AI-Assisted Detection of Dental Caries Using Radiographs for Operative Dentistry Applications

Keywords
Dental caries; Radiographs; Operative dentistry; Artificial intelligence; YOLOv8; Object detection

Authors

Alaa Aljohani*
*Division of Operative Dentistry, General Dentistry Program, Batterjee Medical College, Jeddah, Saudi Arabia alaa.aljohani@bmc.edu.sa

Received-13-05-2026
Revised-20-06-2026
Accepted-23-06-2026
Doi: 10.1922/ejprd.v34i4s.1472

Abstract
Dental caries is still a leading cause of restorative treatment need and radiographs play a pivotal role in the decision-making process for operative dentistry especially when the caries are subtle, anatomically overlapped or proximal. In this study, a preprocessing-integrated detection framework (Caries-YOLOv8-CE) was developed and evaluated, which combined the contrast-limited adaptive histogram equalization with the YOLOv8s-based object detection for radiographic localization of caries. A radiographic dataset of 4,658 images and 6,566 caries annotations was divided into three subsets: training, validation and independent test. The proposed framework is compared with the raw baseline (YOLOv8s) and Faster R-CNN ResNet50-FPN using precision, recall, F1-score, mAP@0.5, mAP@0.5:0.95, qualitative prediction review, and error analysis. Caries-YOLOv8-CE achieved the highest default precision (0.879) and F1-score (0.841), with competitive mAP@0.5 (0.881) and mAP@0.5:0.95 (0.526). The box level error analysis yielded 549 true positive, 171 false positive and 97 false negative results. These results suggest that CLAHE enhanced YOLO based detection could be a valuable assistive method for localizing caries in radiographs, but external validation and clinical reader studies are needed before it can be practically implemented.

1. Introduction

A caries is a progressive oral health burden that is not detected and managed at an appropriate stage can lead to cavitation, pulpal involvement, pain and restorative intervention, and is caused by a mineral imbalance that occurs as a result of the presence of a biofilm [1]. Caries is a public health and service delivery challenge, as well as a biological problem, and is a high prevalence disease worldwide with populations of people at risk of caries being untreated across all age groups and health systems [2]. A diagnostic problem is particularly significant in operative dentistry, as the choice of treatment (monitor, remineralize, seal, restore or replace) relies on accurate lesion detection, establishment and monitoring of the size and activity of the lesion. Radiographs offer a lot of information which can't be gained from visual examination alone, especially for lesions in the proximal areas, for recurrent caries next to restoration margins, and for lesions that overlap with other structures and may not be accessible clinically. Operative intervention should also be guided by a conservative understanding of caries process rather than by radiographic appearance alone, which can result in unnecessary removal of tooth structure and/or delayed diagnosis of caries allowing lesion progression [3]. The tension translates to a demand for reproducible tools to support radiography that will enhance localization, but not overestimate diagnostic certainty.

The increasing relevance of AI to dental image analysis stems from the ability of deep neural networks to learn patterns in images that are correlated with anatomy and disease, from annotated examples [4]. AI systems are

most
beneficial with respect to caries detection when the algorithms provide interpretable localization of suspicious areas, not just image-level classification, for example, to enable efficient review of the radiographs and correlate the algorithmic findings with the operative site. AI models, however, must be assessed with clinical relevance and prudence given the factors that influence the detection of caries on radiographs, such as stage and type of caries, contrast, projection geometry, restoration artifacts, and the variability of these annotations.

AI-aided radiographic caries detection has recently been performed using combined studies, showing promising diagnostic capabilities, but also attributing to a certain degree of heterogeneity in terms of dataset, imaging modalities, model architectures, and reporting [5]. This is further supported by a wide range of evidence from across the field of caries imaging modalities, where AI can aid in the detection of caries, but there remains a need for the field to be further validated and for there to be greater transparency in reporting and a careful distinction between decision support and autonomous diagnosis [6]. The purpose of the present study was to focus on a reproducible object detection workflow for locating radiographic caries instead of on making a clinical diagnosis claim, due to these considerations. This study had the following objectives:

- To develop a CLAHE-enhanced YOLOv8-based framework for radiographic dental caries detection.
- To compare the proposed framework with YOLOv8s Raw Baseline and Faster R-CNN ResNet50-FPN using an independent test set.
- To evaluate the framework using quantitative metrics, qualitative prediction visualization, and box-level and image-level error analysis.

2. Literature Review

In the field of medical image analysis, deep learning has been a game-changer in recent years, as it has been shown to learn features directly from medical images, and has been applied to various tasks such as classification, detection, segmentation, and registration [7]. The key benefit of multilayer neural architectures is that they can learn hierarchical visual representations, but this benefit is highly dependent on the quality of the data, consistency of the data annotations, and external validity [8]. The factors are especially significant in Dentistry where radiography is projected in different ways, exposed for different times, use different sensors, record different anatomy, and different restorative materials are used.

In the field of dentistry, there is a body of evidence that suggests AI has been explored in a variety of diagnostic and workflow-related applications, but several studies have been retrospective in design, small in sample size, and fail to report on the methods and rigor used to validate the applications. For the specific task of caries detection, early convolutional methods demonstrated that neural networks can learn to detect carious patterns on dental radiographs, but also indicated that the success of neural network performance relies on the
a
visibility of carious lesions, the type of image and the availability of experts to provide annotations [9]. Subsequent study on bitewing radiographs showed that deep learning was able to identify radiographically different extensions of the lesions, although subtle lesions and borderline annotations are more difficult, indicating that the visual performance metrics are not enough to interpret the lesions clinically [10].

