ORIGINAL ARTICLE

Diagnosis of interproximal caries lesions with deep convolutional neural network in digital bitewing radiographs

Yusuf Bayraktar[1] · Enes Ayan[2]

Abstract

Objectives This study aimed to investigate the effectiveness of deep convolutional neural network (CNN) in the diagnosis of interproximal caries lesions in digital bitewing radiographs.

Methods and materials A total of 1,000 digital bitewing radiographs were randomly selected from the database. Of these, 800 were augmented and annotated as "decay" by two experienced dentists using a labeling tool developed in Python programming language. The 800 radiographs were consisted of 11,521 approximal surfaces of which 1,847 were decayed (lesion prevalence for train data was 16.03%). A CNN model known as you only look once (YOLO) was modified and trained to detect caries lesions in bitewing radiographs. After using the other 200 radiographs to test the effectiveness of the proposed CNN model, the accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the receiver operating characteristic curve (AUC) were calculated.

Results The lesion prevalence for test data was 13.89%. The overall accuracy of the CNN model was 94.59% (94.19% for premolars, 94.97% for molars), sensitivity was 72.26% (75.51% for premolars, 68.71% for molars), specificity was 98.19% (97.43% for premolars, 98.91% for molars), PPV was 86.58% (83.61% for premolars, 90.44% for molars), and NPV was 95.64% (95.82% for premolars, 95.47% for molars). The overall AUC was measured as 87.19%.

Conclusions The proposed CNN model showed good performance with high accuracy scores demonstrating that it could be used in the diagnosis of caries lesions in bitewing radiographs.

Clinical significance Correct diagnosis of dental caries is essential for a correct treatment procedure. CNNs can assist dentists in diagnosing approximal caries lesions in bitewing radiographs.

Keywords Area under curve · Artificial intelligence · Bitewing radiography · Deep learning · Dental caries

Introduction

Dental caries is one of the most common chronic diseases worldwide and the accurate diagnosis of caries is highly important in clinical practice. For diagnostic purposes, visual-tactile examination is the standard method for detecting dental caries lesions. However, the visual-tactile method may not always detect dental caries lesions in the interproximal region due to wide contact surfaces [1]. To overcome this, bitewing radiographs can be used as an additional method to determine the diagnosis in the interproximal region [2].

Bitewing radiography has higher sensitivity than the visual-tactile method and panoramic radiographs [3–5]. Additionally, digital radiographic systems result in less X-ray exposure [6]. However, accurate caries detection requires at least one clinician to carefully evaluate the radiographic image.

Recently, computer-aided assistance (CAA) systems for dental radiographic imagery has become a popular research area. CAA systems may help dentists, especially those who are less experienced, to make more reliable and accurate assessments of dental caries in bitewing radiographic images. In addition, CAA systems will shorten the time needed for assessment and accelerate the progression to the treatment process.

✉ Enes Ayan
enesayan@kku.edu.tr

[1] Department of Restorative Dentistry, Faculty of Dentistry, Kirikkale University, Kirikkale, Turkey

[2] Department of Computer Engineering, Faculty of Engineering and Architecture, Kirikkale University, Yahsihan, Kirikkale, Turkey

Machine learning (ML) is a subfield of artificial intelligence and consists of powerful algorithms. ML is successfully used in computer-aided diagnosis and assessment systems. However, deep learning, which is a type of artificial neural network, has emerged as a subfield of machine learning in recent years and has replaced classical machine learning methods in computer-aided diagnosis and support systems. Especially, convolutional neural networks (CNNs), which are a type of neural network that inspired by the mammalian visual cortex, have achieved satisfactory results in addressing a variety of medical issues [7, 8]. Unlike classical ML algorithms such as support vector machines [9], K-nearest neighbors [10], and decision trees [11], CNNs can automatically learn hierarchical and contextual features from the dataset through the training procedure [12].

In recent years, CNNs have been successfully employed to clarify various health care problems that use different types of medical imaging. For example, skin lesion segmentation [13], skin cancer classification from dermoscopic images [14], diagnosis of breast cancer from mammograms [15], diagnosis of Alzheimer's from MRI scans [16], and automatic detection and quantization of COVID-19 from chest computed tomography images [17] represent a few recent approaches involving CNNs.

In dentistry, radiology plays an important role in clinical diagnosis. Every year a large number of images are obtained in dental radiology including panoramic, bitewing, periapical, and cephalometric radiographs [18]. Given this vast amount of image data, CNNs seem to have great potential for clinical assessment and diagnosis. Recently, deep learning researchers have started to explore this potential in the field of dental radiology. CNNs have been successfully used to detect periodontal bone loss in periapical radiographs [19], diagnose carious lesions in bitewing radiographs [20], and detect apical lesions in panoramic radiographs [21]. A detailed review describes the use of CNNs in dental radiology [18]. In addition, it is possible to see how the use of artificial intelligence in dentistry could lead to satisfying results, especially in the field of caries diagnosis [20, 22, 23].

