PerioDet: Large-Scale Panoramic Radiograph Benchmark for Clinical-Oriented Apical Periodontitis Detection

Xiaocheng Fang[1,2*], Jieyi Cai[1,2*], Huanyu Liu[1,2], Chengju Zhou[2], Minhua Lu[3], and Bingzhi Chen[1†]

[1] Beijing Institute of Technology, Zhuhai
[2] South China Normal University
[3] Shenzhen University
fangxiaocheng@m.scnu.edu.cn, chenbingzhi@bit.edu.cn

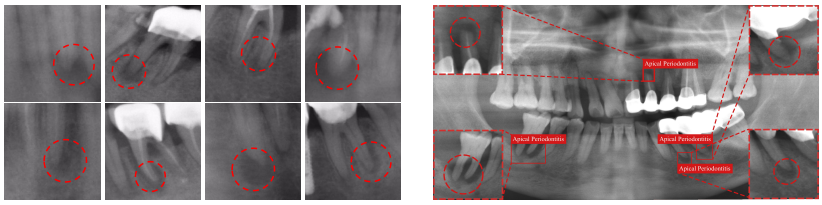
Abstract. Apical periodontitis is a prevalent oral pathology that presents significant public health challenges. Despite advances in automated diagnostic systems across various medical fields, the development of Computer-Aided Diagnosis (CAD) applications for apical periodontitis is still constrained by the lack of a large-scale, high-quality annotated dataset. To address this issue, we release a large-scale panoramic radiograph benchmark called "**PerioXrays**", comprising 3,673 images and 5,662 meticulously annotated instances of apical periodontitis. To the best of our knowledge, this is the first benchmark dataset for automated apical periodontitis diagnosis. This paper further proposes a clinical-oriented apical periodontitis detection (**PerioDet**) paradigm, which jointly incorporates Background-Denoising Attention (BDA) and IoU-Dynamic Calibration (IDC) mechanisms to address the challenges posed by background noise and small targets in automated detection. Extensive experiments on the PerioXrays dataset demonstrate the superiority of PerioDet in advancing automated apical periodontitis detection. Additionally, a well-designed human-computer collaborative experiment underscores the clinical applicability of our method as an auxiliary diagnostic tool for professional dentists. The project is publicly accessible at `https://github.com/XiaochengFang/MICCAI2025_PerioDet`.

Keywords: Computer-Aided Diagnosis · Panoramic Radiograph · Apical Periodontitis Diagnosis · PerioXrays · PerioDet.

1 Introduction

Apical periodontitis is one of the most prevalent oral pathologies, affecting approximately 52% of adults worldwide [23]. Early and accurate diagnosis is essential for the effective management and planning of endodontic treatments to prevent complications [8,14]. Panoramic dental radiography, a standard diagnostic

*Equal contribution.
†Corresponding author: Bingzhi Chen



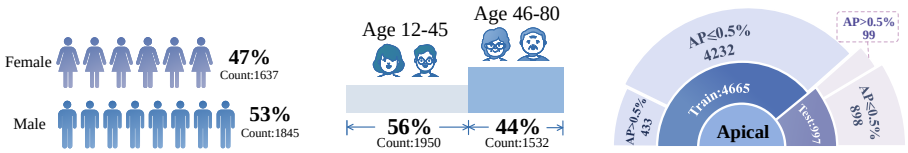
(a) Representative examples (b) Example with manual annotations

Fig. 1: Apical periodontitis in panoramic radiograph from PerioXrays.

tool, provides a comprehensive view of dental structures, aiding in the detection of apical periodontitis [6]. However, the accuracy of diagnosis can be affected by the radiologist's experience and fatigue, particularly during prolonged hours of interpreting panoramic X-rays. This is especially problematic when identifying subtle lesions, as illustrated in Fig. 1a, where apical periodontitis diagnosis remains challenging due to the subtle or ambiguous characteristics of lesions, as well as background noise such as low image quality, poor contrast, and artifacts. Thus, it is crucial to develop an automated system for computer-aided diagnosis of apical periodontitis from panoramic dental X-rays.