However, object-detection methods are of special interest to operative dentistry as they provide spatial predictions which can be checked by a clinician. Object-detection tasks based on bitewings have recently been undertaken with models like the YOLO variants demonstrating the capability of achieving higher average precision and lower rates of false-negatives than some clinician baselines, while simultaneously the presence of artifacts and annotations disagreement still precludes the best performance [11]. Two-stage methods (e.g. Faster R-CNN) are still significant baseline models since they employ RPNs and historically have achieved good localization performance in broad object detection [12]. In the case of small lesions tasks, however, two-stage methods can lead to high number of candidate detections, and hence precision-recall balance becomes a key evaluation concern.

The work for one-stage detection has focused on loss functions and architectures with the aim of enhancing the training process in the presence of foreground-background imbalance [13]. YOLO-family models are now popular due to their efficiency at inference and high detection accuracy, as evidenced by the continuous development of the model from earlier real-time based models to the more recent models such as YOLOv8 [14]. While YOLOv8 implementations offer usable tools for training, validation, inference, and metric reporting, the applications of these models to biomedical tasks remain to be validated, with transparent and detailed reporting of the datasets, and clinically cautious interpretation.

While vital, pre-processing is also a lesser studied aspect of dental AI workflows. Local contrast enhancement technique such as adaptive and contrast-limited histogram (CL) can enhance local contrast in the radiographic image, and recently, the work of dental X-ray has employed adaptive enhancement along CNN-based analysis to enhance image visual discriminability [15]. However, not all metrics will benefit from preprocessing, and it would be possible to improve the visibility of the lesions for some predictions, but worsen the distributions of textures which influence the confidence calibration or localization. Thus, a preprocessing-integrated framework should be judged with respect to a number of indicative metrics such as mAP, precision and recall, some qualitative examples, and error patterns. Additionally, evaluation criteria for object detection help support the use of multiple metrics, as the different metrics, such as mAP, precision, recall, F1-score, and matching thresholds, represent various aspects of detection behavior [16]. This study fills these gaps by applying CLAHE enhanced pre-processing, the

YOLOv8s detection model, baseline comparison, independent test-set evaluation, qualitative assessment, and box-level and image-level error analysis.

3. Methodology

3.1 Study Design and Dataset Access

This study was designed to be a retrospective computation study on AI dental caries localization on radiography images. The purpose was to assess if the preprocessing-integrated object-detection framework could be able to localize suspected carious regions in a reproducible experimental set-up, rather than to replace clinical diagnosis. The X-rays and X-ray annotation files were taken from a publicly available secondary dataset, which is stored in an online Google Drive repository [17]. No primary data collection or direct patient contact or prospective clinical intervention was done. The study thus consisted of a secondary computation analysis of available radiograph data to detect caries using AI.

3.2 Dataset Composition and Label Integrity Checking

The final detection dataset comprised of 4,658 radiographic images and 6,566 caries bounding-box annotations. The dataset was split into training, validation and independent test sets to separate the model optimization and the test of the model performance. There was one class, caries, that was used for the detection task with class ID 0. All the annotations were saved in YOLO bounding-box format, in which the class ID and normalized coordinates of the bounding-box were written in each label line.

Table 1. Dataset Split and Annotation Summary

Split	Images	Caries Annotations	Purpose
Train	3260	4594	Model training
Validation	931	1326	Model selection and threshold tuning
Test	467	646	Independent final evaluation
Total	4658	6566	Complete detection dataset

Table 1 shows the data partition employed in this study. The independent test subset comprised 467 radiographs and 646 caries annotations to assess the final performance on data not seen during training or validation. Label Integrity Check was done prior to model training to check image/labels correspondence, missing label files, label files with no corresponding images, empty label files, and to check invalid label lines and class consistency. After data collection, the data was analyzed for label completeness and the consistency of the annotations before being used for model building as seen in table 2.

Table 2. Dataset Quality and YOLO Label Integrity Summary

Split	Images	Label Files
Train	3260	3260
Validation	931	931
Test	467	467
Overall	4658	4658

As seen in Table 2, all of the images were found to have associated label files and there were no missing labels, unmatched labels, empty label files or invalid label lines. Only one class number which is 0 (caries) was obtained. These checks aid in the data reliability prior to downstream model development.

3.3 Proposed Caries-YOLOv8-CE Workflow

The framework, Caries-YOLOv8-CE, is proposed to fuse the radiographic images with CLAHE and a YOLOv8s object detector for localization of caries in the radiographs. CLAHE has been applied to improve the local contrast of each image without altering its spatial geometry, and the original YOLO bounding-box labels were still valid after the preprocessing.