The use of CNNs in the field of dentistry and especially in caries diagnosis is a new approach. To the best of our knowledge, there is only one study that uses CNNs and bitewing radiographs [20]. Besides, there are also studies in the field of caries diagnosis using deep learning with near-infrared transillumination method [23–25] and digital periapical radiographs [22, 26]. There are also studies investigating the diagnosis of white spot lesions and dental caries lesions in oral photographs using deep learning [27, 28]. Studies in this area can allow for faster and more accurate caries diagnosis in the future. Also, CNNs can contribute as an assistant to the dentist during caries diagnosis [22]. CNNs may also help to prevent possible false or lacking diagnoses in cases that escape the attention of the dentist due to intensive workload

or lack of experience [29]. In addition, although it has not been reported in the literature thus far, these applications may contribute to the clinical education of dental students. This is a novel topic in the field of dentistry and all of these issues are waiting to be investigated.

Unlike previous studies utilized classification [22] and segmentation [20] based CNN models, we modified and trained a real-time object localization and classification CNN model to detect the lesion location and class probability in bitewing radiographs. The aim of this study was to investigate the effectiveness of deep convolutional neural networks (CNNs) in the diagnosis of interproximal caries lesions in digital bitewing radiographs. For this purpose, we trained a YOLO-based CNN model to detect decays in bitewing images.

Materials and methods

Dataset

This study was performed in the Department of Restorative Dentistry, Faculty of Dentistry, Kirikkale University after ethical approval of Kirikkale University Ethics Committee (Date: 08/07/2020, Decision No: 2020.06.21). A total of 1,000 bitewing radiographic images of permanent teeth created between January 2018 and December 2019 were obtained from the faculty's database. The radiographs have been obtained from the Turkish society, mainly in the Central Anatolian Turkish population. Patients' age, gender, and the date of radiographs were also available in the database. However, we did not consider the metadata in classification. All bitewing radiographs were obtained with the same periapical X-ray devices available in our dental clinics at 65 kVp, 7 mA (Gendex Expert DC, Gendex Dental Systems, IL, USA). There is no clear information about the X-ray machine-film distance and the exposure times. It is possible that these parameters are likely different for each radiograph because the radiographs were taken by different staff in different departments. Some bitewing samples from the dataset are shown in Fig. 1. Both dental arches were clearly observed on bitewing images. More than 50% of the crown part of the visible teeth were included in the calculation, while the others were not. Since it was aimed to diagnose primary approximal caries, full crowns, restorations in the approximal surface of the teeth, and mesial or distal sides of third molar teeth were not calculated. The resolution of the obtained bitewing images had varying sizes. Therefore, all resolutions were arranged to 640×480 pixels. After this process, the dataset was split as follows: 80% for training (800 bitewings) and 20% for testing and validation (200 bitewings). Initially, the training dataset (800 bitewings) was consisted of 11,521 approximal surfaces of which 1847 were decayed (lesion prevalence was 16.03%). We did not use any
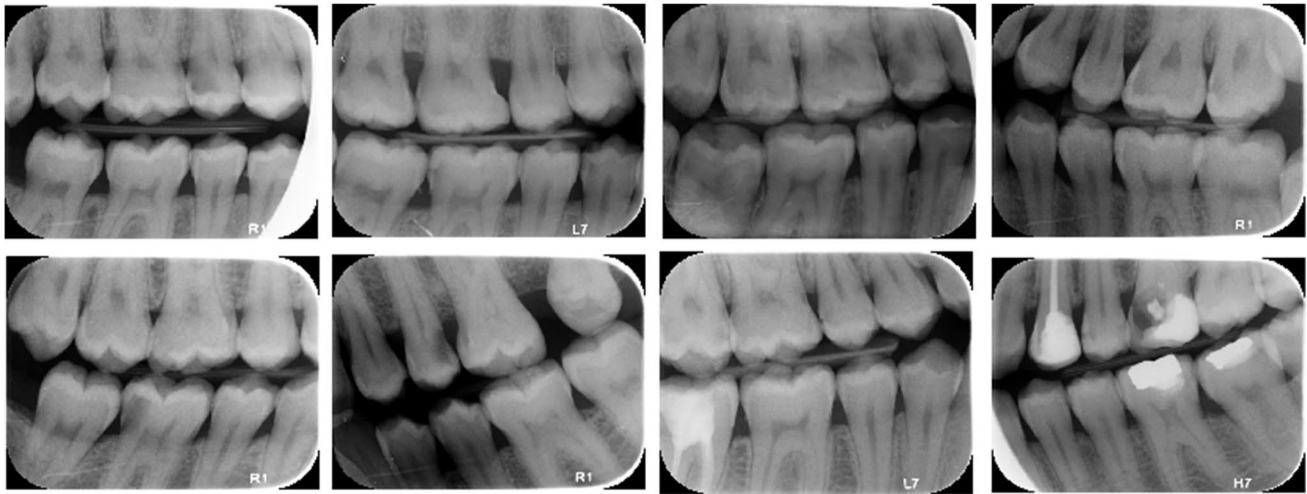
Fig. 1 Representative bitewing radiographs from the dataset

