Supplementary Materials
OralXrays-9: Towards Hospital-Scale Panoramic X-ray Anomaly Detection via Personalized Multi-Object Query-Aware Mining

Bingzhi Chen[1], Sisi Fu[1], Xiaocheng Fang[1], Jieyi Cai[1], Boya Zhang[2,3], Minhua Lu[4] ✉, Yishu Liu[5] ✉
[1]Beijing Institute of Technology, Zhuhai [2]Chinese People's Liberation Army General Hospital
[3]Nankai University [4]Shenzhen University [5]Harbin Institute of Technology, Shenzhen
chenbingzhi@bit.edu.cn, {sisifu5929,fangxiaocheng162,caijieyi055,zboya85}@gmail.com,
luminhua@szu.edu.cn, liuyishu.smile@gmail.com*

1. Details of OralXrays-9 Dataset

X-ray Image Characteristics Panoramic dental imaging utilizes X-ray technology to provide a comprehensive view of the teeth and surrounding oral tissues, facilitating thorough assessments of oral health and the diagnosis of various conditions. Variations in panoramic radiographs arise from differences in natural density, tissue thickness, and pathological changes within the buccal cavity, which affect the transmission of uniform-intensity X-rays [1]. These variations can signal underlying dental issues, such as cavities, periodontal disease, and other oral pathologies. By analyzing these density changes, dental professionals can detect and evaluate potential conditions, thereby informing the development of effective treatment plans. More examples of nine common oral anomalies are shown in Figure 1.

Anomaly Selection The nine selected anomalies in OralXrays-9 represent the most clinically significant and commonly encountered dental conditions in modern dental practice. These conditions were chosen based on their high prevalence, impact on oral health, and their relevance in diagnostic radiology.

Clinical Setting Our dataset was collected from dental departments of multiple hospitals, including dental clinics, and orthodontic centers. These institutions are equipped with advanced panoramic X-ray imaging systems, ensuring high imaging quality and reliable anomaly identification.

Image Processing To optimize computational efficiency and ensure compatibility, a preprocessing pipeline was developed for the OralXrays-9 dataset, following established methods [2]. X-ray images were extracted directly from DICOM files, the standard format for dental radiography, to preserve essential image details. All images were resized to a standard resolution of 1333×800 pixels, ensuring consistency while minimizing information loss. This resizing step facilitated compatibility with deep learning models, preserving critical features necessary for accurate anomaly detection. The final output was formatted as bitmap files, ready for subsequent analysis.

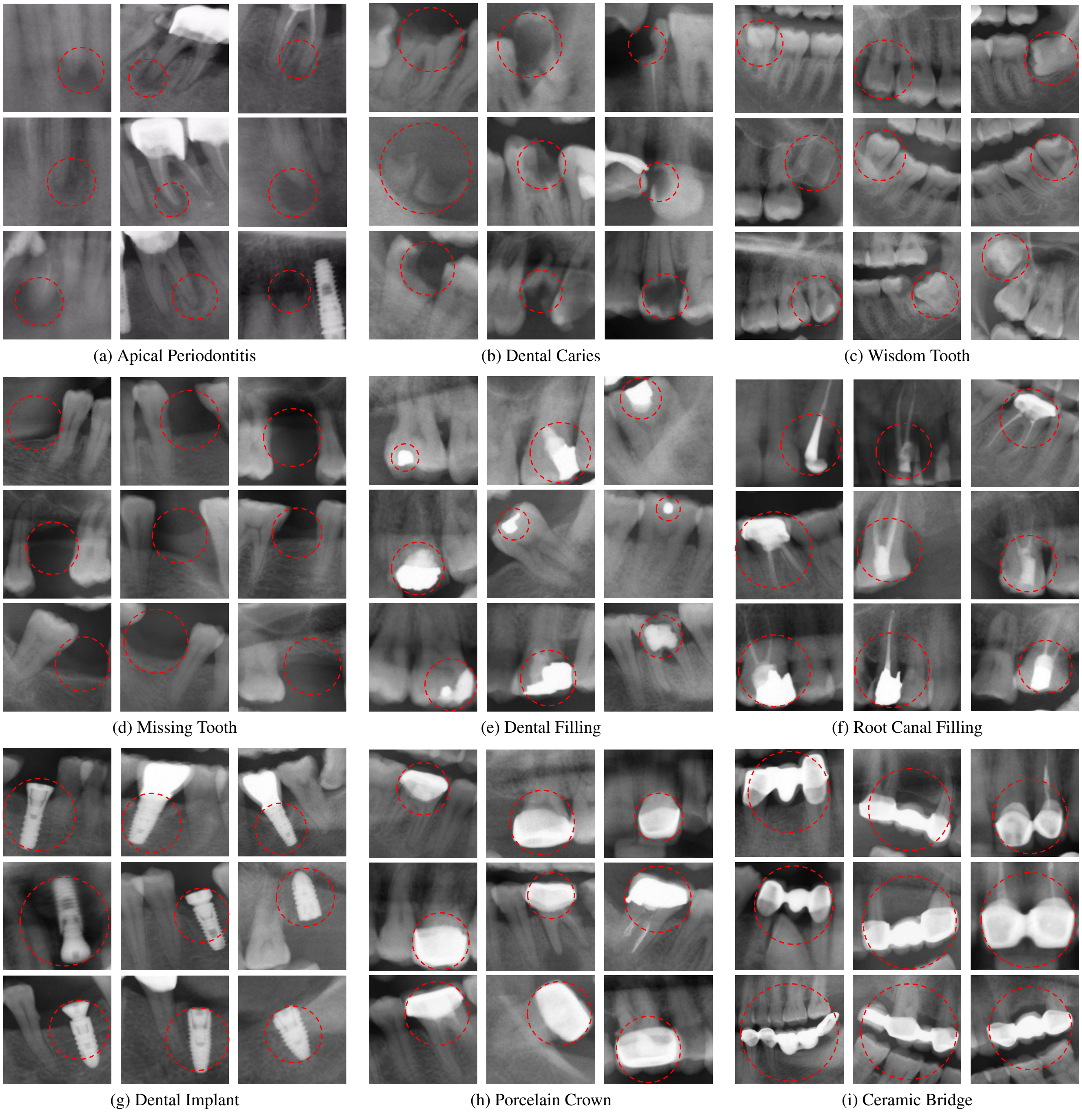
Privacy Protection To protect patient confidentiality, the panoramic X-ray dataset has been meticulously anonymized by removing all personal identifiers, such as names and medical record numbers. Each image is assigned a unique, non-identifiable label to preserve data integrity while strictly complying with privacy regulations and ethical research guidelines, ensuring the dataset is suitable for research purposes without compromising patient privacy.