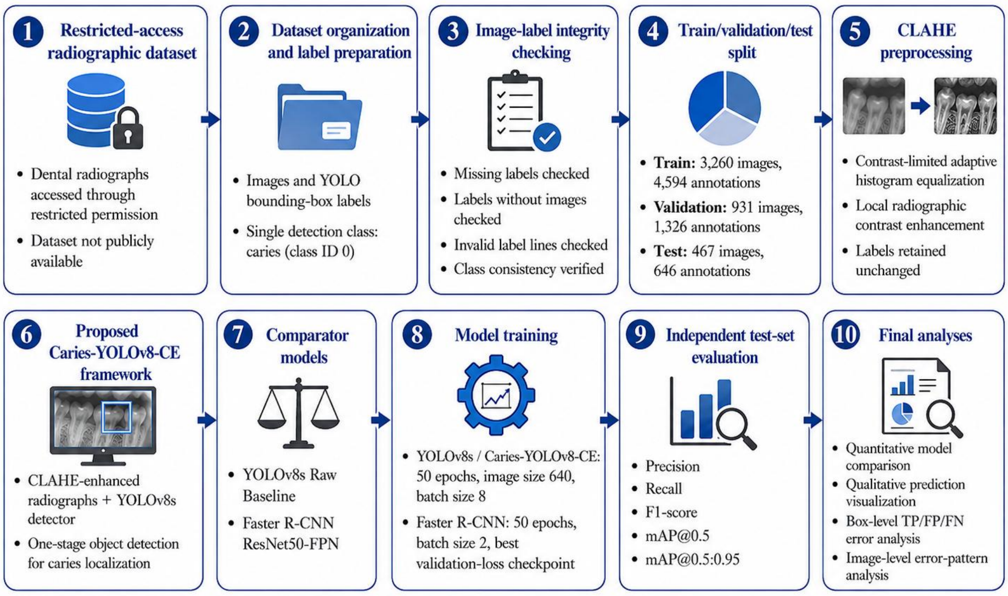

Figure 1. Proposed Caries-YOLOv8-CE Workflow for AI-Assisted Dental Caries Detection.

Figure 1 shows the entire computational process from collecting the X-rays with restricted access to sorting the images, creating the YOLO labels, checking the integrity of the images and labels, splitting the training, validation, and test sets, applying CLAHE preprocessing to the images, training with YOLOv8s, testing on an independent test set, and performing quantitative, qualitative, and error analysis. The contributions of this study are summarized in the figure and it is confirmed that it is a reproducible detection framework that integrates preprocessing.

3.4 Comparator Models and Training Configuration

Two baseline frameworks were compared to the proposed framework: YOLOv8s Raw Baseline and Faster R-CNN ResNet50-FPN Framework. Original radiographs were used as the baseline model for this YOLOv8s Raw baseline model and the detector was directly compared with one-stage detectors. Faster R-CNN ResNet50-FPN was also a two-stage CNN based comparison model that was trained with original radiographs. All YOLO-based models were trained at 50 epochs, an image size of 640 and a batch size of 8. Faster R-CNN trained for 50 epochs with batch size 2 and the best validation-loss checkpoint was used for evaluation at epoch 5.

3.5 Evaluation Strategy and Ethical Considerations

The precision, recall, F1-score, mAP@0.5, and mAP@0.5:0.95 were used for official test evaluation. The Caries-YOLOv8-CE further underwent box-level error analysis by comparing the predicted boxes with the ground truth boxes with a confidence threshold of 0.25 and an IoU threshold of 0.50. This analysis produced the numbers of true positive, false positive, and false negative and was not taken as official mAP based analysis. Image-level error categories were also summarized to differentiate the cases with only false-positive errors, only missed-caries errors, and mixed errors. Ethical approval / permission details should be reported as per the dataset access agreement as the dataset was restricted access and not publicly available.

4. Results

4.1 Radiographic Enhancement and Ground-Truth Annotation Visualization

Representative annotated radiographs are presented to show the spatial target used for object-detection training and evaluation.

a

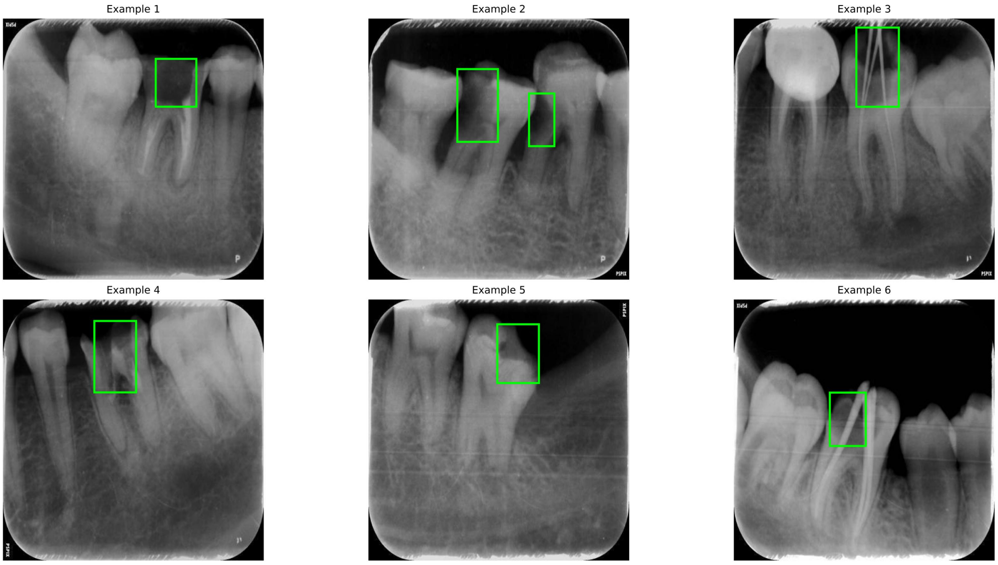

Figure 2. Representative Radiographs with Ground-Truth Dental Caries Annotations

Figure 2 presents representative radiographs with ground-truth bounding boxes marking carious regions. These annotations define the localization target of the detection task and demonstrate that the model was trained to identify spatial caries regions rather than assign only image-level labels. The examples also show variation in lesion location, image contrast, and tooth anatomy, which are important challenges in radiographic caries localization. The visual effect of CLAHE preprocessing is shown in **Figure 3** to demonstrate how the radiographs were modified before training the proposed framework.

