesions in deciduous teeth on panoramic radiographs, compared with the diagnostic accuracy of the students.

**Keywords:** Artificial intelligence, Panoramic radiography, Periapical lesion, Deciduous teeth, Mixed dentition, Clinical decision support systems



## Introduction

Periapical lesions are radiolucent entities that indicate inflammatory conditions around the apex of a tooth root, most commonly resulting from pulp infection. Periapical lesions are common findings in both primary and permanent dentitions, and epidemiological studies in adult populations have reported a prevalence ranging between 34% and 61% in the general population (1). In pediatric populations, radiographic studies have reported lower prevalence rates, generally ranging between approximately 15% and 30% in primary teeth (1, 2). In children, these lesions present a significant clinical challenge, as undiagnosed or inadequately treated cases may result in premature loss of primary teeth, disturbances in the development of permanent tooth buds, and persistent malocclusion (3). Therefore, early and accurate detection plays a crucial role in maintaining children's oral health.

Panoramic radiography is widely used in pediatric dental diagnosis, particularly when intraoral imaging cannot be performed due to limited patient cooperation. However, their interpretation can be challenging due to mixed dentition, anatomical superimpositions, and common imaging artifacts. Such limitations increase the likelihood of missed or misdiagnosed pathologies, including periapical lesions, caries, and cysts. In addition, the lower accuracy of less experienced clinicians or dental trainees in identifying early-stage conditions may further contribute to inaccurate outcomes (2, 4-5)

In recent years, artificial intelligence (AI) has emerged as an important tool in medical and dental diagnostics. Deep learning algorithms have shown strong capability in recognizing complex image patterns beyond human perception, thereby offering reliable and highly precise diagnostic support (6). In dentistry, AI applications have expanded to include caries detection, anatomical landmark segmentation, and prognosis prediction (7-9). However, most existing AI studies have been conducted on adult radiographs, while research on pediatric radiographs remains limited (10). Pediatric imaging poses unique diagnostic challenges that cannot be directly addressed by models trained solely on adult datasets. Therefore, more pediatric-specific AI studies are required to bridge this gap (11).

YOLO (You Only Look Once) is a deep learning algorithm for image-based object detection (12). Multiple variants have been described in the literature (13). In this study, a YOLO-based model was employed to detect periapical lesions on pediatric panoramic radiographs, and its diagnostic performance was evaluated against that of dental students.



**Materials and Methods**

In accordance with the principles outlined in the 1964 Helsinki Declaration concerning medical research ethics, this study was conducted. The research received approval from the Institutional Ethics Committee at Istanbul Okan University, Istanbul, Turkey (Approval number 172/33).

**Patient Selection**

Panoramic radiographs from patients who visited the Istanbul Okan University Faculty of Dentistry between 2022 and 2024 for various indications (infectious conditions, developmental anomalies, evaluation of bony structures, malocclusion, dental trauma, etc.) were retrospectively selected for analysis. Periapical lesions were characterized as round or oval-shaped radiolucency involving the periapical and furcal regions, measuring less than 1 cm in size, associated with deep dentinal caries on the crown, and potentially exhibiting internal or external root resorption. Panoramic radiographs of patients with periapical lesions in maxillary and mandibular deciduous teeth aged between 6 and 12 years were included in the study. Radiographs depicting periapical lesions in deciduous teeth without an underlying permanent tooth bud were excluded from the study to ensure a more robust and representative deep learning model. The study included all periapical lesions, whether or not they affected the underlying permanent tooth germ, that resulted in the destruction of the bone at the apices of deciduous tooth roots.

**Preparation of the Radiographic Data Set**

All panoramic images included in the study were acquired using the Planmeca Promax 2D S3 unit (Planmeca, Helsinki, Finland), with the manufacturer's recommended exposure parameters (66 kVp, 8 mA, 16 s). Radiographs that did not meet the standard exposure parameters (e.g., improper contrast or brightness) were excluded from further analysis. In this study, approximately 400 panoramic images were evaluated by two dentomaxillofacial radiologists via the Picture Archiving and Communication System (PACS). Following a meticulous review, 333 radiographs with 391 periapical lesions were selectively selected in accordance with predefined inclusion criteria (Figure 1).