Dataset Splitting For experimental purposes, our OralXrays-9 dataset is randomly divided into two subsets: a training set comprising 10,000 images and a test set of 2,688 images. The dataset is split at the patient level to ensure no overlap between the training and test sets. Specifically, all images from the same patient are assigned to a single set. The training set contains 65,494 ground truth annotations detailing the presence and locations of oral anomalies, while the test set includes 18,639 annotations for evaluating model performance on unseen data.

2. Algorithm of MOQAM

The training process of MOQAM is outlined in Algorithm 1, where each step is meticulously detailed to maximize clarity and reproducibility.

*Corresponding authors: Yishu Liu and Minhua Lu.



(a) Apical Periodontitis (b) Dental Caries (c) Wisdom Tooth

(d) Missing Tooth (e) Dental Filling (f) Root Canal Filling

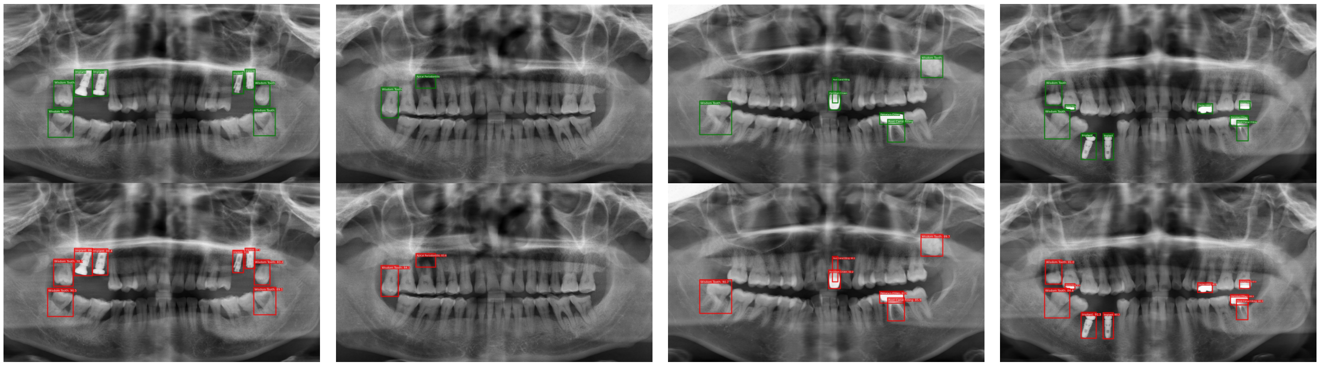
(g) Dental Implant (h) Porcelain Crown (i) Ceramic Bridge

Figure 1. More representative examples of nine common oral anomalies in Panoramic Dental X-rays.