Recent advancements in medical image analysis have been greatly influenced by the development of large-scale datasets [24,20,11,4,3] and deep learning techniques [18,10,5]. Comprehensive datasets covering a wide range of medical conditions and imaging modalities have emerged. For example, the ChestXray14 dataset [24], containing over 112,000 chest X-ray images, has facilitated deep learning applications for detecting a variety of chest conditions. Similarly, the MURA dataset [20], with 40,000 upper limb radiographs, has been pivotal in developing models for diagnosing musculoskeletal diseases, especially bone and joint abnormalities. The RSNA Pediatric Bone Age Challenge dataset [11], consisting of 14,236 pediatric radiographs, aids in training models to assess bone age in children, supporting accurate monitoring of growth and development. These datasets have significantly advanced AI-driven diagnostic systems, demonstrating the potential of deep learning in condition detection, classification, and age assessment. However, progress in AI-based diagnostics for apical periodontitis detection remains limited, mainly due to the absence of large-scale datasets.

Technically, automated detection of apical periodontitis faces two key challenges: **1) Background noise:** The detection of apical periodontitis is often compromised by background noise arising from factors such as low image quality, poor contrast, and various artifacts in panoramic dental X-rays. These disturbances obscure lesion details, making it difficult to distinguish pathological changes from surrounding anatomical structures, thereby hindering accurate detection. **2) Small targets:** Apical periodontitis typically manifests as subtle, localized changes at the root apex with relatively small dimensions in radiographs. This characteristic poses a significant challenge for conventional image



(a) Gender distribution (b) Age distribution (c) Lesion size distribution

Fig. 2: Comprehensive dataset statistics and distribution about PerioXrays.

analysis methods, as precise identification techniques are required to detect even the slightest structural variations.

To address the above issues, we have developed a large-scale, high-quality annotated benchmark dataset, namely "**PerioXrays**" for automated apical periodontitis detection. The dataset comprises 3,673 panoramic X-rays with 5,662 meticulously annotated instances of apical periodontitis. Each annotation was reviewed by four experienced professional dentists to ensure accuracy and reliability. Additionally, we propose a clinical-oriented apical periodontitis (**PerioDet**) paradigm. To mitigate the impact of background noise in panoramic dental X-rays, the Background-Denoising Attention (BDA) module is designed to refine feature representations by modeling channel importance while establishing associations between target and scene features, effectively suppressing irrelevant background interference. To address the challenge of detecting small lesions, we propose the IoU-Dynamic Calibration (IDC) module, which employs an Adaptive IoU Threshold to dynamically adjust positive sample criteria based on lesion size, ensuring a sufficient number of anchors for small lesions. A Dynamic Label Assignment strategy refines anchor-ground truth matching, progressively adjusting sample selection throughout training to enhance localization precision. Our PerioDet paradigm, evaluated on the PerioXrays dataset, achieves state-of-the-art performance and demonstrates clinical potential as an auxiliary diagnostic tool for dentists in human-computer collaborative experiments.

2 PerioXrays Dataset Creation

2.1 Panoramic X-rays Collection

The PerioXrays dataset is a comprehensive collection of panoramic dental X-rays obtained from multiple hospitals, including dental clinics and orthodontic centers, gathered between 2022 and 2024. As shown in Fig. 2a and Fig. 2b, the dataset covers a wide range of apical periodontitis cases across diverse demographic groups, including various age ranges and genders. All images are standardized to a resolution of 1333×800 pixels to ensure consistency while minimizing information loss.

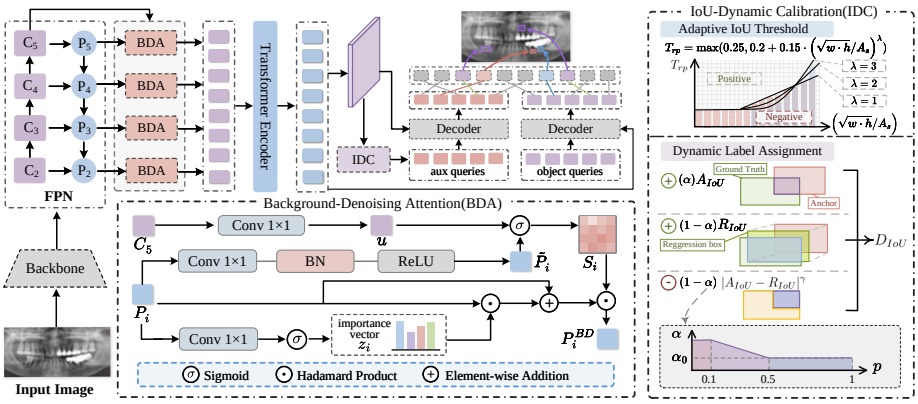


Fig. 3: Illustration of our proposed PerioDet paradigm, which incorporates BDA and IDC mechanisms for automated apical periodontitis detection accurately.