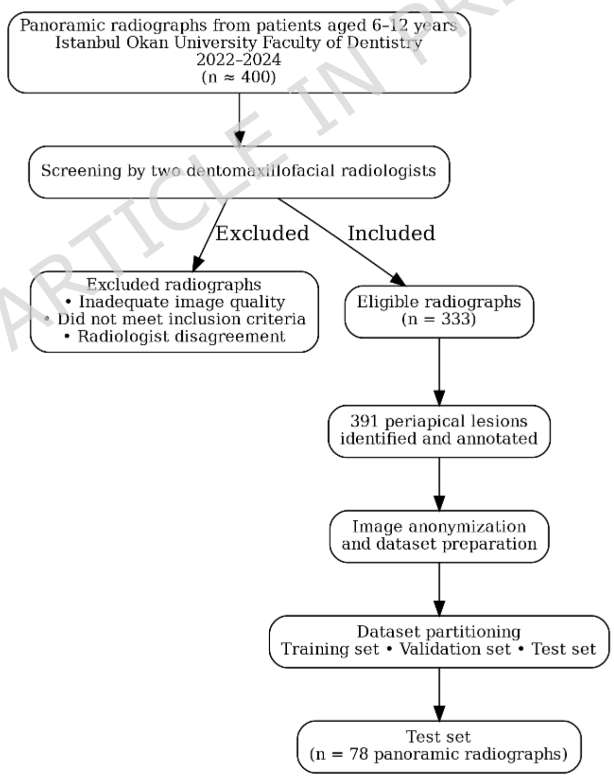



**Figure 1.** The flowchart for the dataset creation process shows the inclusion and exclusion criteria and the number of cases at each stage.

Periapical lesions were diagnosed solely on the basis of panoramic radiographic findings assessed by experienced oral and maxillofacial radiologists, and no clinical, histopathological, or longitudinal follow-up confirmation was available. Disagreements between the radiologists led to the exclusion of the respective radiograph from the study. After selecting the images, the radiologists collaboratively annotated them on the makesense.ai platform. In the annotation software (Figure 2), periapical lesions were delineated using rectangular bounding boxes, and the exported dataset included both image and text information, along with lesion coordinates, for subsequent analysis.

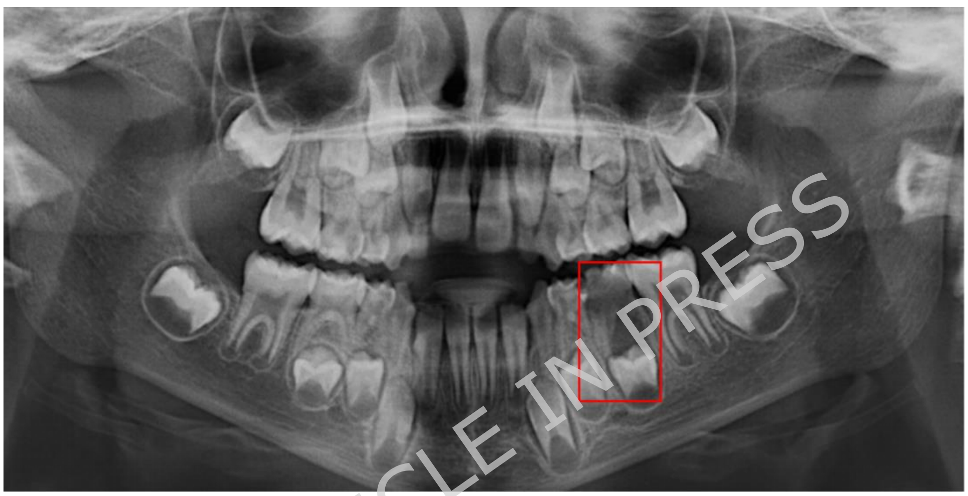


**Figure 2.** A cropped panoramic radiograph shows an example of a deciduous tooth with deep dentinal caries and a wide periapical lesion, with the underlying developing permanent premolar highlighted in the red box. Resorption in the superior cortex of the embedded premolar follicle is noteworthy.

All data were annotated and pre-processed at a resolution of 640 × 640 pixels for the model. While this standard resizing was applied to the entire dataset as preprocessing, the filter-based pipeline functioned strictly as a non-geometric data augmentation strategy for the training subset. The dataset was obtained from a single center and included relatively few cases, as panoramic radiographs are less available in early childhood than in adults. To address this limitation and enhance the model's feature extraction capability, a preprocessing-driven data expansion strategy was employed rather than conventional geometric augmentation (e.g., rotation or flipping). This approach generated 75 additional training images by applying a specialized image processing pipeline to a subset of the original training data. The preprocessing-driven data expansion workflow is summarized in Figure 3.

The expansion of these additional images was strictly based on the following preprocessing operations:

*Grayscale Transformation:* Images were converted to a single-channel grayscale representation using the cv2.cvtColor function to focus on pixel intensity and brightness distribution, reducing computational complexity.

*Noise Reduction:* A Gaussian Blur filter with a 5x5 kernel size was applied to eliminate radiographic noise and prevent the deep learning model from overfitting to imaging artifacts.

*Edge Enhancement:* To highlight the boundaries of periapical lesions, Canny Edge Detection was implemented with optimized dual-threshold values (low threshold = 100, high threshold = 200).

The 'process_image' function in the supplied code snippet focuses on transforming the input image into a greyscale representation, thereby discarding colour information while emphasizing brightness distribution. This transformation is implemented through the 'cv2.cvtColor' function from the OpenCV library. Greyscale conversion is widely employed as an initial step in image processing and computer vision tasks, as it facilitates the extraction of useful features for applications such as object detection and recognition while enhancing robustness against variations in lighting conditions. By incorporating these 75 edge-enhanced and filtered versions into the training pool, the model was forced to prioritize structural contours and density variations, crucial for identifying radiolucent lesions, which may help the model focus on structural radiographic features without introducing synthetic geometric distortions.