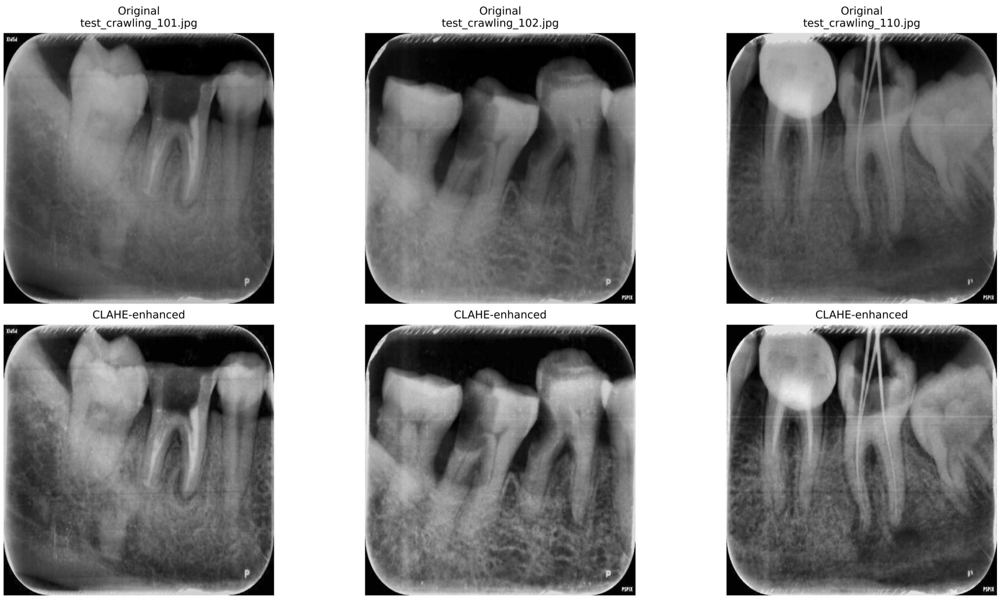

Figure 3. Original versus CLAHE-Enhanced Radiographs

Figure 3 compares original radiographs with their CLAHE-enhanced counterparts. The improved images retain the spatial structure of the radiographs and have higher local contrast. This is crucial, as the proposed Caries-YOLOv8-CE framework would perform CLAHE as a preprocessing step prior to the detection by YOLOv8s, and the original bounding-box label would be retained, since CLAHE is a preprocessing method that alters the intensity of the image while not altering the spatial geometry of the original image.

4.2 Training Behavior and Official Performance of Caries-YOLOv8-CE

The training dynamics of the proposed framework are presented to assess whether model optimization proceeded consistently across epochs.

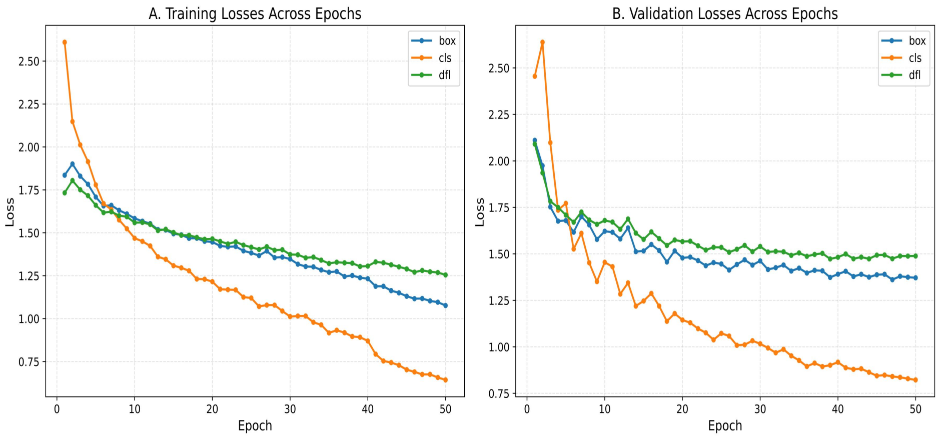

Figure 4. Proposed Caries-YOLOv8-CE Training and Validation Losses

The training and validation loss curves of the proposed Caries-YOLOv8-CE framework are shown in figure 4. The loss curves are downward trending over the epochs, suggesting that the YOLOv8s detector has been optimized using the CLAHE processed radiographs. The decrease in training and validation losses indicates that the model was successful in learning relevant patterns for detection without noticeable convergence failure. The official test-set metric profile and summary of the proposed framework are summarized in Figure 5.

