Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

DentAI: Next-Gen Mobile App for Dental Caries Detection with AI and XR

Shereen Afifi

shereen.moataz@guc.edu.eg

German University in Cairo

Samir ElBatal

German University in Cairo

Nader Rezallah

City University Ajman

Research Article

Keywords: Augmented reality, Deep learning, Dental caries, Extended reality, Mobile health, Object detection, Teledentistry, YOLO

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11556150/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

DentAI: Next-Gen Mobile App for Dental Caries Detection with AI and XR

Shereen Afifi[1*], Samir Sameh ElBatal[1] and Nader Nabil Fouad Rezallah[2]

[1*]Faculty of Media Engineering and Technology, German University in Cairo, New Cairo, 11835, Cairo, Egypt.
[2]Oral and Maxillofacial Radiology, City University Ajman, Ajman, 346, United Arab Emirates.

*Corresponding author(s). E-mail(s): shereen.moataz@guc.edu.eg;
Contributing authors: n.nabil@cu.ac.ae;

Abstract

Dental caries is the most prevalent non-communicable disease worldwide, yet lesions are often diagnosed late because conventional examination depends on clinical visits and on the experience of the examiner. This article presents DentAI, a cross-platform mobile application that detects and classifies dental caries in intraoral photographs and communicates the findings through an augmented-reality-style live camera overlay. A dataset of 900 clinically reviewed intraoral images with 1,504 bounding-box annotations across five lesion classes was curated from two public sources through de-duplication, expert screening, and class balancing. End-to-end YOLOv8 and YOLOv11 detectors and a two-stage YOLOv8 + MobileNetV3 pipeline were compared under identical conditions. YOLOv8 achieved an mAP@0.5 of 72.96%, an mAP@0.5:0.95 of 43.95%, a precision of 69.10%, and a recall of 70.97%; YOLOv11 reached a higher precision (83.43%) and mAP@0.5 (73.90%) but a lower recall (63.54%); the two-stage pipeline reached an mAP@0.5 of only 54.88%. YOLOv8 was deployed because of its higher recall and balanced behavior. The model runs on a Flask backend on AWS EC2 behind a Flutter app that also offers dentist chat, appointment booking, scan history, and OCR-assisted registration. A pilot session at a private dental clinic indicated that the practicing dentist found the application clinically relevant and useful for explaining findings to patients.

Keywords: Augmented reality, Deep learning, Dental caries, Extended reality, Mobile health, Object detection, Teledentistry, YOLO

1 Introduction

Dental caries, commonly known as tooth decay, is the most common non-communicable disease globally; an estimated two billion people suffer from caries of permanent teeth [1]. Although caries is preventable and reversible in its earliest stages, lesions are usually noticed only when they have progressed to cavitation, at which point treatment is more invasive, painful, and costly. Conventional detection relies on visual–tactile examination and radiography, both of which require a clinic visit and are sensitive to the experience of the examiner. Early and hidden lesions are frequently missed, and patients in underserved areas, or those deterred by cost or fear, often present late.

Deep learning has shown strong performance in dental image analysis, with convolutional neural networks (CNNs) reaching high accuracy on radiographs and intraoral photographs [2–5]. In oral radiology, for example, Rezallah et al. [6] developed a two-stage system in which a MobileNetV2 classifier detects periodontal bone loss on panoramic radiographs and a YOLOv8 detector localizes each affected region and grades it as mild, moderate, or severe. The classifier reached an accuracy of 88%, the detector reached a precision of 0.74 and a recall of 0.70, and the results were validated by an oral radiology specialist. Smartphones now carry cameras and processors capable of acquiring and transmitting diagnostic-quality images, and extended reality (XR) makes it possible to present findings directly on the live camera view. Together, these technologies can move caries screening out of the clinic and into the hands of patients.

However, most AI caries-detection systems reported to date remain confined to controlled research settings. They are typically evaluated offline on static images, depend on specialized hardware such as intraoral cameras or radiographic units, and pay little attention to the user experience, to the explainability of results for patients, or to the clinical workflow that must follow a positive finding [3, 5]. Conversely, most dental mobile-health (mHealth) and XR applications focus on hygiene education or anxiety management rather than diagnosis [7].

This article presents *DentAI*, a next-generation mobile application that combines deep-learning-based caries detection with an XR-style live visualization and the services needed to act on a result. The main contributions are as follows:

- A curated dataset of 900 intraoral images with 1,504 bounding-box annotations over five lesion classes, built from two public sources through manual de-duplication, expert clinical screening, and targeted class balancing.
- A controlled comparison of three detection pipelines—end-to-end YOLOv8, end-to-end YOLOv11, and a two-stage YOLOv8 + MobileNetV3 pipeline—trained under identical settings and evaluated quantitatively and qualitatively.
- A cross-platform Flutter application offering image-upload and live camera detection with color-coded, per-class overlays, together with dentist chat, appointment booking, scan history, insurance records, and optical character recognition (OCR)-assisted registration from national ID cards.

Table 1 ICDAS codes for visual caries assessment [8]

Code	Clinical description
0	Sound surface; no evidence of caries after 5 s of air drying
1	First visual change in enamel (white or brown opacity), seen only after drying
2	Distinct visual change in enamel, visible when wet
3	Localized enamel breakdown without visible dentin or shadow
4	Underlying dark shadow from dentin, with or without enamel breakdown
5	Distinct cavity with visible dentin
6	Extensive distinct cavity involving more than half of the surface

- A production-style deployment of the inference and application services on Amazon Web Services (AWS) EC2 and a pilot evaluation with a practicing dentist in a private clinic.

The remainder of this article is organized as follows. Section 2 introduces the clinical and technical background. Section 3 reviews related work and identifies the research gap. Section 4 describes the overall DentAI system and its labeling scheme, and Section 5 describes the dataset. Section 6 presents the detection models, the training setup, and the evaluation metrics. Sections 7 and 8 describe the mobile application, the AR visualization, the backend, and the deployment. Section 9 reports the experimental results, and Section 10 reports the clinical pilot. Section 11 discusses limitations, and Section 12 concludes the article.

2 Background

2.1 Dental caries and the ICDAS scale

Dental caries is a chronic, multifactorial disease in which acids produced by the bacterial fermentation of dietary carbohydrates demineralize the dental hard tissues. Its progression is influenced by the oral microbiome, salivary flow, fluoride exposure, sugar intake, and oral hygiene. Early enamel lesions can be arrested or reversed, whereas untreated lesions progress to irreversible cavitation. Early diagnosis and monitoring are therefore central to preventive dentistry.

To standardize visual diagnosis, the International Caries Detection and Assessment System (ICDAS) grades each tooth surface with a seven-level code from sound enamel (code 0) to extensive cavitation (code 6) [8]. The codes are summarized in Table 1, and clinical examples are shown in Fig. 1. ICDAS provides the clinical reference on which the labeling scheme of DentAI was built (Section 4.2).

2.2 Object detection and the YOLO family

Object detection combines classification with localization: the model must predict, for every object in an image, a bounding box and a class label. This formulation matches intraoral photographs, which usually contain several teeth and may contain several lesions of different types. Single-stage detectors of the YOLO (You Only Look Once)

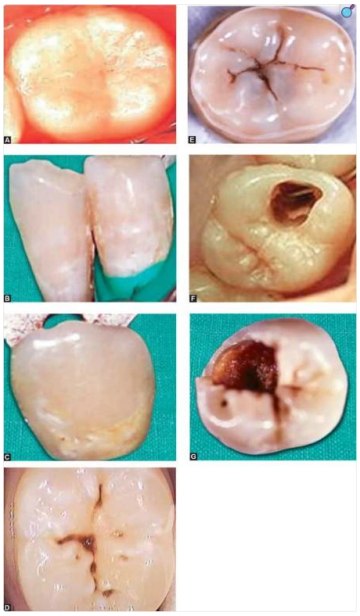


Fig. 1 Clinical examples of ICDAS codes 0–6, shown as panels (A)–(G), from a sound surface to extensive cavitation. Reproduced from [8]

family [9] predict boxes and classes in one forward pass and are widely used where real-time inference is required. The pipeline consists of image pre-processing, CNN feature extraction, dense box and class prediction at several scales, and non-maximum suppression (NMS) to remove duplicate detections.

YOLOv8 [10] introduced an anchor-free, decoupled detection head and C2f feature-extraction blocks. YOLOv11 [11] refines the backbone with C3k2 blocks and adds spatial attention (C2PSA) in the neck, improving the accuracy–latency trade-off on the COCO benchmark [12], as illustrated in Fig. 2. Lightweight classifiers such as MobileNetV3 [13] target mobile hardware with depthwise-separable convolutions and squeeze-and-excitation attention.

2.3 Cross-platform mobile development and XR

Mobile applications can be developed natively (Kotlin/Java for Android, Swift for iOS) or with cross-platform frameworks that compile one codebase for both platforms. Flutter [14] renders its own widgets, offers near-native performance, and provides mature plugins for the camera, image handling, and HTTP communication, which are the core capabilities required by an AI-driven mHealth application.

XR is an umbrella term for augmented reality (AR), which overlays digital content onto the live view of the real world; virtual reality (VR), which immerses the user in a fully synthetic environment; and mixed reality (MR), in which virtual objects are anchored to and interact with the physical scene. In dentistry, XR has been applied mainly to education, surgical training, and anxiety management [15, 16]. DentAI uses

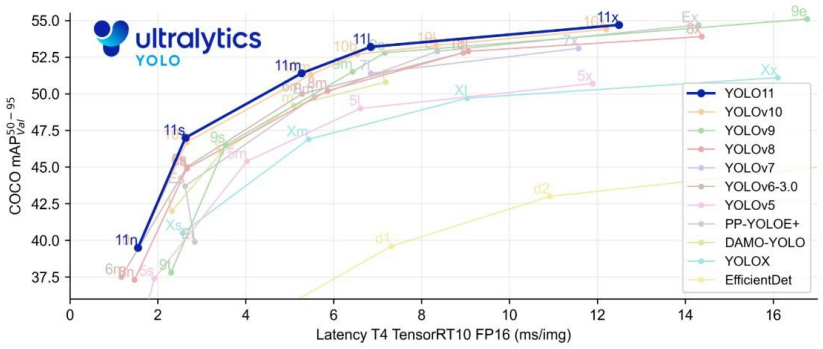


Fig. 2 COCO accuracy (mAP) versus inference latency for YOLO11 and earlier YOLO releases. Reproduced from [11]

AR-style overlays on the live smartphone camera stream to show where the model has detected lesions.

