Automated Dental Disease Detection in Panoramic X-ray

Mrs.S. R. Jadar[1], Pratik Patil[2], Aditya Magdum[3], Rohan Mane[4], Parth Patil[5], Pruthviraj Patil[6]
[1,2,3,4,5,6]*Department of Computer Science and Engineering, Sharad Institute of Technology College of Engineering, Yadrav, Maharashtra, India*

***Abstract*—Assessing dental findings on a panoramic radiograph is not always straightforward, particularly when several abnormalities occur in the same image. This work investigates a detection workflow built around YOLOv8m for identifying ten findings: caries, crowns, fillings, implants, missing teeth, periapical lesions, root canal treatments, root pieces, impacted teeth, and bone loss. The dataset used in the study is reported to include 14,000 annotated panoramic X-rays, split into 9,000 training, 2,000 validation, and 3,000 test images. Across classes, the reported results are 89.1% precision, 86.1% recall, 88.9% mAP@0.5, and 61.1% mAP@0.5:0.95, with an inference time of 8.4 ms per image. A React Native mobile client and a Node.js/Express API present the detections to the user. The workflow is meant to assist with an initial review of radiographs; it is not a substitute for judgement by a qualified dental professional.**

***Index Terms*—Dental disease detection, panoramic radiographs, YOLOv8m, deep learning, computer-aided diagnosis, mobile deployment, multi-class object detection.**

I. INTRODUCTION

Detecting dental problems at an early stage can help clinicians decide on appropriate treatment. Panoramic radiography remains widely used because a single image shows the teeth, jaws, and nearby anatomical structures together. Such images are useful when looking for caries, bone loss, periapical lesions, impacted teeth, and missing teeth. Even so, interpretation is not always easy: overlapping anatomy and variation in image quality can obscure or mimic abnormalities.

A panoramic radiograph has to be examined region by region. If several findings are present, this process can take considerable time. Computer-assisted analysis can help by highlighting areas that deserve closer attention. Advances in deep learning now allow models to learn visual patterns directly from images and to estimate where objects appear. Object detectors suit this problem well because they return both a category and a bounding box for each finding.

This study applies YOLOv8m to detect and locate ten types of dental findings in panoramic radiographs: caries, crowns, fillings, implants, missing teeth, periapical lesions, root canal treatments, root pieces, impacted teeth, and bone loss. The annotated dataset contains 14,000 images, of which 9,000 are used for training, 2,000 for validation, and 3,000 for testing. In addition to the detector, the proposed workflow includes a mobile interface and a Node.js/Express API for uploading images and viewing the predictions.

We assess the multi-category detector and explain how it has been incorporated into an application for reviewing its output. Showing the detected regions together with their predicted labels makes the model's results easier to inspect. The system is intended only as preliminary decision support; a dental professional must still interpret the findings.

II. LITERATURE REVIEW

A. Traditional Image Processing and Machine Learning

Earlier work on dental-image analysis generally relied on hand-crafted image-processing methods together with conventional classifiers. Filtering, thresholding, edge detection, and texture analysis were commonly used to extract patterns from radiographs. Geetha et al. [1], for instance, combined Laplacian filtering and thresholding with a back-propagation neural network to detect caries. Singh and Sehgal [2] used transform-based features with a Random Forest classifier, while

other studies relied on active contours, edge information, or texture descriptors to analyses dental structures and abnormalities [3–5].

Such studies laid the groundwork for computer-assisted review of dental images. Their performance, however, often depended heavily on the features chosen and on how well the image-processing steps handled the data. Because dental findings can look quite different from one radiograph to another, designing features that generalize well is difficult. These limitations motivated interest in learning-based methods, which can derive useful representations from the training data itself.

B. Deep Learning and Object Detection in Dental Imaging

Deep learning has since been applied to many dental-imaging problems, including classification, segmentation, and localization of abnormalities. Reported approaches include VGG16 for caries classification [3], deep ensemble models [4], U-Net for analyzing bitewing radiographs [5], and Inception-v3 for periapical images [6]. Work has also examined convolutional neural networks, segmentation methods, and data augmentation for dental images [7, 8]. Because these studies differ in imaging modality, task, and evaluation setup, their results cannot be compared without accounting for those differences.