2.2 Apical Periodontitis Annotations

In the PerioXrays dataset, a custom annotation tool was used to delineate bounding boxes around instances of apical periodontitis. Each bounding box was precisely adjusted to capture the true extent of the lesion, avoiding both overextension and undersizing. The annotation process followed comprehensive guidelines based on established clinical diagnostic criteria for apical periodontitis [1,13,2]. Fig. 1b shows post-annotation examples from the dataset, highlighting the identified instances of apical periodontitis. To ensure accuracy, each image underwent a rigorous multi-stage review by four experienced professional dentists.

2.3 Dataset Distribution Statistics

The PerioXrays dataset contains 3,673 panoramic dental X-ray images from 3,482 unique patients, with a total of 5,662 annotated instances of apical periodontitis. Fig. 2c shows the distribution of object counts and the area ratio (AR) of labeled instances across the training set and test set. The AR metric represents the proportion of an image occupied by apical periodontitis lesions, indicating their size and spatial coverage. Detecting small lesions is particularly challenging due to their limited visibility and spatial presence. The high prevalence of small objects (AR$\leq$0.5%) underscores the need for advanced detection algorithms to accurately localize apical periodontitis.

3 Methodology

3.1 Background-Denoising Attention

To mitigate the impact of background noise in panoramic dental X-rays, we propose the Background-Denoising Attention (BDA) module, which enhances

feature representations by modeling channel importance while simultaneously capturing associations between target and scene features, effectively suppressing irrelevant background noise.

As shown in Fig. 3, P_i represents the features in the top-down path in Feature Pyramid Network(FPN) [15]. The BDA module generates a new feature map, P_i^{BD}, by weighting P_i using the importance vector z_i and the similarity map S_i. The importance vector z_i is computed as follows,

$$z_i = \text{Sigmoid}\left(\text{Conv}_{1\times1}(P_i)\right), \tag{1}$$

which is designed to prioritize feature channels that capture rich spatial details while suppressing background noise.

The similarity map S_i captures the relationship between the target and scene features, effectively suppressing irrelevant background interference. To compute S_i, two projection functions are employed to learn the target and scene features.

The projection of the target feature involves applying a series of operations to P_i, including a 1×1 convolution for dimensionality reduction, followed by Batch Normalization and ReLU activation to enhance non-linearity and stabilize training. This process is formulated as follows,

$$\tilde{P}_i = \text{ReLU}(\text{BN}(\text{Conv}_{1\times1}(P_i))). \tag{2}$$

Additionally, a one-dimensional scene embedding vector u is computed to capture the contextual information of the scene. This embedding vector is generated by applying a 1×1 convolution operation to the feature map C_5, effectively projecting the scene features into a compact representation while preserving essential contextual details. The computation of u is expressed as follows,

$$u = \text{Conv}_{1\times1}(C_5). \tag{3}$$

With the refined feature map $\tilde{P}_i$ and the scene embedding vector u, the similarity map S_i is calculated as follows,

$$S_i = \text{Sigmoid}(\tilde{P}_i \cdot u). \tag{4}$$

Based on the above analyses, the final output is computed as follows,

$$P_i^{BD} = (1 + z_i) \cdot P_i \odot S_i. \tag{5}$$

3.2 IoU-Dynamic Calibration

To address the problem posed by small targets in apical periodontitis detection, the IoU-Dynamic Calibration(IDC) module mainly benefits from the Adaptive IoU Threshold and Dynamic Label Assignment mechanisms.