3 Related work

3.1 Deep learning for caries detection

Several studies have applied deep learning to intraoral photographs captured by smartphones or intraoral cameras. Thanh et al. [17] compared YOLOv3, Faster R-CNN, RetinaNet, and SSD on smartphone photographs labeled with four severity classes; Faster R-CNN achieved the highest overall accuracy (87.4%), but the sensitivity for early lesions remained below 37%. Park et al. [18] showed that segmenting tooth surfaces with U-Net before classification (ResNet-18) and localization (Faster R-CNN) raised the classification AUC from 0.731 to 0.837. Kühnisch et al. [19] trained MobileNetV2 on 2,417 single-tooth photographs and reached 92.5% accuracy, while Boy et al. [20] obtained 90% accuracy with MobileNetV3-Small on low-cost smartphone images and confirmed its usability in a 50-volunteer field test.

Closer to the present work, Fadilah et al. [21] trained YOLOv8x on 3,221 photographs of schoolchildren labeled with ICDAS codes; the model reached 72.6% precision but only 41.1% recall and 45.8% mAP@0.5, with the weakest performance on early lesions. Frenkel et al. [22] externally validated a web-based classifier on 718 internet photographs (92% detection accuracy), and Zhang et al. [23] evaluated a MobileNetV3–U-Net model on 4,361 teeth of 191 patients in a clinical setting (93.40% accuracy), reporting lower sensitivity for interproximal and buccal lesions. On radiographs, YOLOv8 has reached F1-scores of 82.22% for interproximal caries on bitewings [24] and 92.0% in caries classification [25]. Table 2 summarizes these and related studies [26–28].

Table 2 Summary of representative deep learning studies on caries detection

Ref.	Data and modality	Model(s)	Key results	Main limitation
[17]	1,902 train / 750 test smartphone photos; 4 classes	YOLOv3, Faster R-CNN, RetinaNet, SSD	Faster R-CNN acc. 87.4%; YOLOv3 sens. 74% (cavitated); early-lesion sens. ≤36.9%	One phone model; poor early-lesion sensitivity
[18]	2,348 intraoral camera photos	U-Net + ResNet-18 + Faster R-CNN	AUC 0.731→0.837 with segmentation; localization AP 0.889	Single camera; interproximal lesions not visible
[19]	2,417 single-tooth photos; 3 classes	MobileNetV2	Acc. 92.5%, sens. 89.6%, spec. 94.3%, AUC 0.964	Idealized images; single annotator
[20]	1,200 smartphone photos; binary	MobileNetV3-Small	Acc., precision, and recall 90%; 50-volunteer field test	Sensitive to noise and low resolution
[21]	3,221 photos of schoolchildren; ICDAS D0–D6	YOLOv8x	P 72.6%, R 41.1%, mAP@0.5 45.8%; 4× faster than dentists	Low recall; weak on early stages
[22]	718 internet photos (external validation)	Web-based CNN	Detection acc. 92% (AUC 0.909); localization 97%	Internet-image bias; no histology
[23]	4,361 teeth of 191 patients (clinical)	MobileNetV3 + U-Net	Acc. 93.40%, sens. 81.31%, spec. 95.65%	Low sensitivity on interproximal lesions
[24]	552 bitewing radiographs	YOLOv8	F1 82.22% overall; 96.03% on enamel caries	Single institution and device
[25]	Dental radiographs	YOLOv8 (customized)	Acc. 92.53%, F1 92.0%	Single city; subtle lesions weaker
[26]	Mobile-phone oral photos	YOLOv3	Acc. 89.76%, sens. 91.12%, AUC 0.925	Small dataset; occult caries missed
[27]	Intraoral photos (caries index)	YOLOv5s, DETR	YOLOv5s: P 90.7%, R 85.6%, mAP@0.5 90.7%	Decay only; rural population skew
[28]	232 original smartphone images; 6 classes	YOLOv7	P 96.03%, R 86.53%, mAP@0.5 93.54%	Small dataset; single phone model

acc.: accuracy; sens.: sensitivity; spec.: specificity; P: precision; R: recall; AP: average precision; AUC: area under the ROC curve.

3.2 Mobile health applications for dental care

Patient-oriented dental apps are dominated by oral-health promotion. A scoping review of 45 studies found that most apps targeted toothbrushing and hygiene behavior, and fewer than half addressed diagnosis or teledentistry [7]. A gamified app for mothers of preschool children improved oral-health knowledge and reduced plaque in a randomized trial [29], and the *My Dental Hygiene Care* information system obtained a mean satisfaction of 4.58/5 among 30 dental hygienists [30]. The closest system to DentAI is *AICaries* [31], an Android app with which parents photograph their children's teeth and receive AI-based caries screening, risk scoring, and education; it reached a System Usability Scale score of 78.4, but it was tested at a single site and its diagnoses were not compared against a clinical reference.

Table 3 Comparison of DentAI with related dental mHealth and XR applications

System	AI caries detection	XR	Chat / booking	Evaluation
AICaries [31]	Yes	–	–	Usability; 37 parent–child pairs
My Dental Hygiene Care [30]	–	–	Records only	Satisfaction; 30 hygienists
Gamified app [29]	–	–	–	RCT; 58 mother–child pairs
Dr. Barea [35]	–	AR + VR	–	Pilot; 16 children
VR distraction [33]	–	VR	–	Randomized; 90 adults
DentAI (ours)	**Yes (5 classes)**	**AR**	**Yes / Yes**	Detection metrics; clinic pilot

–: feature not reported in the cited study; RCT: randomized controlled trial.

3.3 XR technologies in dentistry

VR has been used successfully as a non-pharmacological distraction during dental treatment. It lowered anxiety and improved cooperation in 140 children and adolescents with autism spectrum disorder [32], reduced heart rate, stress, and anxiety in a randomized study of 90 adults [33], and is being evaluated for children with special health-care needs [34]. *Dr. Barea* [35] combines a 360° VR clinic tour with AR scanning of dental instruments for Arabic-speaking children, reducing dental-fear scores in a pilot with 16 children. Reviews highlight AR guidance and VR simulation in dental surgery and education [15, 16]. None of these systems uses XR to visualize diagnostic findings.

Table 3 contrasts DentAI with these applications.

3.4 Research Gaps

The literature reveals three gaps. First, most caries models are trained on small, single-center, or single-device datasets and are evaluated offline, so their behavior in real clinical workflows is largely unknown [3, 5]. Second, AI screening apps rarely connect a positive finding to the next clinical step, such as consulting or booking a dentist. Third, XR in dentistry has been used for education and anxiety management, but not to show patients the location and type of detected lesions. DentAI addresses these gaps by combining multi-class caries detection, live AR-style visualization, and teledentistry services in a single deployed application that was piloted in a clinic.

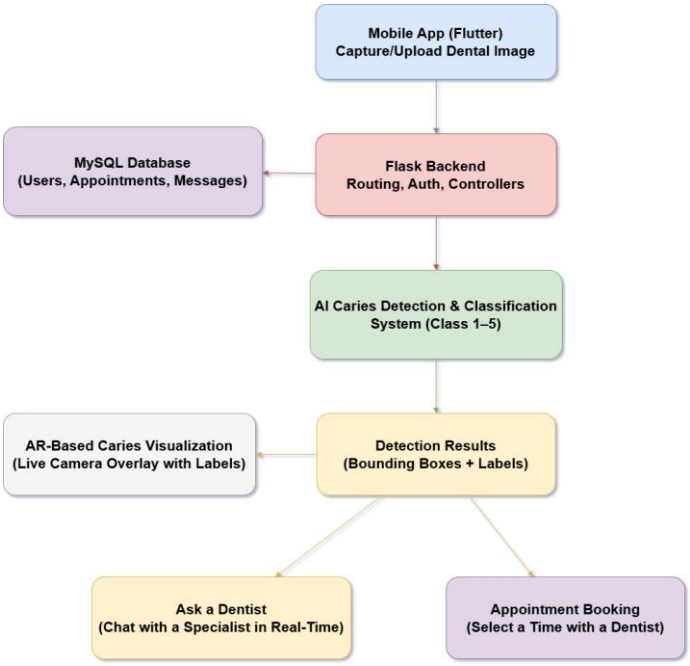


Fig. 3 Overall DentAI system architecture, from image capture in the Flutter app through backend inference to AR visualization, dentist consultation, and appointment booking

4 The DentAI system

4.1 System overview

DentAI was initially conceived as an app that detects caries in intraoral photographs. Following consultation with the clinical co-author, an oral and maxillofacial radiologist, the scope was extended to classifying lesions into five clinically meaningful classes and to supporting the steps that follow a diagnosis. Fig. 3 shows the resulting architecture. The Flutter app captures or uploads an intraoral image and sends it to a Flask backend, which handles routing, authentication, and persistence in a MySQL database. The backend runs the caries detection model and returns bounding boxes with class labels. The results are then used in three ways: they are drawn as an AR-style overlay on the live camera view or on the uploaded image, they can be shared with a dentist through the *Ask a Dentist* chat, and they can lead to an appointment booking.

The application provides the following functional modules:

- **Caries detection and classification** from an uploaded photograph or from the live camera stream, with per-class counts.
- **Live AR visualization** of detected lesions as color-coded boxes on the camera preview.
- **Ask a Dentist**, an in-app messaging service through which a scan can be shared with a chosen dentist.