3. Clinical Interpretation and Evaluation

We have designed a human-computer collaborative diagnosis experiment to evaluate the clinical applicability of our framework as an auxiliary diagnostic tool. A total of 8 professional dentists participated in this experiment, with their detailed professional backgrounds provided in Table 1.

We randomly selected 100 panoramic X-ray images from the OralXrays-9 dataset as the test set. The participants were asked to make two consecutive diagnostic judgments for each image: As presented in Table 2, the recall rate increased significantly from 68.1% to 95.5%, while the average image processing time decreased from 60 seconds per image to 26 seconds per image. These results underscore the clinical applicability of MOQAM, demonstrating its effectiveness in real-world medical diagnostics.



(a) (b) (c) (d)

Figure 2. llustration of bounding box predictions generated by the proposed MOQAM framework.

Algorithm 1 The Proposed MOQAM Algorithm

Input: Panoramic X-ray dataset $\mathcal{D}$, Transformer-based model parameters Θ, Base IoU threshold t_b, Growth exponent p, Weight of the CB-SCR Loss α;
Initialize: DI-RPN and CB-SCR modules;
Output: Bounding boxes and confidence scores for oral anomaly detection.

for $t = 0$ to *max-step* **do**
    Extract features via ResNet-50;
    Apply Distribution-Aware Mechanism;
    Compute Gaussian ERF;
    Rank samples by KL divergence and select top-k positive samples;
    Apply IoU-Aware Mechanism;
    Employ one-many label assignment strategy to increase positive samples;
    Compute DI-RPN loss $\mathcal{L}_{\text{DI-RPN}}$;
    Map feature representations into hyperspherical space;
    Apply class-balanced regularization;
    Compute CB-SCR loss $\mathcal{L}_{\text{CB-SCR}}$;
    Compute total loss $\mathcal{L}_{\text{MOQAM}} = \mathcal{L}_{\text{DI-RPN}} + \alpha \cdot \mathcal{L}_{\text{CB-SCR}}$;
end for

Total	Professional Title			Hospital Level		Employment Years			
	Junior	Mediate	Senior	City	National	1-5	6-10	11-15	16-20
8	2	3	3	3	5	1	2	3	2

Table 1. Basic information of the participant group.

Methods	Dentists w/o MOQAM	Dentists w/ MOQAM
Recall	68.1%	95.5% (27.4% ↑)
Efficiency	≈ 60 Secs/Img	≈ 26 Secs/Img (34 Secs ↓)

Table 2. Comparison of the diagnostic performance.

4. Discussion on Limitations

More visualization results are shown in Figure 2. Due to the inherent limitations of imaging principles, panoramic dental X-ray images exhibit spatial two-dimensional overlap, which introduces noise interference in the detection of oral anomalies. As illustrated in Figure 3, while our MOQAM framework effectively addresses many challenges, it struggles to completely eliminate noise shadow effects, potentially leading to missed detections and false positives in bounding box predictions. This limitation highlights the need for further refinement in handling overlapping structures in dental radiographs.

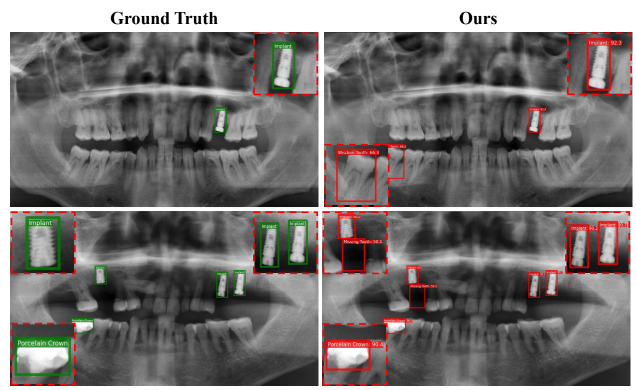


Figure 3. The ground truth of oral anomalies and the inaccurate bounding box prediction results from the MOQAM framework.

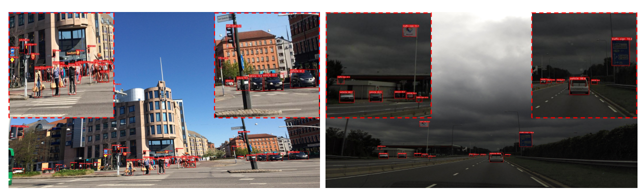


Figure 4. Visualization of detection results on natural datasets.

5. Visualization on Natural Datasets

The qualitative visualizations are presented in Figure 4.

References

[1] Anton Sklavos, Daniel Beteramia, Seth Navinda Delpachitra, and Ricky Kumar. The panoramic dental radiograph for emergency physicians. *Emergency Medicine Journal*, 36(9):565–571, 2019. 1

[2] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *CVPR*, pages 2097–2106, 2017. 1