



King Abdulaziz University

Faculty of Computing and Information Technology

Professional Master in Artificial Intelligence

Enhanced Dental Diagnostics: Caries Detection and Sizing in Panoramic X-Ray Image

Research extracted from a master's thesis in computer science

By: Shahad Saad Ibrahim Alhawiti

Supervised by: Dr. Arwa Basbrain

Abstract

This study introduces a comprehensive framework that utilizes deep learning techniques for the automated detection and segmentation of dental caries in panoramic X-ray images. The proposed system is designed to assist dental practitioners by highlighting regions of decay and accurately outlining their boundaries, thereby facilitating more efficient and precise diagnoses. A two-stage model, Faster R-CNN, is employed for object detection due to its effectiveness in identifying carious lesions of varying sizes with enhanced localization accuracy. For the task of semantic segmentation, the DeepLabV3+ architecture is implemented to generate detailed pixel-level masks of the affected regions, providing a more granular understanding of the decay patterns. The overall approach encompasses image preprocessing, model development, and performance evaluation using established quantitative metrics. The experimental results indicate the capability and robustness of the proposed method in improving diagnostic accuracy. These findings underscore the potential of this framework to serve as a valuable clinical aid and pave the way for the future integration of AI-powered dental analysis tools into modern dental practice.

Key Words : Enhanced Dental Diagnostics - Caries Detection - Sizing in Panoramic - X-Ray Image.

ملخص :

تقدم هذه الدراسة إطاراً شاملاً يستخدم تقنيات التعلم العميق للكشف الآلي عن تسوس الأسنان وتقسيمة في صور الأشعة السينية البانورامية، وصمم النظام المقترح لمساعدة أطباء الأسنان من خلال تحديد مناطق التسوس وتحديد حدودها بدقة، مما يسهل تشخيصاً أكثر كفاءة ودقة، ويستخدم نموذج Faster R-CNN ثنائي المرحلتين للكشف عن الأجسام نظراً لفعاليتها في تحديد الأفات التسوسية ذات الأحجام المختلفة بدقة تحديد موقع مُحسّنة. ولغرض التجزئة الدلالية، طبقت بنية DeepLabV3+ لإنشاء أقنعة مفصلة على مستوى البكسل للمناطق المتأثرة، مما يوفر فهماً أكثر تفصيلاً لأنماط التسوس، ويشمل النهج الشامل المعالجة المسبقة للصور، وتطوير النموذج، وتقييم الأداء باستخدام مقاييس كمية معتمدة، وتُشير النتائج التجريبية إلى قدرة ومتانة الطريقة المقترحة في تحسين دقة التشخيص، وتؤكد هذه النتائج على إمكانات هذا الإطار ليكون بمثابة مساعدة سريرية قيمة وتمهد الطريق للتكامل المستقبلي لأدوات تحليل الأسنان المدعومة بالذكاء الاصطناعي في ممارسة طب الأسنان الحديثة.

الكلمات المفتاحية : تشخيص الأسنان - كشف التسوس - تحديد الحجم في صورة بانورامية - صورة بالأشعة السينية.

1.1 Introduction

Oral health is a critical aspect of overall well-being, and dental decay remains one of the most widespread oral health problems affecting people globally. Despite advancements in medical technology, the prevalence of dental decay continues to rise, placing significant strain on healthcare systems worldwide. As the need for more accurate and efficient diagnostic methods grows, modern technologies, especially AI, have the potential to revolutionize dental care.

By the year 2017, tooth decay emerged as the most common illness around the globe, impacting more than 3.5 billion people [22, 19]. Even with major progress made in medical technology, the incidence has not diminished, continuing to impose a substantial burden on healthcare systems worldwide. Machine learning, especially neural networks, has experienced extraordinary advancements in the last ten years, achieving performance levels beyond those of humans in numerous tasks since 2015 [15]. The rate of errors in deep learning models has been reduced considerably [33], resulting in their extensive use across multiple industries, such as medical imaging. In certain areas, deep learning models have exceeded human abilities, notably in the detection of breast cancer [38] and diabetes [21].

1.2 Problem Statement

Dental caries stands out as one of the most prevalent dental health concerns globally, impacting individuals across various age groups. Conventional detection techniques, which depend on visual examinations and the manual analysis of X-ray images, frequently result in mistakes and variable interpretations. This may result in delayed treatments, as well as underestimation of the decay. The challenges are aggravated by the cumbersome nature of manual detection especially in a busy clinical environment where dental caries may not be visible to the naked eye. Thus, there is an imminent need for improved, accurate, and reliable means for caries detection and diagnosis. AI, especially in the context of deep learning models, provides an opportunity to revolutionize dental diagnostics in particular, by increasing the accuracy of caries assessments, by minimizing human error, and by simplifying the diagnosis procedure.

1.3 Aim and Objectives

1.3.1 Aim

This project aims to develop an intelligent framework that utilizes artificial intelligence to analyze panoramic dental X-ray images to detect cavities and measure their severity using deep learning and computer vision techniques. The proposed framework aims to enhance detection and diagnostic accuracy compared to traditional methods.