Table 4 Caries classes used for annotation in DentAI

Class	Lesion site and appearance
1	Pits and fissures of occlusal surfaces of molars and premolars, and similar grooves on anterior teeth
2	Proximal (interdental) surfaces of posterior teeth; often hidden without instruments or radiographs
3	Proximal surfaces of incisors and canines, without involvement of the incisal edge
4	Proximal surfaces of anterior teeth involving the incisal edge; affects function and appearance
5	Cervical third (near the gum line) of the buccal or lingual surface of any tooth

- **Appointment booking** with search, filtering, and real-time availability of time slots.
- **Patient history**, in which scans and results are stored so that dentists can review previous findings.

4.2 Lesion labeling scheme

A seven-level scheme mirroring ICDAS codes 0–6 was initially planned. After reviewing the collected images with the clinical co-author, Classes 0 and 6 were found to be rare or poorly represented, so the final scheme uses five classes (Class 1 to Class 5) that are common in practice and visually distinguishable in photographs. The classes are defined primarily by the anatomical site of the lesion, as summarized in Table 4.

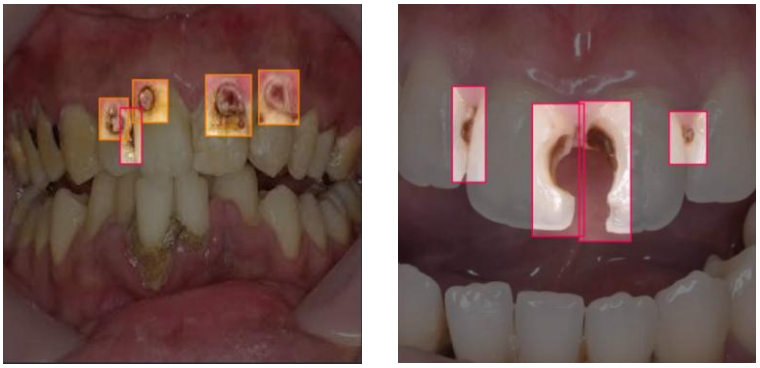
5 Dataset

5.1 Collection and filtering

The dataset used in this study was developed from two publicly available intraoral photographic datasets hosted on Roboflow: The Tooth Caries Detection dataset by D7i [**36**], comprising 2,250 images, and the DentalMate caries dataset [**37**], comprising 5,400 images. These datasets were selected based on their relevance to dental caries detection and the availability of intraoral photographic images.

The dataset preparation process involved two stages of filtering. Initially, duplicate images, previously augmented versions, and images of insufficient quality were manually excluded, resulting in approximately 1,500 original high-resolution images. Subsequently, the remaining images underwent clinical screening and verification in collaboration with the clinical co-author, a specialist in oral and maxillofacial radiology. Images demonstrating unrelated oral conditions, unclear dental anatomy, or inadequate diagnostic quality were excluded.

Following this process, 750 clinically reviewed intraoral photographs were retained for annotation and subsequent model development. The images were obtained exclusively from the publicly available datasets, without direct patient recruitment or collection of additional clinical images at the authors' institutions. Representative examples are presented in **Figure 4**.



(a) (b)

Fig. 4 Intraoral images containing multiple carious lesions of different classes, from the source dataset [36]

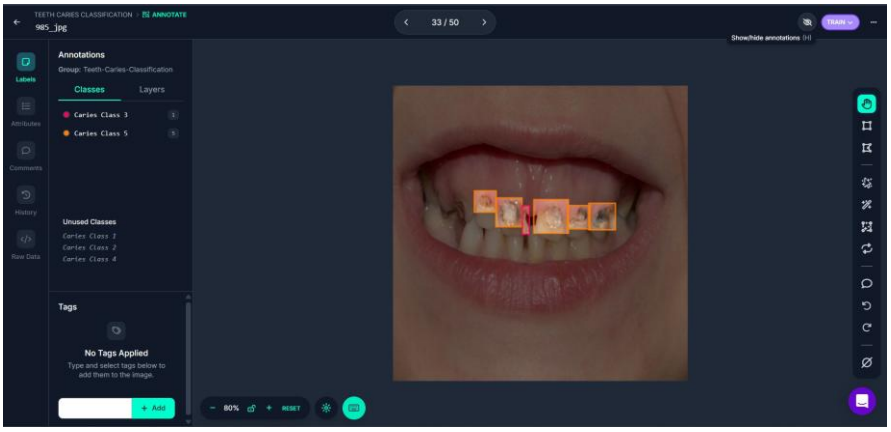


Fig. 5 Example of an intraoral image annotated in Roboflow, with each lesion labeled by its caries class

5.2 Annotation

All 750 images were annotated manually in Roboflow by drawing a bounding box around every visible lesion and assigning one of the five classes of Table 4. Annotation was carried out in consultation with the clinical co-author to ensure clinical accuracy and consistency. An annotated example is shown in Fig. 5.

5.3 Class balancing

The initial 750 images contained 1,298 annotated lesions, with a clear imbalance: Class 1 was over-represented and Class 4 was under-represented (Table 5). Targeted augmentation was applied to images containing the minority Classes 2 and 4, and the number of Class 1 annotations was slightly reduced to limit overfitting to the dominant

Table 5 Distribution of caries annotations before and after balancing

Class	Initial (750 images)	Final (900 images)
Class 1	377	310
Class 2	234	298
Class 3	289	291
Class 4	107	314
Class 5	291	291
Total	1,298	1,504

Table 6 Data augmentation operations

Operation	Setting and purpose
Flip	Random horizontal and vertical flips (capture orientation)
90° rotation	Multiples of 90° (camera angle)
Free rotation	Uniform in $[-20^{\circ}, +20^{\circ}]$
Brightness	Uniform in $[-20\%, +20\%]$ (illumination)
Gaussian blur	Up to 2.5 pixels (focus errors)
Gaussian noise	Additive noise (low-quality sensors)
Multiplier	3 augmented copies per training image

class. The balanced dataset contains 900 images and 1,504 annotations, with every class represented by 291 to 314 lesions.

5.4 Preprocessing and augmentation

A preprocessing pipeline implemented with Pillow, NumPy, and OpenCV applied three operations to every image: (i) auto-orientation, which corrects rotation using the EXIF metadata written by phones and intraoral cameras; (ii) resizing to 640×640 pixels, the input size of the detectors; and (iii) contrast stretching between the 2nd and 98th intensity percentiles, which improves the visibility of subtle lesions in dark or low-contrast images. Pixel intensities were then normalized to $[0, 1]$.

To improve generalization, the training data were further expanded with the Albumentations library [38], generating three augmented copies of each training image. The transformations, listed in Table 6, simulate variations in capture orientation, illumination, focus, and sensor quality without distorting the anatomy. In addition, mosaic augmentation was enabled during detector training.

6 Caries detection models

A classification network assigns a single label to a whole image and cannot localize or separate multiple lesions. Two detection-based designs were therefore investigated: an end-to-end design in which a YOLO model localizes and classifies every lesion in one pass, and a two-stage design in which detection and classification are decoupled.

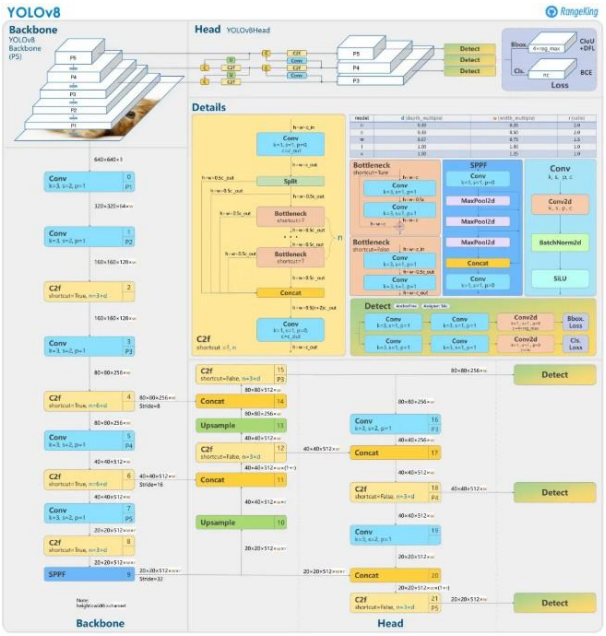


Fig. 6 YOLOv8 backbone and detection head, showing the convolutional and C2f blocks used for multi-scale feature extraction. Adapted from [39]

6.1 YOLOv8

YOLOv8 [10] uses a convolutional backbone built from C2f blocks and an anchor-free head that predicts boxes and class scores directly at three scales (Fig. 6). Predictions from the different feature maps are decoded, concatenated, and filtered with NMS to produce the final detections.

6.2 YOLOv11

YOLOv11 [11] consists of a backbone with C3k2 blocks for efficient feature extraction; a neck with spatial pyramid pooling (SPPF) for multi-scale context and C2PSA attention blocks that focus on informative regions; and a head that detects objects at three scales (Fig. 7).

6.3 MobileNetV3

MobileNetV3 [13] is a lightweight CNN designed for low-latency inference on mobile devices. Each block (Fig. 8) expands the input with a 1×1 pointwise convolution, captures spatial information with a 3×3 depthwise convolution, reweights channels with a squeeze-and-excitation module, and projects back to a lower dimension with a second 1×1 convolution, using hard-swish activations. In DentAI, MobileNetV3 classifies cropped lesion regions into the five caries classes.

