**ISSN 1840-4855**
**e-ISSN 2233-0046**

**Original scientific article**
**http://dx.doi.org/10.70102/afts.2025.1834.777**

# PROXIMAL CARIES DETECTION USING YOLOV11 IN NEAR-INFRARED LIGHT TRANSILLUMINATION IMAGES

Asma Alatawi[1*], Wdaee Alhalabi[2], Hani Nassar[3], Arwa Basbrain[4], Hattan Jamalellail[5], Mohammed Alsadat[6]

*[1*] PhD Student, Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia. e-mail: asmasuliaman@outlook.com, orcid: https://orcid.org/0000-0001-9198-0849*
*[2]Professor, Immersive virtual reality research group, Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia. e-mail: wsalhalabi@kau.edu.sa, orcid: https://orcid.org/0000-0002-4505-7268*
*[3]Professor, consultant, and chairman, Department of Restorative Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia. e-mail: hnassar@kau.edu.sa, orcid: https://orcid.org/0000-0002-6241-1118*
*[4]Assistant Professor, Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. e-mail: abasabreen@kau.edu.sa, orcid: https://orcid.org/0000-0002-0752-366X*
*[5]General Practitioner, Ministry of National Guard Health Affairs, King Abdulaziz Medical City, Jeddah, Saudi Arabia. e-mail: hjamalellail@gmail.com, orcid: https://orcid.org/0000-0002-5706-1125*
*[6]General Practitioner, University Dental Hospital, King Abdulaziz University, Jeddah, Saudi Arabia. e-mail: malsadat@kau.edu.sa, orcid: https://orcid.org/0000-0001-8223-6583*

***Received: September 20, 2025; Revised: October 29, 2025; Accepted: December 02, 2025; Published: December 30, 2025***

## SUMMARY

The study will be conducted to complement the diagnosis of proximal caries lesions that are difficult to improve due to their location between the teeth, as shown in Near-Infrared Light Transillumination (NILT) photographs. It was proposed to enhance caries detection with a semantic segmentation model based on YOLOv11, which is more specific in detecting mesial and distal caries lesions, which were not adequately explored in previous literature. The model was trained on a sample of 440 augmented grayscale NILT images collected from 17 patients at the Faculty of Nursing, King Abdulaziz University, Jeddah, Saudi Arabia. These pictures were categorized into five groups, i.e., enamel, dentin, background, mesial caries, and distal caries. The significance of the classes was addressed by optimizing hyperparameters and class weights for mesial and distal caries. The YOLOv11 model had an overall Dice coefficient of 87 and the highest scores of enamels (80) and dentin (89) due to their low prevalence in the dataset (mesial and distal caries, respectively). However, the model was deemed very specific, and its negative predictive value was 0 in all classes. The findings indicate that a well-represented class is the most effective with the model, whereas mesial and distal caries need further improvement. Future work will focus on improving the model's performance on underrepresented classes through data augmentation and class-balanced training, ultimately enhancing its clinical applicability for the intricate, non-invasive detection and diagnosis of caries.

Key words: *proximal caries, near-infrared light transillumination (NILT), deep learning, YOLOv11, semantic segmentation, dental imaging, caries detection, pixel-level segmentation, hyperparameter optimization, mesial and distal caries detection.*

## INTRODUCTION

Proximal caries lesions develop between adjacent teeth and, unfortunately, are very difficult to detect because of their location. Such lesions are frequently overlooked during standard clinical assessment, leading to disease progression and more invasive intervention [1]. It is essential to diagnose proximal caries early to avoid additional tooth destruction and costly restorative procedures. Traditional techniques, including visual examination and exploration with fingers, are often used to identify caries. Yet, they cannot be used because they cannot reach the environment between the proximal parts [3]. Despite their usefulness, bitewing radiographs expose patients to ionizing radiation and are less applicable to some groups of people, e.g., children and pregnant women.

Conversely, Digital Imaging Fiber Optic Transillumination (DiFOTI) and Near-Infrared Light Transillumination (NILT) have proven to be radiation-free and offer better detection features. NILT, especially, enables light to penetrate deep into the tooth, providing an additional non-invasive method for detecting caries at early stages when they are not yet visible on x-rays or during clinical examination [5][6]. The NILT-based DIAGNOcam device has demonstrated potential for identifying both the presence and location of caries in posterior teeth by visualizing them with transmitted light [7]. The method is appropriate when proximal caries is identified early, but its accuracy depends on image interpretation, especially for mesial and distal caries [8].

Dental image analysis has also become an area where deep learning models, especially Convolutional Neural Networks (CNNs), are used [4]. CNNs are highly accurate across a wide range of diagnostic tasks, including caries detection [9][10]. Previously made research, however, tended to concentrate on generic caries detection without making a distinction between mesial and distal lesions that are clinically important in treatment planning and diagnosis [11]. In addition, there is the class imbalance whereby some types of lesions, such as mesial and distal caries, are underrepresented, thus reducing the effectiveness of such models in real-world settings [12].

The current paper presents a new method of using the semantic segmentation of NILT images that involves the application of the state-of-the-art deep learning model, YOLOv11 [2]. The distinction of mesial and distal caries is the major innovation of this work that has not been sufficiently researched previously. The distinction of these types of caries makes the study more accurate in the localization of lesions and increases the accuracy of the detection at an early stage. Moreover, the hyperparameter optimization and class prioritization methods are also utilized to address the problem of class imbalance and to enhance the segmentation accuracy, particularly in the underrepresented mesial and distal areas [13][14]. Another significant point of this research is that the correct annotation of images and the possibility of automating the segmentation process provided by YOLOv11 results in the valuable assistance of dental professionals in caries diagnosis of proximal caries [15][16].

The literature on combining deep learning with dental imaging has been growing recently, but most studies have been restricted to radiographies or have not addressed minor details in the localization of caries. Alternatively, the proposed research is NILT images in particular, which offer a rare benefit over conventional techniques and rely on infrared light to identify caries non-invasively. The obtained results indicate that even with a small dataset, YOLOv11 has the potential to identify proximal caries lesions successfully and offers substantial benefits over traditional algorithms, which is a promising direction to use in the future in clinical practice.