1.3.2 Objectives

To fulfill this aim, the project will pursue the following objectives:

1. **Detection:** Design a model that reliably flags potential carious lesions in panoramic X-ray images by drawing bounding boxes around them.
2. **Segmentation:** Use a semantic segmentation model to create accurate pixel-level masks of areas impacted by caries.
3. **Evaluate** the detection and segmentation models using standard performance metrics such as Intersection over Union (IoU), Average Precision (AP).

this project aims to deliver a comprehensive answer for the early detection of dental caries of tooth decay through the incorporation of cutting-edge AI methods within dental practices. The system will boost the identification of caries and their severity evaluation, leading to increased precision in diagnosis and aiding in more informed healthcare choices.

1.4 Organization of the Document

This paper is divided into sections designed to give a thorough grasp of the research question, the methods used, and the results obtained. The breakdown of the document is:

- Chapter 1 – Introduction: Presents a general view of tooth decay, the issue being addressed, the research goals and aims, plus how the document is organized.
- Chapter 2 – Background: Explores dental anatomy, what causes and how to identify cavities, the tasks of computer vision, and the place of AI and deep learning models in diagnosing dental issues.
- Chapter 3 – Related Work: Analyzes past research and leading approaches for spotting and outlining cavities using deep learning methods.
- Chapter 4 – Methodology: Outlines the data collection utilized, the steps taken for data

preparation, the models implemented for detection and segmentation, as well as the configuration for training.

- Chapter 5 – Results: Provides an overview of the experiments carried out, the assessment criteria utilized, and emphasizes the results achieved in the detection and segmentation methods implemented in the research.
- Chapter 6 – challenges , limitations , Conclusion and Future Work : Examines the primary difficulties faced throughout the research and highlights the limitations associated with the effectiveness of the model. Summarizes the overall findings, highlights the efficiency of the suggested system, and proposes possible avenues for enhancing lesion identification and the implementation of the system in healthcare environments.

This research contributes to AI-driven dental healthcare innovations by improving the detection and assessment of dental caries.

2.1 Results

This segment outlines the results derived from the two main elements of the suggested system: object detection and semantic segmentation. Each element was examined separately with relevant performance indicators to evaluate the model's proficiency in processing panoramic dental X-ray images successfully.

During the object detection phase, the model's effectiveness was judged on its capability to identify possible cavities through the application of bounding boxes. Assessment metrics like AP and AR were applied across various IoU thresholds and object dimensions to gauge the model's precision and reliability amid diverse detection scenarios.

In the segmentation phase, the aim was to attain accurate pixel-wise classification to clearly define the specific shape and size of the carious areas. The segmentation model underwent evaluation utilizing metrics such as PA, Mean Intersection over Union (mIoU), and loss figures for both training and validation datasets.

Collectively, these findings deliver an in-depth understanding of the system's capabilities and its potential to aid in automated dental diagnosis by utilizing radiographic information.

2.1.1 Object Detection Results

In the initial phase of the system, techniques for object detection were utilized to recognize and pinpoint areas of decay in panoramic dental X-ray images. The goal of this stage was to create bounding boxes around possible dental abnormalities utilizing the Faster R-CNN architecture. Using COCO-style metrics, the model's efficacy was evaluated numerically, with an emphasis on precision and recall at various IoU thresholds and object dimensions.

Table 5.1: Object Detection Performance Metrics

Metric	Value
AP@[IoU=0.50:0.95]	0.227
AP@[IoU=0.50]	0.583
AP@[IoU=0.75]	0.095
AP@[IoU=0.50:0.95] (small)	0.0
AP@[IoU=0.50:0.95] (medium)	0.159
AP@[IoU=0.50:0.95] (large)	0.292
AR@[IoU=0.50:0.95] maxDets=1	0.01
AR@[IoU=0.50:0.95] maxDets=10	0.158
AR@[IoU=0.50:0.95] maxDets=100	0.397
AR@[IoU=0.50:0.95] (small)	0.0
AR@[IoU=0.50:0.95] (medium)	0.322
AR@[IoU=0.50:0.95] (large)	0.468

The table object detection results demonstrate the model's overall effectiveness based on multiple evaluation metrics. mAP across IoU thresholds from 0.50 to 0.95 settled at 0.227, suggesting a moderate degree of detection proficiency. Specifically, the average precision when the IoU was 0.50 was notably higher at 0.583, indicating the model's effectiveness when there is a reasonable overlap between the predicted boxes and the actual ground truth. However, under more stringent criteria, the precision decreased sharply, with AP@0.75 dropping to 0.095.

When analyzing performance based on object dimensions, it became clear that the model struggled significantly with smaller objects (AP = 0.0) but exhibited greater proficiency in detecting medium-sized (AP = 0.159) and large (AP = 0.292) items. In terms of recall, the model recorded an AR of 0.397 with the allowance of up to 100 detections, whereas stricter parameters (such as maxDets=1 and maxDets=10) led to diminished recall rates (0.01 and 0.158, respectively). Like the precision, recall saw an improvement as the size of the objects rose, achieving AR values of 0.322 and 0.468 for medium and large objects, accordingly. Together, these findings imply that while the model excels at recognizing moderate and large areas of concern, it faces difficulties when addressing smaller or less prominent dental elements.