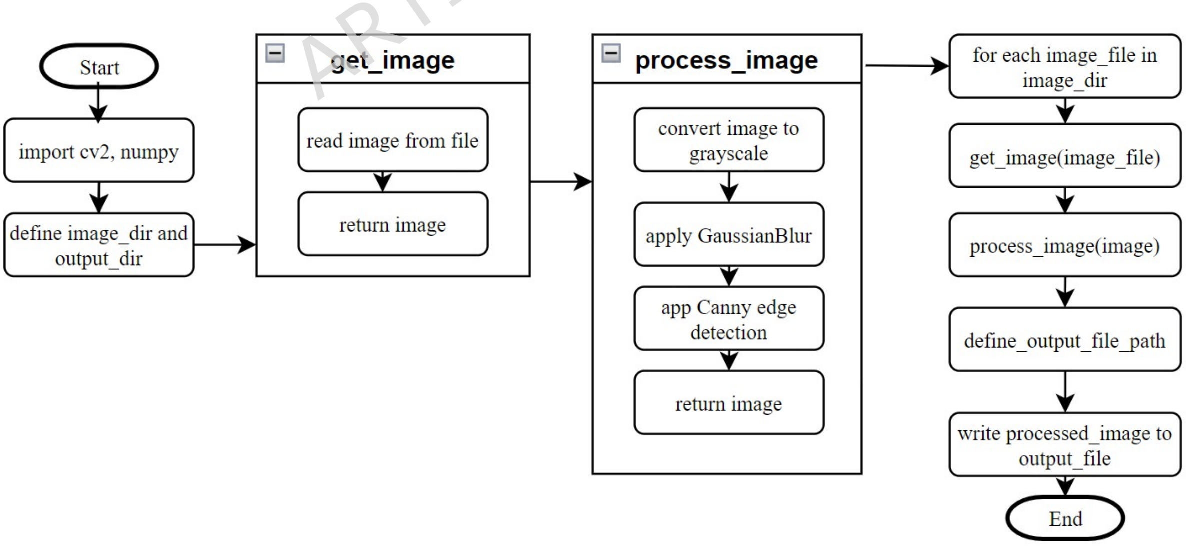



**Figure 3.** Feature-based preprocessing expansion steps.

**YOLO Architecture**

YOLO algorithms partition an input image into an N×N grid, where each grid cell predicts bounding boxes and object classes. A key advantage of YOLO lies in its speed, as the model processes the entire image in a single pass, outperforming many other object detection methods. The YOLOv7 architecture extends prior versions by integrating RepConv, a technique designed to optimize convolution layer outputs and reduce accuracy loss. Furthermore, by performing multi-scale object detection within a single inference, YOLOv7 achieves high performance, making it particularly effective for real-time applications.

**Model Parameters and Evaluations**

To ensure a systematic and unbiased evaluation of the YOLOv7 model, the initial dataset of 333 panoramic radiographs was divided into three subsets: a training set, a validation set and an independent test set. This split was performed at image level to prevent information leakage and ensure that radiographic data from the same patient did not appear in more than one subset. To prevent data leakage a single panoramic radiograph per patient was included in the dataset. Furthermore subsets were separated using anonymized patient IDs, ensuring no overlap between the training, validation, and testing phases. The training set was used to optimize the model's weight parameters, while the validation set was used for hyperparameter tuning and to monitor the learning process and reduce the risk of overfitting. The independent test set was reserved exclusively for the final performance evaluation and comparison with dental students.



The final training set consisted of 302 images, including 227 original radiographs and 75 processed images. In contrast, the validation and test sets were kept unchanged and contained only original images, with 28 and 78 radiographs, respectively, to preserve the clinical authenticity of the evaluation. The detailed distribution of images and instances across these subsets is presented in Table 1.

**Table 1:** Distribution of the Dataset

| Subset | Original Data | Processed Data | Total Data |
|---|---|---|---|
| Training | 227 | 75 | 302 |
| Validation | 28 | 0 | 28 |

| | | | |
|---|---|---|---|
| Test | 78 | 0 | 78 |
| Total | 333 | 75 | 408 |

Model training was conducted in Google Colab Notebook using a Tesla T4 GPU with 12.7 GB of RAM. The YOLOv7 model was implemented in Python, with hyperparameters set to a learning rate of 0.01, a batch size of 8, and 2000 epochs. The training process aimed to automate the detection of periapical lesions. Training was executed for a fixed duration of 2000 epochs without an early stopping mechanism, as the validation curves demonstrated steady and stable convergence until the end of the run. The optimal model weights were selected based on the highest validation fitness score, which automatically evaluates the peak mAP@0.5 and mAP@[0.5:0.95] performance during training.