**Key Contribution**

- This study distinguishes between mesial and distal caries lesions in NILT images, which has not been adequately undertaken in previous studies.

- The research paper proves the usefulness of applying YOLOv11, which is a modern deep-learning model, in the semantic segmentation of dental structures in NILT images to enhance the detection of caries.
- To overcome the imbalance in classes, specifically similar classes, the hyperparameter optimization and the use of class weights were used to improve the performance of the models.
- Segmentation of enamel and dentin in the model was highly accurate (80 percent and 89 percent), thus guaranteeing good detection of these well-represented features in images of NILT.
- The research presents an encouraging, non-invasive tool for early detection of proximal caries, which will aid in the reduction of invasive treatment and the improvement of patient outcomes in clinical practice.

## LITERATURE REVIEW

Detection of proximal caries, especially in the spaces between teeth, is a big problem in dental diagnostics [17]. The conventional techniques, like visual inspection and radiography, cannot easily identify caries at its initial stages in locations that are not between the teeth [18][19]. Although bitewing radiographs are helpful, they expose one to ionizing radiation, which is dangerous, especially in children and expectant women. This has seen the rise of non-invasive procedures such as Near-Infrared Light Transillumination (NILT) to provide a safer method of caries detection in its initial stages. NILT operates on the principle of light penetration through the dental features, where it has the ability to visualize the caries that do not show in the normal imaging procedures [20].

The recent developments in deep learning demonstrated promise in the image analysis of dentists, especially through the detection of caries. Convolutional Neural Networks (CNNs) are extensively used in detecting caries and have proven to be effective in detecting lesions in radiographs and intraoral images with high accuracy. Nonetheless, most of these methods are aimed at the assessment of caries without the differentiation between mesial and distal caries, which are vital in the diagnosis and treatment [21][22]. These lesions are frequently ignored in terms of clinical significance, including the segmentation of these lesions.

Many studies have tried to employ machine learning architectures like CNNs to recognize dental caries in different types of imaging, including radiographs and NILT. However, most of these studies do not consider the issues of imbalance in classes, where the underrepresented classes, such as mesial and distal caries, are affected. Not many works have investigated the idea of semantic segmentation in dental imaging, and even fewer have applied such a model as YOLOv11 to the problem. Recently, the YOLOv11, which is efficient and accurate in object detection, has been applied to segmentation tasks, and it provides a potential solution to these issues in detecting dental caries.

Moreover, other studies have also emphasized the application of deep learning to caries detection, yet few have combined the techniques to combat the problem of class imbalance, which is essential to enhance the performance of models that run on a dataset that lacks representation of some classes, such as mesial and distal caries [23][24]. With the help of more sophisticated training methods, including class-balanced loss functions and data augmentation, this research will be targeted at streamlining the performance of the model and guaranteeing the accurate identification of caries lesions in the well-represented classes as well as in those that are underrepresented.

## MATERIALS AND METHODS

Caries lesions near the midline within the oral cavity were termed mesial, whereas those situated further from the midline were referred to as distal. Dentists must ascertain the location and orientation of the caries lesions. After acquiring the NILT images, it is evident that the image concentrates on the region of interest of a single tooth, thereby precluding the determination of the caries direction (mesial or distal) from the image alone. Thus, the significance of employing a CNN-based segmentation technique in deep learning is emphasized.

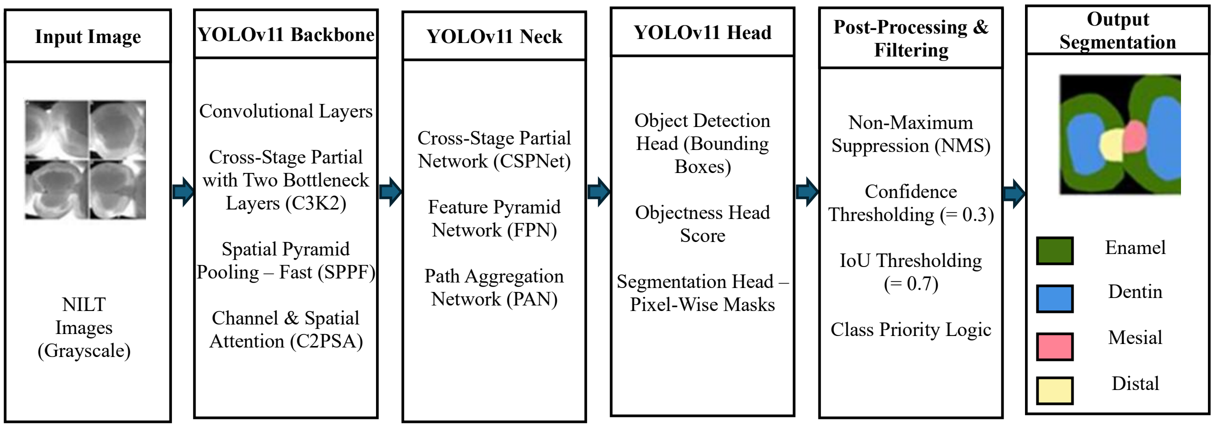



Figure 1. YOLOv11 Segmentation pipeline for proximal caries detection in near-infrared light transillumination (NILT) images

Figure 1 shows the workflow diagram of the YOLOv11 segmentation pipeline that is used to identify proximal caries with the help of the NILT images. The pipeline is broken down into several steps, including input image processing, YOLOv11 backbone features extraction, object detection, mask generation, post-processing and filtering, and final segmentation map of the output. Among the significant steps is the process of classifying the prediction of enamel, mesial, and distal caries. Post-processing steps are used to refine the results of the segmentation process and enhance the diagnostic accuracy.

In this study, the capabilities of YOLOv11x-seg are utilized for segmentation. The process begins with dataset collection, followed by the annotation process. Subsequently, hyperparameters are adjusted, the model is trained, and the results obtained are shown in Figure 2 for further illustration.