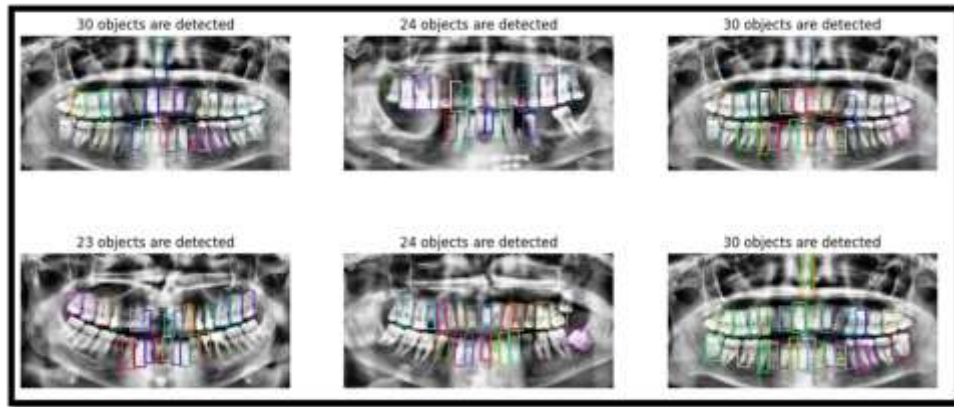


Figure 5.1: Detection Result

Figure 5.1 Sample predictions from the object detection model applied to panoramic dental radiographs. Each colored bounding box indicates a detected tooth, and the total number of detections per image is displayed above. We used all 7,754 photos from the original dataset to train the YOLOv8 model. Although it took a long time to train, the results were disappointing and fell short of our expectations for detection accuracy. We thus rebuilt and reorganized the collection, choosing a more manageable sample of excellent and properly annotated photos. Following that, we used the Faster R-CNN model, which resulted in noticeably better performance. Crucially, these choices were informed by our earlier YOLOv8 tests; the knowledge we acquired from those trials helped us better comprehend the data and improve our strategy as a whole.

2.1.2 Segmentation Result

The subsequent phase of the pipeline concentrated on semantic segmentation, aiming for detailed pixel-level categorization of dental components. Employing the DeepLabV3+ framework, the system was trained to differentiate between tooth areas and the surrounding space in comprehensive panoramic X-ray visuals. The assessment was carried out using both accuracy and overlap-related measures to evaluate the system's proficiency in generating accurate and reliable segmentation masks. The main performance metrics obtained throughout the training and validation phases are listed in the table below.

The DeepLabV3Plus-based semantic segmentation model showcased impressive capabilities throughout the training, validation, and testing stages. After completing training over 10 epochs, the model recorded the following performance indicators: The results presented in Table

Table 5.2: Segmentation Performance Metrics

Metric	Training	Validation
PA	97.2%	95.3%
mIoU	83.8%	79.7%

demonstrate the effectiveness of the DeepLabV3Plus segmentation model. During training, the model achieved a high PA of 97.2%, indicating that the vast majority of pixels were correctly classified. On the validation set, the model maintained a strong PA of 95.3%, reflecting good generalization to unseen data. mIoU scores were also notable, with 83.8% in training and 79.7% in validation. These values highlight the model's ability to accurately delineate the overlap between the predicted and ground truth masks, a critical factor in medical image segmentation tasks.

these metrics confirm that the model performed reliably and with a high degree of accuracy on both training and validation datasets.