**Diagnostic Evaluation by Dental Students**

To compare the performance of the YOLOv7 model with human observers, a subset of 78 panoramic radiographs was selected from the independent test dataset. This subset consisted of 33 healthy images and 45 images containing periapical lesions originating from deciduous teeth.

The images were evaluated by 10 final-year (intern) dental students. Each student independently assessed the same set of 78 panoramic images, which were presented in a randomized order. The students were instructed to determine the presence or absence of periapical lesions and to record the corresponding tooth numbers. Based on these evaluations, each student's responses were classified as true positives (TP), true negatives (TN), false positives (FP), or false negatives (FN). Using these values, performance metrics including accuracy, sensitivity, specificity, precision, and F1-score were calculated separately for each student.

All observations were conducted under the researcher's supervision, using the same medical monitor (Radioforce MX216, EIZO Corporation, Ishikawa, Japan), in dim lighting during a single session. Each session lasted approximately 30 minutes. In the YOLOv7 algorithm, only diagnostic accuracy, rather than completion time, was compared between the model and the students.

**Performance Evaluation Metrics**

A confusion matrix was utilized alongside the procedures and metrics outlined in Table 2 and Equation (14) to evaluate the model's performance. To ensure full interpretability of the reported metrics, the YOLOv7 model's technical detection performance was evaluated at the lesion level. A True Positive (TP) represents the correct detection of a lesioned tooth, whereas a False Positive (FP) corresponds to the erroneous identification of a lesion in a healthy tooth. Conversely, a False Negative (FN) indicates failure to detect an existing lesion, and a True Negative (TN) denotes correct recognition of a lesion-free tooth. Based on FP, TN, and FN, the following metrics were calculated:

**Table 2:** The evaluation criteria to measure the performance of the model.

| Criteria | Formulas |
|---|---|
| Accuracy Rate (ACC) | (TP+TN)/(TP+TN+FP+FN) |
| Sensitivity - True Positive Rate (TPR) | TP/(TP+FN) |
| Specificity - True Negative Rate (TNR) | TN/(TN+FP) |
| F1 Score | 2*((PPV*TPR)/(PPV+TPR)) |
| False Positive Rate (FPR) | FP/(FP+TN) |
| False Negative Rate (FNR) | FN/(FN+TP) |
| Positive Predictive Value - Precision (PPV) | TP/(TP+FP) |
| Error Rate (ERR) | (FP+FN)/(TP+TN+FP+FN) |

Additionally, mAP was computed as a key performance metric, reflecting the model's prediction accuracy across multiple Intersection-Over-Union (IoU) thresholds through precision-recall analysis. Detection performance of the YOLOv7 model was evaluated at the lesion level using bounding-box matching with an Intersection-over-Union (IoU) threshold of 0.5. For comparison with the dental students, image-level diagnostic decisions, including lesion localization according to tooth position, were used.

$$mAP = \frac{1}{n}\sum_{1}^{n} AP_k$$

$$Where:$$

$$AP = \sum_{0}^{n}(recalls(i) - Recalls(i+1)).Precision(i)$$

$$n: number\ of\ classes, i: tresholds$$

The performance of YOLOv7 in detecting periapical lesions on the test dataset was compared with that of intern dentistry students. Ten students were provided with the same test dataset images, presented in random order, and asked to detect lesions on pediatric panoramic radiographs. Their responses were coded as TP, TN, FP, and FN.

**Statistical Analysis**

For each panoramic radiograph, performance metrics were calculated for the YOLOv7 model and compared with the corresponding mean values from the 10 students, yielding paired observations at the image level.

Since the data did not follow a normal distribution, a nonparametric paired approach was adopted. The Wilcoxon signed-rank test was used to assess the statistical significance of differences in diagnostic accuracy between the YOLOv7 model and the students. Diagnostic accuracy was selected as the primary outcome measure and was the only metric included in the statistical comparison. Sensitivity, specificity, precision (PPV), F1-score, false positive rate, false negative rate, and error rate were calculated and reported descriptively.

