Žiga Bizjak, Tina Robič

Advancing Dental Caries Detection in Panoramic X-rays: A Two-Step Deep Learning Approach

Abstract. Caries, preventable but common oral disease, can cause tooth pain and tooth loss. Early detection is key for timely treatment. To improve caries detection, we employed a two-step deep learning algorithm, specifically adapting the YOLOv8 computer vision model to analyse panoramic X-rays. Firstly, the Teeth Enumeration Model (TEM) is trained to enumerate teeth following the Fédération Dentaire Internationale system. Subsequently, the Caries Detection Model (CDM), built upon the TEM, focuses on detecting caries lesions. We used clinically relevant metrics to evaluate the model, including lesion-level sensitivity, patient-level specificity, and number of false positives per image. The TEM demonstrated strong performance, achieving a high mAP0.5 value of 0.95, indicating accurate tooth labelling. The subsequent CDM exhibited satisfactory lesion-level sensitivity in both internal (0.70) and external (0.67) validations. A comparison with previous studies, including the DENTEX challenge winner, highlights the superiority of the proposed approach. This research contributes to the advancement of dental caries detection through a robust algorithm and emphasises the potential of YOLOv8 in automating tooth labelling and caries detection.

Key words: dental caries; deep learning; computer vision; algorithms; DENTEX database.

Odkrivanje zobnega kariesa na ortopantomogramih: pristop v dveh korakih z uporabo globokega učenja

Povzetek. Karies, ki je preprečljiva, a pogosta bolezen ustne votline, lahko povzroči bolečine in izgubo zob. Njegovo zgodnje odkrivanje je ključno za pravočasno zdravljenje. Za izboljšanje odkrivanja kariesa na panoramskih rentgenskih posnetkih smo uporabili dvostopenjski algoritem globokega učenja, ki temelji na prilagojenem modelu YOLOv8. V prvem koraku smo naučili in uporabili model za označevanje zob (TEM) po sistemu Mednarodne zveze zobozdravnikov. V drugem koraku pa smo na osnovi modela TEM naučili še model za odkrivanje kariesa (CDM), ki je osredotočen na odkrivanje karioznih lezij. Za ocenjevanje modela smo uporabili klinično relevantne mere, vključno z občutljivostjo na ravni lezije, specifičnostjo na ravni pacienta in številom lažno pozitivnih primerov na sliko. Model TEM je pokazal močno uspešnost, saj je dosegel visoko vrednost mAP0.5 (0,95), kar kaže na natančno označevanje zob. Model CDM je pokazal zadovoljivo občutljivost na ravni lezije tako v notranjih (0,70) kot v zunanjih (0,67) validacijskih množicah. Primerjava s prejšnjimi študijami, vključno z zmagovalcem izziva DENTEX, kaže na superiornost predlaganega pristopa. Raziskava prispeva k napredku pri avtomatskem odkrivanju zobnih kariesov z uporabo robustnega algoritma in poudarja potencial YOLOv8 pri avtomatizaciji označevanja zob in odkrivanju kariesa.

Ključne besede: karies; globoko učenje; računalniški vid; algoritmi; YOLOv8, podatkovna zbirka DENTEX.

Introduction

Although dental caries is entirely preventable and treatable, it continues to persist as a prevalent source of tooth pain and eventual tooth loss. The timely detection of caries is crucial for effective intervention. Employing methods such as visual inspection and probing is essential for accurately identifying visible cavities at an early stage.[1] In dentistry, imaging methods are crucial for diagnosing dental issues effectively. One commonly used technique is X-ray radiography, which provides detailed images of the teeth and surrounding structures. Among these methods, panoramic imaging, which is also used in this study, is particularly valuable because it offers a broad view of the entire mouth area in a single image. However, despite its usefulness, panoramic imaging may not always reveal hidden lesions or problems. On the other hand, bitewing radiography, which focuses on specific areas, is highly sensitive but may not capture all aspects of a dental caries, especially those concealed beneath the surface.[2]

In the field of dentistry, the interpretation of panoramic radiographs is vital, requiring thorough and extensive training.[3] Recent study shows that experienced dentists outperform novices in accurately assessing caries lesions.[4] The utilization of machine learning, especially convolutional neural network (CNN) models, is gaining traction in automating the detection of dental diseases. These advanced algorithms showcase robust capabilities in accurately segmenting dental structures and identifying lesions. By leveraging these technologies, dentists can improve their ability to detect all caries accurately, thereby minimising the risk of overlooking any cavities during diagnosis.[5]