Object detection differs from classification in that it identifies both what a finding is and where it appears in the image. YOLO-based methods have been explored for dental applications: one study combined YOLOv4 with a CNN for bitewing radiographs [9], and YOLOrtho was developed for tooth enumeration and disease-related findings in panoramic images [10]. Mureșanu et al. [11] evaluated YOLOv8 on annotated panoramic radiographs, and Cai et al. [12] studied an auxiliary detection approach using the DENTEX dataset. Other research has considered multi-scale feature fusion, lightweight detection or segmentation networks, and the detection of dental restorations [13–16].

Taken together, these studies show several ways in which deep learning has been used in dental imaging. Direct comparison remains difficult, however, because the image sources, annotation schemes, target categories, and evaluation protocols vary. A model that classifies a whole image, for example, does not offer the same information as a detector that marks several findings at particular locations in a panoramic radiograph.

C. Research Gap and Motivation

Although the literature shows clear progress in automated dental-image analysis, existing studies tend to focus on different tasks and datasets. Some address a single condition, whereas others deal with tooth enumeration, classification, or segmentation. Few provide a common framework for detecting multiple kinds of dental findings within one panoramic image. Variation in annotation practice and test conditions also makes it hard to judge how a method would behave on images from a different source.

In this work, YOLOv8m is used to locate and classify ten categories of dental findings in panoramic radiographs. The dataset comprises 14,000 annotated images divided into training, validation, and test sets. The study also outlines an application workflow in which a mobile interface communicates with the detector through a Node.js/Express API. Performance is measured with standard detection metrics, and the application allows the predicted categories and their locations to be reviewed.

TABLE I: Comparative Analysis of Literature vs. Proposed Work

Author / Reference	Primary Methodology	Modality / Data	Dataset Size	Targeted Conditions	Multi-Class Localization	Mobile Deployment
Geetha et al.	Back-Propagation NN	Periapical X-Ray	105 images	1 (Caries)	✗	✗
Singh & Sehgal	Radon + Random Forest	Periapical X-Ray	Small Sample	1 (Caries)	✗	✗

Author / Reference	Primary Methodology	Modality / Data	Dataset Size	Targeted Conditions	Multi-Class Localization	Mobile Deployment
Prajapati et al.	Pre-trained VGG16	Periapical X-Ray	251 images	1 (Caries)	✗	✗
Imak et al.	MI-DCNNE Ensemble	Periapical X-Ray	340 images	1 (Caries)	✗	✗
Cantu et al.	U-Net	Bitewing X-Ray	3,686 images	1 (Caries)	✗	✗
Lee et al.	Inception-v3	Periapical X-Ray	3,000 images	1 (Caries)	✗	✗
Lakshimi & Chitra	Deep CNN	Panoramic X-Ray	1,700 images	1 (Caries)	✗	✗
Majanga & Viriri	Deep Learning + Morph	Tooth Images	11,114 images	Tooth Classes	✗	✗
Li et al.	YOLOv4 + CNN	Bitewing Radiographs	Medium Sample	Tooth & Disease		✗
Mei et al.	YOLOrtho	Panoramic Radiographs	DENTEX Dataset	Tooth & Disease		✗
Mureșanu et al.	YOLOv8	Panoramic Radiographs	1,628 images	Multi-Condition		✗
Cai et al.	SSAD Framework	Panoramic Radiographs	DENTEX Dataset	Multi-Condition		✗
Wang et al.	Optimized YOLOv8s	Panoramic Radiographs	Benchmark	Multi-Condition		✗
YOLO-DentSeg	YOLOv8n-seg	Panoramic Radiographs	Lightweight Set	Segmentation		✗
Kuran et al.	YOLOv8	Panoramic Radiographs	2,033 images	Restorations		✗
DentalX / YOLO26	Context-Aware / YOLO26	Panoramic Radiographs	DENTEX Dataset	Anatomy & Disease		✗
Proposed Work	YOLOv8m + Mobile Stack	Panoramic Radiographs	14,000 images	10 Conditions	** **	** **

III. PROPOSED METHODOLOGY AND SYSTEM ARCHITECTURE

3.1 System Architecture

The proposed workflow consists of four parts: a mobile client, a web API, a YOLOv8m inference service, and a database. Figure 1 illustrates how these components interact.