image enhancement or pre-processing methods to improve the bitewing images. The 800 bitewing radiographs were augmented four times using rotation, scaling, zooming, and cropping operations. As a result of augmentation, we obtained 3,200 bitewing radiographs that were used for in the training of the YOLO-based CAA system. Any formal sample size was not calculated in this exploratory descriptive study such as a previous study [27].

Study design

A schematic representation of the proposed CAA system is shown in Fig. 2. The proposed YOLO-based CAA system for the detection of dental caries in the bitewing radiographs consists of three main stages: feature extraction via convolutional layers, caries detection with a confidence score, and fully connected layers for output of the location of caries lesion.

Data labeling

The caries lesions in the sample bitewing radiographs in the study were labeled by two expert dentists via consensus to create a fuzzy gold standard (reference standard). Each of the expert dentists performing the labeling had more than 10 years of working experience in the restorative dentistry
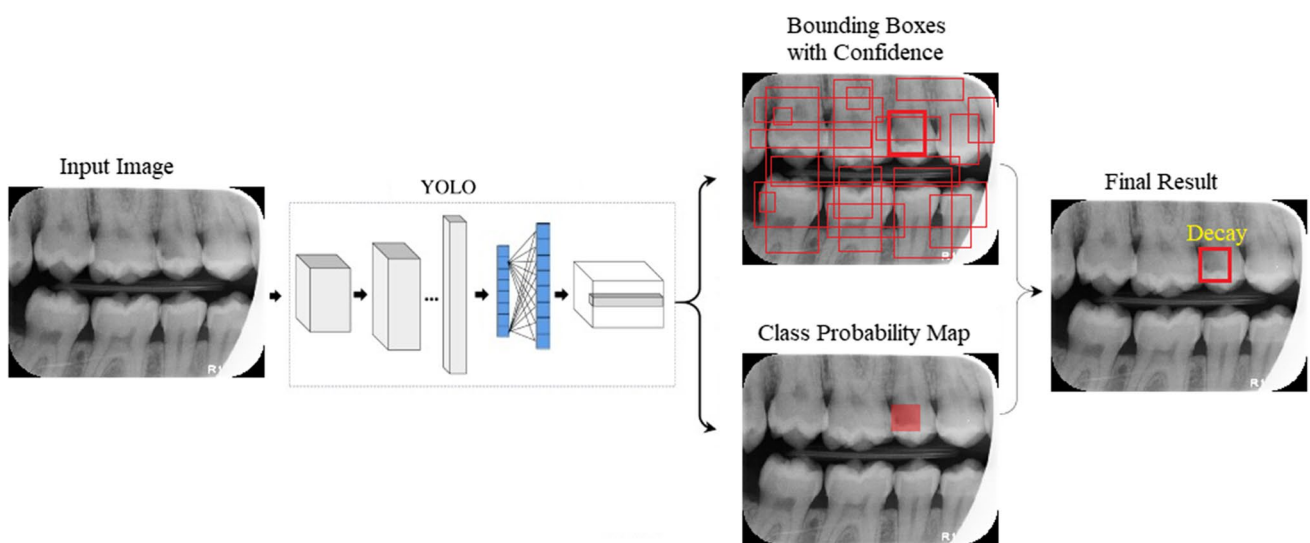
Fig. 2 Scheme of the proposed YOLO-based CAA system

department. During labeling, both expert dentists sat next to each other and after the labeling decision, the area to be labeled was marked by Y.B. In the labeling process, a bounding box was drawn by the experts around each caries lesions to provide YOLO training process requirements. This labeling process was carried out using a tagging tool developed using Python programming language. There are five label parameters for each bounding box. The box parameters are provided in Fig. 3.