Innovations in medical imaging deep learning encompass the use of specialised models such as U-Net and CNNs, which have been applied to achieve precise detection of conditions like alveolar bone loss, apical cysts, and caries lesions.[6] However, a recent literature review has highlighted a notable gap in research attention towards caries detection specifically in panoramic X-rays, indicating an area that requires further exploration and study.[7]

A notable challenge in deep learning approaches is the insufficient availability of training data.[6,8] The DENTEX challenge successfully addressed this by introducing a publicly accessible dataset of panoramic X-rays, hierarchically annotated with quadrant-enumeration-diagnosis information.

Our study makes several contributions: firstly, we modify a well-known computer vision model to detect cavities in panoramic X-rays. Secondly, we utilise a two-step learning approach. Thirdly, we assess the model's effectiveness with a real-world clinical dataset. Finally, we introduce novel evaluation metrics inspired by clinical practices to improve assessment within this field.

Methods

In the realm of computer vision, a multitude of robust models, such as U-Net[5] for segmentation and YOLO (You Only Look Once)[9] for object detection, have demonstrated exceptional proficiency in addressing tasks involving 2D images. Specifically, U-Net stands out as a seminal segmentation model within the medical imaging domain, renowned for its efficacy in segmenting structures of interest within images. On the other hand, YOLO represents an end-to-end object detection framework, designed to swiftly and accurately predict class labels for individual pixels or bounding boxes encompassing objects of interest within an image. These models, through their distinct architectures and methodologies, have significantly advanced the capabilities of computer vision systems in various domains, including medical imaging and general object recognition tasks.
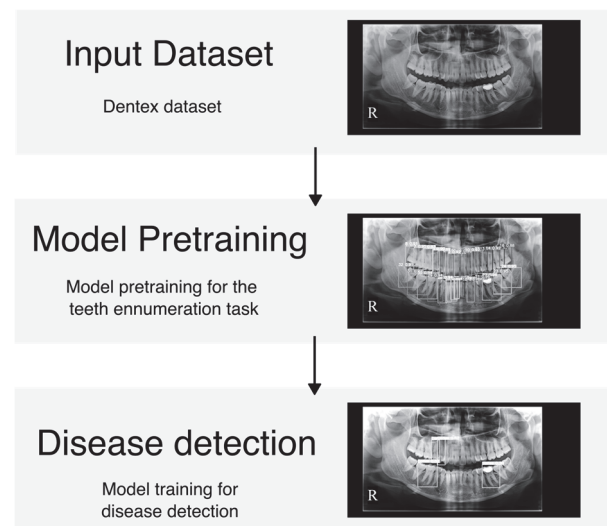

Figure 1 Our two-step deep learning approach using YOLOv8 for precise dental caries detection in panoramic X-rays.

Our approach consists of two main stages: tooth enumeration and disease detection (Figure 1). Firstly, we pretrain the model on a large dataset of dental images with annotated tooth locations to accurately localise teeth. This pretraining process enhances the

model's ability to recognise variations in tooth morphology and appearance. Following pretraining, we proceed to disease detection, where we assess the presence of pathologies within each identified tooth bounding box. This methodology aims to improve the accuracy and reliability of dental condition diagnosis.

Data

In this study we used two separate datasets, one was publicly available dataset DENTEX which was used for training and validation of our model, and second was external on-site dataset that was used for external validation.