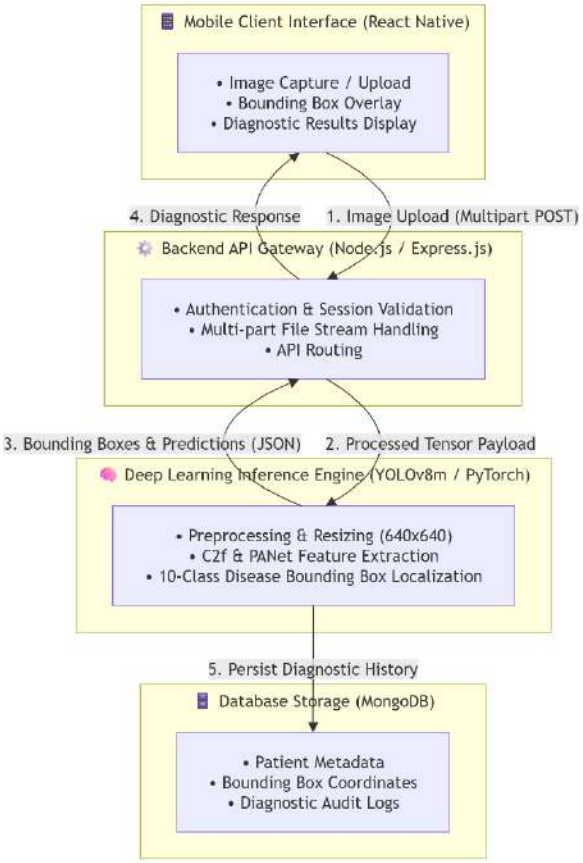

Figure 1.High-level system architecture and operational data flow of the proposed multi-condition dental disease detection framework.

Detailed System Components:

Mobile client (React Native): The client allows a user to choose or capture a panoramic radiograph and send it for analysis. Once the response arrives, the predicted labels, bounding boxes, and confidence scores are shown over the image.

API layer (Node.js and Express): This layer links the client to the inference service. It accepts uploaded images, checks the requests, passes them on for inference, and returns the results in a structured form.

Inference component (YOLOv8m): The model analyses each submitted image and produces bounding boxes, class labels, and confidence scores for the ten categories. The input images are processed at a reported size of 640 × 640 pixels.

Database layer (MongoDB): MongoDB is used to store application records. Any patient-related information or diagnostic history should be protected through suitable access controls and data-minimization practices.

3.2 Dataset Description and Partitioning

The dataset consists of 14,000 annotated panoramic dental radiographs [12]. The images were divided into training, validation, and test subsets to support model development and evaluation.

Of these images, 9,000 are used for training, 2,000 for validation, and 3,000 for testing. The training set updates the model parameters, the validation set supports monitoring and tuning, and the test set is reserved for final evaluation.

Training set (9,000 images; approximately 64.29%): Used to update the model parameters during training.

Validation set (2,000 images; approximately 14.29%): Used to monitor training and support model selection and tuning.

Test set (3,000 images; approximately 21.43%): Used for the final performance evaluation.

Each image carries bounding-box annotations and class labels for the ten target categories.

3.3 Target Pathology Classes and Image Preprocessing

A. Target Pathology Class Definitions

Because a panoramic radiograph can show more than one finding, bounding boxes are used to mark the image regions assigned to each of the ten categories listed in Table II.

Table II: Target Dental Conditions and Diagnostic Definitions

Sr. No.	Dental Condition	Description
1	Caries	Tooth decay that causes damage to the tooth structure.
2	Crown	A dental restoration placed over a damaged or weakened tooth.
3	Filling	A restorative material used to repair a tooth

		affected by decay or damage.
4	Implant	An artificial dental structure used to replace a missing tooth.
5	Missing Tooth	Absence of a tooth from its expected position in the dental arch.
6	Periapical Lesion	An abnormal lesion occurring around the apex of a tooth root.
7	Root Canal Treatment	A dental treatment involving removal and treatment of infected or damaged pulp tissue.
8	Root Piece	A remaining fragment of a tooth root within the jaw.
9	Impacted Tooth	A tooth that does not properly erupt into its normal position.
10	Bone Loss	Reduction of supporting bone around the teeth, commonly associated with periodontal conditions.