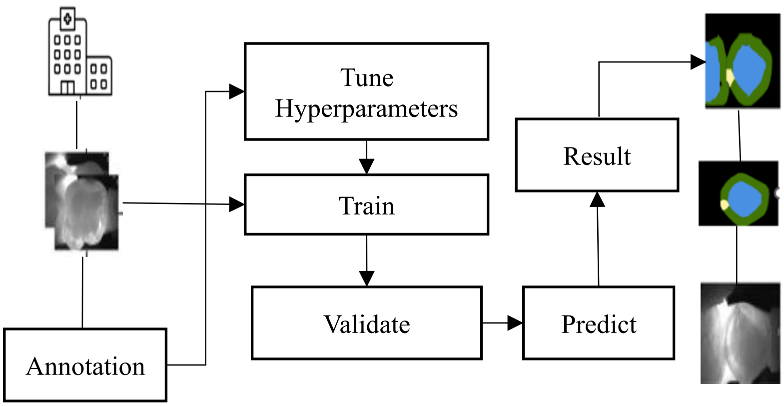



Figure 2. Workflow of the proposed dental image segmentation pipeline

## Data Collection

The dataset was built from images obtained using DIAGNOcam (KaVo, Biberach, Germany) in collaboration with the Faculty of Dentistry of King Abdulaziz University, Jeddah, Saudi Arabia, under ethical approval no. (305-10-21). Two dentists acquired images of 17 patients aged 18 to 55 years. The collection comprised 200 NILT images, approximately 109 of which were proximal caries cataloged with patient and image identifiers. The teeth were rinsed with water and dried using compressed air to eliminate distractions in the photographs, while the dental light was off to prevent interference. The camera position focused on the region of interest, particularly the area that requires examination above the occlusal surface or proximal caries. Teeth with restorations or orthodontic treatment were not included in this study. In addition, teeth that are chipped or pitted are not included because they appear as caries in greyscale images.

**Data Annotation**

Image analysis was performed by three experienced dentists who identified the caries site (mesial or distal) with a consensus agreement approach. The examiners set together and annotated the images; in cases of disagreement between examiners, a consensus method was adopted. Images were segmented into five categories (Background, Enamel, Dentin, and Mesial or Distal) utilizing the online pixel annotation application in the Supervisly (2024) AI-powered data labelling platform. This facilitates the labelling process for object detection and segmentation, producing datasets in many formats, including YOLOv8, YOLOv5, and COCO. The Enamel and Dentin layers were meticulously outlined pixel by pixel using a polygon tool. Subsequently, the locations of caries (Mesial or Distal) were identified and annotated accordingly, if present.

The determination of the caries lesion location, whether Mesial or Distal, varies based on the direction of the captured image. Specifically, the orientation of the mesial and distal surfaces depends on the arm position. When the device arm is positioned to the right, the mesial surface is also oriented to the right, whereas the distal surface is oriented on the opposite side, as shown in Figure 3. A. Conversely, when the device arm is positioned to the left, the mesial surface is oriented to the left, with the distal surface on the opposite side, as shown in Figure 3. B. To ensure accurate identification of the caries' location, it is essential to monitor the direction of the device arm during labelling.

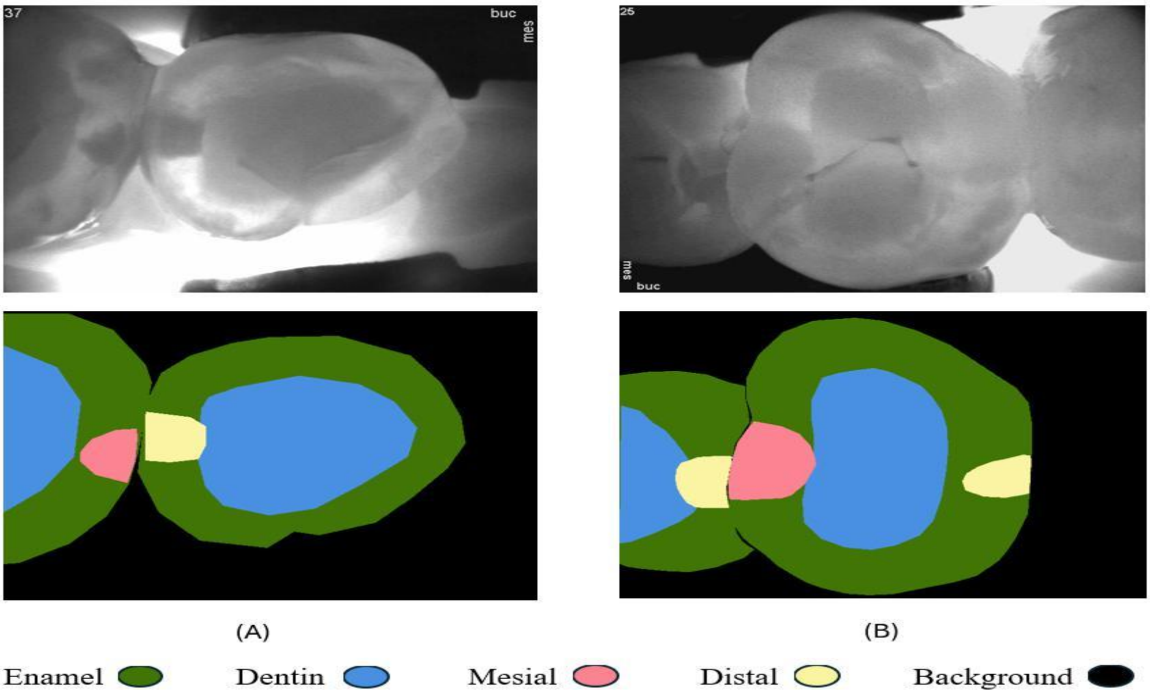



Figure 3. Location of caries based on arm direction, (A) image of teeth with the imaging arm positioned to the right, and (B) image of teeth with the imaging arm positioned to the left