Adaptive IoU Threshold. Traditional IoU thresholds struggle to capture small lesions due to their sensitivity to positional deviations[7]. To address this limitation, we propose an Adaptive IoU Threshold that dynamically adjusts the

positive sample thresholds based on the object's area, ensuring sufficient anchors for small lesions. The Adaptive IoU Threshold is defined as follows,

$$T_{rp} = \max(0.25, 0.2 + 0.15 \cdot \left(\sqrt{w \cdot h}/A_s\right)^{\lambda}), \tag{6}$$

where w and h denote the width and height of the object, and A_s represents the minimal area definition which is adjusted for different datasets. T_{rp} is the reference threshold for positive samples, while λ controls the growth rate of T_{rp}.

Dynamic Label Assignment. To enhance label assignment for small lesions, we propose a Dynamic Label Assignment that employs Dynamic IoU(DIoU) to adaptively adjust the sample selection during training, refining the matching process between anchors and ground truth (GT). DIoU is defined as follows,

$$D_{IoU} = \alpha \cdot A_{IoU} + (1-\alpha) \cdot R_{IoU} - (1-\alpha) \cdot |A_{IoU} - R_{IoU}|^{\gamma}, \tag{7}$$

where A_{IoU} represents the IoU of anchor and GT, R_{IoU} represents the IoU of regression box and GT, and α and γ are hyperparameters used for weighting.

Our Dynamic Label Assignment strategy follows the following procedure. First, we calculate T_{rp} based on Eq. 6. Subsequently, the DIoU of each anchor is computed during label assignment, and anchors are designated as positive samples when DIoU is greater than T_{rp}. To ensure the stability of training, the influence of R_{IoU} on sample selection is gradually increased with training iterations. The adjustment of α at each stage follows the criteria:

$$\alpha(p, \alpha_0) = \begin{cases} 1, & 0 \le p < 0.1 \\ \left(\frac{\alpha_0 - 1}{0.5 - 0.1}\right)(p - 0.1) + 1, & 0.1 \le p < 0.5 \\ \alpha_0, & p \ge 0.5 \end{cases} \tag{8}$$

where $\mathrm{p} = \frac{\mathrm{epoch}}{\mathrm{epochs}}$, with epoch representing the current training epoch and epochs the total number of epochs. The hyperparameter α_0 controls α.

4 Experiments

4.1 Experiment Setup

To evaluate our method, we compare various object detection baselines on the PerioXrays dataset, including **CNN-based** and **Transformer-based** approaches. The dataset was randomly split into a training set of 3,000 images and a test set of 673 images, with the division made at the patient level to ensure that all images from a single patient were assigned to the same set. We use Average Precision (AP) [16] as the primary metric, along with AP_{50} and AP_{75} at IoU thresholds of 0.5 and 0.75. Objects are categorized into small (S), medium (M), and large (L) sizes for detailed performance analysis. During training, we used AdamW as the optimizer with a batch size of 2, a 1× schedule (12 epochs), an initial learning rate of 2e-4, and a weight decay of 0.0001. All baseline models were trained with a **ResNet-50 backbone** for fair comparison.

Table 1: Comparison with detection approaches on the PerioXrays.

Models	Methods	Ref.	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
CNN	Faster RCNN [21]	NIPS'15	46.1	76.4	50.1	23.6	44.7	53.1
	ATSS [28]	CVPR'20	47.2	80.2	50.1	33.0	46.3	54.4
	YOLOX [9]	CVPR'21	46.7	79.5	48.4	28.3	45.2	55.1
	Sparse RCNN [22]	CVPR'21	47.0	80.1	49.5	36.9	45.1	54.3
	RFLA [25]	ECCV'22	47.5	80.3	50.6	36.4	46.1	55.0
	CFINet [26]	ICCV'23	47.7	80.6	51.0	37.2	46.4	55.6
Transformer	Deformable-DETR [29]	ICLR'20	38.4	75.8	35.1	21.7	36.6	46.6
	Conditional-DETR [19]	ICCV'21	42.6	78.4	41.0	23.3	41.1	49.7
	DAB-DETR [17]	ICLR'22	44.0	78.5	45.0	33.4	43.1	50.0
	DINO [27]	ICLR'23	48.9	81.2	51.3	20.4	47.6	56.9
	Co-DINO [30]	ICCV'23	49.9	81.5	53.4	36.8	48.5	57.7
	Salience-DETR [12]	CVPR'24	50.3	81.7	53.7	37.0	48.2	57.8
	Ours	-	**53.5**	**84.2**	**55.6**	**42.3**	**50.9**	**58.7**

Table 2: Ablation studies.