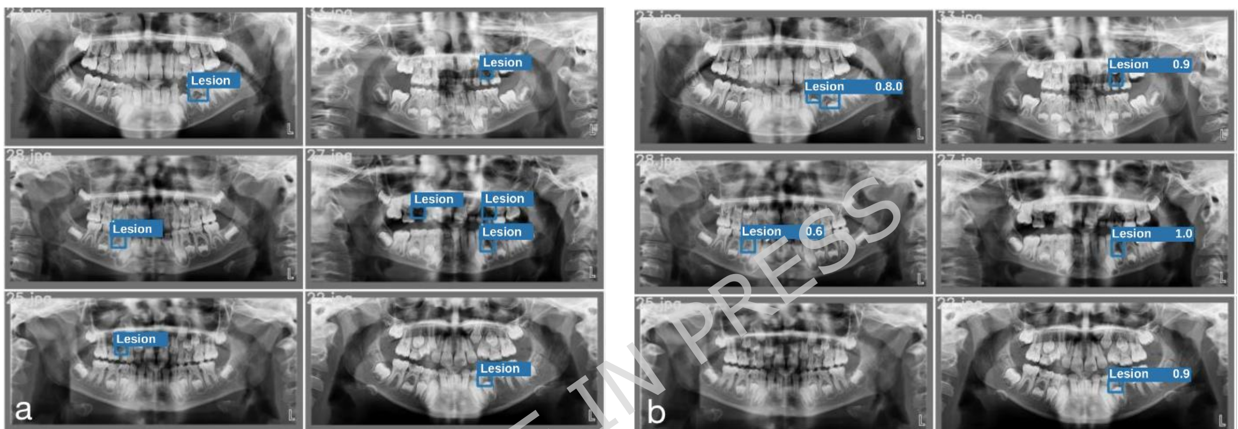
A significance level of $p < 0.05$ was adopted. Statistical analysis was carried out using SPSS Statistics 29.0.2.0 (IBM, Armonk, NY, USA) and Microsoft Excel.

**Results**

The training process was performed on Google Colab using Python. Figure 4 illustrates the training and validation curves of the YOLOv7 model, including box loss, objectness loss, precision, recall, and mean Average Precision (mAP). Model performance improved consistently over the 2000 epochs, with final precision and recall values of 0.906 and 0.736, respectively. At convergence, the model achieved an mAP@0.5 of approximately 0.8 and a mAP@[0.5:0.95] of around 0.3, indicating reliable lesion detection performance. Two radiologists independently reviewed all radiographs, and retrospective evaluation demonstrated a high absolute percentage interobserver agreement rate of 97.2%, with absolute intraobserver agreement rates of 96.4% and 98.3% for the first and second observers, respectively, demonstrating high descriptive consistency in the reference standard.

**Figure 4.** Loss functions (box and objectness), precision, recall, and mean average precision (mAP) scores of the YOLOv7 model's training and validation sets.

Representative examples are shown in Figure 5, where expert annotations (a) and YOLOv7 predictions (b) are compared on the same test images. In this case, experts labeled 8 true positive lesion sites across 6 panoramic radiographs, while YOLOv7 correctly identified 5 and missed 1. This example highlights the model's ability to approximate expert performance, although it does not represent the full dataset distribution.



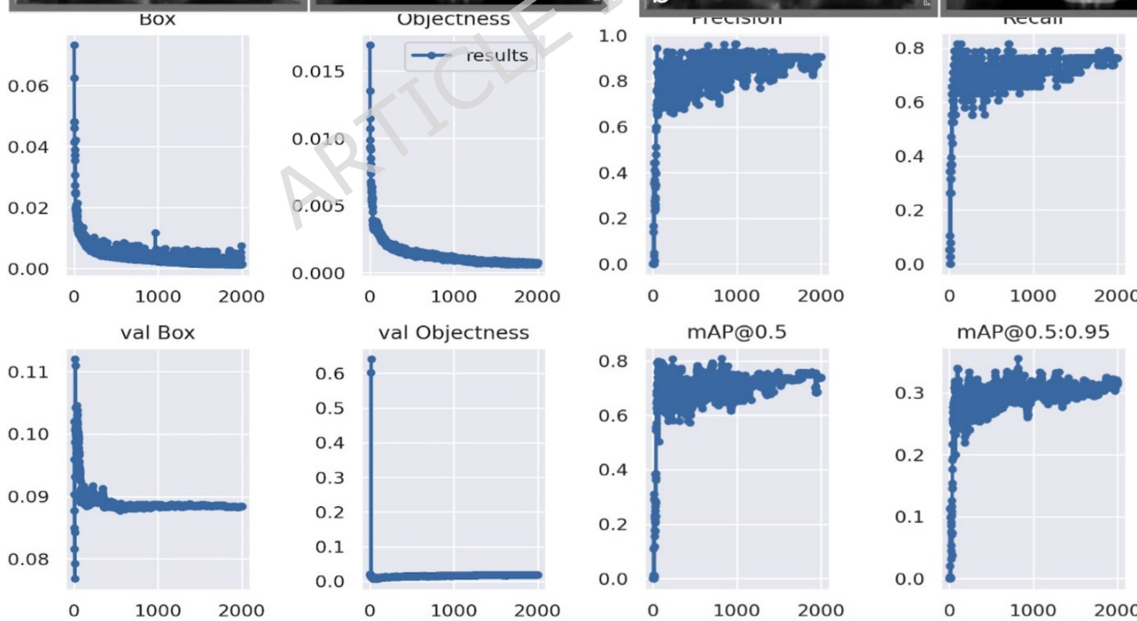






**Figure 5.** The blue bounding boxes on the panoramic radiographs (a) show the oral radiologist labels for the test data set, while (b) shows the YOLOv7 predictions for the test data set. 8 TP apical lesion sites were labeled by oral radiologists on 6 panoramics, and YOLOv7 performed well with 5 TP, 2 FN, and 1 FP.