**Hyperparameters for Model Training**

The training of the YOLOv11 model was done in a thorough manner to ensure that the hyperparameters were optimised to ensure that the model gave the best possible segmentation results. The learning rate was used to regulate the rate at which the model can change its weights at each optimization step, and auto selection is used to select the learning rate dynamically to achieve optimal performance. The batch size was a parameter that indicated the size of the samples that the model used before it updated its weights, and some of the values that were tested were 4, 8, and 16. An increased batch size usually increases stability at the expense of memory. Training of the model was done between 50 and 200 epochs, with a higher number of iterations generally resulting in improved performance, but the chances of overfitting increase with a high number of iterations. To be optimized, it was decided to use the AdamW optimizer, which is able to better address non-stationary data and is less time-consuming in terms of convergence. The sizes of the input images were reduced to 224x224 or 320x320 pixels, and larger images with more details were used, at the expense of incurring higher computational costs. Also, augmentation methods such as horizontal flipping, scaling, and rotation were used to increase the generalization of the model by creating diverse versions of training data. Having executed the search algorithm, we received the (30) results presented in Figure 4, the best iteration (11), the mAP50 (0.73) with the epochs (189) and the batch size (4), the optimizer (AdamW).

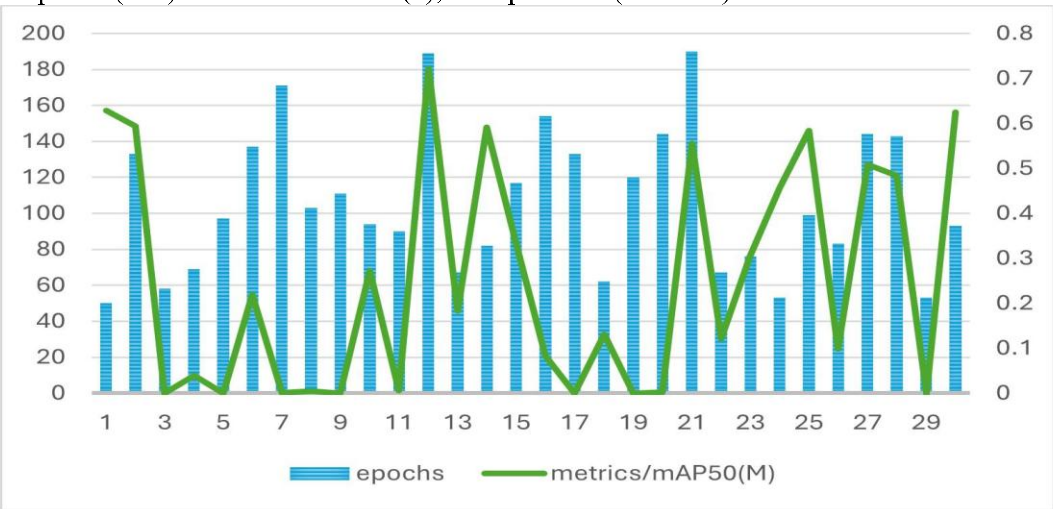



Figure 4. Chart for 30 iterations in tuning hyperparameters, the blue represents the number of epochs, and green represents the result of metrics mAP50 for the mask

**Augmentation**

Augmentation techniques were applied: horizontal flipping, random brightness adjustment, and in-plane rotation. As a result, a total of 200 images containing all classes were augmented, producing 440 samples, in addition to the original dataset.

**YOLOv11**

The model version is YOLOv11x-seg, called from the Ultralytics package and Google Colab with a T4 GPU. The dataset was split into 70% training and 30% validation. Training was done using arguments that were tuned as described above; it took two hours, and the best weight was used to predict the mask for each segment. Prediction setting was confidence (0.3) and IoU (0.7) after several tests and random choices. A lower confidence threshold considers more potential detection for small areas (mesial and distal), and a high IoU allows more overlap before eliminations, which reduces false suppression.

**Evaluation Metrics**

For modal evaluation, the mean average precision across classes (mAP) was computed, which measures how the model detects objects, as determined using Equation 1. The value of (mAP50) means IoU overlap of at least 50% with the actual object to be considered correct.
Mean average precision (mAP) is an estimation of the overall performance of the model across all classes. It is characterized by the average of the average precision (AP), per class.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k \quad (1)$$

$$n: number\ of\ classes$$
$$AP: average\ pressure\ for\ class\ k$$

The case can be as follows: n = the total number of classes (in your case, 5: Enamel, Dentin, Background, Mesial, and Distal).

Dice coefficient calculated to evaluate the overlap between the predicted segmentation and the ground truth, according to the following Equation 2. Also, Sensitivity (Recall), Specificity, Positive Predictive

Value (PPV or Precision), and Negative Predictive Value (NPV) were computed to measure the proportion between predicted and ground truth using Equations 3-6.

$$Dic = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{2}$$

Where TP is the number of actual positives (caries pixels that had been correctly predicted), FP is the false positives (predicted caries pixels that are false). FN = false negative (pixels that were not recognized as caries).

The sensitivity or recall is determined as the percentage of true caries lesions, which are correctly detected by the model:

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{3}$$

Where TP is the true positives (correct caries detection), FN is the false negatives (Caries missed).

Having a high sensitivity means that the model can identify most of the actual caries lesions, reducing the failure to identify lesions.

The specificity is used to determine the percentage of correct non-carrier regions in all the real negatives.

$$Specificity = \frac{TN}{(TN + FP)} \tag{4}$$

Where, TN is the number of true negatives (correctly identified non-caries regions). FP is the false positives (incorrectly predicted caries regions).

Having a high specificity means that the model is effective in differentiating caries and non-caries areas to minimize false positive outcomes.

The proportion of the number of predicted positive samples (caries lesions) which are true positives is called the positive predictive value (PPV) or precision:

$$PPV = \frac{TP}{(TP + FP)} \tag{5}$$

Where TP is the true positives. FP is the number of false positives.

It is the negative predictive value (NPV), which is the fraction of the sampled negative samples (non-caries regions) who are, in fact, negative:

$$NPV = \frac{TP}{(TN+FN)} \tag{6}$$

with TP (True Positives) being the number of pixels that are correctly identified as positive, FP (False Positives) being the number of pixels that are incorrectly considered positive, FN (False Negatives) being the number of pixels that the model has failed to identify as positive, and TN (True Negatives). All these metrics are the assessment of the full capacity of the model to correctly divide dental structures with minimal false detections, which are useful in determining reliable models in the case of imbalanced classes of dental data.