BDA	IDC	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
✗	✗	49.9	81.5	53.4	36.8	48.5	57.7
✓	✗	51.7	83.1	54.7	40.6	49.8	58.0
✗	✓	52.3	83.5	55.2	41.1	50.3	58.3
✓	✓	53.5	84.2	55.6	42.3	50.9	58.7

Table 3: Diagnostic comparison.

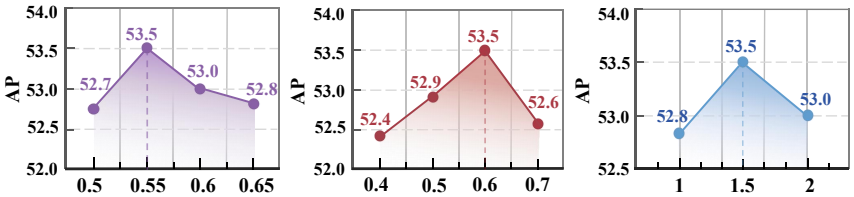
Methods	w/o PerioDet	w PerioDet
Precision	73.1%	92.5%(19.4% ↑)
Recall	74.3%	96.1%(21.8% ↑)
Efficiency	≈28 s/img	≈13 s/img(15s↓)

4.2 Comparisons with State-of-The-Arts

We evaluate the performance of our method on the PerioXrays dataset, where it surpasses all baseline approaches. As shown in Table 1, PerioDet achieves a 5.8% improvement in AP compared to leading CNN-based models such as CFINet (53.5% vs. 47.7%) and a 3.2% improvement over state-of-the-art Transformer-based models like Salience-DETR (53.5% vs. 50.3%). Notably, PerioDet excels in detecting small lesions, with a 5.1% improvement in AP_S.

4.3 Ablation Studies and Parameter Analysis

In our ablation studies, we systematically assess the contribution of each component in the PerioDet method using the PerioXrays dataset to validate its effectiveness. As shown in Table 2, the results demonstrate that all modules work synergistically, reinforcing each other and confirming the overall effectiveness of the method. Additionally, we analyze the impact of three critical hyperparameters, λ, α_0, and γ, as defined in Eq. 6, Eq. 7 and Eq. 8. Fig. 4 shows that our PerioDet achieves optimal performance with $\lambda = 0.55$, $\alpha_0 = 0.6$, and $\gamma = 1.5$.



(a) Parameter λ (b) Parameter α_0 (c) Parameter γ

Fig. 4: Parameter analysis (%) on the PerioXrays dataset.

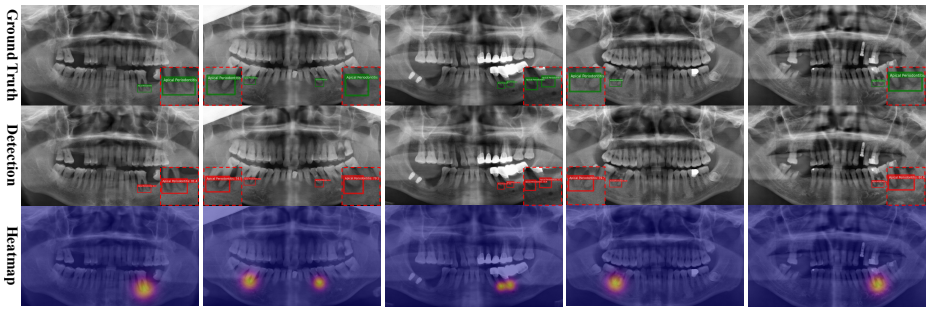


Fig. 5: Visualization of detections and heatmaps on the PerioXrays dataset.