YOLO architecture

You only look once (YOLO) is a successful deep learning model with real-time object detection tasks [30]. It can detect multiple objects simultaneously in real time from an image or image sequences. In classical object detection algorithms such as R-CNN [31], the image is divided into a certain number of regions and different CNNs are used to classify each region. However, YOLO uses a single CNN and the entire image to calculate the class possibilities of the objects in the image and the potential ROIs of related classes. The design architecture of YOLO, even though it has a much simpler structure, can achieve very fast and accurate object detection and classification results. There are improved versions of YOLO such as YOLOV2 [32], YOLOV3 [33], and YOLOV4 [34]. In this study, we utilized a modified version of YOLOV3 to detect dental caries in bitewing images.

It will be useful to briefly describe YOLO's detection steps to understand the success of the study. The detection process of YOLO starts by dividing the input image into $N \times N$ non-overlapping grid cells. Every grid cell is responsible for detecting possible objects in the input image where the center of the object falls. Therefore, each grid cell creates a number of bounding boxes (B) around the possible object area and a box confidence score for

whether these boxes contain objects. The box confidence is obtained by multiplying the probability that the area in the related grid cell is an object by the value of intercession over union (IOU) as in the following formula:

$$\text{Confidence} = \text{Prob(object)} \times \text{IOU}_{\text{predicted}}^{\text{ground truth}}$$

The box confidence represents the absence or presence of an object in the suspected area in the bounding box. If the box confidence score value is zero, it means there is no object in the relevant grid cell. So, bounding boxes with zero confidence are excluded. In addition to the box confidence score, each bounding box has four other parameters. These are the center points (x,y) and width, height (w,h) values of the bounding box. Also, YOLO calculates the class probabilities of objects (C). Each grid cell predicts C conditional class probabilities ($\text{Prob(Class}_i|\text{object)}$). After adding the class probabilities to the output vector, the new output vector is $N \times N \times (B \times 5 + C)$. The conditional probability obtained for classes is calculated according to the formula below.

$$\begin{aligned}\text{Confidence score} &= \text{Prob(Class}_i|\text{object)} \times \text{Confidence} \\ &= \text{Prob(Class}_i) \times \text{IOU}_{\text{predicted}}^{\text{ground truth}}\end{aligned}$$

In this study, YOLOV3 was trained using the entire bitewing image and its ROI information. For the training phase, we prepared the training data with their bounding box information. The bounding box information contains the center coordinates (x, y), width, height (w, h) values, and class label of dental caries in the bitewing images. We use 106 convolutional layers architecture with different kernel sizes, residual blocks, skip connections, activation functions, and fully connected layers for caries detection in bitewing images. YOLOV3 uses DarkNet-53 CNN model for feature extraction and stacked with 53 more layers for detection operation. Instead of using pooling layers, additional convolutional layers with a stride of 2 are used to downsample the feature maps. In this way, low-level features are preserved in the feature maps. The low-level features help improve the ability to detect small caries lesions in bitewings. We set the input size of the network as 608×608 pixels. In our experiments, we divided the bitewing image into 19×19 (N) grid cells with a size of 32×32 pixels. The detections in YOLOV3 are obtained at layers 82, 94, and 106 for various lesion scales such as large, medium, and small. Thus, we modified the original detection kernels at the 82nd, 94th, and 106th layers according to the (b*(5 + c)) formula for our problem. YOLOV3 predicts 3 bounding boxes (B) for each grid cell. The b in the formula represents the bounding box count, while the c represents the number of classes that will be detected. So, we set the output kernels as 18 at layers 82, 94, and 106. The architecture details are presented in Fig. 4.