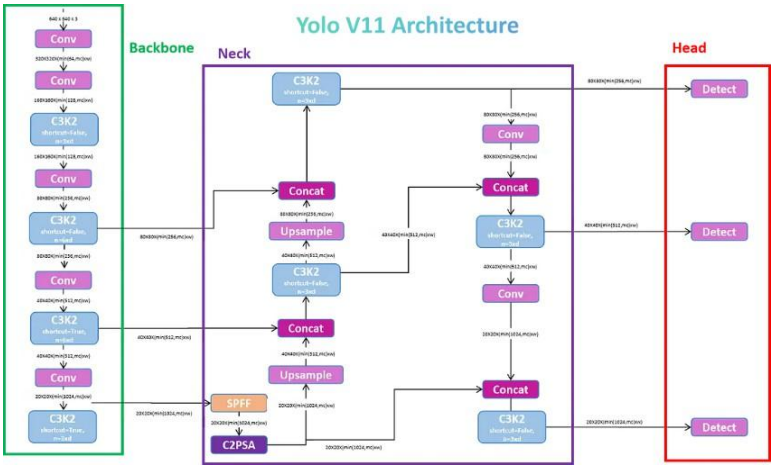


Fig. 7 YOLOv11 architecture with the backbone (feature extraction), neck (feature fusion and attention), and head (three-scale detection). Adapted from [40]

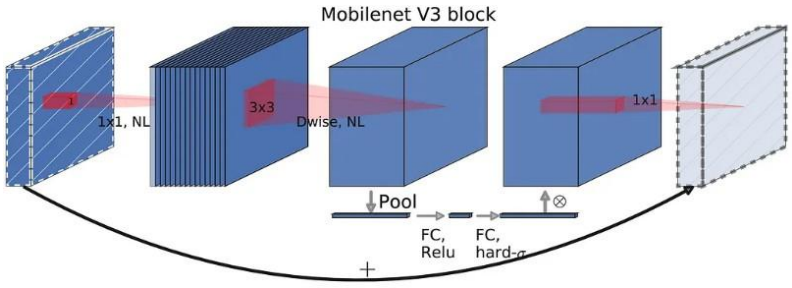


Fig. 8 MobileNetV3 block with pointwise and depthwise convolutions and a squeeze-and-excitation module for channel-wise reweighting. Adapted from [41]

6.4 Implementation 1: end-to-end detection

In the first implementation, YOLOv8 and YOLOv11 are each trained directly on the five-class annotations. At inference time, an image received from the app is passed through the detector, which returns every lesion as a box with its class and confidence score. If no lesion is found, the app is informed that the teeth appear healthy; otherwise, an annotated image or the list of detections is returned (Fig. 9a). This choice was motivated by recent reports of YOLOv8 on dental images [21, 24, 25].

6.5 Implementation 2: two-stage detection and classification

In the second implementation, all five classes are merged into a single *caries* class, and YOLOv8 is trained purely for lesion localization. Every predicted box is then cropped from the original image and classified by MobileNetV3 into Classes 1–5 (Fig. 9b). To train the classifier, every annotated box in the dataset was cropped and stored in

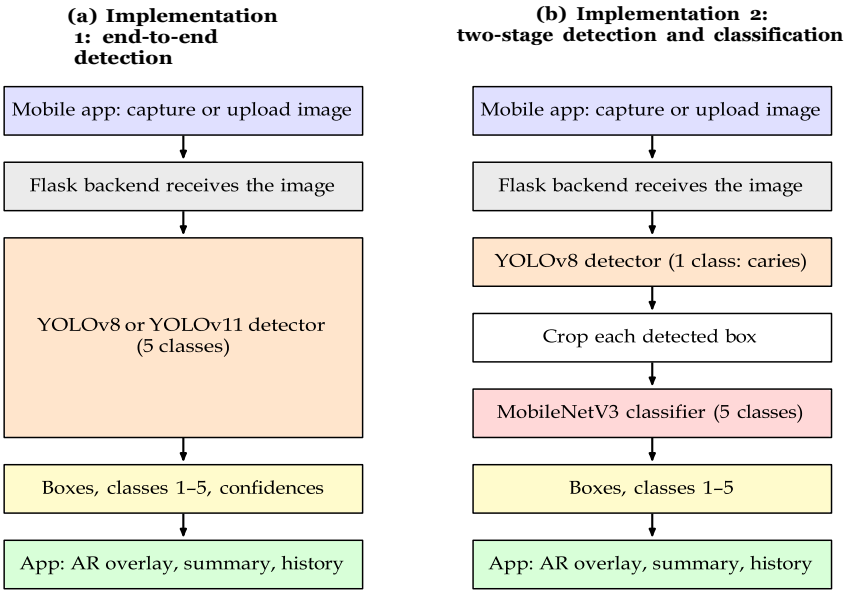


Fig. 9 The two caries detection pipelines compared in this work. In both pipelines, an image with no detection returns a "healthy" result to the app

a class-specific folder. This design, motivated by the robustness of MobileNetV3 on smartphone images [20] and by the two-stage MobileNet and YOLOv8 system used in oral radiology [6], allows each stage to be optimized independently.

6.6 Training setup

All models were trained with the Ultralytics framework [10] on a local NVIDIA RTX 4060 GPU, starting from COCO-pretrained small variants (yolov8s.pt and yolo11s.pt). The dataset was split with stratification into 70% training, 20% validation, and 10% test data. Both end-to-end detectors used identical settings (Table 7), and training budgets of 50, 100, 150, and 200 epochs were explored progressively. A custom script automated data loading, model initialization, training, evaluation, and result visualization so that all pipelines were compared under the same conditions.

6.7 Evaluation metrics

A prediction is matched to a ground-truth box when their intersection over union,

$$\mathrm{IoU}(B_p, B_g) = \frac{|B_p \cap B_g|}{|B_p \cup B_g|}, \tag{1}$$

exceeds a threshold τ and the classes agree. Counting true positives (TP), false positives (FP), and false negatives (FN), precision, recall, and the F1-score are

$$P = \frac{TP}{TP + FP}, \qquad R = \frac{TP}{TP + FN}, \tag{2}$$

Table 7 Training configuration

Setting	Value
Framework	Ultralytics YOLO (Python)
Hardware	NVIDIA RTX 4060 GPU (CUDA)
Initialization	COCO-pretrained yolov8s / yolo11s
Input size	640×640 pixels
Epoch budgets	50, 100, 150, 200 (explored progressively)
Data split	70% / 20% / 10% (train / validation / test), stratified
Online augmentation	Mosaic, ratio 0.5
Classes	5 (Implementation 1); 1 detector class + 5 classifier classes (Implementation 2)

$$F_1 = \frac{2PR}{P + R}. \tag{3}$$

The average precision of class k is the area under its precision–recall curve $p_k(r; \tau)$, and the mean average precision averages it over the $K = 5$ classes:

$$\mathrm{AP}_k(\tau) = \int_0^1 p_k(r; \tau)\, dr, \tag{4}$$

$$\mathrm{mAP@}\tau = \frac{1}{K} \sum_{k=1}^{K} \mathrm{AP}_k(\tau). \tag{5}$$

Following the COCO protocol [12], results are reported as mAP@0.5 and as mAP@0.5:0.95, the average of Eq. (5) over $\tau \in \{0.50, 0.55, \ldots, 0.95\}$, which rewards tighter localization. Training and validation losses (box, classification, and distribution focal losses), learning-rate schedules, and confusion matrices were also monitored.

7 Mobile application and AR visualization

7.1 Technology stack and architecture

The DentAI app was implemented with Flutter [14] and Dart so that a single codebase serves Android and iOS. The main packages are listed in Table 8. The code follows the Model–View–ViewModel (MVVM) pattern: models/ holds the data classes (users, doctors, time slots) that decode backend JSON; screens/ holds top-level pages; widgets/ holds reusable UI components; providers/ holds state management, API calls, and authentication; and assets/ holds images, icons, and insurance logos. The logo, icon, and several screens were designed with a professional UX designer to give the app a consistent visual identity.

7.2 Application walk-through

Fig. 10 shows the main screens of the app. Three onboarding screens introduce caries detection, appointment booking, and *Ask a Dentist*. At sign-up, the user chooses the patient or doctor role and uploads a photograph of the Egyptian national ID