The DENTEX dataset comprises panoramic X-rays from three institutions, reflecting diverse clinical practices. It includes three hierarchically organised types of data: (a) 693 quadrant-labelled X-rays, (b) 634 tooth-labelled X-rays with quadrant and enumeration classes, and (c) 1005 fully labelled X-rays. The dataset includes a total of 2681 teeth affected by caries. All annotations are meticulously crafted by three dental experts, ensuring high-quality and accurate data for dental research. Figure 2 displays annotated panoramic X-rays from the DENTEX dataset, featuring tooth labels and disease bounding boxes essential for training our two-step dental caries detection model. The dataset lacks additional information, including scanner details, patient age, and sex.
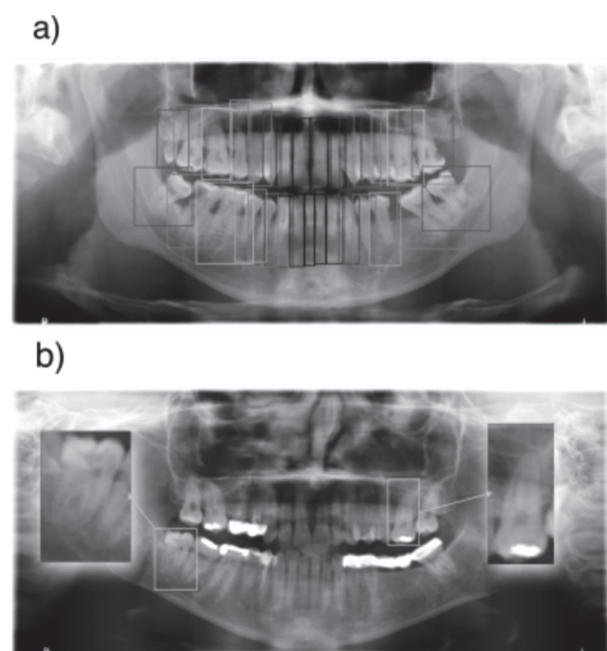

Figure 2 DENTEX Dataset – annotated panoramic X-rays showcasing tooth labels (a) and disease bounding boxes (b).

The second dataset is an external dataset comprising 85 panoramic X-rays from 44 women and 41 man. All of those were collected on the VA TECH FLEX 3D during routine clinical inspections. The median age of the individuals in this external dataset was 54. The dataset comprised 20 individuals without caries and 65 individuals with caries, totalling 93 teeth affected by caries.

YOLO architecture

The YOLO model[9] is a real-time object detection model known for its efficiency and accuracy. Unlike traditional methods that apply region proposal networks (R-CNN) or sliding window approaches, YOLO reframes object detection as a single regression problem, streamlining the detection process into a single network evaluation.

The model architecture is a unified convolutional neural network (CNN) model designed for efficient object detection. It typically takes a 448×448 pixel input image and processes it through a series of convolutional layers for feature extraction. The standard YOLO model includes 24 convolutional layers followed by 2 fully connected layers. The convolutional layers use 1×1 and 3×3 filters with max pooling to capture spatial hierarchies and reduce dimensionality. The final layers convert the extracted features into detection results, producing both bounding box coordinates and class probabilities. This design leverages global context and spatial positioning to achieve rapid and accurate detection, consolidating the entire process into a single network evaluation without the need for multiple stages or region proposals.

Two step deep learning approach

To facilitate the detection of caries, we devised a two-step deep learning approach leveraging YOLOv8. Our methodology commenced with the training of the Teeth Enumeration Model (TEM), which was tailored to adhere to the Fédération Dentaire Internationale (FDI) system for tooth enumeration. Each tooth was allocated a distinct bounding box characterised by coordinates (x, y, w, h), where x and y represented the centre of the box, while w and h denoted its width and height, respectively. While we integrated basic augmentation techniques to refine training, we intentionally omitted flipping in this step due to the significant influence of image orientation on tooth enumeration accuracy. This phase involved rigorous training of the TEM over 2000 epochs, utilising a dataset comprising 551 images sourced from the DENTEX dataset.

The pretrained TEM formed the cornerstone for our subsequent task: caries detection. Initially, the DENTEX dataset provided four labels – caries, deep caries, periapical lesion, and impacted. However, for the purposes of our study, we narrowed our focus exclusively to the labels for caries and deep caries. Once again, each instance of caries was outlined by a bounding box specified by (x, y, w, h), where x and y denoted the centre of the box, w indicated its width, and h specified its height. Augmentation for the Caries Detection Model (CDM) was comprehensive, incorporating techniques such as rotation, scaling, shear, flip, and others. The CDM underwent extensive training for 2000 epochs, utilising the pre-trained architecture from the preceding phase and the same set of 551 images from the DENTEX dataset, encompassing a total of 2112 teeth affected by caries. During training, we ensured consistency in the train/test split for both the CMD and TEM models. A schematic overview of our methodology is provided in Figure 1.