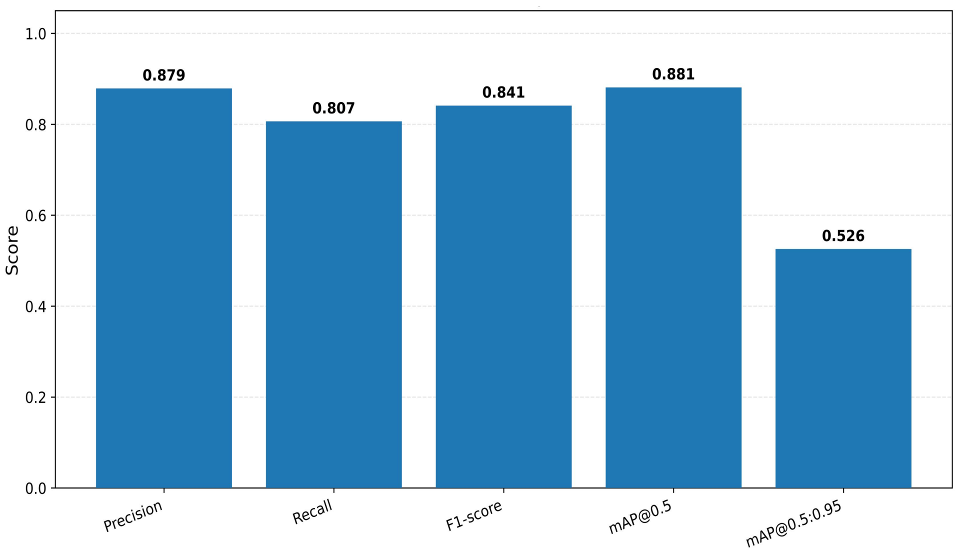

Figure 5. Official Test Metric Profile of Proposed Caries-YOLOv8-CE

Figure 5 presents the official test-set performance of Caries-YOLOv8-CE, with precision of 0.879, recall of 0.807, F1-score of 0.841, mAP@0.5 of 0.881, and mAP@0.5:0.95 of 0.526. The results indicate that the proposed framework achieved strong precision-oriented detection while maintaining competitive localization performance across IoU thresholds. The training and evaluation configuration of the proposed and comparator models is presented in **Table 3**.

Table 3. Model Training and Evaluation Configuration

Model	**Detector Type**	**Input**	**Epochs**	**Image Size**
YOLOv8s Raw Baseline	One-stage detector	Original radiographs	50	640
Caries-YOLOv8-CE	One-stage detector with CLAHE preprocessing	CLAHE-enhanced radiographs	50	640
Faster R-CNN ResNet50-FPN	Two-stage CNN detector	Original radiographs	50	Original image tensor

Table 3 shows that the YOLOv8s Raw Baseline and the proposed Caries-YOLOv8-CE framework were a

trained under comparable YOLO-based settings, whereas Faster R-CNN used a configuration appropriate for its two-stage detector implementation. The table supports reproducibility and clarifies that the proposed framework differs from the raw YOLOv8s baseline primarily through CLAHE-enhanced preprocessing rather than through a newly designed detector architecture.

4.3 Comparative Model Performance

The official independent test-set comparison is presented in **Table 4** to evaluate whether the proposed preprocessing-integrated framework improved performance relative to the baseline detectors.

Table 4. Final Model Comparison on Independent Test Set

Model	Detector Type	Input	Precision	Recall	F1-Score	mAP@0.5	mAP@0.5:0.95
YOLOv8s Raw Baseline	One-stage detector	Original radiographs	0.866	0.813	0.838	0.889	0.528
Faster R-CNN ResNet50-FPN	Two-stage CNN detector	Original radiographs	0.540	0.845	0.659	0.780	0.417
Caries-YOLOv8-CE	One-stage detector with CLAHE preprocessing	CLAHE-enhanced radiographs	0.879	0.807	0.841	0.881	0.526

Table 4 demonstrates that Caries-YOLOv8-CE achieved the highest default precision and F1-score among the evaluated models. The YOLOv8s Raw Baseline retained marginally higher mAP@0.5 and mAP@0.5:0.95 values, indicating that CLAHE preprocessing improved precision-oriented behavior and F1-score but did not produce a universal gain across all localization metrics. Faster R-CNN achieved relatively high recall but substantially lower precision, suggesting a tendency toward excessive false-positive detections in this one-class radiographic detection task. The main comparative advantage of the proposed framework is visually highlighted in **Figure 6**.

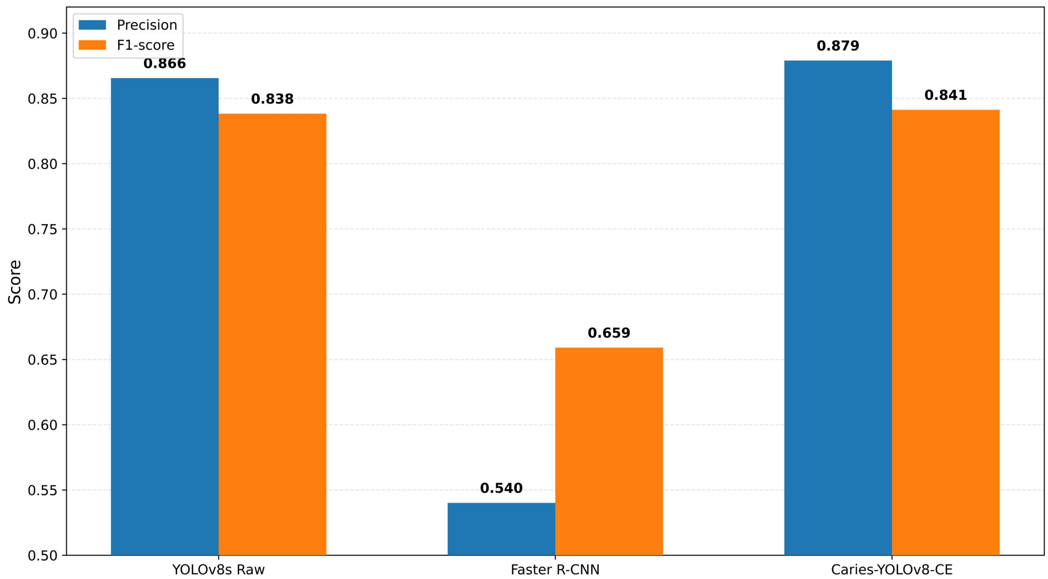

Figure 6. Proposed Caries-YOLOv8-CE Advantage in Precision and F1-Score

Figure 6 compares precision and F1-score across the evaluated models. The figure highlights that Caries-YOLOv8-CE achieved the strongest default precision and F1-score, which are important for reducing unnecessary false-positive emphasis while maintaining balanced detection behavior. The figure should be interpreted alongside **Table 4**, because the raw YOLOv8s baseline retained slightly higher mAP values.