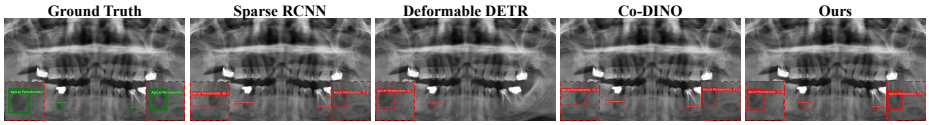


Fig. 6: Comparison of predicted bounding boxes across various models.

4.4 Human-Computer Collaborative Experiment

We designed a novel human-computer collaborative experiment to evaluate the clinical applicability of our method as an auxiliary diagnostic tool. Six professional dentists participated, including two junior, two mediate, and two senior professionals. None had prior exposure to the data, ensuring an unbiased evaluation. We randomly selected 100 panoramic X-ray images from PerioXrays as the test set, and participants made two consecutive diagnostic judgments for each image. As shown in Table 3, these results highlight the clinical applicability of PerioDet, demonstrating its effectiveness in real-world medical diagnostics.

4.5 Visualization Results

We visualize detections and heatmaps. As shown in Fig. 5, PerioDet accurately identifies apical periodontitis and effectively localizes the lesions. The attention distribution further emphasizes PerioDet's ability to capture relevant features,

highlighting its enhanced sensitivity to regions of interest. Additionally, Fig. 6 demonstrates that PerioDet achieves higher confidence scores and successfully detects apical periodontitis that other models miss.

5 Conclusion

In this paper, we release a large-scale panoramic radiograph benchmark called "PerioXrays". Moreover, we propose a clinical-oriented apical periodontitis detection (PerioDet) paradigm which jointly incorporates Background-Denoising Attention (BDA) and IoU-Dynamic Calibration (IDC) mechanisms to address the challenges posed by background noise and small targets. Extensive experiments on the PerioXrays dataset, along with a well-designed human-computer collaborative experiment, establish a strong foundation for automated apical periodontitis detection in dental healthcare.

Acknowledgments. This work was supported in part by the Guangdong Basic and Applied Basic Research Foundation (No. 2025A1515010225), and in part by the National Natural Science Foundation of China (No. 62302172).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Abbott, P.V.: Classification, diagnosis and clinical manifestations of apical periodontitis. Endodontic topics **8**(1), 36–54 (2004)
2. Armitage, G.C.: Clinical evaluation of periodontal diseases. Periodontology 2000 **7**(1), 39–53 (1995)
3. Chen, B., Fu, S., Fang, X., Cai, J., Zhang, B., Lu, M., Liu, Y.: Oralxrays-9: Towards hospital-scale panoramic x-ray anomaly detection via personalized multi-object query-aware mining. In: CVPR. pp. 15570–15579 (2025)
4. Chen, B., Fu, S., Liu, Y., Pan, J., Lu, G., Zhang, Z.: Cariesxrays: Enhancing caries detection in hospital-scale panoramic dental x-rays via feature pyramid contrastive learning. In: AAAI. vol. 38, pp. 21940–21948 (2024)
5. Chen, B., Liu, Y., Zhang, Z., Lu, G., Kong, A.W.K.: Transattunet: Multi-level attention-guided u-net with transformer for medical image segmentation. IEEE Transactions on Emerging Topics in Computational Intelligence (2023)
6. Choi, H.R., Siadari, T.S., Kim, J.E., Huh, K.H., Yi, W.J., Lee, S.S., Heo, M.S.: Automatic detection of teeth and dental treatment patterns on dental panoramic radiographs using deep neural networks. Forensic Sciences Research **7**(3), 456–466 (2022)
7. Dong, C., Duoqian, M.: Control distance iou and control distance iou loss for better bounding box regression. Pattern Recognition **137**, 109256 (2023)
8. Estrela, C., Leles, C.R., Hollanda, A.C.B., Moura, M.S., Pécora, J.D.: Prevalence and risk factors of apical periodontitis in endodontically treated teeth in a selected population of brazilian adults. Brazilian dental journal **19**, 34–39 (2008)