Evaluation metrics

Conventional evaluation metrics like precision, recall, and area under ROC curve (TPR/FPR) are insufficient for capturing the nuances of this complex task. Drawing inspiration from the evaluation proposed for intracranial aneurysm detection,[10] we utilised lesion-level sensitivity, patient-level specificity, and the number of false positives per image (FPs/image) as our primary evaluation criteria.

Lesion-level sensitivity is calculated by determining the proportion of correctly identified caries instances relative to the total number of caries instances within the dataset. Meanwhile, patient-level specificity evaluates the model's precision in labelling healthy subjects, achieved by dividing the accurately labelled scans by the total number of scans devoid of caries. It's imperative to include healthy subjects in our evaluation to accurately gauge the model's capacity to identify cases without caries. Correctly labelled healthy scans are those without any false positive findings.

Furthermore, the number of FPs/image metric offers insights into the rate of false positive caries detections across the entire dataset, normalised by the total number of scans portraying caries. This metric is particularly valuable for understanding the occurrence of false positives and complements patient-level specificity, especially when analysing scans of healthy subjects.

To ensure compatibility with existing literature, we also employ Mean Average Precision at IoU 0.5 (mAP$_{0.5}$). This metric calculates the mean average precision at a specified intersection-over-union (IoU) threshold of 0.5, providing a comprehensive evaluation of the precision-recall curve across different confidence thresholds:

$$mAP_{0.5} = \frac{1}{N}\sum_{i=1}^{N} AP_i,$$

where (AP_i) is the average precision for class (i) at IoU 0.5, and (N) is the number of classes. Classes in our case are deep and normal caries.

Results

The results are separated for both steps, results of pretrained model TEM and CDM, however it's worth noting that the primary focus lies on the evaluation of the CDM. The TEM exhibited robust performance, achieving a mAP$_{0.5}$ score of 0.9471, underscoring its accuracy in labelling various classes. Precision and recall values on the validation dataset reached 0.91 and 0.92, respectively, further highlighting the model's efficacy. Visual representations of the pretrained model's results are showcased in Figure 3, providing a tangible illustration of its performance.
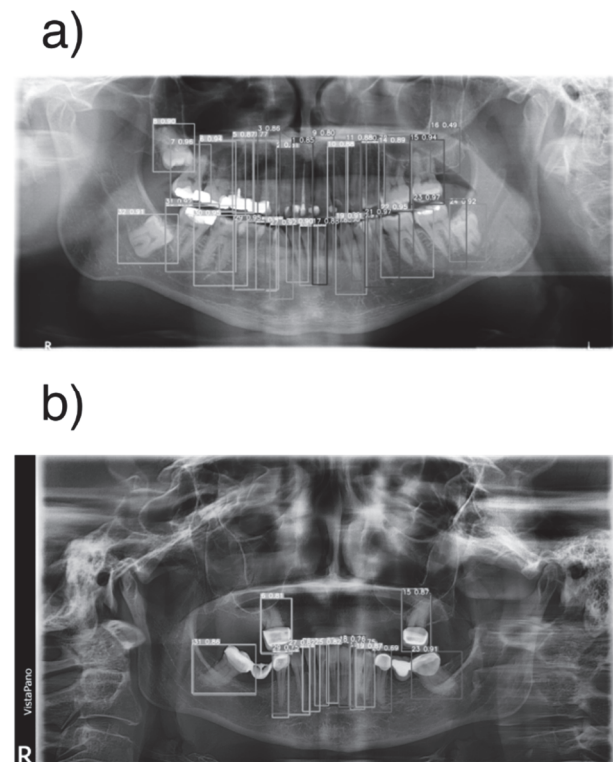
Figure 3 Visual results of the Teeth Enumeration Model (TEM) demonstrating accurate tooth labeling in diverse scenarios, including (a) a complete set of teeth and (b) a case with multiple missing teeth, showcasing the adaptability of the deep learning model.