4.4 Qualitative Prediction and Error Analysis

Qualitative prediction examples are introduced to examine how the proposed framework localizes carious regions on representative test radiographs.

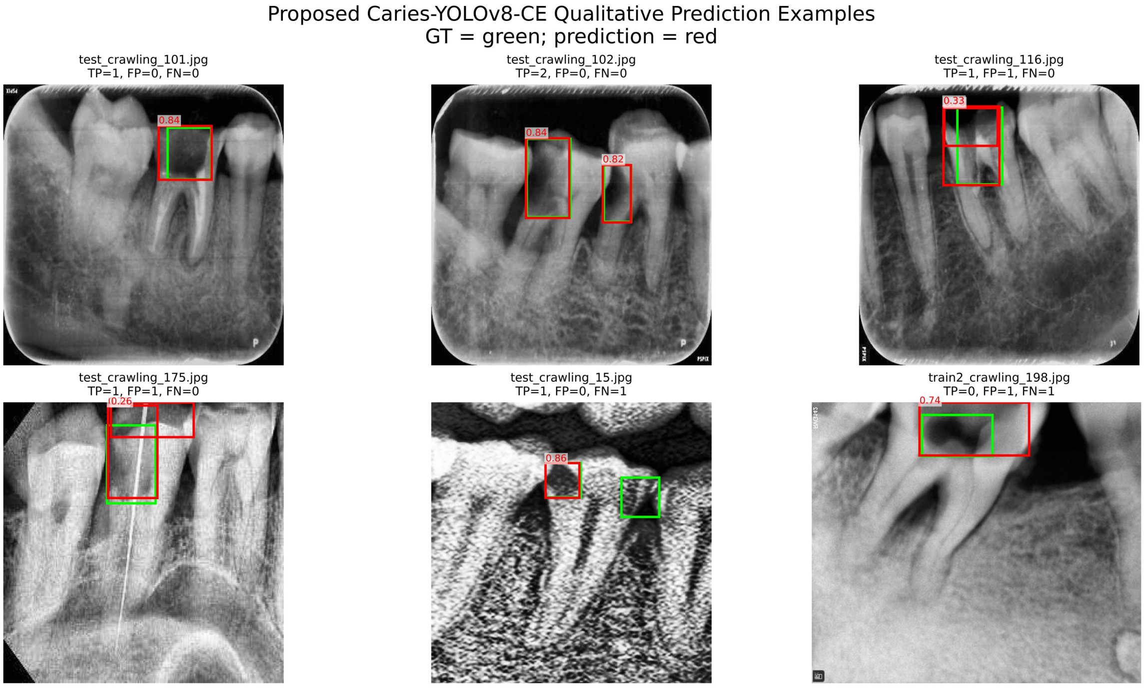

Figure 7. Proposed Caries-YOLOv8-CE Qualitative Prediction Examples

Figure 7 shows the qualitative results of the proposed Caries-YOLOv8-CE framework for some representative images. Correct detection, false positive prediction, missed lesion and mixed-error cases can be visually inspected using ground truth boxes and the predicted boxes. This qualitative review is vital as, by itself, aggregate data do not necessarily reflect if model errors are present in clinically significant or in anatomically difficult areas. To quantify the number of true and false positives and negatives, explicit prediction-to-ground-truth matching is done with the box-level detection-error profile shown in Figure 8.

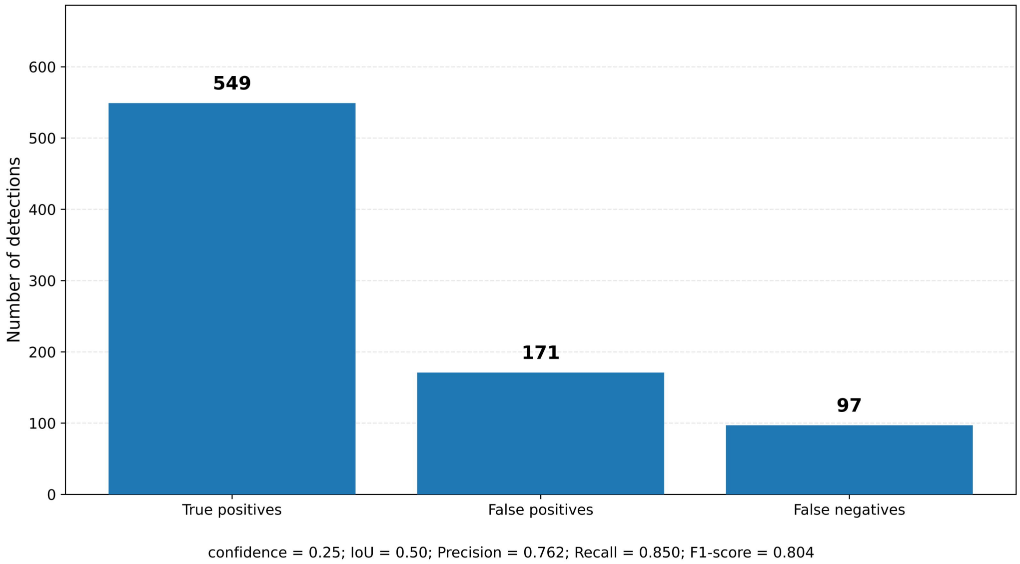

Figure 8. Code-Computed Caries-YOLOv8-CE Test-Set Error Summary

Figure 8 shows the code-computed box-level error summary for Caries-YOLOv8-CE at confidence threshold 0.25 and IoU threshold 0.50. The analysis identified 549 true positives, 171 false positives, and 97 false negatives, corresponding to precision of 0.762, recall of 0.850, and F1-score of 0.804. This analysis is separate from the official mAP-based evaluation and provides a more directly interpretable summary of matched detection outcomes. Image-level error categories are summarized in **Table 5** to complement the box-level counts shown in **Figure 8**.