Compared with 10 final-year dental students, YOLOv7 achieved superior diagnostic performance (Table 3). The model reached an accuracy rate of 88.0%, compared to 76.3% for the students. Sensitivity (76.1% vs. 55.2%) and specificity (99.8% vs. 97.3%) were both higher for the model. Precision (PPV) was nearly perfect (99.8% vs. 95.3%), while the F1 score improved markedly from 69.9% to 86.4%. Error rate was reduced by half (12.0% vs. 23.7%). Importantly, false positive and false negative rates also decreased substantially (0.2% vs. 2.7% and 23.9% vs. 44.8%, respectively). The Wilcoxon signed-rank test demonstrated a statistically significant difference in diagnostic accuracy between the YOLOv7 model and the student group ($p = 0.014$). These findings indicate that YOLOv7 achieved higher diagnostic accuracy than final-year dental students under the conditions of this study.

**Table 3:** Results of student and YOLOv7 test set evaluation based on the performance criteria

| Performance Metrics | Students Results | YOLOv7 Algorithm Results |
|---|---|---|
| Accuracy Rate (ACC) | 76.3% | 88.0% |
| Sensitivity (TPR) | 55.2% | 76.1% |
| Specificity (TNR) | 97.3% | 99.8% |
| F1 Score | 69.9% | 86.4% |
| False Positive Rate (FPR) | 2.7% | 0.2% |
| False Negative Rate (FNR) | 44.8% | 23.9% |
| Positive Predictive Value - Precision (PPV) | 95.3% | 99.8% |
| Error Rate (ERR) | 23.7% | 12.0% |

## Discussion

Dentomaxillofacial radiology plays a central role in diagnosing and managing dental conditions, yet image interpretation can be time-consuming and often requires specialist knowledge, particularly in mixed-dentition cases (15). Recent advances in AI have extended applications to caries detection, periodontal assessment, orthodontic planning, and lesion analysis (16). In pediatric dentistry, however, progress has been limited due to radiation constraints, patient cooperation issues, and imaging artifacts (17). While existing literature has addressed AI applications such as tooth numbering, caries prediction, and age estimation, very few studies have focused on periapical lesions in the mixed dentition period (11, 18). Lee et al. (19) demonstrated that more than half of children aged 6–12 years required endodontic treatment, underscoring the clinical need for improved diagnostic approaches.

Periapical radiographs are generally considered the most sensitive imaging modality for detecting apical lesions; however, their routine use in pediatric patients may be limited by practical challenges, including patient cooperation and positioning difficulties. Consequently, panoramic radiography is often a more feasible imaging modality in pediatric dental practice, despite its relatively lower sensitivity in detecting small or early-stage lesions (1, 3). The present study evaluated the performance of YOLOv7 in detecting periapical lesions on pediatric panoramic radiographs. Our findings indicate that the model can serve as a decision-support tool by highlighting suspicious regions and assisting clinicians in identifying them during panoramic evaluation. Additionally, such tools may provide pedagogical value by functioning as interactive training platforms for dental students, offering immediate feedback on radiographic interpretation. In practical settings, such systems could be integrated into digital radiology platforms to automatically flag suspicious regions during routine panoramic evaluation, assisting clinicians in identifying lesions that may require closer examination or further imaging (16, 17).

Periapical lesions in children are diagnostically challenging because they are often small and lack distinctive radiographic features. Manual interpretation is labor-intensive and subjective, whereas deep learning methods offer objective, reproducible alternatives (4). Previous studies have shown comparable benefits of AI in pediatric imaging. Ahn et al. (20) reported high diagnostic accuracy for mesiodens detection using ResNet-based models, whereas Beser et al. (21) found that YOLOv5 had difficulty with tooth segmentation in mixed dentition, illustrating the complexity of pediatric images. Similarly, Chen et al. (4) demonstrated that a convolutional neural network achieved an F1 score of 82% for proximal caries detection, compared with 57% among graduate students, and Cho et al. (22) showed that CNNs outperformed dermatologists in distinguishing malignant from benign lip lesions. Consistent with these reports, our YOLOv7 model outperformed final-year dental students in terms of accuracy, sensitivity, specificity, precision, and F1 score.

Despite the promising performance of YOLOv7, false positives and false negatives should also be considered. False-positive detections may lead to unnecessary examinations or increased clinician workload, whereas false-negative results may result in missed periapical lesions and delayed diagnosis or treatment. In the present study, YOLOv7 demonstrated lower false positive and false negative rates than the students, suggesting improved diagnostic consistency. Nevertheless, AI-based systems should currently be considered supportive decision-aid tools rather than standalone diagnostic systems, and final clinical interpretation should remain under the supervision of experienced clinicians.