Subsequently, the CDM underwent a training phase leveraging the insights gained from the TEM model. Upon evaluation using the internal validation dataset, the CDM exhibited a noteworthy lesion-level sensitivity of 0.704, coupled with a patient-level specificity (TNR) of 0.44, and a calculated number of false positives per image at 0.8. Transitioning to an external dataset sourced from a private clinic, the CDM maintained a competitive lesion-level sensitivity of 0.670, while holding a consistent TNR of 0.42. Additionally, a marginal increase in the false positives per image was observed. Remarkably, our achieved sensitivities on both internal and external datasets surpassed those of the DENTEX challenge winner, who reported a sensitivity of 0.54 while utilising the same dataset for training and validation, as outlined in Table 1. It's important to highlight that despite the promising results published in literature, none of the authors provided insights into clinically significant metrics such as FPs/image or TNR.

Figure 4 provides a comprehensive visual representation of the outcomes generated by the CDM, illustrating various scenarios encountered during the detection process: (a) successful detection of all caries instances is depicted by green bounding boxes, indicating accurate identification; (b) in one scenario, a single caries instance was missed, denoted by a black bounding box, while seven others were correctly identified; (c) another scenario showcases the detection of two caries instances with accurate identification, while two were missed (highlighted in black bounding boxes), alongside one false positive detection depicted in a white bounding box; (d) lastly, a scenario where all caries instances were successfully detected, with only one false positive detection highlighted in white, providing a nuanced understanding of the model's performance across different scenarios.

Table 1 Comparative analysis of our findings against state-of-the-art results.

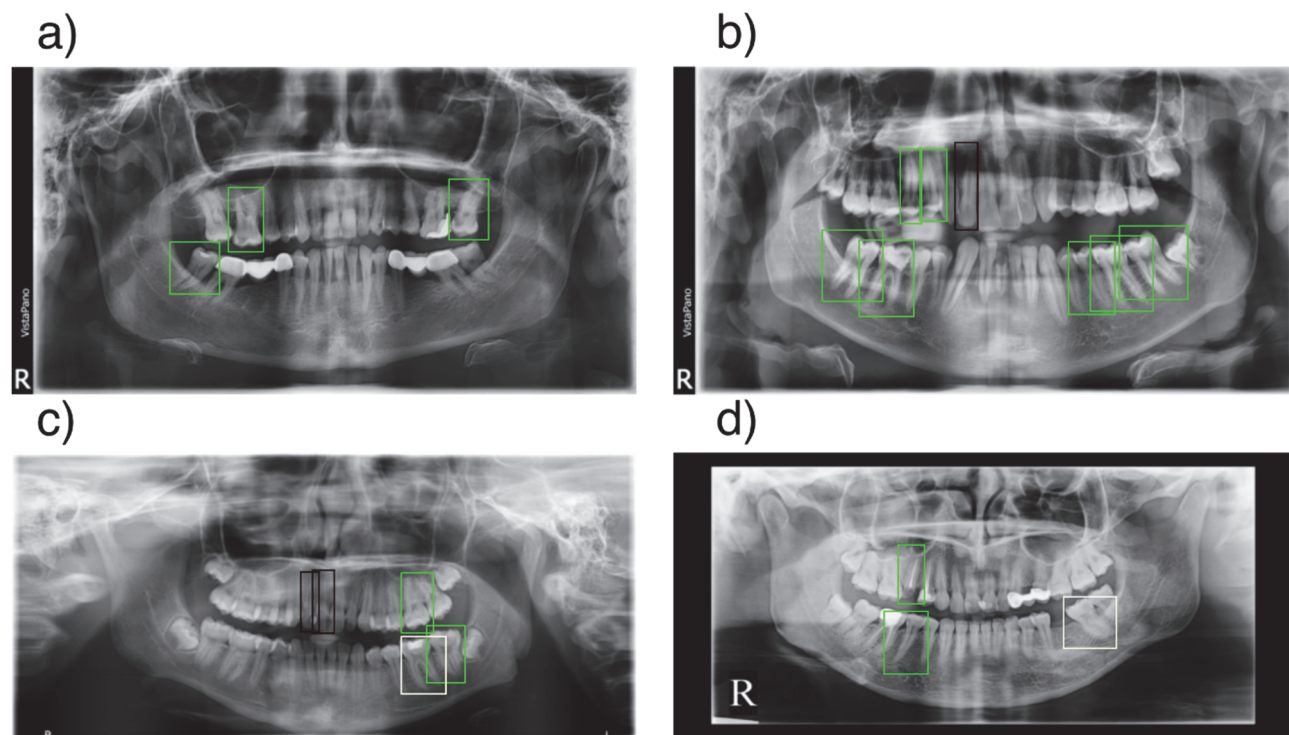| Method	Sensitivity	TNR	FPs/image
PaxNet	0.51	NR	NR
He et al. (DENTEX winner)	0.54	NR	NR
CDM (internal validation)	0.70	0.44	0.80
CDM (external validation)	0.64	0.42	0.92

Legend: NR- not reported

Figure 4 Visual results of the caries detection model (CDM) depicting the model's performance with green bounding boxes representing True Positives (TP), white for False Positives (FP), and black for False Negatives (FN) in detecting dental caries on panoramic X-rays.