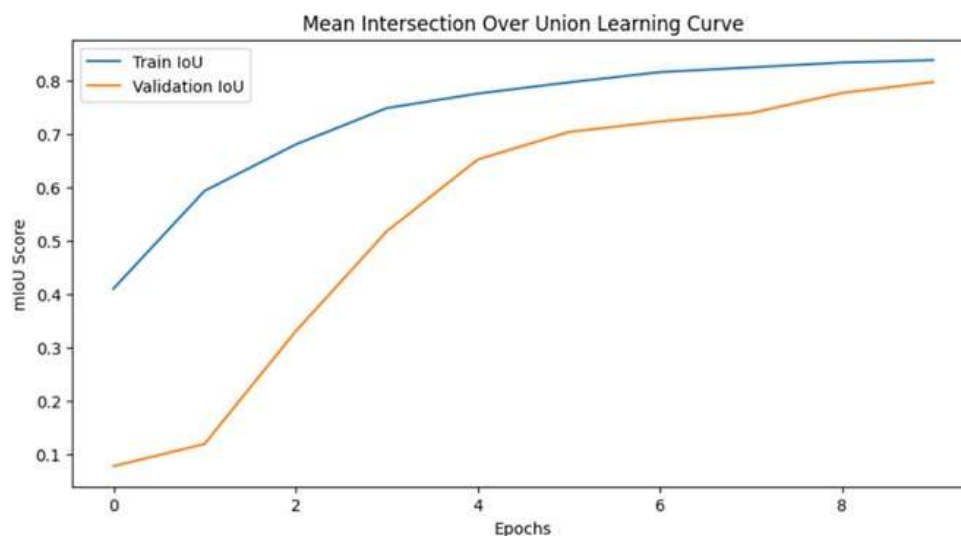


Figure 5.2: IoU curve

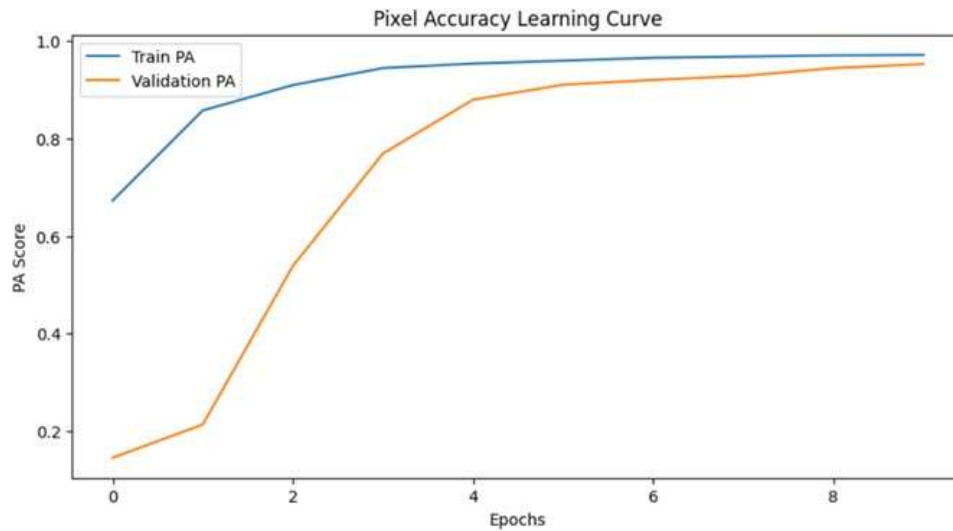


Figure 5.3: Accuracy curve

Figure 5.3 : illustrates PA learning curve for both the training set and validation set over a span of 10 epochs. The PA curve for training displays a swift climb from around 68% to close to 99%, signifying a significant enhancement in accuracy for pixel classification. Likewise, the validation PA curve progressively ascends from roughly 13% to beyond 95%, showcasing the model's successful generalization to new, unseen data. The steady upward movement and the narrow gap between the training and validation curves in the later epochs highlight the model's strength and its capability to effectively differentiate between tooth regions and the background in panoramic X-ray images.

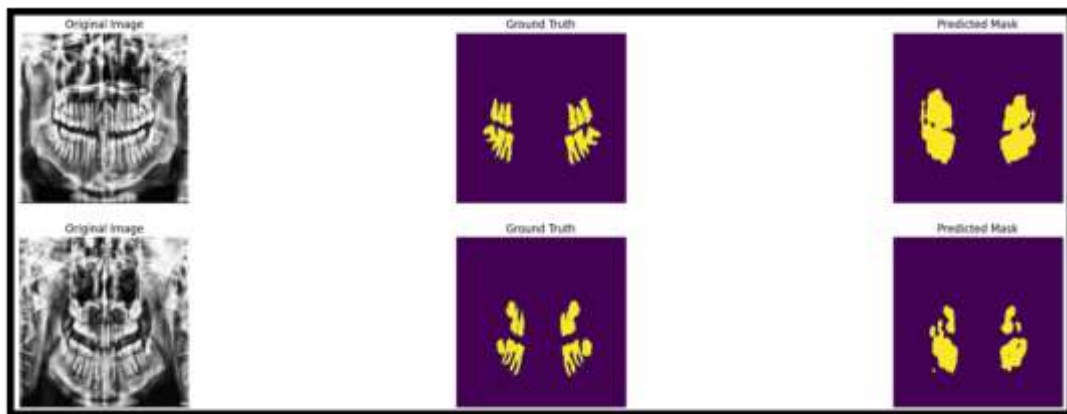


Figure 5.4: Segmentation output

Figure 5.4: Illustration of the outcomes from semantic segmentation on panoramic dental X-rays. Every row showcases the initial image, the related annotated ground truth mask, and the mask

derived from the model's predictions. The masks generated by the model closely match the ground truth, demonstrating the model's capability to accurately recognize and segment distinct tooth structures.

2.2 Experiment

2.2.1 YOLOv8n

In this first experiment, our goal was to evaluate how well a simplified object detection model could recognize dental cavities in panoramic X-rays. For this purpose, we utilized the YOLOv8n architecture, which is a more efficient version of the YOLOv8 detection framework, to pinpoint cavities in the X-ray visuals. The collection included annotated radiographic images prepared for YOLO, where each occurrence of decay was marked with bounding boxes. To suit restricted processing power, the model underwent training for 30 epochs with a batch size of 2, and the input images were scaled to 320 by 320 pixels.

During the training phase, there was a steady decline in box loss, class loss, and distribution focal loss, suggesting effective model learning. As time progressed, metrics saw enhancement, highlighted by rising precision and mAP figures.

Upon completion of the training, the model delivered the following results on the validation dataset.

Table 5.3: Validation Results

Metric	Value
Precision	0.289
Recall	0.124
mAP@0.5	0.425
mAP@0.5–0.95	0.194

The model's results on the validation set are displayed in Table 5.3, where it displays a precision of 0.289, recall of 0.124, mAP@0.5 of 0.425, and mAP@0.5–0.95 of 0.194. These outcomes demonstrate the model's low recall and middling detection accuracy.

Table 5.4: Test Results

Metric	Value
--------	-------

mAP@0.5	0.425
mAP@0.75	0.140
mAP@0.5–0.95	0.194

The model's performance on the test set is summarized in Table 5.4, where it obtained a lower mAP@0.75 of 0.140 and a higher mAP@0.5 of 0.425. With a decrease in precision at higher thresholds, the overall mAP across IoU thresholds from 0.5 to 0.95 was 0.194, suggesting a respectable detection capability.