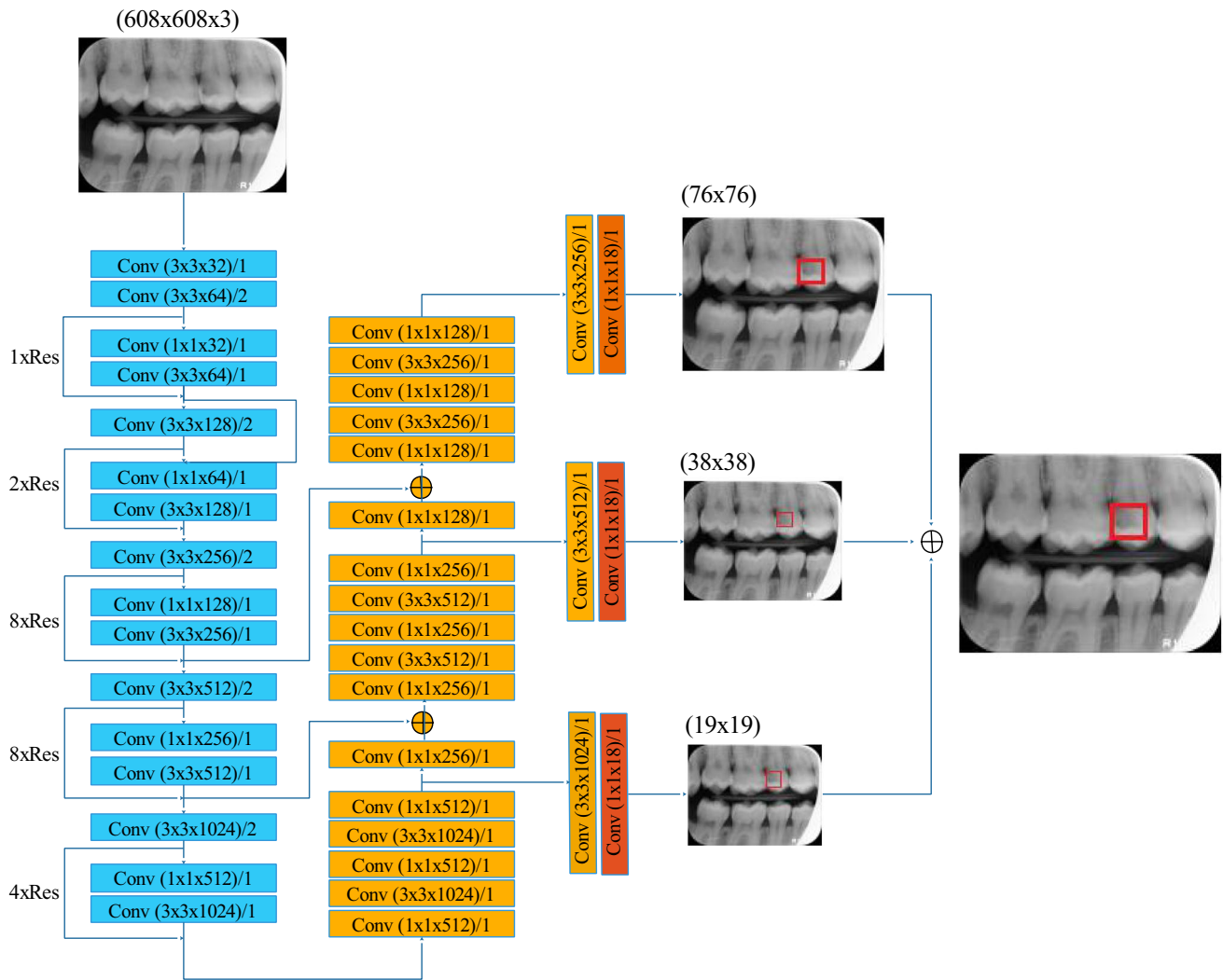**Fig. 3** Labeling of decays for YOLO training

Fig. 4 Architecture of the YOLO-based CAA system

Training and testing

In this study, we trained our proposed CAA model with an augmented dataset. The dataset was randomly divided into training (80%) and testing-validation (20%) groups. We used the holdout validation method throughout the training process. The separated data for the training were augmented using online and offline augmented strategy. The model parameters were optimized by using only training data to avoid any bias. Then, the final model performance was calculated using only the test dataset (20% of the data). Transfer learning strategy is very effective and useful in training deep neural networks [35]. The initial weights of the model were obtained from the pre-trained DarkNet-53 on ImageNet. The weights in the layers are updated utilizing the back-propagation algorithm used during the training. The modified YOLOV3 model trained through to 5,000 epochs (24 h) with 64 batch size and 0.001 learning rate. The entire process was

conducted using a workstation with Nvidia 1080Ti graphics card, two Xenon processor, and 32 GB RAM. All implementations were developed using Python and C + + programming languages on the Ubuntu 14.04 operating system.

Caries detection process and performance evaluation metrics

In the study, we used objective evaluation metrics to evaluate the performance of our YOLO-based CAA system. Figure 5 shows our evaluation strategy during the testing phase of the system. As shown in Fig. 5, if the box class probability score of detected bounding boxes was less than a determined threshold value (0.50), the related boxes were considered undetected caries lesions. In contrast, if the box class score was an equal or higher probability than the threshold, it was regarded in the detection step. In the detection step, if the IOU value of bounding boxes was greater than 50%, the