Discussion

Automatic detection plays a pivotal role in the realm of dental caries detection on panoramic X-rays. It provides substantial benefits by simplifying and standardising the evaluation process. By minimising reliance on subjective manual inspection methods, these systems ensure a more objective and consistent evaluation process. Furthermore, automatic detection effectively mitigates intra- and inter-rater variability, which can lead to uniform assessment of caries across diverse clinical environments and among different clinicians. Moreover, these systems can function as invaluable secondary reviewers, carefully checking to make sure no tooth caries is missed. In our endeavour to automate dental caries detection, we adapted the renowned computer vision model YOLOv8[9] for analysis of panoramic X-rays. We employed a two-step learning process and rigorously evaluated the trained model both on internal validation and on external clinical dataset. Additionally, we introduced innovative clinically inspired evaluation metrics to enhance the assessment of dental diseases.

Our deep learning method, trained in two steps, concentrates on identifying teeth and detecting caries, leverages the capabilities of established models such as YOLOv8. The trained Teeth Enumeration Model (TEM) demonstrated impressive performance, boasting a high mean Average Precision at Intersection over Union 0.5 ($mAP_{0.5}$) score of 0.9471, indicating precise labelling across most teeth. Despite teeth labelling not being the primary focus of our study, its outstanding performance is noteworthy. Following this, the subsequent Caries Detection Model (CDM), trained based on the TEM model, showed satisfactory sensitivity in detecting lesions on the internal validation dataset. However, there was a slight decline in sensitivity observed when applying the model to an external validation dataset from a private clinic, highlighting the challenges associated with generalising model performance across varying clinical environments. Notably, our approach achieved higher sensitivity on both internal (0.704) and external (0.670) datasets compared to the winner of the DENTEX challenge[11] (0.543), who utilised the same dataset.

Training a model to count teeth before using it for disease detection is crucial for several reasons. First, by exposing the model to a large dataset of dental images with annotated tooth locations, it becomes adept at recognising and accurately locating teeth, laying a strong foundation for subsequent disease detection tasks. This pretraining phase is essential for ensuring the model can effectively handle the diverse variations in tooth morphology and appearance, influenced by factors such as age, dental health, and imaging conditions. Through exposure to diverse examples during pretraining, the model develops enhanced generalization capabilities, allowing it to adeptly adapt to a spectrum of scenarios encountered during disease detection. Overall, pretraining for teeth enumeration fortifies the model's capacity to accurately detect and localise teeth within dental images, establishing a robust framework for subsequent tasks such as disease detection. This thorough method leads to better results and trustworthiness in diagnosing dental problems, which helps both dentists and patients.

Our methodological innovation extends to the comprehensive integration of advanced metrics, encompassing lesion-level sensitivity, patient-level specificity, and number of false positives per image (FPs/image), facilitating a multifaceted evaluation of our model's performance in clinical contexts. In contrast to previous studies such as PaxNet,[12] which predominantly prioritised high sensitivity without commensurate consideration for specificity or false positive rates, our investigation underscores the imperative of harmonising sensitivity with specificity and number of FPs/image for pragmatic clinical applicability. While high sensitivity remains a main attribute, its optimization necessitates a delicate equilibrium with specificity to prevent unwarranted false positives that could compromise the clinical efficacy of the model. Notably, a model's utility is predicated on its capacity to detect fewer than two false positives per image, thereby ensuring precise diagnoses devoid of undue alarm. Studies that prioritise high specificity based on teeth characteristics might unintentionally increase the number of false positives. For example, even with a claimed 90 % specificity,[13] there could still be around three false positives per image, which ultimately undermines the usefulness of these metrics in real clinical settings. In our study, although our sensitivity is slightly lower, we ensure that the rate of false positives per image remains below 1 through careful attention to detail. By carefully navigating the delicate balance between sensitivity and specificity, we uphold the clinical reliability of our model. This nuanced approach reflects our unwavering commitment to developing a practical and effective tool for detecting dental caries in real-world clinical environments.