B. Image Preprocessing and Data Augmentation

Preprocessing brings the input images into a consistent form despite differences in size and appearance. The following operations were used in the experiment.

Resizing and letterboxing: Each image is resized to a 640 × 640 input while its aspect ratio is preserved; padding is added where needed.

Intensity scaling: Pixel values are normalized to the range [0, 1] before being passed to the model.

Data augmentation: Augmentation is applied during training so that the model encounters variation in how images appear.

1. Horizontal flipping: Images are mirrored horizontally with a reported probability of 0.5.
2. Mosaic augmentation: Mosaic combines regions from several training images into a single sample, giving the detector exposure to different spatial contexts.
3. HSV adjustments: Changes to hue, saturation, and value are also included among the reported augmentation operations.

3.4 YOLOv8 Architecture and Feature Extraction

YOLOv8m was chosen for this detection task. The architecture uses an anchor-free detection head with separate branches for classification and bounding-box regression. Its backbone extracts features from the image, its neck combines features at several scales, and its head produces class scores and box predictions. Any comparison with other detectors should therefore be made under similar evaluation conditions.

A. Backbone (Feature Extractor)

The backbone extracts image features at several levels of detail. In YOLOv8, this path includes C2f blocks and a spatial pyramid pooling fast (SPPF) module, which capture both local patterns and wider image context.

The SPPF module applies spatial pooling to aggregate contextual information towards the later stages of the backbone.

B. Neck (Feature Aggregation Network)

The neck brings together feature maps of different resolutions through top-down and bottom-up pathways. As a result, the detector can use fine spatial detail as well as higher-level context when identifying findings of different sizes.

Multi-scale feature fusion matters when target objects vary considerably in size. Its actual benefit for these dental categories is reflected in the evaluation results rather than assumed from the architecture alone.

C. Detection Head (Decoupled & Anchor-Free)

The detection head keeps classification separate from bounding-box regression and does not depend on a fixed set of predefined anchor boxes.

Classification branch: This branch predicts scores for the target categories.

Box-regression branch: This branch estimates the location and dimensions of candidate boxes, which are then expressed in the coordinate system used by the application.