**Algorithm 1: YOLOv11-based Proximal Caries Detection Model**

**Input:** NILT images with annotations for Enamel, Dentin, Background, Mesial Caries, and Distal Caries.

**Data Preprocessing**

Collect and augment NILT images (flip, rotate, adjust brightness).

Normalize and resize images (e.g., 224x224 pixels).

**Data Annotation**

Annotate regions as Enamel, Dentin, Background, Mesial Caries, and Distal Caries.

**Model Setup**

Use YOLOv11 for feature extraction with pre-trained weights.

Add a segmentation layer to predict pixel-level caries regions.

**Training**

Tune hyperparameters (learning rate, batch size).

Apply class weighting to handle class imbalance (for mesial/distal caries).

Use AdamW optimizer for efficient training.

**Evaluation**

Evaluate using metrics: Dice coefficient, mAP, sensitivity, specificity, precision.

**Inference & post-processing**

Apply model to new images for caries detection.

Refine predictions with post-processing (remove small irrelevant detections).

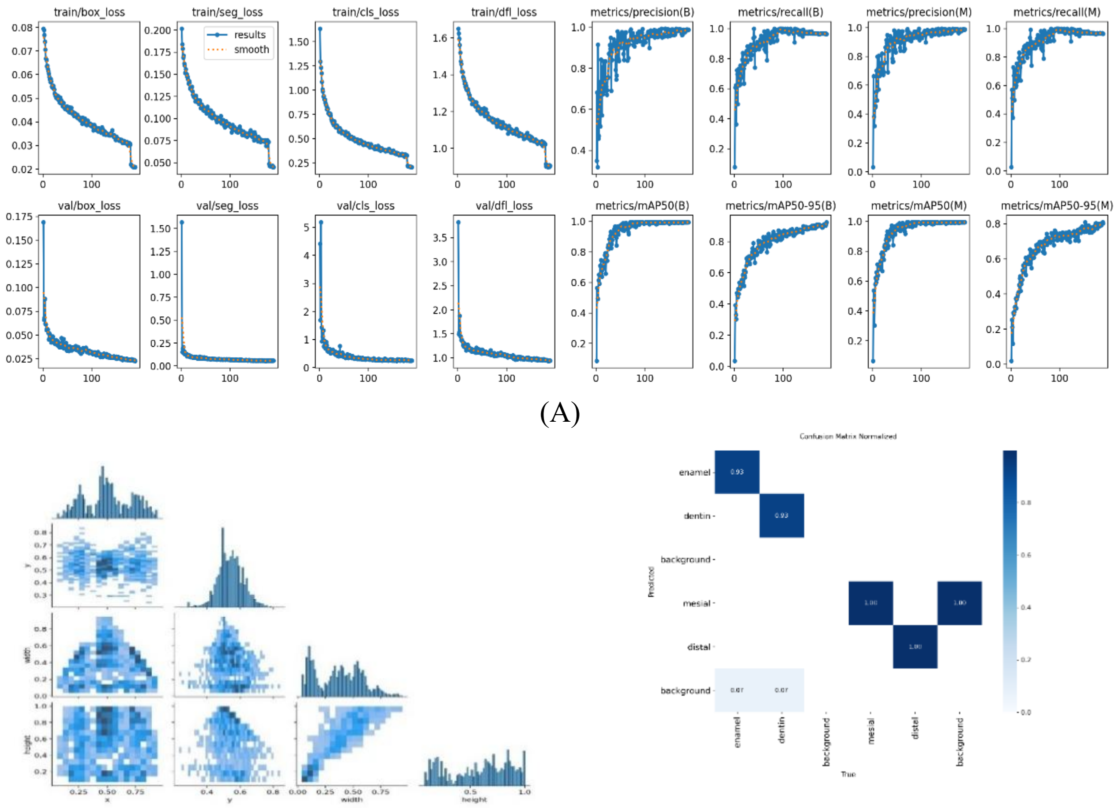
**Output:** Display predicted caries masks over original images.

Algorithm 1 is based on YOLOv11 to identify caries at a proximal level in NILT images. It would entail image gathering, image augmentation, and image annotation, and then model training using a class balance. Hyperparameters are optimized and AdamW optimizer is employed. Performance of the model is assessed and refinement of the prediction is done through post-processing, which presents the segmented caries masks to the original images.

RESULTS

**Model Training and Performance**

To thoroughly assess the performance of the YOLOv11 model on the dental dataset, we analyzed the dynamics of training, distribution of dataset, and results of segmentation through visual summaries as shown in Figure 5. Figure 5.A shows the training and validation performance of YOLOv11, where all the loss curves (bounding box, segmentation, classification, distribution focal loss) generally decrease throughout the epochs, which shows that the convergence is stable during the training period. Figure 5.B is a pairwise correlagram of normalized bounding box, which illustrates the spatial distribution of the instances annotated in the dataset. The normalized confusion matrix of the YOLOv11 segmentation outputs of the enamel, dentin, background, mesial, and distal classes is shown in the third panel (Figure 5.C). The model was highly classified as correct on enamel and dentin (93% both) showing that these are highly reliable classes to segment. It is noteworthy that the mesial and distal classes obtained 100 percent correct classification at prediction proving the model to identify these structures correctly even though they were underrepresented in the dataset. The percentage of misclassified background enamel and dentin pixels was small (about 7 per cent each) which implies that occasionally fine edges or thinner parts can be ignored. On balance, the confusion matrix confirms that the model has a high competence in differentiating key dental structures with a great level of specificity, and the possibility to mitigate the misclassification of minor boundaries in the future work can be achieved with the help of specific additions and class-balanced training adjustments.