Table 8 Technology stack of the DentAI application and backend

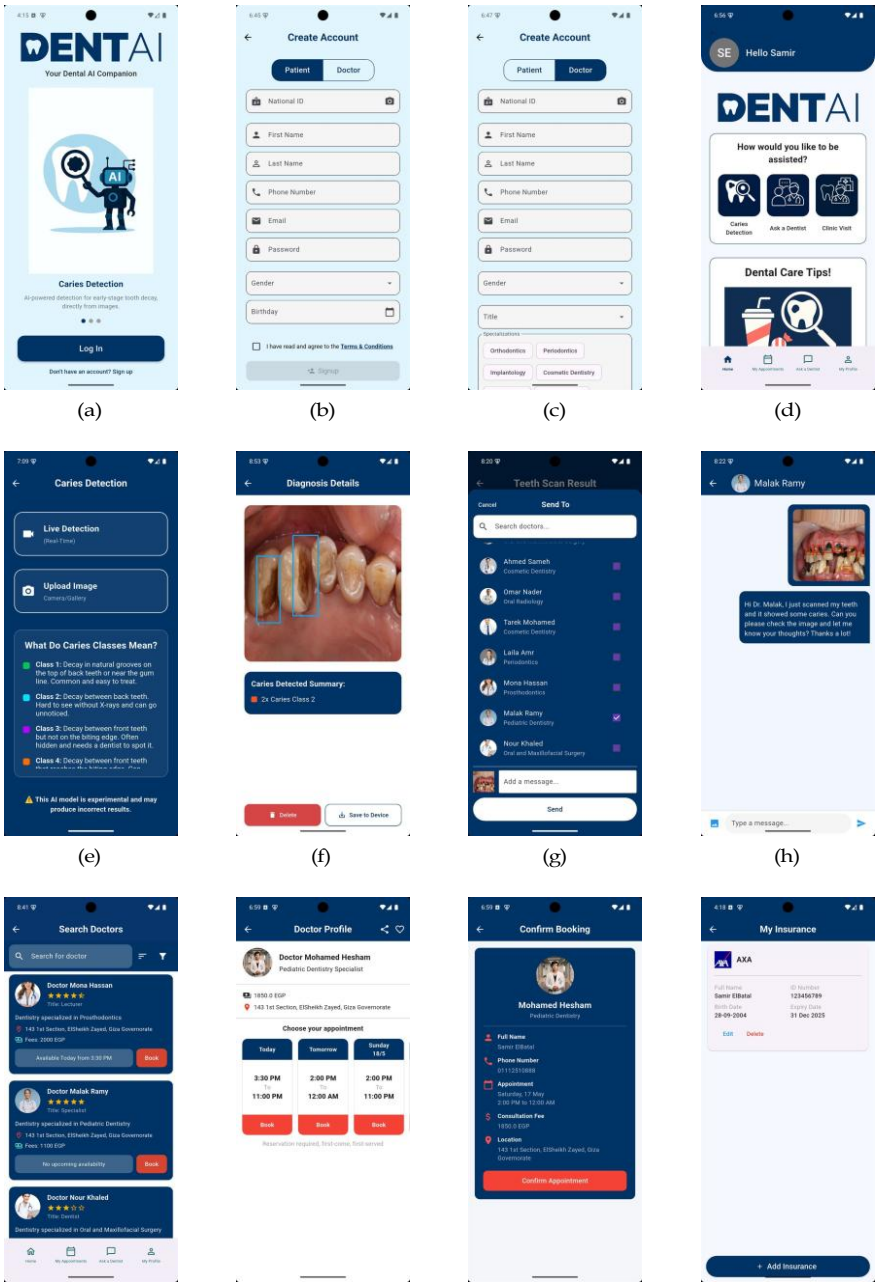
Component	Technology (version)
Mobile application	
UI framework	Flutter SDK (3.7.0), Dart
State management	Provider (6.1.2)
Networking	HTTP (1.3.0)
Local storage	Shared Preferences (2.5.2)
Image capture	Camera (0.11.1), Image Picker (1.0.4)
Saving results	Image Gallery Saver Plus (4.0.1)
Backend and AI service	
Web framework	Flask (3.1.1), Flask-CORS, Gunicorn
Authentication	Flask-JWT-Extended (4.7.1)
Database	MySQL on Aiven, via Flask-SQLAlchemy (3.1.1)
AI model	YOLOv8, Ultralytics (8.3.139), PyTorch (2.7.0), TorchVision (0.22.0)
OCR	Google Cloud Vision (3.10.1), OpenCV, Pillow
Text parsing	Large language model via the Groq API; deep-translator
Media storage	Cloudinary (1.44.0)
Hosting	AWS EC2, Ubuntu 22.04 LTS

card; an OCR service pre-fills the name, national ID number, birth date, and gender, so a patient only needs to add a phone number, e-mail address, and password. Doctors additionally enter their title, specializations, consultation fee, clinic address, biography, and years of experience. Password recovery uses a six-digit code sent by e-mail.

The home screen greets the user by name and offers three entry points: *Caries Detection*, *Ask a Dentist*, and *Clinic Visit*. It also shows a scrollable set of dental-care tips to encourage prevention. The detection screen offers *Live Detection* and *Upload Image* and explains every caries class, so that the user can interpret the result. After an upload, the app shows the annotated image with per-class counts and offers to *Save* the scan to the user's history or *Share* it with a dentist chosen by specialization, together with a short message. Conversations continue in the chat module. For an in-person visit, the user can search for doctors by name or specialty and filter by gender, title, fee, and availability (e.g., today or tomorrow), choose a time slot on the doctor's profile, and confirm the booking. The profile section manages account data, insurance records (the provider, ID number, expiry date, and a photograph of the card), and the history of scans. In doctor mode, dentists can review shared scans, reply to patients, and see upcoming appointments.

7.3 Live AR caries visualization

A first prototype used Unity to overlay a 3D jaw model on the camera feed and mark lesions on the corresponding virtual teeth (Fig. 11). This approach was abandoned for two reasons: the Flutter and Unity versions available at the time were incompatible, and registering the virtual teeth to the real teeth in the live feed proved technically difficult. The final design implements the AR layer natively in Flutter, as described



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

(e) (f) (g) (h)

(i) (j) (k) (l)

Fig. 10 DentAI user interface. (a) Onboarding. (b) Patient sign-up with OCR pre-filled fields. (c) Doctor sign-up with professional fields. (d) Home screen. (e) Detection entry with class definitions. (f) Saved scan with detection result. (g) Sharing a scan with a dentist. (h) *Ask a Dentist* chat. (i) Doctor search. (j) Doctor profile with time slots. (k) Booking confirmation. (l) Insurance records

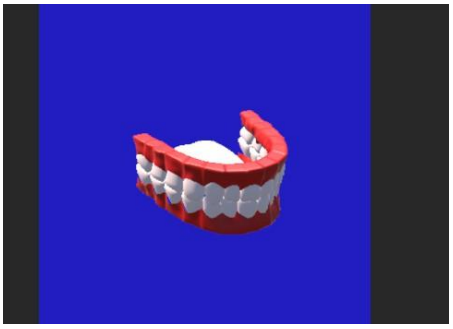


Fig. 11 Unity-based 3D jaw prototype for visualizing caries on a virtual model. It was not deployed because of Flutter–Unity integration limitations

Algorithm 1 Live AR caries detection loop

Require: camera stream S, detection endpoint E, class colors $\kappa_1, \ldots, \kappa_5$, sampling period $T = 1$ s

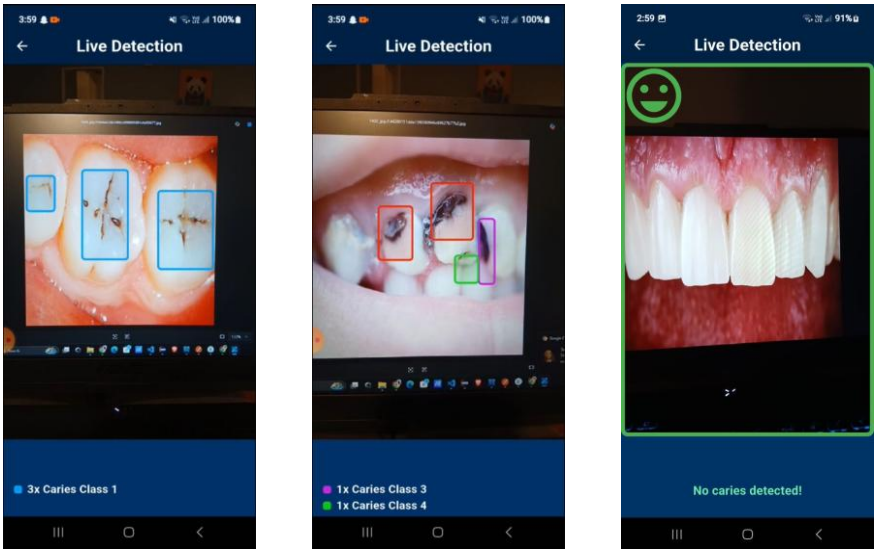
while the live view is open **do**
  $I \leftarrow$ frame captured from S every T
  send I to E with an HTTP POST request
  Server: $\mathsf{D} \leftarrow \{(b_i, c_i, s_i)\} = \text{YOLOv8}(I)$
  Server: return D as JSON
  if $\mathsf{D} = \emptyset$ **then**
    draw a green frame and show "No caries detected!"
  else
    for all $(b_i, c_i, s_i) \in \mathsf{D}$ **do**
      scale b_i to the preview size; draw it with color κ_{c_i}
    end for
    show counts $n_k = |\{i : c_i = k\}|$ for $k = 1, \ldots, 5$
  end if
end while

in Algorithm 1. The camera plugin streams the preview, one frame is sent to the backend every second, and the backend returns the boxes, classes, and confidences as JSON without rendering an image. The app then draws a rectangle for every detection, colored by class, and updates a per-class summary below the preview. When no lesion is found, the preview is framed in green with the message "No caries detected!" (Fig. 12). This design avoids the memory and battery cost of a 3D engine and is easy to distribute.

8 Backend architecture and deployment

8.1 Services

The backend is a Flask application organized into routes/ (API endpoints), controllers/ (request handling), services/ (AI inference, image processing,



(a) (b) (c)

Fig. 12 Live detection with the AR-style overlay. (a), (b) Lesions are marked with boxes colored by class, and a per-class summary is shown below the preview. (c) Healthy result with a green frame

and utilities), models/ (SQLAlchemy object-relational mapping classes), and middlewares/ (JSON Web Token (JWT) authentication and role validation). Its core services are as follows:

- **Authentication and authorization** with JWT-protected endpoints for patients and doctors.
- **Caries detection.** The best YOLOv8 checkpoint is loaded by the server and exposed through two endpoints. The *static* endpoint returns an annotated image and a per-class summary for uploaded photographs; the *live* endpoint returns only box coordinates, classes, and confidences so that the app can render the overlay itself. Inference runs efficiently on CPU, allowing deployment on low-cost cloud instances.
- **Messaging**, which stores consultation threads, messages, and exchanged images for every doctor–patient pair.
- **Scheduling**, which verifies slot availability, books appointments, and records their status.
- **OCR registration.** The ID-card photograph is sharpened and denoised with OpenCV and Pillow, raw text is extracted with the Google Cloud Vision API, a large language model accessed through the Groq API parses the text into structured fields (name, address, birth date, and national ID number), and Arabic names and digits are translated and normalized with the deep-translator library.