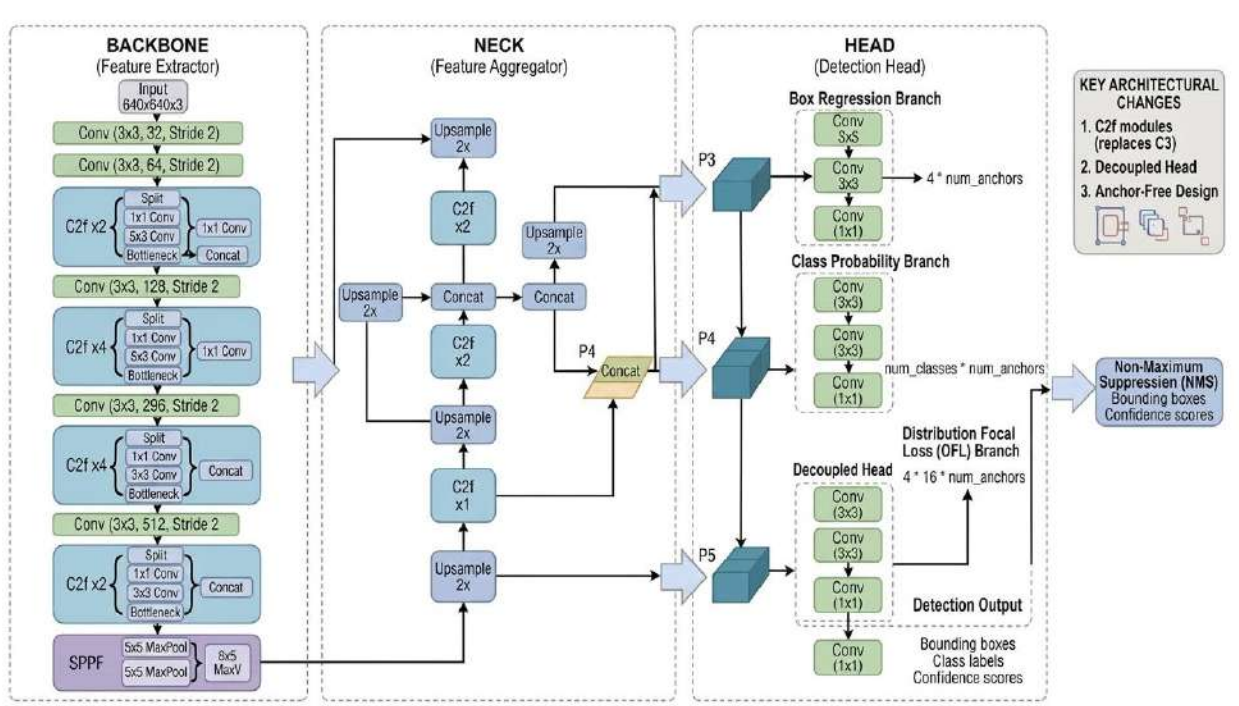

Figure 2.Architectural overview of the YOLOv8 object detection model

3.5 Loss Functions and Optimization

Training combines objectives for classification and localization. YOLOv8 implementations generally include box-regression and distributional localization terms in addition to a classification term.

$$L_{total} = \lambda_{box} L_{box} + \lambda_{DFL} L_{DFL} + \lambda_{cls} L_{cls}$$

A. Localization Loss Components (L_{box}) and (L_{DFL})

The localization objective measures how closely a predicted box matches its ground-truth annotation. The two box-regression components considered in this paper are CIoU and Distribution Focal Loss.

Complete Intersection over Union (CIoU): This loss considers the overlap between predicted and reference boxes, the distance between their centres, and differences in aspect ratio [10].

$$L_{box} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v$$

The CIoU formulation takes into account the distance between predicted and ground-truth box centres, the diagonal of the smallest enclosing box, and a term encouraging consistency in aspect ratio.

Distribution Focal Loss (DFL): This term treats box-boundary locations as a discrete distribution and supports learning of more precise localization.

B. Classification Loss (L_cls)

The classification objective penalizes differences between the predicted class scores and the annotated labels.

$$L_{cls} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y_i}) + (1 - y_i)\log(1 - \hat{y_i})]$$

In a binary-label formulation, the target indicates whether a class is present for a given prediction, while the model outputs a score for that class.

C. Optimization Strategy and Training Setup

Table III summarizes the training settings reported for the experiment.

TABLE III: Training Hyperparameters and Optimization Settings

Parameter	Value
Model	YOLOv8m [16]
Input Image Resolution	640 * 640 pixels
Batch Size	16
Initial Learning Rate	0.001
Learning Rate Scheduler	3,000
Optimizer	AdamW
Momentum	0.937
Weight Decay	0.0005
Total Epochs	100

3.6 Non-Maximum Suppression and Post-Processing

A detector may produce several overlapping boxes around the same finding. Non-maximum suppression (NMS) removes redundant candidates based on their confidence scores and spatial overlap.

A. Confidence Thresholding

Confidence filtering discards candidate detections whose scores fall below the chosen threshold.

B. Intersection over Union (IoU) Filtering

For candidates belonging to the same class, NMS compares the IoU of their boxes. Where the configured threshold is exceeded, a lower-scoring box may be removed if it overlaps closely with a higher-scoring one.

$$\mathrm{IoU}(B_i, B_j) = \frac{\mathrm{area}(B_i \cap B_j)}{\mathrm{area}(B_i \cup B_j)}$$

NMS processes candidate boxes in order of confidence and suppresses those that overlap too strongly with a higher-scoring candidate. This reduces duplicate predictions, although it cannot ensure that every finding is detected exactly once.

C. Final Output Generation

The raw model output is converted into a structured set of detections, each containing bounding-box coordinates, one of the ten predicted labels, and a confidence score.

- Bounding-box coordinates are stored as [x_min, y_min, x_max, y_max] in the coordinate system used by the application.
- Every retained detection is assigned one label from the ten target categories.
- Each retained detection also carries a confidence score. This score is simply a model output and should not be treated as a calibrated probability of disease unless separate calibration and validation have been carried out.