Fig. 5 Evaluation approach of the proposed system
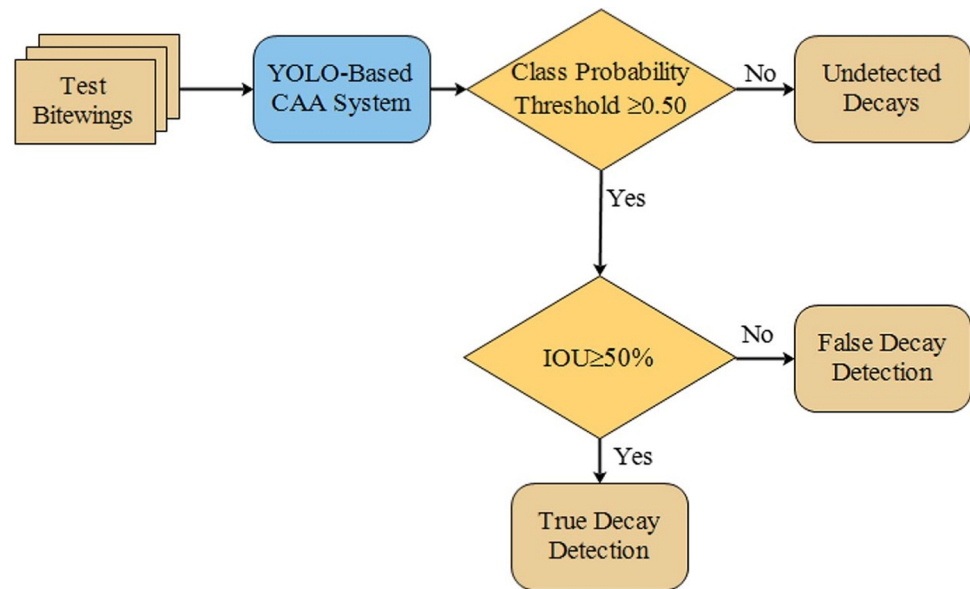
related boxes were considered true detections; otherwise, it was considered a false detection.

To evaluate the proposed model quantitatively, the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) diagnosis of the CNNs were calculated for each bitewing image. Then, the accuracy, sensitivity, specificity, PPV, NPV, and AUC of the CNN were calculated.

Results

A total of 2700 approximal surfaces were tested (1323 for premolars, 1377 for molars) in 200 bitewing radiographs. The 375 of the approximal surfaces had carious lesions (195 for premolars, 180 for molars), while 2325 were sound (1130 for premolars, 1195 for molars). The lesion prevalence for testing dataset was 13.89%. However, the training dataset consisted of 800 radiographs with 11,521 approximal surfaces (5648 for premolars, 5873 for molars). The 1847 of the approximal surfaces had carious lesions (936 for premolars, 911 for molars). The lesion prevalence for training dataset was 16.03%.

The overall accuracy of the model was 94.59% (94.19% for premolars, 94.97% for molars), the sensitivity was 72.26% (75.51% for premolars, 68.71% for molars), the specificity was 98.19% (97.43% for premolars, 98.91% for molars), the PPV was 86.58% (83.61% for premolars, 90.44% for molars), and the NPV was 95.64% (95.82% for premolars, 95.47% for molars). The diagnosing performance of the CNN is shown in Table 1.

The ROC curve of the CNN is shown in Fig. 6. The area under the curve (AUC) was measured as 87.19 (Fig. 6).

Table 1 The accuracy, sensitivity, specificity, PPV, and NPV values of the CNN

	Accuracy	Sensitivity	Specificity	PPV	NPV
Overall	94.59%	72.26%	98.19%	86.58%	95.64%
Premolar	94.19%	75.51%	97.43%	83.61%	95.82%
Molar	94.97%	68.71%	98.91%	90.44%	95.47%
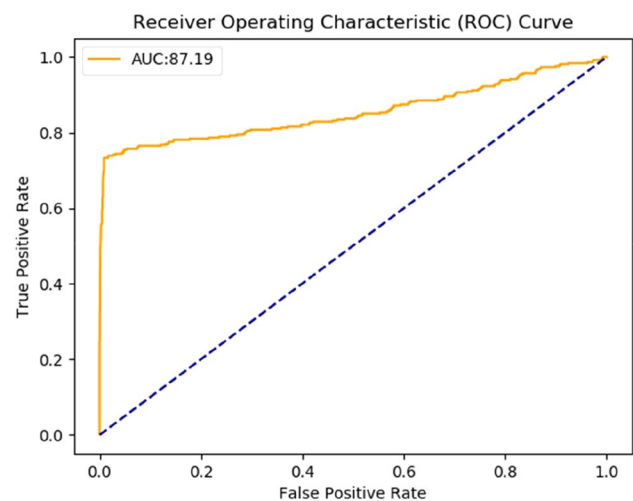
Fig. 6 The ROC curve of the proposed model

Discussion

In this study, the effectiveness of the proposed CNN model for the diagnosis of interproximal caries lesions in digital

bitewing radiographs was evaluated. Bitewing radiography is a more useful method than other radiographic techniques for diagnosing approximal caries lesions [3]. Additionally, the use of CNNs in dental caries diagnosis is a new field in dentistry. To the best of our knowledge, presently there is only one report on the diagnosis of approximal caries lesions in bitewing radiographs using deep learning. In that study, the overall accuracy was 80%, sensitivity was 75%, and specificity was 83%. The PPV and NPV were 70% and 86%, respectively [20].