In summary, our investigation significantly advances the domain of dental caries detection by harnessing the power of deep learning, effectively tackling longstanding challenges in conventional diagnostic methods. The promising outcomes of our study point

towards notable enhancements in automated disease diagnosis, signalling a transformative shift in clinical practice. Furthermore, our findings highlight the imperative for standardised and enhanced metric reporting in dental imaging studies, facilitating improved comparability among research findings and driving forward the collective understanding of the field. As we look towards the future, it becomes increasingly vital for research endeavours to prioritise the establishment of robust evaluation standards, thereby laying the groundwork for more effective clinical applications and ultimately improving patient outcomes.

References

1. Gill J: Dental caries: the disease and its clinical management, third edition. *Br Dent J* 2016; 221(8): 443-443. https://doi.org/10.1038/sj.bdj.2016.767

2. Kaur R, Sandhu RS, Gera A, Kaur T: Edge detection in digital panoramic dental radiograph using improved morphological gradient and MATLAB. In: *2017 International conference on smart technologies for smart nation (SmartTechCon)*; Bengaluru, India 2017: IEEE; 793-797. https://doi.org/10.1109/SmartTechCon.2017.8358481

3. Wirtz A, Mirashi SG, Wesarg S: Automatic teeth segmentation in panoramic X-Ray images using a coupled shape model in combination with a neural network. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G (eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Granada 2018: Springer; 712-719. https://doi.org/10.1007/978-3-030-00937-3_81

4. Geibel MA, Carstens S, Braisch U, Rahman A, Herz M, Jablonski-Momeni A: Radiographic diagnosis of proximal caries—influence of experience and gender of the dental staff. *Clin Oral Invest* 2017; 21(9): 2761-2770. https://doi.org/10.1007/s00784-017-2078-2

5. Ronneberger O, Fischer P, Brox T: U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF (eds.), *Medical image computing and computer-assisted intervention – MICCAI 2015*. Munich 2015: Springer; 234-241. https://doi.org/10.1007/978-3-319-24574-4_28

6. Hamamci IE, Er S, Simsar E et al.: Diffusion-based hierarchical multi-label object detection to analyze panoramic dental x-rays. *arXiv* 2023. https://doi.org/10.48550/arXiv.2303.06500

7. Reyes LT, Knorst JK, Ortiz FR, Ardenghi TM: Machine learning in the diagnosis and prognostic prediction of dental caries: a systematic review. *Caries Res* 2022; 56(3): 161-170. https://doi.org/10.1159/000524167

8. Hamamci IE, Er S, Simsar E et al.: DENTEX: an abnormal tooth detection with dental enumeration and diagnosis benchmark for panoramic x-rays. *arXiv* 2023. https://doi.org/10.48550/arXiv.2305.19112

9. Redmon J, Divvala S, Girshick R, Farhadi A: You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas 2016: IEEE; 779-788. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html (23. 6. 2024)

10. Bizjak Ž, Špiclin Ž: A Systematic review of deep-learning methods for intracranial aneurysm detection in CT angiography. *Biomedicines* 2023; 11(11): 2921. https://doi.org/10.3390/biomedicines11112921

11. He L, Liu Y, Wang L: Intergrated segmentation and detection models for dentex challenge. *arXiv* 2023. https://doi.org/10.48550/arXiv.2308.14161

12. Haghanifar A, Majdabadi MM, Haghanifar S, Choi Y, Ko SB: PaXNet: Tooth segmentation and dental caries detection in panoramic X-ray using ensemble transfer learning and capsule classifier. *Multimed Tools Appl* 2023; 82(18): 27659-27679. https://doi.org/10.1007/s11042-023-14435-9

13. Lin XJ, Zhang D, Huang MY, Cheng H, Yu H: [Evaluation of computer-aided diagnosis system for detecting dental approximal caries lesions on periapical radiographs]. *Zhonghua Kou Qiang Yi Xue Za Zhi*. 2020; 55(9): 654-660. https://doi.org/10.3760/cma.j.cn112144-20200209-00040