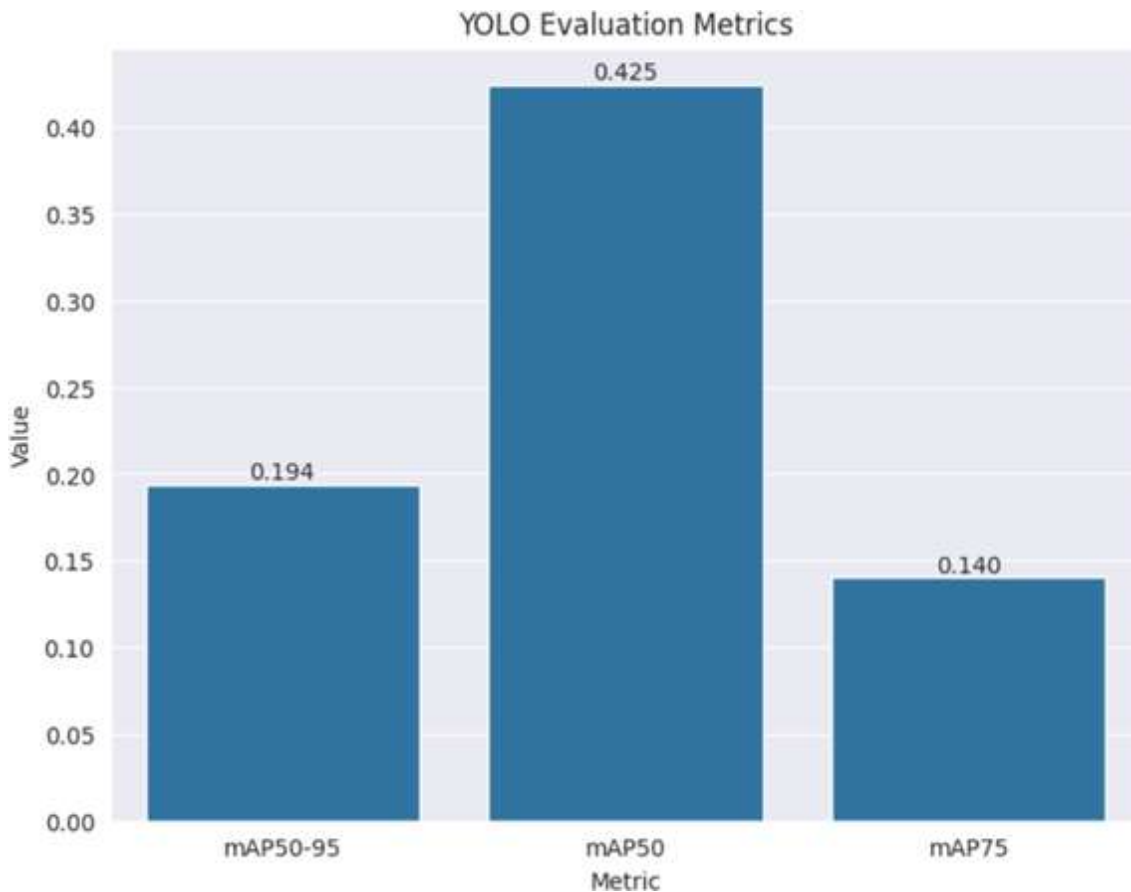


Figure 5.5: YOLOv8n Evaluation Metrics

Figure 5.5 19 Though the model successfully identified some instances of tooth decay, the overall effectiveness was limited. The relatively low recall (0.124) indicates that the system often missed a considerable number of true positive detections. Additionally, while the mAP@0.5 (42.5%) may seem adequate for a lightweight model, the mAP@0.5:0.95 (19.4%) shows restricted generalization across different IoU thresholds, which is crucial for medical tasks where accurate localization is essential.

Though the model successfully identified some instances of tooth decay, the overall effectiveness was limited. The relatively low recall (0.124) indicates that the system often missed a consider-

able number of true positive detections. Additionally, while the $mAP@0.5$ (42.5%) may seem adequate for a lightweight model, the $mAP@0.5:0.95$ (19.4%) shows restricted generalization across different IoU thresholds, which is crucial for medical tasks where accurate localization is essential.

The performance observed can be linked to multiple elements:

- A dataset that could be restricted or unbalanced, potentially hindering the model's capacity to learn various forms of tooth decay.
- A smaller input dimension (320×320), which might result in a loss of significant visual information in radiographs.

In conclusion, this experiment illustrates that YOLOv8n can serve as a rapid and resource-efficient option for initial detection of dental cavities. However, the model's effectiveness, especially regarding recall and $mAP@0.5:0.95$, is not yet optimal for use in clinical settings.

2.2.2 YOLOv8l

The initial experiment with the YOLOv8n model revealed certain constraints—particularly its insufficient recall and restricted generalization when dealing with IoU thresholds—necessitating the adoption of a more sophisticated model to boost detection capabilities. Therefore, YOLOv8l was chosen for the next experiment due to its more extensive and capable architecture, designed to capture faint visual cues of dental caries that the smaller model might overlook. The aim was to enhance both sensitivity and accuracy in localization, which are essential in clinical settings, while also assessing whether a more intricate model could provide significant improvements in performance. To ensure uniformity, the same dataset utilized in the prior experiment—which contained annotated dental X-ray images formatted for YOLO—was utilized once again. The training phase lasted for 30 epochs, employing a batch size of 2 and adjusting input images to 320×320 pixels.

During the training, a steady decline in box loss, classification loss, and distribution focal loss was noted, suggesting effective model alignment. The performance metrics demonstrated consistent advancements in precision, recall, and mAP, as shown in the training charts.