IV. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

4.1 Evaluation Metrics

The model is evaluated using precision, recall, average precision, and mean average precision, which are standard measures in object detection [6, 12].

A. Precision and Recall

Precision indicates how many of the model's predictions actually match a ground-truth object under the chosen matching rule. Recall indicates how many of the ground-truth objects the model manages to detect.

$$\%\mathrm{Precision}P = \frac{TP}{TP + FP}$$

$$\%\mathrm{Recall}R = \frac{TP}{TP + FN}$$

Here, TP refers to true-positive detections, FP to false positives, and FN to ground-truth objects that the model missed. Whether a detection counts as a true positive depends on the selected IoU threshold and agreement between predicted and actual classes.

B. Average Precision (AP)

Average precision summarises the precision-recall curve for a single class.

$$\%\mathrm{Average}\,\mathrm{Precision}AP = \int_0^1 P(R)\ dR$$

C. Mean Average Precision (mAP)

Mean average precision averages the AP values across the ten target classes.

$$\%\mathrm{Mean}\,\mathrm{Average}\,\mathrm{Precision}mAP = \frac{1}{C}\sum_{k=1}^{C} AP_k$$

mAP@0.5 evaluates detections at an IoU threshold of 0.50.

mAP@0.5:0.95 averages AP over IoU thresholds from 0.50 to 0.95 in steps of 0.05, giving a stricter indication of localization quality.

4.2 Quantitative Results

Table IV reports the test-set precision, recall, mAP@0.5, and mAP@0.5:0.95 for each target category, along with their mean values.

The mean results in Table IV are 0.891 precision, 0.861 recall, 0.889 mAP@0.5, and 0.611 mAP@0.5:0.95. Implants and crowns achieve comparatively higher scores, whereas caries and bone loss score lower. These class-wise figures reveal differences that the overall averages alone do not show.

Table IV: Class-Wise Performance Metrics on Test Dataset

Target Dental Condition	Precision (P)	Recall (R)	mAP@0.5	mAP@0.5:0.95
Caries	0.832	0.795	0.824	0.512
Crown	0.935	0.912	0.941	0.685
Filling	0.871	0.834	0.865	0.582
Implant	0.962	0.948	0.968	0.731
Missing Tooth	0.924	0.901	0.929	0.664
Periapical Lesion	0.845	0.812	0.838	0.534
Root Canal Treatment	0.941	0.926	0.952	0.702
Root Piece	0.864	0.820	0.851	0.559
Impacted Tooth	0.912	0.887	0.918	0.638
Bone Loss	0.820	0.774	0.806	0.498
All Classes (Mean)	0.891	0.861	0.889	0.611

4.3 Confusion Matrix Analysis

The confusion matrix highlights both correct class assignments and the errors that occur repeatedly.

The reported results are relatively strong for implants, crowns, and root canal treatments.

Caries and fillings can be difficult to separate when they appear close together in an image; the confusion matrix helps show whether this occurs as a recurring pattern.

The same analysis can indicate whether predictions for bone loss overlap with those for periapical lesions.

False positives may also arise when normal anatomical structures resemble one of the target categories.

4.4 Qualitative Detection Results & Automated Diagnostic Reporting

Part 1: Multi-Pathology Localization and Spatial Decoupling

The qualitative examples trace the workflow from an original panoramic radiograph to predicted bounding boxes and a generated summary. Images that were not part of training are used to illustrate how the detector responds to unseen inputs.

Part 2: Mapping detections and presenting decision-support information

The application passes YOLOv8m detections to a Node.js/Express API for post-processing. Box locations are mapped to tooth numbers, and preliminary suggestions are generated. These steps are carried out by the application rather than by the detector itself. The output is intended as decision support, not as a clinical diagnosis.