(A)

(B) (C)

Figure 5. (A)Training and validation loss, precision, recall, and mAP curves for the YOLOv11 model on the dental dataset, demonstrating stable convergence during training. (B) pairwise correlogram of normalized bounding box parameters (center x, center y, width, height). (C) normalized confusion matrix for the YOLOv11 segmentation outputs across enamel, dentin, background, mesial, and distal classes

**Segmentation Performance by Class**

Table 1 represents the segmentation performance of the YOLO model on enamel, dentin, background, distal, and mesial classes with mAP50and Dice coefficients of 0.75and 0.86, respectively, showing great segmentation performance on the dataset. In the cases of the enamel and dentin classes, the model scored high in Dice scores (0.80 and 0.89) and mAP50 value (0.78 and 0.82) indicating the correct delineation of these dominant structures with sensitivity of more than 0.83 and 0.96 and PPV of more than 0.77 and 0.82 respectively. Background class was also performing strongly with Dice of 0.89 and high mAP50 of 0.91. On the underrepresented distal and mesial classes, the model scored lower Dice scores (0.65 and 0.53) and mAP50 (0.79 and 0.43), the difficulty of this task can be explained by the lack of information on small structures. The high specificity (>0.99) and NPV (>0.99) of all the classes however means that the model is doing a proper job in differentiating between these structures and non-targeted areas with a very small number of false positives not in the classes of interest.

Figure 6 shows a spider chart of the Dice coefficient and mAP50 when comparing between the dental dataset in the enamel, dentin, background, mesial, and distal classes. As shown by the chart, the model had the highest Dice and mAP50 measures in the dentin and background classes, indicating uniform and precise segmentation of these well-represented structures. The enamel class also showed good performance and this indicates the capability of the model to draw boundaries of the teeth in an effective manner.

Table 1. Statistical indicators for the considered classes enamel, dentin, mesial, distal and the average

| Classes | Dice | mAP50 | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|---|---|
| Enamel | 0.80 | 0.78 | 0.83 | 0.92 | 0.78 | 0.94 |
| Dentin | 0.89 | 0.82 | 0.96 | 0.92 | 0.82 | 0.99 |
| Background | 0.89 | 0.91 | 0.87 | 0.93 | 0.91 | 0.90 |
| Mesial | 0.65 | 0.79 | 0.55 | 0.99 | 0.79 | 0.99 |
| Distal | 0.53 | 0.43 | 0.71 | 0.99 | 0.43 | 0.99 |

Conversely, the mesial and distal classes, though registering lower Dice score, still registered competitive mAP scores, which means that the model can identify these underrepresented classes very well, but perhaps with further improvement, the segmentation quality can be enhanced. By and large, the spider chart effectively compares the performance of class-wise segmentation and detection in terms of efficiency of the YOLOv11 model and offers the prospects of improvement by introducing a more intricate augmentation strategy and class-balanced training to boost the segmentation of the less represented dental structures.

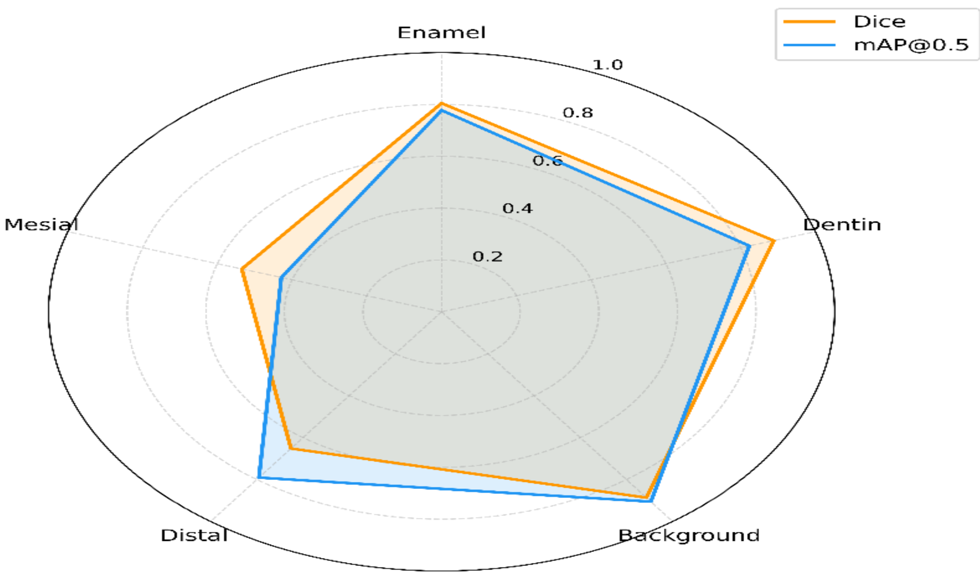



Figure 6. Spider chart illustrating the dice coefficient and mAP50 for enamel, dentin, background, mesial, and distal classes on the dental dataset

Figure 7 shows the final results showing the original NILT images, masked versions of the original image, and the YOLOv11 predicted mask. On the whole, YOLOv11 correctly labeled enamel and dentin most of the time, which proves a high segmentation quality of these highly represented classes. However, the consistency in the prediction of mesial and distal regions were lesser with the misclassification and false positives occurring between the two classes. Also, there were a few instances of the predicted cavities in the output masks, where the ground truth did not have such predictions, which indicates contained false positives, and there were other areas, which had true negatives. These inconsistencies are mostly credited to the limitations of the dataset, especially, the underrepresentation of mesial and distal classes, which impacts the learning by the model of subtle differences between these structures in comparison to enamel and dentin.

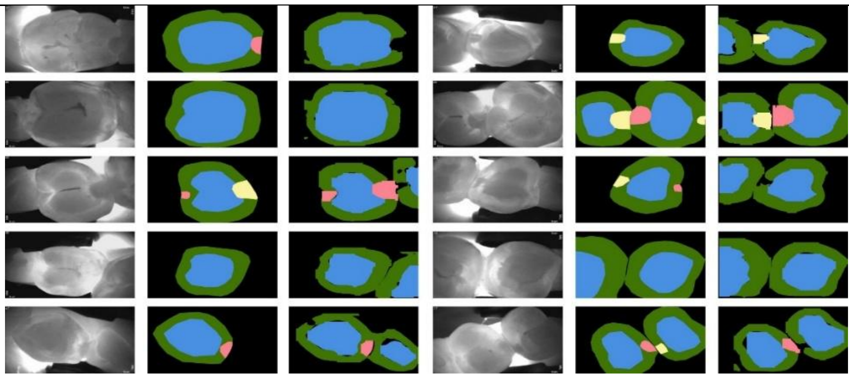


Figure 7. Original NILT images, ground truth images, and predicted from YOLOv11 model

**Clinical Observations**

Segmentation accuracy is also affected by the quality of input images and consistency of annotations, which indicates the relevance of accurate labeling in medical image analysis training datasets. It is worth noting that the analysis revealed that caries were more commonly found on mesial surfaces as compared to distal surfaces in the dataset. This observation can be attributed to the clinical concept that the mesial surfaces experience a consistent low level of occlusal forces in biting and chewing, which results in a progressive enamel erosion and susceptibility to proximal caries.