Table 5.5: Training Epoch Metrics

Metric	Value
Precision	0.290
Recall	0.125
mAP@0.5	0.107
mAP@0.5:0.95	0.0337

The model's performance during training is displayed in Table 5.5, which displays a precision of 0.290 and a recall of 0.125. The overall mAP spanning IoU thresholds from 0.5 to 0.95 was 0.0337, whereas the mAP@0.5 was 0.107. These numbers show poor detection performance during training, especially when IoU criteria are more stringent.

Table 5.6: Test Set Evaluation

Metric	Value
Precision (P)	0.762
Recall (R)	0.0818
mAP@0.5	0.425
mAP@0.75	0.140
mAP@0.5:0.95	0.194

The model's evaluation on the test set is summarized in Table 5.6, where it obtained a high precision of 0.762, suggesting reliable positive predictions. But at 0.0818, the recall was noticeably low, indicating low sensitivity. Moderate localization performance was indicated by the mAP scores, which were 0.425 at IoU 0.5, 0.140 at 0.75, and 0.194 across the 0.5–0.95 range.

Figure 5.6 displays these evaluation findings. The mAP@0.5 aligns with the results of the YOLOv8n experiment (0.425), suggesting that the larger model maintained its detection effectiveness at this specific IoU threshold. Nonetheless, significant improvements at elevated IoU thresholds or recall rates were not evident.

The visuals of model outputs indicate that YOLOv8l managed to identify several cases of tooth decay with high confidence levels. However, several false negatives were noted, implying that the model might still be underfitting or limited by the dataset's constraints.

Even though YOLOv8l is structurally more advanced than YOLOv8n, this experiment's outcomes did not demonstrate substantial performance enhancements in comparison to the lighter

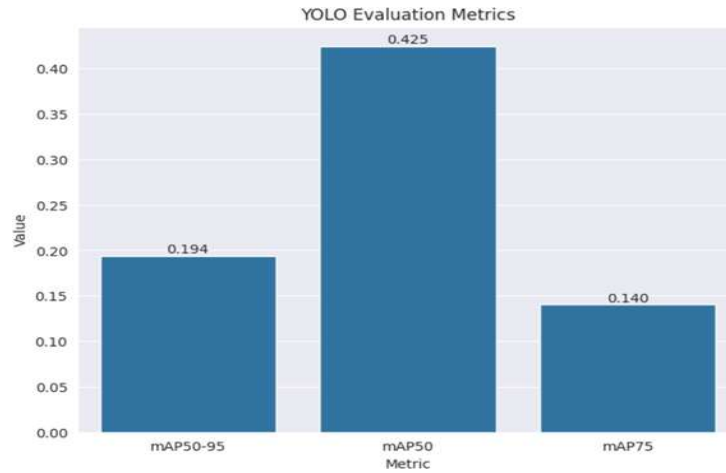


Figure 5.6: YOLOv8l evaluation metrics

model. While test set precision improved, recall remained low, and the overall $mAP@0.5:0.95$ was still restricted. Therefore, simply enlarging model dimensions without improving the quality or resolution of the dataset may be inadequate for achieving clinical-grade detection of dental carie.

Challenges and Limitations

3.1 Challenges

3.1.1 Detection Challenges

Throughout the creation and implementation of the suggested method for detecting and segmenting caries in panoramic dental X-ray images, several challenges and limitations were encountered:

- Even though the more sophisticated and bigger YOLOv8l model was anticipated to provide better outcomes, it did not maintain consistently high precision. Both YOLOv8n and YOLOv8l encountered difficulties in identifying carious lesions of different sizes, especially when there were variations in image quality or anatomy.
- This irregularity highlighted essential shortcomings in their capacity to generalize across various clinical situations, where both detection sensitivity and localization accuracy are essential.

- As a result, these obstacles resulted in the implementation of a more resilient two-stage detection framework and initiated a thorough restructuring of the dataset to improve the model's effectiveness and dependability.
- Although the Faster R-CNN framework demonstrated superior effectiveness relative to single-stage detectors, it nonetheless exhibited various shortcomings. The model faced challenges in precisely identifying small carious lesions or those with minimal contrast, which frequently occur in panoramic dental imaging.
- These difficulties underscore the necessity for additional refinement, especially in improving the model's capacity to manage faint or less noticeable lesions.

3.1.2 Limitations

Significant limitations of this research are related to:

- Dependence on a single dataset acquired from an online source (Roboflow), which lacked critical demographic details including patient age, gender, and clinical history.
- The lack of this metadata hampers the model's capacity to extend its predictions to various populations and actual clinical settings.
- The comparison of models was confined to just two detection frameworks—YOLOv8 and Faster R-CNN—without investigating more sophisticated or contemporary models like DETR or those using transformers, which might have provided enhanced detection precision and resilience.

These shortcomings underscore the necessity for additional research that includes more varied datasets and a wider assortment of detection frameworks to guarantee the scalability and practical relevance of the proposed system.

3.2 Future Work

3.2.1 Research Directions

To boost the efficiency and practical use of the suggested framework, upcoming efforts might concentrate on the following areas:

- Enhance the identification of very tiny carious lesions, which existing models struggle to detect efficiently. This could entail utilizing images with higher resolution and improving the precision of the annotations.
- Open up more detection architectures (including the leading models such as DETR, YOLOv9 and the transformer based detector) for experiment and assessment on more challenging tasks.
- Use ensemble learning methods that combine strengths of multiple models to enhance robustness and accuracy across multiple lesion and imaging types.
- Broaden the dataset by integrating clinically verified images from multiple sources, as well as incorporating metadata details like the patient's age, gender, and dental background, to enhance generalizability and clinical applicability.
- Adopt explainable AI techniques (XAI) to help dental professionals make clear decisions and encourage the use of these approaches in actual clinical settings.
- Create an intuitive graphical interface (GUI) that enables dental practitioners to engage with the detection and segmentation system with greater ease. This interface can showcase X-ray images, emphasize areas of decay, and offer clinical measurements, enhancing the system's usability and applicability in practical settings.