9. Ge, Z., Liu, S., Wang, F., Li, Z., Sun, J.: Yolox: Exceeding yolo series in 2021. arXiv preprint arXiv:2107.08430 (2021)
10. Greenwald, N.F., Miller, G., Moen, E., Kong, A., Kagel, A., Dougherty, T., Fullaway, C.C., McIntosh, B.J., Leow, K.X., Schwartz, M.S., et al.: Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nature biotechnology **40**(4), 555–565 (2022)
11. Halabi, S.S., Prevedello, L.M., Kalpathy-Cramer, J., Mamonov, A.B., Bilbily, A., Cicero, M., Pan, I., Pereira, L.A., Sousa, R.T., Abdala, N., et al.: The rsna pediatric bone age machine learning challenge. Radiology **290**(2), 498–503 (2019)
12. Hou, X., Liu, M., Zhang, S., Wei, P., Chen, B.: Salience detr: Enhancing detection transformer with hierarchical salience filtering refinement. In: CVPR. pp. 17574–17583 (2024)
13. Huumonen, S., Ørstavik, D.: Radiological aspects of apical periodontitis. Endodontic topics **1**(1), 3–25 (2002)
14. Karunakaran, J., Abraham, C.S., Karthik, A.K., Jayaprakash, N.: Successful nonsurgical management of periapical lesions of endodontic origin: A conservative orthograde approach. Journal of pharmacy & bioallied sciences **9**(Suppl 1), S246 (2017)
15. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S.: Feature pyramid networks for object detection. In: CVPR. pp. 2117–2125 (2017)
16. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: ECCV. pp. 740–755. Springer (2014)
17. Liu, S., Li, F., Zhang, H., Yang, X., Qi, X., Su, H., Zhu, J., Zhang, L.: Dab-detr: Dynamic anchor boxes are better queries for detr. In: ICLR (2021)
18. Lu, L., Dercle, L., Zhao, B., Schwartz, L.H.: Deep learning for the prediction of early on-treatment response in metastatic colorectal cancer from serial medical imaging. Nature communications **12**(1), 6654 (2021)
19. Meng, D., Chen, X., Fan, Z., Zeng, G., Li, H., Yuan, Y., Sun, L., Wang, J.: Conditional detr for fast training convergence. In: ICCV. pp. 3651–3660 (2021)
20. Rajpurkar, P., Irvin, J., Bagul, A., Ding, D., Duan, T., Mehta, H., Yang, B., Zhu, K., Laird, D., Ball, R.L., et al.: Mura: Large dataset for abnormality detection in musculoskeletal radiographs. arXiv preprint arXiv:1712.06957 (2017)
21. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. In: NeurIPS. vol. 28 (2015)
22. Sun, P., Zhang, R., Jiang, Y., Kong, T., Xu, C., Zhan, W., Tomizuka, M., Li, L., Yuan, Z., Wang, C., et al.: Sparse r-cnn: End-to-end object detection with learnable proposals. In: CVPR. pp. 14454–14463 (2021)
23. Tibúrcio-Machado, C., Michelon, C., Zanatta, F., Gomes, M.S., Marin, J.A., Bier, C.A.: The global prevalence of apical periodontitis: a systematic review and meta-analysis. International endodontic journal **54**(5), 712–735 (2021)
24. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: CVPR. pp. 2097–2106 (2017)
25. Xu, C., Wang, J., Yang, W., Yu, H., Yu, L., Xia, G.S.: Rfla: Gaussian receptive field based label assignment for tiny object detection. In: ECCV. pp. 526–543. Springer (2022)
26. Yuan, X., Cheng, G., Yan, K., Zeng, Q., Han, J.: Small object detection via coarse-to-fine proposal generation and imitation learning. In: ICCV. pp. 6317–6327 (2023)

27. Zhang, H., Li, F., Liu, S., Zhang, L., Su, H., Zhu, J., Ni, L.M., Shum, H.Y.: Dino: Detr with improved denoising anchor boxes for end-to-end object detection. arXiv preprint arXiv:2203.03605 (2022)
28. Zhang, S., Chi, C., Yao, Y., Lei, Z., Li, S.Z.: Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection. In: CVPR. pp. 9759–9768 (2020)
29. Zhu, X., Su, W., Lu, L., Li, B., Wang, X., Dai, J.: Deformable detr: Deformable transformers for end-to-end object detection. In: ICLR (2020)
30. Zong, Z., Song, G., Liu, Y.: Detrs with collaborative hybrid assignments training. In: ICCV. pp. 6748–6758 (2023)