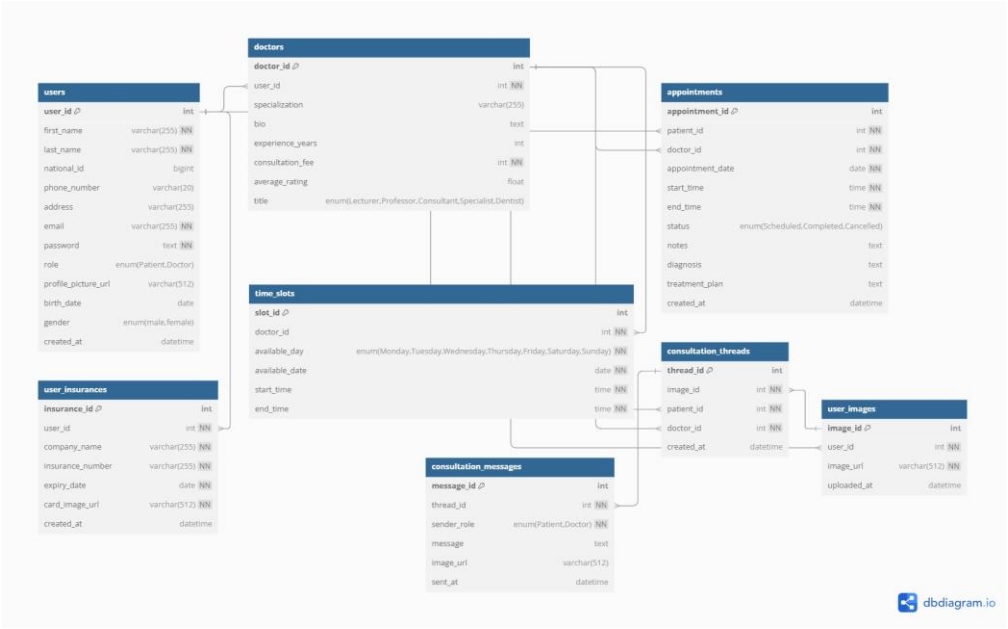


Fig. 13 Enhanced entity–relationship diagram of the DentAI database

8.2 Database design

The relational schema, designed in MySQL Workbench and hosted on the Aiven cloud service, is shown in Fig. 13. The users table stores common account data and the role; doctors extends it with specialization, experience, fee, title, biography, and rating; appointments links patients and doctors with date, time, status, diagnosis, and treatment plan; time_slots stores doctor availability; user_images stores uploaded scans; user_insurances stores insurance details; and consultation_threads and consultation_messages store the chat history, linked to the shared image. Uploaded images are stored in Cloudinary.

8.3 Cloud deployment

The backend was deployed on an AWS EC2 instance running Ubuntu 22.04 LTS. The instance was initially a t3.medium and was later upgraded to a t3.large to better support YOLO inference and concurrent API calls. The project was copied to the server over SSH, dependencies were installed in a Python 3.10 virtual environment, and the application was served by Gunicorn on port 5000. The security group allowed SSH only from the developer's address. The deployed API, which serves detection, authentication, chat, appointment, and insurance endpoints, was verified with curl and Postman before the mobile app was connected to it.

Table 9 Performance of the three pipelines on five-class caries detection (%)

Metric	YOLOv8	YOLOv11	YOLOv8 + MobileNetV3
mAP@0.5	72.96	**73.90**	54.88
mAP@0.5:0.95	**43.95**	43.64	23.54
Precision	69.10	**83.43**	39.64
Recall	**70.97**	63.54	32.84
F1-score	70.02	**72.11**	32.86

The best value in each row is shown in bold.

9 Experimental results

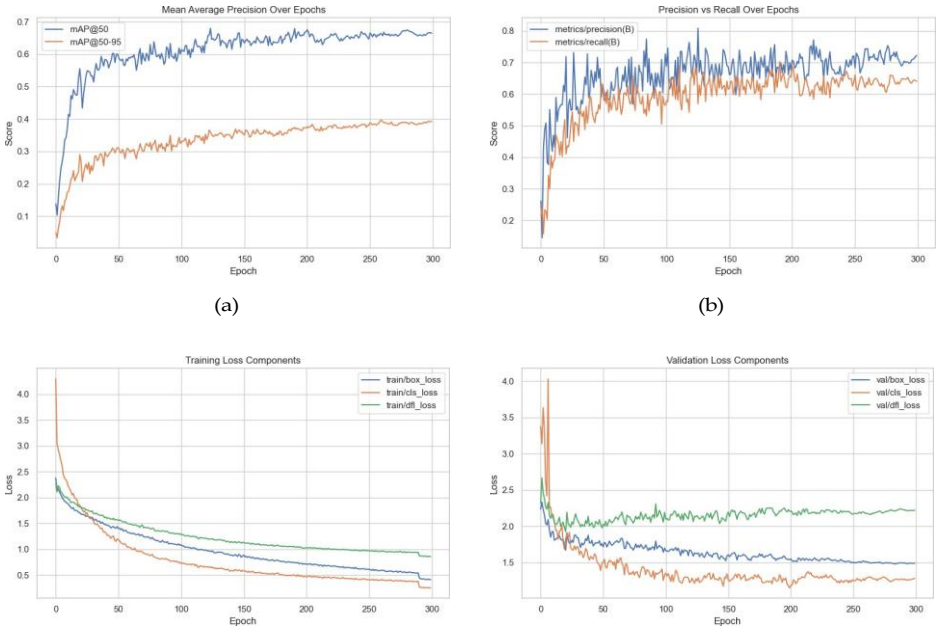
9.1 Quantitative comparison

Table 9 compares the three pipelines, which were trained and evaluated on the same dataset under the same conditions. The two end-to-end detectors perform similarly and far better than the two-stage pipeline. YOLOv11 obtains the highest precision (83.43%), mAP@0.5 (73.90%), and F1-score (72.11%). YOLOv8 obtains the highest recall (70.97%, 7.43 points above YOLOv11) and the highest mAP@0.5:0.95 (43.95%), and it is the only model whose precision and recall are balanced. The differences in mAP between the two YOLO models are below one percentage point and should not be over-interpreted given the size of the test split. The two-stage YOLOv8 + MobileNetV3 pipeline is markedly worse on every metric (mAP@0.5 of 54.88% and F1-score of 32.86%).

Fig. 14 shows the training dynamics of YOLOv8. The validation mAP rises steeply during the first 50 epochs and then plateaus, precision and recall stabilize, and the training and validation losses decrease smoothly without divergence, indicating stable convergence and no severe overfitting.

9.2 Per-class analysis

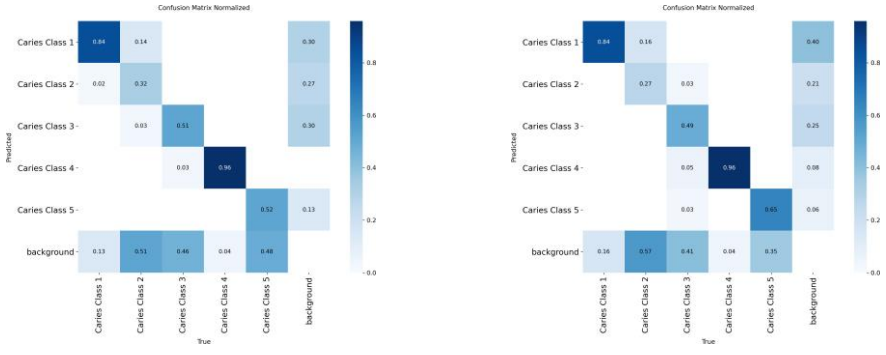
The normalized confusion matrices of the two end-to-end detectors (Fig. 15) reveal a consistent pattern across classes. Class 4 is detected almost perfectly by both models (0.96), and Class 1 is detected well (0.84). In contrast, Class 2 is the most difficult class (0.32 for YOLOv8 and 0.27 for YOLOv11), followed by Class 3 (0.51 and 0.49) and Class 5 (0.52 and 0.65). Almost all errors are missed detections, in which the lesion is predicted as background, rather than confusions between caries classes. This pattern matches the clinical nature of the classes: Class 2 lesions lie on the proximal surfaces of posterior teeth, which are partly hidden in a photograph, whereas Class 4 lesions involve the incisal edge of anterior teeth and have a conspicuous outline. The same difficulty with interproximal lesions has been reported for other photograph-based systems [18, 23]. Background regions falsely detected as lesions are spread over Classes 1–3, with Class 1 receiving the largest share for YOLOv11 (0.40), which suggests that stains in pits and fissures are occasionally mistaken for lesions.



(a) (b)

(c) (d)

Fig. 14 YOLOv8 training dynamics. (a) Validation mAP@0.5 and mAP@0.5:0.95 per epoch. (b) Validation precision and recall. (c) Training losses. (d) Validation losses



(a) (b)

Fig. 15 Normalized confusion matrices (columns: true class; rows: predicted class). (a) YOLOv8. (b) YOLOv11

9.3 Qualitative comparison

Fig. 16 shows the predictions of the three pipelines, as displayed in the app, on one test image of each class. The two end-to-end models agree with each other and with the expected class on all five examples. They occasionally produce overlapping boxes



Fig. 16 Predictions of the three pipelines on one test image per class, as rendered in the DentAI app. Each panel lists the detected lesions per class below the image. The end-to-end detectors (top and middle rows) recover the expected class, whereas the two-stage pipeline (bottom row) often assigns a different class

on the same lesion, and they differ slightly in the number of lesions counted (e.g., six versus seven Class 5 lesions in the last column). The two-stage pipeline usually localizes the same regions but frequently assigns the wrong class: Class 1 lesions become Class 4 or Class 3, Class 2 lesions become Class 3, and Class 5 lesions become Classes 3 and 4. On another Class 4 test image, the two-stage pipeline detected no lesion at all.

These errors explain the low scores of the two-stage pipeline in Table 9. Because the DentAI classes are defined by anatomical site, the correct label depends on context outside the lesion itself, such as whether the tooth is anterior or posterior and whether the lesion is proximal, occlusal, or cervical. Cropping the lesion before classification removes this context, and localization errors propagate to the classifier. End-to-end detectors, whose features cover the whole image, retain this context.