3.2.2 Conclusion

This research introduced a framework utilizing deep learning for the automated detection and segmentation of dental caries within panoramic X-ray images. The project focused on two primary objectives: detecting objects and performing semantic segmentation.

- During the detection phase, both YOLOv8n and YOLOv8l were tested initially but showed subpar results, especially regarding the accuracy of localization and recall rates. Consequently, the Faster R-CNN model was utilized, thanks to its superior accuracy and

capacity to identify lesions of different sizes, although it faced challenges with detecting smaller cavities.

- For the segmentation aspect, DeepLabV3+ with a ResNet34 backbone was employed, demonstrating exceptional ability in accurately marking the precise borders of the decayed areas. The model achieved impressive pixel-level accuracy and IoU, signifying trustworthy segmentation results.

In summary, the suggested system was effective in recognizing and segmenting carious lesions, with particular success noted for moderate to larger cavities. Nonetheless, additional enhancements are needed to better detect very small lesions, increase computational efficiency, and improve the model's adaptability. The results affirm the promise of deep learning in progressing automated dental diagnostics and set the stage for broader practical uses in clinical environments.

Bibliography

- [1] Alharbi, S. S., AlRugaibah, A. A., Alhasson, H. F. and Khan, R. U. [2023], 'Detection of cavities from dental panoramic x-ray images using nested u-net models', *Applied Sciences* **13**(23), 12771.
- [2] Asci, E., Kilic, M., Celik, O., Cantekin, K., Bircan, H., Bayrakdar, I. and Orhan, K. [2024], 'A deep learning approach to automatic tooth caries segmentation in panoramic radiographs of children in primary dentition, mixed dentition, and permanent dentition', *Children* **11**, 690.
- [3] Aslam, A. and Curry, E. [2020], 'A survey on object detection for the internet of multimedia things (iomt) using deep learning and event-based middleware: Approaches, challenges, and future directions', *Sensors* **20**(20), 5796.
- [4] Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., Zhang, Z., Cheng, D., Zhu, C., Cheng, T., Zhao, Q., Li, B., Lu, X., Zhu, R., Wu, Y., Dai, J., Wang, J., Shi, J., Ouyang, W., Loy, C. C. and Lin, D. [2019], 'Mmdetection: Open mmlab detection toolbox and benchmark'.
- [5] Code, P. W. [2023], 'Weight decay explained'.
URL: <https://paperswithcode.com/method/weight-decay>
- [6] contributors, W. [2025], 'Loupe'. Online Reference.
URL: <https://en.wikipedia.org/wiki/Loupe>
- [7] *Convolutional layer* [2023].
URL: https://en.wikipedia.org/wiki/Convolutional_neural_networkConvolutionalayer
- [8] *Dental Caries: The Disease and Its Clinical Management* [2015], BLACKWELL PUBL. Online Reference.
URL: https://www.ebook.de/de/product/23695989/dental_caries_the_disease_and_its_clinical_management
- [9] *Dental X-rays* [2023]. Online Reference.
URL: <https://my.clevelandclinic.org/health/articles/11199-dental-x-rays>
- [10] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. and Houlsby, N. [2020], 'An image is worth 16x16 words: Transformers for image recognition at scale'.
- [11] Everingham, M. et al. [2010], 'The pascal visual object classes (voc) challenge', *IJCV* **88**(2), 303–338.
- [12] *F-score* [2023].
URL: <https://en.wikipedia.org/wiki/F-score>

- [13] Haghanifar, A., Majdabadi, M. M. and Ko, S.-B. [2020], ‘Paxnet: Dental caries detection in panoramic x-ray using ensemble transfer learning and capsule classifier’, *arXiv* .
- [14] He, K., Zhang, X., Ren, S. and Sun, J. [2015a], ‘Deep residual learning for image recognition’.
- [15] He, K., Zhang, X., Ren, S. and Sun, J. [2015b], Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, in ‘Proceedings of the IEEE International Conference on Computer Vision (ICCV)’.
- [16] Huang, X., Shen, Z. and Sun, Y. [2023], ‘Artificial intelligence in oral health: A scoping review’, *Heliyon* **9**(9), e19796.
- [17] *Intersection over Union (IoU) Definition* [2023].
URL: <https://encord.com/glossary/iou-definition/>
- [18] Ioffe, S. and Szegedy, C. [2015], Batch normalization: Accelerating deep network training by reducing internal covariate shift, pp. 448–456.
URL: <https://proceedings.mlr.press/v37/ioffe15.html>
- [19] James, D. and Abate, S. L. [2018], ‘Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017’, *The Lancet* **392**(10159), 1789–1858.
- [20] Jordan, J. [2023], ‘Semantic segmentation losses’.
- [21] Kamble, T. P. and Patil, S. T. [2016], ‘Diabetes detection using deep learning approach’, *International Journal for Innovative Research in Science Technology* .
- [22] Kassebaum, N. J., Bernabé, E., Dahiya, M., Bhandari, B., Murray, C. J. L. and Marcenes, W. [2015], ‘Global burden of untreated caries: A systematic review and metaregression’, *J Dent Res* **94**(5), 650–658.
URL: <https://doi.org/10.1177/0022034515573272>
- [23] Kunt, L. [2023], ‘Dental caries detection from bitewing x-ray images’, *Dental* .
- [24] Labs, V. [2023a], ‘Coco dataset guide: Everything you need to know’.
URL: <https://www.v7labs.com/blog/coco-dataset-guide>
- [25] Labs, V. [2023b], ‘Intersection over union (iou) guide’.
URL: <https://www.v7labs.com/blog/intersection-over-union-guide>
- [26] Labs, V. [2023c], ‘Mean average precision (map) explained: Everything you need to know’.
URL: <https://www.v7labs.com/blog/mean-average-precision>
- [27] Lal, V., Lara, J. S. and Olegário, I. C. [2022], ‘Visual examination for caries detection using merged-icdas stages’, *Journal of the Irish Dental Association* **68**(6), 318–319.
- [28] Li, Y. and Yang, T. [2018], Word embedding for understanding natural language: A survey, in S. Srinivasan, ed., ‘Guide to Big Data Applications’, Springer International Publishing, Cham, pp. 83–104.
URL: https://doi.org/10.1007/978-3-319-53817-4_4