Table 5. Image-Level Caries-YOLOv8-CE Error-Pattern Summary

Error Category	Images	Total GT Boxes	Total Predicted Boxes	Total TP	Total FP	Total FN	Percentage of Test Images
Correct detection without FP/FN	309	375	375	375	0	0	66.17
False-positive only	78	109	208	109	99	0	16.70
Missed-caries only	25	53	25	25	0	28	5.35
Mixed FP and FN errors	55	109	112	40	72	69	11.78

As can be seen from table 5, 309 images out of 467 test images, 66.17% of test images were detected correctly without false positive or negative. All the rest of the images were of cases with only false positive, with only missed caries and with mixed error cases. This image-level analysis, in addition to the one shown in Figure 8, provides a distribution of detection errors for test images; it is not a collection of absolute box-level counts. The qualitative examples, box-level analysis, and image-level summary offer a more comprehensive understanding of the practical detection behavior of the proposed model.

5. Discussion

This study was aimed at creating and testing the performance of an AI-supported framework named Caries-YOLOv8-CE for the localization of dental caries in radiograms. The results show that the CLAHE preprocessing and YOLOv8s detection exhibited better default precision and F1 score compared to the baseline model (raw YOLOv8s), and the baseline model showed a slight increase in mAP. This is consistent with evidence more generally that human-level performance is possible in specific imaging tasks with deep learning systems, and that performance is crucially dependent on the task under consideration, the quality of the dataset and the design of the validation process [18]. Thus the most important result of this study is not that the proposed framework is superior in every aspect, but the more detailed and nuanced detection profile of the proposed framework under the official test setting.

Comparisons with Faster R-CNN are also interesting. The higher recall of the faster R-CNN indicated that there were a higher number of false-positive detections in this one-class radiographic setting, as opposed to the two-stage detector. The clinical importance of this behavior is that false positives would cause an increase in the amount of files that must be reviewed and false negatives would lead to possibly missing carious areas when formulating a plan for operation. Evidence reviews of AI imaging studies highlight that these compromises should be clearly disclosed, and not encapsulated in an overall performance metric [19]. In this study, both official metrics and explicit TP/FP/FN analysis are used, which can be used to distinguish the localization performance from the actual error patterns.

The framework with CLAHE seems to have had an effect of moving the model towards precision oriented detection. This is reasonable as local contrast enhancement could enhance the contrast between radiographic intensity changes of caries, especially if the caries is subtle or within a complex tooth structure. But as can be seen, the raw YOLOv8s baseline has a slightly higher mAP, indicating that in some cases the preprocessing does not have a positive effect on the evaluation. In medical AI, it is crucial to not only have interpretability but also to be reproducible, as improvements in one metric can come at the cost of other metrics like calibration, generalization, or failure-mode behavior [20]. Thus, the results of the current study should be taken as an indicator of the usefulness of CLAHE in this workflow, rather than an indicator of the superiority of contrast enhancement for all tasks in dental radiographic detection.

The clinical implication of the suggested framework for operative dentistry is that it can be applied to to identify areas of possible caries during a radiographic review. This localization can draw the eye of the clinician to those areas which need further investigation, especially in proximal and visually hard to see areas. However, AI-generated content should always be secondary to clinical assessment, patient history, visual and tactile examination of the lesions, and consideration of their activity, so as not to replace the process of making decisions to restore. The challenge of translating the AI from retrospective experiments to clinical practice has to be addressed by considering workflow integration, building trust among physicians and other stakeholders, shifting of datasets, and consequences of errors [21]. Thus, the presented model is not meant to be a diagnostic device that can be used in the clinic but rather a prototype that can be used as a tool for decision making.

There are also limitations to this study which impact upon interpretations. Only restricted access to the dataset was allowed and it was not publicly available, thus the possibility of independent replication by external groups was limited. The study relied on a single source of the data set and did not assess any external multicenter validation and only evaluated one class of detection. This model provided localization of the lesions through bounding boxes instead of pixel-

a

wise segmentation; therefore, regions of suspicion were estimated instead of boundaries of the lesions. The prospective clinical validation study was not done, nor was the dentist-AI reader study or evaluation of the impact on treatment decision. Recommendations for reporting AI diagnostic accuracy studies highlight the importance of reporting these limitations as the performance of the AI in the study may not apply to the real-world setting [22].

Further studies are needed that involve external multicenter validation, prospective clinical evaluation, dentist-AI reader studies, lesion-severity classification, a comparison of different types of radiographs, explainable AI outputs, and an integration into operative dentistry workflows. Further research should also focus on thresholds optimization per clinical use case, since screening use cases might be more interested in sensitivity, while the operative decision support use cases might want to have more precise thresholds. In line with the general principles of AI translation in healthcare, the development of these algorithms must also consider factors such as reproducibility, interpretability, fairness, and assessment in actual clinical decision-making processes [23].

6. Conclusion

This study introduced and tested Caries-YOLOv8-CE, a CLAHE enhanced YOLOv8s framework for AI-assisted radiographic localization of dental caries, in operative dentistry. The framework achieved the greatest default precision and F1 score compared to the other models evaluated, and competitive mAP results when compared to the raw YOLOv8s baseline model using the restricted-access dataset consisting of 4658 radiographs and 6566 annotations. These results indicate that the use of local contrast enhancement in combination with object detection using YOLO can enhance the accuracy of the detection of suspected carious areas on dental X-rays. Thus, the framework can be used as a useful tool for radiographic review and operative planning for treatment, especially in cases where there are subtle or anatomically overlapping lesions. The study is, however, limited by the fact that it uses a single source dataset, does not undergo external validation, is a one class bounding-box design and that the prospective clinician-reader evaluation is missing. The framework should be validated in independent institutions, imaging protocols, and in different severities of lesions and in real clinical workflows in the future.