YOLO is a one-stage object detection architecture that balances high accuracy with computational efficiency, making it particularly attractive for medical imaging tasks. Earlier versions have been widely used in dentistry to detect caries, cysts, tumors, and pulpal calcifications (23). In this study, YOLOv7 was selected because it offered stable performance and robust documentation at the time of implementation (13). Since then, newer versions such as YOLOv10 and YOLOv11 have been introduced, with incremental improvements on benchmark datasets like COCO (24, 25). Gašparović et al. (26) noted that performance differences between YOLO versions are often modest and dataset-dependent. In this context, Ayhan et al. (27) proposed RCFLA-YOLO for automated root canal filling assessment on periapical radiographs. Although the clinical targets differ, this development demonstrates the expanding versatility of advanced YOLO models in dental education. Given the specific challenges of pediatric panoramic radiographs, YOLOv7 represented a reasonable and pragmatic choice. The single-center reliance on Planmeca Promax hardware introduces a domain shift risk, necessitating multicenter external validation prior to clinical translation. Future research should assess newer versions in larger, multicenter datasets to determine whether reported gains translate into meaningful improvements in clinical practice.

This study has several strengths. It addresses an underexplored area—AI applications in pediatric panoramic radiographs—highlighting their potential for both clinical support and education. Furthermore, the direct comparison with dental trainees underscores AI's potential pedagogical role in dental radiology training.

Nevertheless, several limitations should be acknowledged. First, the dataset was relatively small and obtained from a single institution, which may limit the generalizability of the findings to broader clinical populations. Although feature-based preprocessing expansion was applied to partially mitigate this limitation, the model was trained and evaluated using images from a single-center dataset. Therefore, future studies involving larger datasets collected from multiple institutions are necessary to validate the robustness and clinical applicability of AI-based detection systems in pediatric dental radiology. In addition, future research should explore the applicability of newer YOLO architectures (e.g., YOLOv10 and YOLOv11) and assess whether improvements reported in benchmark datasets translate into clinically meaningful gains in pediatric imaging tasks. AI systems should be considered as decision-support tools rather than standalone diagnostic systems. In addition, student evaluations were conducted in a single-session setting, which may not fully reflect the conditions of routine clinical interpretation. Additionally, low-quality and diagnostically uncertain radiographs were excluded during retrospective dataset construction, potentially introducing spectrum bias.

**Conclusion**

Within the limitations of this retrospective single-center internal validation study, the YOLOv7 algorithm demonstrated promising performance for detecting periapical lesions in children with mixed dentition. Compared with final-year dental students, the algorithm achieved higher accuracy, sensitivity, specificity, precision, and F1 score while also demonstrating lower false positive and false negative rates. These findings indicate that deep learning methods may contribute to more timely and accurate identification of periapical lesions in pediatric panoramic radiographs. Incorporating such algorithms into clinical workflows may assist radiographic evaluation and support clinical decision-making in pediatric dentistry; however, further multicenter validation studies are required before routine clinical implementation can be considered. However, the relatively small size of the test dataset may limit the generalizability of the present findings, and further validation with larger multicenter datasets is required. Nevertheless, despite the promising performance of the model, AI systems should currently be considered supportive tools to assist clinical decision-making rather than standalone diagnostic systems.

## Declarations

### Ethics approval and consent to participate

This retrospective study was approved by the Istanbul Okan University Institutional Ethics Committee (Approval No: 172/33, Date: 10.01.2024). Given the retrospective design and the use of fully anonymized radiographs, the requirement for informed consent was waived by the committee.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This research received no external funding.

### Authors' contributions

Conceptualization: (C.B., F.Y.); Methodology: (C.B., PS.M., P.AG., Y.K.); Data curation: (C.B., F.Y.); Formal analysis: (PS.M., P.AG., Y.K.); Writing—original draft: (PS.M., P.AG., C.B.); Writing—review & editing: all authors. All authors read and approved the final manuscript.

### Availability of data and materials

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Clinical trial registration**

Clinical trial number: not applicable.