- [29] Lin, T.-Y., Goyal, P., Girshick, R., He, K. and Dollár, P. [2017], ‘Focal loss for dense object detection’, *arXiv* **1708.02002**.
- [30] *Loss Functions - Cross-Entropy* [2023].
URL: https://ml-cheatsheet.readthedocs.io/en/latest/loss_functions.html
- [31] Mărginean, A. C., Muresanu, S., Hedeşiu, M. and Diosan, L. [2024], ‘Teeth segmentation and carious lesions segmentation in panoramic x-ray images using cariseg, a networks’ ensemble’, *Heliyon* **10**, e30836.
- [32] Padilla, R., Passos, W. L., Dias, T. L. B., Netto, S. L. and da Silva, E. A. B. [2021], ‘A comparative analysis of object detection metrics with a companion open-source toolkit’, *Electronics* **10**(3), 279.
- [33] Papers with Code [2023], ‘Browse the state-of-the-art in machine learning’. Online Reference.
URL: <https://paperswithcode.com/sota>
- [34] *Precision and recall* [2023].
URL: https://en.wikipedia.org/wiki/Precision_and_recall
- [35] PUBL, B. [2015], ‘The disease and its clinical management’, *Denta Caries* .
URL: https://www.ebook.de/de/product/23695989/dental_caries_the_disease_and_its_clinical_management.
- [36] Redmon, J., Divvala, S. K., Girshick, R. B. and Farhadi, A. [2015], ‘You only look once: Unified, real-time object detection’, *CoRR* **abs/1506.02640**.
URL: <http://arxiv.org/abs/1506.02640>
- [37] Ren, S., He, K., Girshick, R. and Sun, J. [2015], ‘Faster r-cnn: Towards real-time object detection with region proposal networks’.
- [38] Rodriguez-Ruiz, A., Lång, K., Gubern-Merida, A., Broeders, M., Gennaro, G., Clauser, P., Helbich, T. H., Chevalier, M., Tan, T., Mertelmeier, T., Wallis, M. G., Andersson, I., Zackrisson, S., Mann, R. M. and Sechopoulos, I. [2019], ‘Stand-alone artificial intelligence for breast cancer detection in mammography: Comparison with 101 radiologists’, *JNCI: Journal of the National Cancer Institute* **111**(9), 916–922.
- [39] Ronneberger, O., Fischer, P. and Brox, T. [2015a], ‘U-net: Convolutional networks for biomedical image segmentation’, *arXiv* **1505.04597**.
URL: <https://arxiv.org/abs/1505.04597>
- [40] Ronneberger, O., Fischer, P. and Brox, T. [2015b], ‘U-net: Convolutional networks for biomedical image segmentation’.
- [41] Scikit-learn [2023], ‘Plot precision-recall curve’.
URL: https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html
- [42] Sharma, R., Saqib, M., Lin, C. T. and Blumenstein, M. [2022], ‘A survey on object instance segmentation’, *SN Computer Science* **3**(499).
- [43] Strassler, H. and Pitel, M. [2014], ‘Using fiber-optic transillumination as a diagnostic aid in dental practice’, *Compendium of continuing education in dentistry (Jamesburg, N.J. : 1995)* **35**, 80–88.

- [44] Sturdevant [2019], *Art and Science of Operative Dentistry*, Elsevier.
- [45] Tan, M. and Le, Q. V. [2019], 'Efficientnet: Rethinking model scaling for convolutional neural networks', *CoRR* **abs/1905.11946**.
URL: <http://arxiv.org/abs/1905.11946>
- [46] Tan, M., Pang, R. and Le, Q. V. [2020], Efficientdet: Scalable and efficient object detection, in 'Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)'.
- [47] Tenyks [2023], 'Mean average precision (map): Common definition, myths and miscon- ceptions'.
- [48] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. and Polosukhin, I. [2017], 'Attention is all you need'.
- [49] Wu, Y. and He, K. [2018], 'Group normalization'.
- [50] Yin, K. [2020], 'Sign language translation with transformers'.
- [51] Zou, Z., Shi, Z., Guo, Y. and Ye, J. [2019], 'Object detection in 20 years: A survey', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **43**(3), 1070–1088.