DISCUSSION

The caries is inaccessible and this is the reason why the proximal caries are still hard to detect. The addition of deep learning to identify proximal caries enhances efficiency and may result in improved patient outcomes because of the possibility to take action earlier. Proximal caries detection using YOLOv11 was also encouraging in this study, and the hyperparameter optimization reduced the amount of time that was taken to determine the best parameters to use to run this model. The model obtained was found to possess good detection and segmentation capabilities of dental structures in the NILT images.

The evaluation results revealed that the segmentation of enamel and dentin and background classes was high and mAP50 and Dice scores are lower in the case of distal and mesial classes. This bias can be explained directly by the fact that the classes distal and mesial regions are underrepresented and so limits the exposure of the model, and consequently it ability to learn fine-grained boundaries of these smaller structures. Specificity and Negative predictive value of these classes are however high indicating that the model is useful in minimizing false positives but it has the potential of missing true positives in such areas.

In order to address this lack of balance, training can oversample the distal and mesial classes to expose the model to more of these under-represented structures without preferentially learning all of the classes. More variability of the dataset of the dataset can be enhanced further with more advanced variability strategies, such as random distal and mesial cropping, and synthetic data generation, which will enhance the generalization capability of the model. Also, class-balanced loss functions such as focal loss or class-weighted Dice loss could be used to increase sensitivity and accuracy to distal and mesial structures and permit stable and dependable segmentation using all the anatomical structures of NILT dental imaging. The second step in the future would be to implement the methods of mitigating these class imbalances in the training pipeline to achieve refinement of detection and classification of distal and mesial classes, and overall system performance to clinical and research applications of automated caries detection and dental structure analysis.

To make a comparison of our results with the literature published earlier, it is necessary to state that the various studies could vary in the model structures, number of data sets, and/or evaluation processes. Using the example of Casalegno et al. [21], they trained their segmentation model on 217 DIAGNOcam images and achieved a mean Intersection over Union (mIoU) of 72.7, our model achieved a mean IoU of 76, which is superior. However, their dataset and model cannot be benchmarked as such as they are not available. Additionally, the method of object detection was employed in both papers written by Holtkamp et al. [22] and Schwendicke et al. with a score of 78 and 74 respectively. Due to the differences between object detection and semantic or instance segmentation in the intent of the task and evaluations, the direct comparisons are to be regarded with a grain of salt.

## CONCLUSIONS

In this paper, a deep learning-based model, YOLOv11, has been used to segment and detect proximal caries lesions on Near-Infrared Light Transillumination (NILT) images. The overall result of the model was high Dice coefficient (87%), and the well-represented classes (Enamel 80% and Dentin 89%). These areas were highly sensitive and specific, meaning that the model is effective in the detection of these structures and a few false positives. Conversely, Mesial (65%) and Distal (53%) caries lesions that were underrepresented in the dataset had lower performance and lower Dice scores leading to the misclassification of Distal caries lesions as Background. Small dataset size and lack of diversity also restricted the performance of the model as it could not generalize to different caries lesions and population. These lessons point at the difficulties of using deep learning models on clinical data with class imbalances and small samples. The research that should be undertaken in the future is to enhance the performance of the model to Mesial and Distal caries by increase in the dataset with more diversified samples. Augmentation of data based on smaller lesions and the use of class-balanced loss functions may help to enhance the segmentation of underrepresented classes. Also, other deep-learning architectures should be tested, which may also help in increasing the detection accuracy. The clinical implication of the study is also meaningful, as the non-invasive nature of the tool is automated, and it is used to detect proximal caries in an early stage, which could hardly be detected by conventional methods. The method has the potential to minimize the invasive nature of the treatments, particularly among the radiation sensitive population, including children and the elderly. Embarking on AI-based solutions in the field of dentistry may result in achieving more accurate diagnosis, enhanced patient outcomes, and simplified clinical procedures.


## ACKNOWLEDGEMENT

The authors would like to express their gratitude to King Abdulaziz University for providing the necessary resources and support for this research. The authors also would like to thank Dr. Sulafah Mugharbil, Dr. Shereen Jifri, and Dr. Amjad Azab for their valuable contribution.

## FUNDING SOURCES

The project was funded by KAU Endowment (WAQF) at king Abdulaziz University, Jeddah, Saudi Arabia. The authors, therefore, acknowledge with thanks WAQF and the Deanship of Scientific Research (DSR) for technical and financial support.


## AUTHOR CONTRIBUTIONS

Asma Alatawi: Methodology, Software, Visualization, Resources, Validation, Writing - Original Draft. Wadee Alhalabi: Supervision, Funding acquisition, Methodology, Validation, Writing - Review & Editing. Hani Nassar: Supervision, Methodology, Conceptualization, Formal analysis, Resources, Data Curation, Investigation, Writing - Review & Editing. Arwa Basbrain: Supervision, Methodology, Validation, Writing - Review & Editing. Hattan Jamalellail: Resources, Formal analysis, Data Curation, Investigation, Writing - Review & Editing. Mohammed Alsadat: Resources, Formal analysis, Data Curation, Investigation, Writing - Review & Editing.

## ETHICS STATEMENT

This study was approved by the Faculty of Dentistry of King Abdulaziz University, Jeddah, Saudi Arabia, under ethical approval no. (305-10-21).