Table 10 Indicative comparison with YOLO-based studies on intraoral photographs (%)

Study	Model	Classes	P	R	mAP@0.5
Fadilah et al. [21]	YOLOv8x	7 (ICDAS)	72.6	41.1	45.8
Adnan et al. [27]	YOLOv5s	Decay	90.7	85.6	90.7
Rahman et al. [28]	YOLOv7	6 (conditions)	96.03	86.53	93.54
DentAI (ours)	YOLOv8s	5 (site)	69.10	**70.97**	72.96
DentAI (ours)	YOLO11s	5 (site)	83.43	63.54	73.90

P: precision; R: recall.

9.4 Model selection

YOLOv8 was selected for deployment. In a screening application, a missed lesion (false negative) is more harmful than a false alarm, which will be checked by a dentist through the chat or booking features. YOLOv8 offers the highest recall and the best localization under strict IoU thresholds, together with balanced precision and recall, while its mAP@0.5 is within one point of YOLOv11. Its visual outputs were also consistent across the clinical scenarios examined.

9.5 Comparison with related work

Table 10 places the results alongside other YOLO-based studies on intraoral photographs. Because the datasets, class definitions, and protocols differ, the comparison is indicative only. DentAI reaches a considerably higher recall and mAP@0.5 than the seven-class ICDAS model of Fadilah et al. [21] at similar precision. Studies with fewer or coarser classes [27, 28] report higher scores, reflecting the greater difficulty of fine-grained multi-class labeling.

10 Clinical pilot evaluation

To assess usability under real conditions, the application was piloted with a practicing dentist. The preliminary clinical usability evaluation was conducted by the clinical co-author, a practicing dentist and specialist in oral and maxillofacial radiology. The evaluation focused on the clinical relevance, functionality, and usability of the DentAI application. As the evaluation was undertaken by a member of the research team, no independent dentist was recruited for this assessment.

The dentist used the app hands-on and provided feedback, describing DentAI as follows:

- **A useful and simple dental-care tool** that can detect and categorize caries.
- **An innovative and engaging application,** highlighting in particular the live-camera AR view, whose class-colored boxes make findings both clear and engaging for patients.
- **A clinically relevant solution,** supported by appointment booking and direct communication through *Ask a Dentist*.

The dentist also suggested several improvements: guiding users to capture up to six views (front, left, right, upper and lower jaws, and close-ups of problem areas) to improve diagnostic coverage; adding short instructional videos for dental students and

junior dentists; adding interactive 3D models of teeth and jaws to explain conditions to patients; and adding localization, starting with Arabic.

11 Discussion and limitations

The results show that a small-variant, single-stage detector trained on fewer than one thousand curated photographs can localize and classify five caries classes with a precision and recall of around 70%. Such a detector can be served cheaply from a CPU cloud instance and integrated into a complete patient–dentist workflow. The comparison also shows that, for site-defined classes, the decoupled detection–classification design is inferior to end-to-end detection. Nevertheless, several limitations must be considered when interpreting the results:

- **Dataset size and provenance.** The 900 images originate from two public Roboflow collections, and their demographic and device characteristics are unknown. Performance on photographs taken by lay users with ordinary phones, which may be blurred, poorly lit, or obscured by saliva, has not yet been quantified.
- **Possible leakage between splits.** Class balancing added augmented variants of existing images to form the 900-image set that was subsequently split. Variants of the same source image may therefore appear in both the training and test splits, which could inflate the scores, especially for the heavily augmented Class 4. Future experiments should split by source image before any augmentation.
- **Reference standard.** The labels were created by one annotator in consultation with the clinical co-author, without inter-rater agreement analysis or radiographic or histological confirmation.
- **Interproximal lesions.** As with other photograph-based systems, Class 2 lesions remain hard to detect. Photographs cannot replace radiographs for hidden lesions, so DentAI should be positioned as a screening and communication aid rather than a diagnostic device.
- **Clinical evaluation.** The pilot involved a single dentist in a single clinic and was qualitative. A prospective study that compares app findings with clinical examination on consecutive patients is required.
- **Deployment.** Live detection depends on a network connection and a server round trip, and its latency was not measured. The pilot deployment served the API over plain HTTP; production use requires TLS, hardened authentication, and compliance with health-data protection regulations.

12 Conclusion and future work

This article presented DentAI, a mobile application that brings AI-based dental caries detection and AR-style visualization to standard smartphones and connects the results to dentist consultation and appointment booking. On a curated dataset of 900 intraoral images with five lesion classes, the end-to-end YOLOv8 and YOLOv11 detectors clearly outperformed a two-stage YOLOv8 + MobileNetV3 pipeline. YOLOv8 was deployed for its higher recall and balanced performance (mAP@0.5 of 72.96% and recall of 70.97%). The complete system, comprising a Flutter app, a Flask backend on

AWS EC2, and a MySQL database, was piloted in a dental clinic, where the practicing dentist confirmed its usefulness for explaining findings to patients and for accelerating consultations.

Future work will follow several directions:

- **Full oral assessment.** The detection scope will be extended to plaque, calculus, other lesions, and malocclusion, using a guided six-view capture protocol so that the system can generate a report for the whole mouth.
- **Larger and more diverse data.** Images will be collected across age groups, devices, and clinical sites, with multi-rater labels and source-level data splits.
- **Richer AR.** Detections will be mapped onto an interactive 3D dental model, and the models will be exported to on-device formats to reduce latency and remove the network dependency of live detection.
- **Education and accessibility.** VR tutorial modules will be developed for dental students, and the app will support Arabic, voice commands, and audio feedback.
- **Integration and validation.** The app will be integrated directly with insurance providers for real-time validation, and multi-clinic prospective studies will be conducted.

Declarations:

Ethics Approval and Consent to Participate

This study utilized publicly available intraoral photographic datasets obtained from the Tooth Caries Detection dataset by D7i and the DentalMate caries dataset. The images were used for the development, training, validation, and evaluation of artificial intelligence models for dental caries detection.

The datasets underwent screening, annotation, and clinical verification by the research team, including the clinical co-author, a specialist in oral and maxillofacial radiology. No patients were directly recruited for image collection, and no additional patient images or clinical records were obtained from City University Ajman Dental Clinics.

The preliminary clinical usability evaluation was conducted by the clinical co-author as part of the research team's assessment of the DentAI application. No independent dentist was recruited for this evaluation.

The study procedures were undertaken with consideration of the relevant ethical principles of the Declaration of Helsinki, particularly those concerning the responsible use of human-derived data, privacy, confidentiality, and informed consent.

Consent for Publication

Not Applicable.

Funding

This research was entirely self-funded by the authors. All expenses associated with the development, implementation, and evaluation of the Dent-AI application were covered through the authors' personal financial contributions. No external funding, institutional financial support, or research grants were received.

Availability of Data and Materials
The datasets used in this study were obtained from two publicly available sources: the Tooth Caries Detection dataset by D7i and the DentalMate caries dataset, as cited in the manuscript.
Following image screening, clinical verification, annotation, and class balancing, a curated dataset comprising 900 intraoral photographs and 1,504 annotated carious lesions was developed.
The curated dataset, annotations, and supporting data generated during this study are available from the corresponding author upon reasonable request, subject to the applicable data-sharing conditions of the original sources.

Authors' Contributions
Shereen Afifi: Conceptualization, methodology, project supervision, technical guidance, validation, and manuscript review and editing.
Samir Sameh ElBatal: Software development, artificial intelligence model implementation, data collection and preprocessing, image annotation, model training and evaluation, mobile application development, data visualization, formal analysis, and writing of the original manuscript draft.
Nader Nabil Fouad Rezallah: Clinical conceptualization, clinical methodology, expert evaluation and verification of intraoral photographs, validation of dental caries annotations and classifications, interpretation of clinical findings, preliminary clinical usability evaluation of the DentAI application, assessment of its clinical relevance, and manuscript review and editing.
All authors contributed to the research, critically reviewed the manuscript, and approved the final version for submission.

Acknowledgements
The authors acknowledge the developers and contributors of the publicly available Tooth Caries Detection dataset by D7i and the DentalMate caries dataset. These resources supported the development, training, and evaluation of the artificial intelligence models presented in this study.

Competing Interests
The authors declare that they have no competing interests.

References

[1] World Health Organization: Oral health: Key facts. Fact sheet (2023). https://www.who.int/news-room/fact-sheets/detail/oral-health

[2] Khanagar, S.B., Alfouzan, K., Awawdeh, M., Alkadi, L., Albalawi, F., Alfadley, A.: Application and performance of artificial intelligence technology in detection, diagnosis and prediction of dental caries (DC)—a systematic review. Diagnostics **12**(5), 1083 (2022) https://doi.org/10.3390/diagnostics12051083

[3] Moharrami, M., Farmer, J., Singhal, S., Watson, E., Glogauer, M., Johnson, A.E.W., Schwendicke, F., Quinonez, C.: Detecting dental caries on oral photographs using artificial intelligence: A systematic review. Oral Dis. **30**(4), 1765–1783 (2024) https://doi.org/10.1111/odi.14659

[4] Al-Khalifa, K.S., Ahmed, W.M., Azhari, A.A., Qaw, M., Alsheikh, R., Alqudaihi, F., Alfaraj, A.: The use of artificial intelligence in caries detection: A review. Bioengineering **11**(9), 936 (2024) https://doi.org/10.3390/bioengineering11090936

[5] Negi, S., Mathur, A., Tripathy, S., Mehta, V., Snigdha, N.T., Adil, A.H., Karobari, M.I.: Artificial intelligence in dental caries diagnosis and detection: An umbrella review. Clin. Exp. Dent. Res. **10**(4), 70004 (2024) https://doi.org/10.1002/cre2.70004