The current study demonstrated promising results in terms of the evaluated parameters. While the sensitivity results of the premolar teeth were higher than molar teeth, an opposite effect was observed in terms of specificity. The overall specificity value was also found to be higher than the other parameters. Bitewing radiography can detect sound approximal surfaces more successfully [36]. According to findings of this study, the use of CNNs in diagnosis of approximal caries lesions in bitewing radiographs demonstrates a higher specificity value than sensitivity, similar to the study by Cantu et al. [20]. The higher specificity values of bitewing radiography were also reported in previous studies [3, 36, 37]. In a previous study, Lee et al. [22] reported higher sensitivity and lower specificity values for premolars in periapical radiographs in contrast to findings of our study. In the study by Lee et al., the periapical radiographic images were cropped, and the cropped images consisted of only one tooth for classification. In our study, no cropping procedure was applied to the bitewing radiographs.

Cantu et al. [20] reported higher accuracy and sensitivity of the CNN than the mean results of seven dentists. The accuracy scores of the CNN and the dentists were 0.80 and 0.71, respectively. The CNN was reported as more sensitive than the dentists (0.75 and 0.36, respectively). Besides, dentists showed lower sensitivity values in diagnosing initial enamel lesions. Although the sensitivity values of dentists increased for advanced lesions, none of them could reach a higher sensitivity score than CNN [20]. However, the mean specificity value of the seven dentists (91%) was higher than the specificity value (83%) of the CNN. In our study, the evaluated parameters, except the sensitivity value, were found to be higher than the study mentioned above. However, it should be clearly known that the segmentation method used by Cantu et al. and our real-time decay localization approach are different problems in computer vision. We observed that the values in this study were higher than the average values of the seven dentists in the study mentioned. On the other hand, clinical experience of the dentist directly affects the accuracy of the caries diagnosis [38]. Using CNNs in caries diagnosis can provide high accuracy regardless of the dentist's clinical experience. However, high accuracy score may not indicate that a method is always useful. If deep learning interprets each approximal surface as

"decayed," it is possible to reach 100% accuracy. However, this is unacceptable. Therefore, the sensitivity and specificity scores have been reported in this study as reported in previous studies [20, 23].

Although not in bitewing radiographs, there are also studies in which deep learning is used in the diagnosis of caries [23–26]. In a previous study about caries detection in near-infrared light transillumination images using CNNs, Schwendicke et al. reported satisfying results to detect caries lesions [23]. Casalegno et al. reported that deep learning showed promising results to increase accuracy in caries detection and to support dentists' diagnoses [25]. Askar et al. also reported satisfying accuracy about detecting white spot lesions on dental photography using deep learning [27]. In another study conducted on oral photographs, Zhang et al. reported that the deep learning model showed promising results in detecting dental caries [28]. The results of our study on bitewing radiographs are also promising. Holtkamp et al. reported that deep learning models trained and tested in vivo for caries detection in near-infrared light transillumination images showed higher accuracy (0.78 ± 0.04) than in vitro (0.64 ± 0.15) [24]. The study of Holtkamp et al. is important in terms of training criteria of CNNs that directly affect the results. Considering the studies above, it can be said that using deep learning to detect dental caries shows promising results.

The AUC was measured as 87.19% in this study. The results are found to be promising and CNNs can be used for the diagnosis of caries in bitewing radiographs. CNNs can also assist dentists in diagnosing caries and can even be used in dental education.

In this study, two experienced dentists with expertise in restorative dentistry labeled carious lesions according to a fuzzy gold standard method [39]. Since the validity of the results cannot be evaluated with a histological "gold standard," the decisions and possible mistakes of the annotators affect the results in such a designed study. The diagnostic performance of the CNN and dentists was not compared. Also, this was a relatively small dataset. These are the limitations of the study.