DATA AVAILABLE

The data will be available under the request.

REFERENCES

[1] Shen L, Yang X, Dong C, Tang X, Zheng S, Wang T, Wang L, Yang F, Zheng Y. Near-infrared light reflection for the early detection of proximal caries in posterior teeth: an in vivo study. BMC Oral Health. 2025 Jan 26;25(1):139.

[2] Bharathi S, Senthilarasi M, Hari K. Key frame extraction based on real-time person availability using YOLO. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Application. 2023 June 30;14(2):31–40. https://doi.org/10.58346/JOWUA.2023.I2.003

[3] Schwendicke F, Elhennawy K, Paris S, Friebertshäuser P, Krois J. Deep learning for caries lesion detection in near-infrared light transillumination images: A pilot study. Journal of dentistry. 2020 Jan 1; 92:103260. https://doi.org/10.1016/j.jdent.2019.103260

[4] Rajan D, Sharma N. Risk Evaluation of Dental Erosion Using the Basic Erosive Wear Examination Scoring Method. Clinical Journal for Medicine, Health and Pharmacy. 2025 Sep 30;3(3):16-31.

[5] Patel J, Vannemreddy A, Goh YJ, Francis Y, Anthonappa R. Evaluation of near-infrared digital imaging transillumination compared with bitewing radiography for proximal caries detection in children. International Journal of Paediatric Dentistry. 2025 Jan;35(1):108-17. https://doi.org/10.1111/ipd.13205

[6] Mahizha SI, Annrose J, Mano Christaine Angelo J, Domilin Shyni I, veda Giri GV. Deep convolutional neural networks for early detection of interproximal caries using bitewing radiographs: A systematic review. Evidence-Based Dentistry. 2025 Mar 21:1-9.

[7] Öncel AA, Ayan E. Convolutional neural network applications in dental diagnostics: a focused review on bitewing-based caries detection. Archives of Computational Methods in Engineering. 2025:1–21.

[8] Faizan Ahmed SM, Ghori MH, Khalid A, Nooruddin A, Adnan N, Lal A, Umer F. Annotated intraoral image dataset for dental caries detection. Scientific data. 2025 Jul 25;12(1):1297.

[9] Adnan A, Rizwan MT, Attaullah HM, Basheer S, Quasim MT. Deep Architectural Classification of Dental Pathologies Using Orthopantomogram Imaging. Computers, Materials and Continua. 2025 Oct 23;85(3):5073-91. http://dx.doi.org/10.32604/cmc.2025.068797

[10] Thakur GK, Thakur A, Kulkarni S, Khan N, Khan S. Deep learning approaches for medical image analysis and diagnosis. Cureus. 2024 May 2;16(5). https://doi.org/10.7759/cureus.59507

[11] ForouzeshFar P, Safaei AA, Ghaderi F, Hashemikamangar SS. Dental Caries diagnosis from bitewing images using convolutional neural networks. BMC Oral health. 2024 Feb 10;24(1):211.

[12] Thanh MT, Van Toan N, Ngoc VT, Tra NT, Giap CN, Nguyen DM. Deep learning application in dental caries detection using intraoral photos taken by smartphones. Applied Sciences. 2022 May 29;12(11):5504. https://doi.org/10.3390/app12115504

[13] Musri N, Christie B, Ichwan SJ, Cahyanto A. Deep learning convolutional neural network algorithms for the early detection and diagnosis of dental caries on periapical radiographs: A systematic review. Imaging science in dentistry. 2021 Jul 13;51(3):237. https://doi.org/10.5624/isd.20210074

[14] Kühnisch J, Meyer O, Hesenius M, Hickel R, Gruhn V. Caries detection on intraoral images using artificial intelligence. Journal of dental research. 2022 Feb;101(2):158-65. https://doi.org/10.1177/00220345211032524

[15] Pérez de Frutos J, Holden Helland R, Desai S, Nymoen LC, Langø T, Remman T, Sen A. AI-Dentify: deep learning for proximal caries detection on bitewing x-ray-HUNT4 Oral Health Study. BMC Oral Health. 2024 Mar 18;24(1):344.

[16] Chen X, Guo J, Ye J, Zhang M, Liang Y. Detection of proximal caries lesions on bitewing radiographs using deep learning method. Caries research. 2022 Dec 28;56(5-6):455-63. https://doi.org/10.1159/000527418

[17] Bayraktar Y, Ayan E. Diagnosis of interproximal caries lesions with deep convolutional neural network in digital bitewing radiographs. Clinical oral investigations. 2022 Jan;26(1):623-32.

[18] Lee S, Oh SI, Jo J, Kang S, Shin Y, Park JW. Deep learning for early dental caries detection in bitewing radiographs. Scientific reports. 2021 Aug 19;11(1):16807.

[19] Dayı B, Üzen H, Çiçek İB, Duman ŞB. A novel deep learning-based approach for segmentation of different type caries lesions on panoramic radiographs. Diagnostics. 2023 Jan 5;13(2):202. https://doi.org/10.3390/diagnostics13020202

[20] Holtkamp A, Elhennawy K, Cejudo Grano de Oro JE, Krois J, Paris S, Schwendicke F. Generalizability of deep learning models for caries detection in near-infrared light transillumination images. Journal of clinical medicine. 2021 Mar 1;10(5):961. https://doi.org/10.3390/jcm10050961

[21] Chandana RK, Ramachandra AC. Real time object detection system with YOLO and CNN models: A review. arXiv Prepr. arXiv2208. 2022 Jul 17;773.

[22] Hussain M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection. Machines. 2023 Jun 23;11(7):677. https://doi.org/10.3390/machines11070677

[23] Warreth A. Dental caries and its management. International journal of dentistry. 2023;2023(1):9365845. https://doi.org/10.1155/2023/9365845

[24] Michou S, Vannahme C, Bakhshandeh A, Ekstrand KR, Benetti AR. Intraoral scanner featuring transillumination for proximal caries detection. An in vitro validation study on permanent posterior teeth. Journal of Dentistry. 2022 Jan 1;116:103841. https://doi.org/10.1016/j.jdent.2021.103841