[6] Rezallah, N.N.F., Sherif, G., Abdelkarim, A.Z., Afifi, S.: Enhancing periodontal bone loss diagnosis through advanced AI techniques. Appl. Sci. **15**(12), 6832 (2025) https://doi.org/10.3390/app15126832

[7] Väyrynen, E., Hakola, S., Keski-Salmi, A., Jämsä, H., Vainionpää, R., Karki, S.: The use of patient-oriented mobile phone apps in oral health: Scoping review. JMIR mHealth uHealth **11**, 46143 (2023) https://doi.org/10.2196/46143

[8] Gugnani, N., Pandit, I.K., Srivastava, N., Gupta, M., Sharma, M.: International caries detection and assessment system (ICDAS): A new concept. Int. J. Clin. Pediatr. Dent. **4**(2), 93–100 (2011) https://doi.org/10.5005/jp-journals-10005-1089

[9] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, pp. 779–788 (2016)

[10] Jocher, G., Chaurasia, A., Qiu, J.: Ultralytics YOLOv8. Software (2023). https://github.com/ultralytics/ultralytics

[11] Ultralytics: YOLO11: Overview and key features. Online documentation (2024). https://docs.ultralytics.com/models/yolo11/

[12] 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: Proc. Eur. Conf. Comput. Vis. (ECCV), Zurich, Switzerland, pp. 740–755 (2014)

[13] Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q.V., Adam, H.: Searching for MobileNetV3. In: Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Seoul, South Korea, pp. 1314–1324 (2019)

[14] Google: Flutter: Build apps for any screen. Software framework (2025). https://flutter.dev

[15] Lin, P.-Y., Chen, T.-C., Lin, C.-J., Huang, C.-C., Tsai, Y.-H., Tsai, Y.-L., Wang, C.-Y.: The use of augmented reality (AR) and virtual reality (VR) in dental surgery education and practice: A narrative review. J. Dent. Sci. **19**, 91–101 (2024) https://doi.org/10.1016/j.jds.2024.10.011

[16] Mune, B., Hegde, R., Shigli, A., Gawali, P., Jadhav, G., Mohite, V.: Exploring the potential of augmented reality and virtual reality for dental education: A scoping review. Int. J. Dent. Med. Sci. Res. **6**(3), 234–237 (2024) https://doi.org/10.35629/6018-0603234237

[17] Thanh, M.T.G., Van Toan, N., Ngoc, V.T.N., Tra, N.T., Giap, C.N., Nguyen, D.M.: Deep learning application in dental caries detection using intraoral photos taken by smartphones. Appl. Sci. **12**(11), 5504 (2022) https://doi.org/10.3390/app12115504

[18] Park, E.Y., Cho, H., Kang, S., Jeong, S., Kim, E.-K.: Caries detection with tooth

surface segmentation on intraoral photographic images using deep learning. BMC Oral Health **22**, 573 (2022) https://doi.org/10.1186/s12903-022-02589-1

[19] Kühnisch, J., Meyer, O., Hesenius, M., Hickel, R., Gruhn, V.: Caries detection on intraoral images using artificial intelligence. J. Dent. Res. **101**(2), 158–165 (2022) https://doi.org/10.1177/00220345211032524

[20] Boy, A.F., Akhyar, A., Arif, T.Y., Syahrial, S.: Development of an artificial intelligence model based on MobileNetV3 for early detection of dental caries using smartphone images: A preliminary study. Adv. Sci. Technol. Res. J. **19**(4), 109–116 (2025) https://doi.org/10.12913/22998624/200308

[21] Fadilah, R.P.N., Rikmasari, R., Akbar, S., Setiawan, A.S.P.P.: Examination of new clinical dental caries in school children using real intra oral photos with artificial intelligence model YOLO-V8x. Research Square (preprint) (2024) https://doi.org/10.21203/rs.3.rs-5657504/v1

[22] Frenkel, E., Neumayr, J., Schwarzmaier, J., Kessler, A., Ammar, N., Schwendicke, F., Kühnisch, J., Dujic, H.: Caries detection and classification in photographs using an artificial intelligence-based model—an external validation study. Diagnostics **14**(20), 2281 (2024) https://doi.org/10.3390/diagnostics14202281

[23] Zhang, J.-W., Fan, J., Zhao, F.-B., Ma, B., Shen, X.-Q., Geng, Y.-M.: Diagnostic accuracy of artificial intelligence-assisted caries detection: A clinical evaluation. BMC Oral Health **24**, 1095 (2024) https://doi.org/10.1186/s12903-024-04847-w

[24] Bayati, M., Alizadeh Savareh, B., Ahmadinejad, H., Mosavat, F.: Advanced AI-driven detection of interproximal caries in bitewing radiographs using YOLOv8. Sci. Rep. **15**, 4641 (2025) https://doi.org/10.1038/s41598-024-84737-x

[25] Uddin, S.M.Z., Aslam, M.I., Moinuddin, M.: Dental caries classification using YOLOv8. J. Popul. Ther. Clin. Pharmacol. **31**(6), 2570–2586 (2024) https://doi.org/10.53555/jptcp.v31i6.6983

[26] Ding, B., Zhang, Z., Liang, Y., Wang, W., Hao, S., Meng, Z., Guan, L., Hu, Y., Guo, B., Zhao, R., Lv, Y.: Detection of dental caries in oral photographs taken by mobile phones based on the YOLOv3 algorithm. Ann. Transl. Med. **9**(21), 1622 (2021) https://doi.org/10.21037/atm-21-4805

[27] Adnan, N., Ahmed, S.M.F., Das, J.K., Aijaz, S., Sukhia, R.H., Hoodbhoy, Z., Umer, F.: Developing an AI-based application for caries index detection on intraoral photographs. Sci. Rep. **14**, 26752 (2024) https://doi.org/10.1038/s41598-024-78184-x

[28] Rahman, R.B., Tanim, S.A., Alfaz, N., Shrestha, T.E., Miah, M.S.U., Mridha, M.F.: A comprehensive dental dataset of six classes for deep learning based object

detection study. Data Brief **57**, 110970 (2024) https://doi.org/10.1016/j.dib.2024.110970

[29] Zolfaghari, M., Shirmohammadi, M., Shahhosseini, H., Mokhtaran, M., Mohebbi, S.Z.: Development and evaluation of a gamified smartphone mobile health application for oral health promotion in early childhood: A randomized controlled trial. BMC Oral Health **21**(1), 18 (2021) https://doi.org/10.1186/s12903-020-01374-2

[30] Harapan, I.K., Laut, D.M., Bidjuni, M., Ratuela, J.E., Ningrum, N., Yuliana, N.M.: Analysis of satisfaction towards the application of dental health information system. J. Dent. Hyg. Ther. **6**(1), 17–24 (2025) https://doi.org/10.36082/jdht.v6i1.2142

[31] Al-Jallad, N., Ly-Mapes, O., Hao, P., Ruan, J., Ramesh, A., Luo, J., Wu, T.T., Dye, T., Rashwan, N., Ren, J., Jang, H., Mendez, L., Alomeir, N., Bullock, S., Fiscella, K., Xiao, J.: Artificial intelligence-powered smartphone application, AICaries, improves at-home dental caries screening in children: Moderated and unmoderated usability test. PLOS Digit. Health **1**(6), 0000046 (2022) https://doi.org/10.1371/journal.pdig.0000046

[32] Al Kheraif, A.A., Adam, T.R., Wasi, A., Alhassoun, R.K., Haddadi, R.M., Alnamlah, M.: Impact of virtual reality intervention on anxiety and level of cooperation in children and adolescents with autism spectrum disorder during the dental examination. J. Clin. Med. **13**(20), 6093 (2024) https://doi.org/10.3390/jcm13206093

[33] Ledwoń, A., Dębski, P., Jędrusik, P., Mielcarska, S., Misiolek, H., Meisner, M., Łopacińska, M., Skucha-Nowak, M.: An attempt to use virtual reality as a tool to reduce patient anxiety during dental treatment. J. Clin. Med. **13**(22), 6832 (2024) https://doi.org/10.3390/jcm13226832

[34] Wu, W., Le May, S., Hung, N., Fortin, O., Genest, C., Francoeur, M., Guingo, E., St-Arneault, K., Sylfra, A., Vu, A.K., Carmel, J., Lessard, L., Cara-Slavich, S., De Koven, K., Paquette, J., Hoffman, H., Asselin, M.-E.: Effects of a virtual reality game on children's anxiety during dental procedures (VR-TOOTH): Protocol for a pilot randomized controlled trial. JMIR Res. Protoc. **12**, 49956 (2023) https://doi.org/10.2196/49956

[35] Alabduljabbar, R., Almutawa, M., Alkathiri, R., Alqahtani, A., Alshamlan, H.: An interactive augmented and virtual reality system for managing dental anxiety among young patients: A pilot study. Appl. Sci. **13**(9), 5603 (2023) https://doi.org/10.3390/app13095603

[36] D7i: Tooth caries detection dataset. Roboflow Universe (2024). https://universe.roboflow.com/d7i/tooth-caries-test-21

[37] Dental Data2: DentalMate caries detection dataset. Roboflow Universe (2025).

https://universe.roboflow.com/dental-data2/dental-dataset-87ne6

[38] Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A.: Albumentations: Fast and flexible image augmentations. Information **11**(2), 125 (2020) https://doi.org/10.3390/info11020125

[39] Solawetz, J., Francesco: What is YOLOv8? A complete guide. Roboflow Blog (2024). https://blog.roboflow.com/what-is-yolov8/

[40] Rao, N.: YOLOv11 explained: Next-level object detection with enhanced speed and accuracy. Medium (2024). https://medium.com/@nikhil-rao-20/yolov11-explained-next-level-object-detection-with-enhanced-speed-and-accuracy-2dbe2d376f71

[41] Polukhin, A.: MobileNet architectures. Medium (2022). https://medium.com/@pandrii000/mobilenet-architectures-17fe7406d794