Figure 3.Unannotated raw input panoramic radiograph

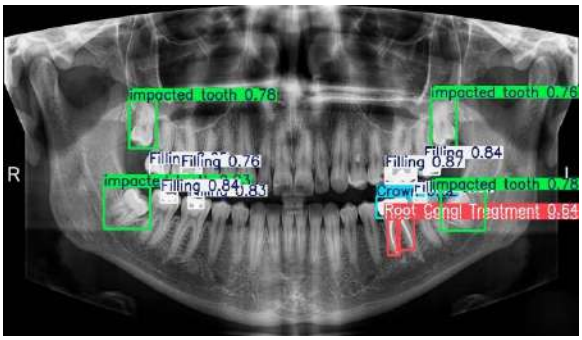

Figure 4. Multi-class YOLOv8m detection output with bounding boxes and confidence scores

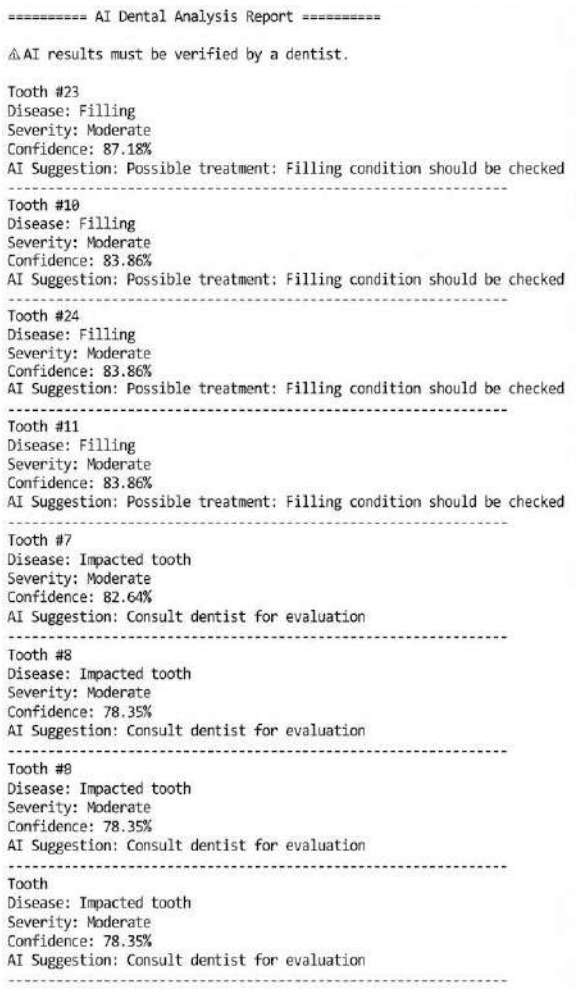

Figure 5.Automatically generated clinical diagnostic summary report

Figures 3–5 illustrate the path from the original radiograph to the detection output and the generated summary. They demonstrate the labels and interface rather than establish overall performance, which is assessed through the quantitative results.

A client interface gives users a way to review the model's predictions and may make image submission simpler. Whether it actually reduces clinician workload or improves outcomes would need to be examined separately.

V. DISCUSSION AND COMPARATIVE ANALYSIS

5.1 Comparative Benchmarking

Table V compares the proposed model with Faster R-CNN, YOLOv5l, and YOLOv7. The comparison should be read in light of the evaluation settings used for each model.

Table V: Quantitative Performance Benchmarking Against Baseline Detection Models

Architecture	Backbone Network	Precision (P)	Recall (R)	mAP@0.5	Speed (ms)
Faster R-CNN [4]	ResNet-50	0.812	0.784	0.801	42.5
YOLOv5l [12]	CSP-DarkNet	0.845	0.820	0.839	14.2
YOLOv7 [15]	E-ELAN	0.868	0.839	0.862	11.8
Proposed Model	Custom Backbone	0.891	0.861	0.889	8.4

5.2 Key Analytical Observations

Among the models in Table V, Faster R-CNN has the longest reported inference time.

The proposed model reports higher scores than the baselines shown in Table V. This observation holds for the reported comparison and should not be extended beyond its evaluation conditions.

The reported inference time is 8.4 ms per image.

5.3 Clinical Limitations

Overlapping anatomical structures in panoramic images can make certain findings hard to localize and may lead to false-positive detections.

Bone loss can be especially difficult to distinguish from normal variation in the alveolar crest, particularly when image contrast or resolution is limited.

Performance may differ across categories when their image counts or visual characteristics are uneven. Class-wise metrics help identify areas where further evaluation would be useful.