Data Availability

This secondary set of radiographic data is openly available in the Google Drive repository referenced in this manuscript [17]. The data set includes the radiographic images in a dataset and the annotation files for the training, validation and independent testing of the models. The corresponding author may make available processed experimental outputs, such as tables, outputs from trained models, and manuscript figures upon reasonable request. A Google Drive location that was used in this study was a https://drive.google.com/drive/folders/14Km_y3JvuJesQtbDPFqSZhu2Zgv7xzS9?usp=sharing.

References

[1] N. B. Pitts *et al.*, “Dental caries,” *Nat Rev Dis Primers*, vol. 3, no. 1, p. 17030, May 2017, doi: 10.1038/nrdp.2017.30.

[2] W. H. Organization, *Global oral health status report: towards universal health coverage for oral health by 2030*. World Health Organization, 2022.

[3] F. Schwendicke *et al.*, “Managing Carious Lesions: Consensus Recommendations on Carious Tissue Removal,” *Adv Dent Res.*, vol. 28, no. 2, pp. 58–67, May 2016, doi: 10.1177/0022034516639271.

[4] F. Schwendicke, W. Samek, and J. Krois, “Artificial Intelligence in Dentistry: Chances and Challenges,” *J Dent Res*, vol. 99, no. 7, pp. 769–774, Jul. 2020, doi: 10.1177/0022034520915714.

[5] D. Albano *et al.*, “Artificial intelligence for radiographic imaging detection of caries lesions: a systematic review,” *BMC Oral Health*, vol. 24, no. 1, p. 274, Feb. 2024, doi: 10.1186/s12903-024-04046-7.

[6] S. Arzani *et al.*, “Examining the diagnostic accuracy of artificial intelligence for detecting dental caries across a range of imaging modalities: An umbrella review with meta-analysis,” *PLOS ONE*, vol. 20, no. 8, p. e0329986, Aug. 2025, doi: 10.1371/journal.pone.0329986.

[7] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[8] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[9] J.-H. Lee, D.-H. Kim, S.-N. Jeong, and S.-H. Choi, “Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm,” *Journal of Dentistry*, vol. 77, pp. 106–111, Oct. 2018, doi: 10.1016/j.jdent.2018.07.015.

[10] A. G. Cantu *et al.*, “Detecting caries lesions of different radiographic extension on bitewings using deep learning,” *Journal of Dentistry*, vol. 100, p. 103425, Sep. 2020, doi: 10.1016/j.jdent.2020.103425.

[11] J. Pérez de Frutos *et al.*, “AI-Dentify: deep learning for proximal caries detection on bitewing x-ray - HUNT4 Oral Health Study,” *BMC Oral Health*, vol. 24, no. 1, p. 344, Mar. 2024, doi: 10.1186/s12903-024-04120-0.

[12] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[13] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal Loss for Dense Object Detection,” presented at the Proceedings of the IEEE International Conference on Computer Vision,

ejprd.org- Published by Riset Publishing Services LLC.

2017, pp. 2980–2988. Accessed: Jul. 09, 2026. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2017/html/Lin_Focal_Loss_for_ICCV_2017_paper.html

[14] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680–1716, Dec. 2023, doi: 10.3390/make5040083.

[15] Y. Wang, L. Sun, Y. Zhang, D. Lv, Z. Li, and W. Qi, "An Adaptive Enhancement Based Hybrid CNN Model for Digital Dental X-ray Positions Classification," May 01, 2020, *arXiv*: arXiv:2005.01509. doi: 10.48550/arXiv.2005.01509.

[16] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Jul. 2020, pp. 237–242. doi: 10.1109/IWSSIP48289.2020.9145130.

[17] "Dental caries radiographic dataset for AI-assisted detection," Google Drive dataset repository, 2026. [Online]. Available: https://drive.google.com/drive/folders/14Km_y3JvuJesQtbDPFqSZhu2Zgv7xzS9

[18] X. Liu *et al.*, "A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis," *The Lancet Digital Health*, vol. 1, no. 6, pp. e271–e297, Oct. 2019, doi: 10.1016/S2589-7500(19)30123-2.

[17] M. Nagendran *et al.*, "Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies," *BMJ*, p. m689, Mar. 2020, doi: 10.1136/bmj.m689.

[20] M. Ghassemi, L. Oakden-Rayner, and A. L. Beam, "The false hope of current approaches to explainable artificial intelligence in health care," *The Lancet Digital Health*, vol. 3, no. 11, pp. e745–e750, Nov. 2021, doi: 10.1016/S2589-7500(21)00208-9.

[21] C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, "Key challenges for delivering clinical impact with artificial intelligence," *BMC Med*, vol. 17, no. 1, p. 195, Dec. 2019, doi: 10.1186/s12916-019-1426-2.

[22] V. Sounderajah *et al.*, "Developing specific reporting guidelines for diagnostic accuracy studies assessing AI interventions: The STARD-AI Steering Group," *Nat Med*, vol. 26, no. 6, pp. 807–808, Jun. 2020, doi: 10.1038/s41591-020-0941-1.

[23] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, Jan. 2019, doi: 10.1038/s41591-018-0316-z.