**References**


1. Segura-Egea JJ, Martín-González J, Castellanos-Cosano L. Endodontic medicine: Connections between apical periodontitis and systemic diseases. Int Endod J. 2015;48(10):933-51.
2. Kalaoglu EE, Alyamac S. Prevalence and factors of peri-apical lesions in primary molars. Int Dent J. 2024;74:S87-8.
3. Yang NY, et al. Increased interleukin 1α and interleukin 1β expression is involved in the progression of periapical lesions in primary teeth. BMC Oral Health. 2018;18:1-6.
4. Chen X, Guo J, Ye J, Zhang M, Liang Y. Detection of proximal caries lesions on bitewing radiographs using deep learning method. Caries Res. 2022;56(5-6):455-63.
5. Bahadir HS, Keskin NB, Çakmak EŞK, Güneç G, Cesur Aydin K, Sari SK. Comparison of diagnoses made by dentistry students and by artificial intelligence dentists. J Dent Educ. 2025;89(7):1165-73.
6. Sohrabniya F, Hassanzadeh-Samani S, Ourang SA, Jafari B, Farzinnia G, Gorjinejad F, et al. Exploring a decade of deep learning in dentistry: A comprehensive mapping review. Clin Oral Investig. 2025;29(2):1-11.
7. Park EY, Cho H, Kang S, Jeong S, Kim EK. Caries detection with tooth surface segmentation on intraoral photographic images using deep learning. BMC Oral Health. 2022;22(1):573.
8. Oh S, Kim YJ, Kim J, Jung JH, Lim HJ, Kim BC, et al. Deep learning-based prediction of osseointegration for dental implant using plain radiography. BMC Oral Health. 2023;23(1):208.
9. Yuce F, Buyuk C, Bilgir E, et al. Deploying a novel deep learning framework for segmentation of specific anatomical structures on cone-beam CT. Oral Radiol. 2025. doi:10.1007/s11282-025-00831-4.
10. Turosz N, Chęcińska K, Chęciński M, Lubecka K, Bliźniak F, Sikora M. Artificial intelligence (AI) assessment of pediatric dental panoramic radiographs (DPRs): A clinical study. Pediatr Rep. 2024;16(3):794-805. doi:10.3390/pediatric16030067.
11. Laborie LB, Lee J, Antram E, et al. Performance of adult-trained artificial intelligence models in paediatric imaging: A scoping review. Eur Radiol. 2026. doi:10.1007/s00330-026-12354-5.
12. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 779-

88.

13. Hussain M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines. 2023;11(7):677. doi:10.3390/machines11070677.
14. Zeren MT, Arslankaya S, Altuntaş Y, Cam N, Kırelli Y, Özdemir MH. Doctors versus YOLO: Comparison between YOLO algorithm, orthopedic and traumatology resident doctors and general practitioners on detection of proximal femoral fractures on X-ray images with multi methods. 2024. doi:10.1142/S0218213023500562.
15. Srivastav S, Chandrakar R, Gupta S, Babhulkar V, Agrawal S, Jaiswal A, et al. ChatGPT in radiology: The advantages and limitations of artificial intelligence for medical imaging diagnosis. Cureus. 2023;15(7):e37546142.
16. Ding H, Wu J, Zhao W, Matinlinna JP, Burrow MF, Tsoi JK. Artificial intelligence in dentistry: A review. Front Dent Med. 2023;4:1085251.
17. Ramgopal S, Sanchez-Pinto LN, Horvat CM, Carroll MS, Luo Y, Florin TA. Artificial intelligence-based clinical decision support in pediatrics. Pediatr Res. 2023;93(2):334-41.
18. Bonfanti-Gris M, Herrera A, Salido Rodríguez-Manzaneque MP, Martínez-Rus F, Pradies G. Deep learning for tooth detection and segmentation in panoramic radiographs: A systematic review and meta-analysis. BMC Oral Health. 2025;25(1):1280.
19. Lee SM, Yu YH, Karabucak B. Endodontic treatments on permanent teeth in pediatric patients aged 6-12 years old. J Dent Sci. 2023;18(3):1109-15.
20. Ahn Y, Hwang JJ, Jung YH, Jeong T, Shin J. Automated mesiodens classification system using deep learning on panoramic radiographs of children. Diagnostics (Basel). 2021;11(8):1477. doi:10.3390/diagnostics11081477.
21. Beser B, Reis T, Berber MN, Topaloglu E, Gungor E, Kılıc MC, et al. YOLOv5-based deep learning approach for tooth detection and segmentation on pediatric panoramic radiographs in mixed dentition. BMC Med Imaging. 2024;24(1):172.
22. Cho SI, Sun S, Mun JH, Kim C, Kim SY, Cho S, et al. Dermatologist-level classification of malignant lip diseases using a deep convolutional neural network. Br J Dermatol. 2020;182(6):1388-94.
23. Diwan T, Anirudh G, Tembhurne JV. Object detection using YOLO: Challenges, architectural successors,

datasets and applications. Multimed Tools Appl. 2023;82(6):9243-75.

24. Wang CY, Yeh IH, Liao HYM. YOLOv10: Real-time end-to-end object detection. arXiv [preprint]. 2024. arXiv:2405.14458.
25. Ultralytics. YOLO11 release notes. 2024. Available from: https://docs.ultralytics.com. Accessed 10 May 2026.
26. Gašparović B, Mauša G, Rukavina J, Lerga J. Evaluating YOLOv5, YOLOv6, YOLOv7, and YOLOv8 in underwater environment: Is there real improvement? In: 2023 8th International Conference on Smart and Sustainable Technologies (SpliTech); 2023. p. 1-4.
27. Ayhan M, Kayadibi İ, Aykanat B. RCFLA-YOLO: a deep learning-driven framework for the automated assessment of root canal filling quality in periapical radiographs. BMC Med Educ. 2025 Jul 1;25(1):894. doi:10.1186/s12909-025-07483-2.