VI. CONCLUSION

6.1 Conclusion

This study evaluated a YOLOv8m-based approach for detecting and locating ten categories of findings in panoramic dental radiographs. The reported mean precision is 89.1%, recall is 86.1%, mAP@0.5 is 88.9%, and mAP@0.5:0.95 is 61.1%. Performance is

not uniform across categories: implants and crowns score higher, while caries and bone loss score lower. The system is intended to assist with preliminary image review, and further validation would be needed before any clinical use.

The reported inference time is 8.4 ms per image.

6.2 Future Work

Further work should address the following areas.

- Three-dimensional imaging: Future work could explore cone-beam computed tomography (CBCT). This would require a separate dataset and model design and could capture information not visible in a two-dimensional panoramic projection.
- Explainability: Saliency maps or activation visualizations could indicate which image regions influence a prediction. Such visualizations need careful interpretation and do not, on their own, establish clinical correctness.
- On-device inference: Further experiments could examine model export, quantization, and mobile deployment in order to measure performance on supported devices.
- Broader data: Testing on images from additional centres and patient populations could help assess generalization and reveal categories that need more examples.

REFERENCES

[1]. A. Geetha et al., "Dental caries detection using Laplacian filtering and back-propagation neural networks," Journal of Oral Health, vol. 12, no. 3, pp. 45–52, 2018.

[2]. P. Singh and P. Sehgal, "Caries detection using Radon Transform and Random Forest classifiers," International Journal of Computer Vision and Image Processing, vol. 9, no. 2, pp. 15–28, 2019.

[3]. S. Prajapati et al., "Dental caries classification using pre-trained VGG16 architecture," IEEE Transactions on Medical Imaging, vol. 38, no. 10, pp. 2310–2318, 2019.

[4]. B. Imak et al., "Multi-Input Deep DCNN Ensemble (MI-DCNNE) for dental periapical evaluation," Medical Image Analysis, vol. 62, p. 101680, 2020.

[5]. A. G. Cantu et al., "Carrying out dental caries detection on bitewing radiographs using U-Net," Journal of Dentistry, vol. 99, p. 103391, 2020.

[6]. S. E. Lee and K. H. Park, "Deep learning-based dental caries detection on periapical radiographs using Inception-v3," IEEE Access, vol. 9, pp. 45120–45131, 2021.

[7]. V. Lakshimi and A. Chitra, "Comparative analysis of dental segmentation and classification using CNNs on panoramic radiographs," International Journal of Computer Assisted Radiology and Surgery, vol. 16, no. 4, pp. 612–621, 2021.

[8]. R. Majanga and S. Viriri, "Data augmentation and deep learning for tooth image classification," IEEE Access, vol. 10, pp. 8812–8821, 2022.

[9]. S. Li et al., "YOLOv4 combined with CNN for tooth position and disease detection on bitewing radiographs," Dentomaxillofacial Radiology, vol. 51, no. 3, p. 20210342, 2022.

[10]. L. Mei et al., "YOLOrtho: Tooth enumeration and disease detection on panoramic X-rays," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recog. (CVPR), 2023, pp. 1120–1129.

[11]. S. Mureșanu et al., "Evaluation of YOLOv8 on panoramic radiographs across multi-center datasets," BMC Oral Health, vol. 24, no. 1, p. 105, 2024.

[12]. X. Cai et al., "Self-Supervised Auxiliary Detection (SSAD) framework on DENTEX dataset," IEEE Transactions on Medical Imaging, vol. 43, no. 2, pp. 512–523, 2024.

[13]. P. Wang et al., "Multi-scale feature fusion in optimized YOLOv8s for dental disease detection," BMC Oral Health, vol. 24, no. 1, p. 214, 2024.

[14]. O. Kuran et al., "YOLOv8 for detecting and segmenting dental restorations in mixed dentition," Dentomaxillofacial Radiology, vol. 53, no. 1, p. 20230182, 2024.

[15]. DentalX / YOLO26 Research Group, "Context-Aware YOLO architecture for panoramic radiograph multi-condition detection," Journal of Digital Imaging, vol. 37, no. 2, pp. 310–322, 2025.

[16]. YOLO-DentSeg Team, "Lightweight YOLOv8n-seg segmentation models for low-latency panoramic radiograph processing," IEEE Access, vol. 13, pp. 10214–10225, 2025.