In Fig. 7, some of the successful and unsuccessful detections are shown. Deep learning-based methods need large amounts of training data to achieve successful results [40]. More bitewing radiographic images could be used in this study. However, obtaining large-scale data in the medical field is a laborious, time-consuming process with some ethical challenges [41]. Therefore, in this study, the data augmentation method was used to increase the training data. The data augmentation method acquires new data from raw data with the help of various image processing methods (rotation, histogram equalization, scaling, cropping, etc.). Using large datasets could yield near-perfect results in terms of caries diagnosis using CNNs. However,
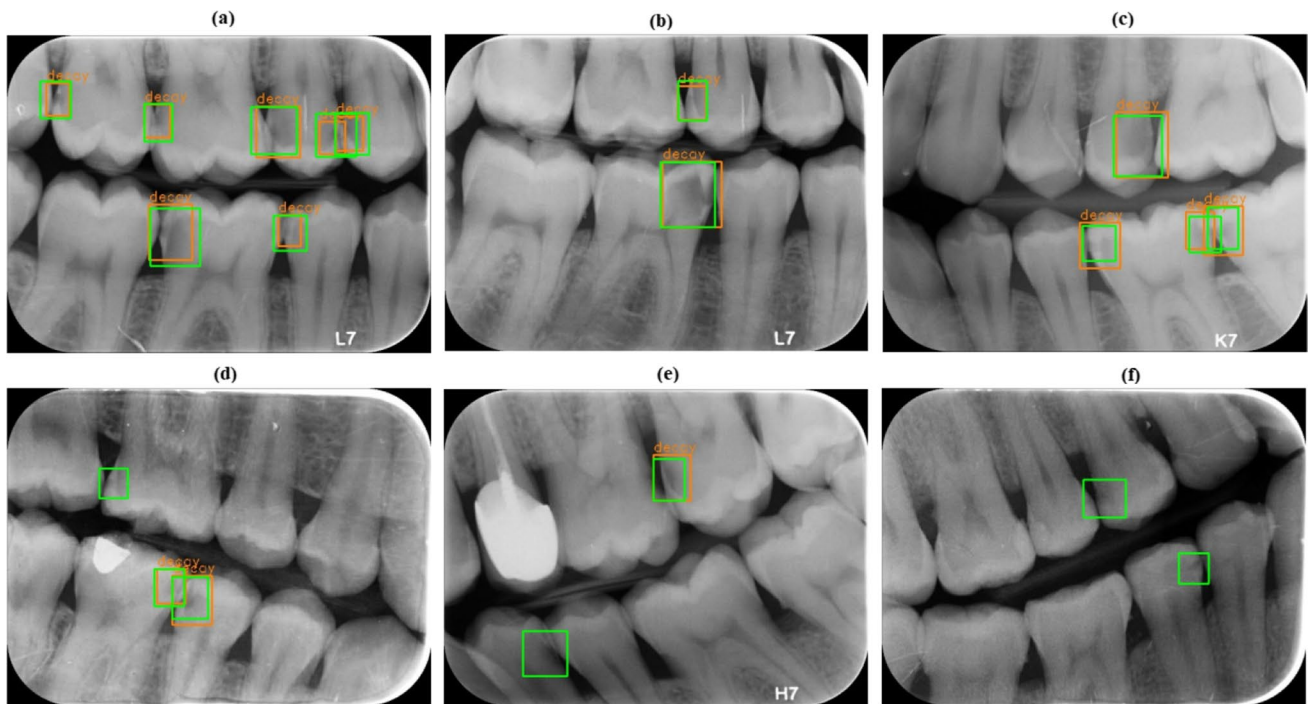
Fig. 7 The green box represents the ground truth, while the orange box represents the detections of proposed CNN model. **a**, **b**, and **c** Successful detections. **d**, **e**, and **f** Successful and false negative detections (only green boxes alone)

in this study, carious lesions were not classified as "enamel caries" or "dentin caries." This is one of the limitations of the study.

Conclusion

In this paper, we proposed a YOLO-based CAA system for the detection of caries lesions in bitewing images. Our test results show promising and feasible outcomes in terms of detecting caries lesions in bitewing images. The proposed CAA system achieved an accuracy of more than 90% in diagnosing approximal caries lesions without the need for a dentist. The next step is to collect more bitewing images and improve the model detection accuracy. Within the limitations of this study, we conclude that:

– The CNNs can be used in the diagnosis of caries lesions in bitewing radiographs.
– Using CNNs provides more than 90% accuracy in diagnosing approximal caries lesions without the need for a dentist.

Declarations

References

1. Selwitz RH, Ismail AI, Pitts NB (2007) Dental caries. Lancet 369:51–59. https://doi.org/10.1016/S0140-6736(07)60031-2
2. Pitts NB, Stamm JW (2004) International Consensus Workshop on Caries Clinical Trials (ICW-CCT)—final consensus statements: agreeing where the evidence leads. J Dent Res 83(Spec No C):C125-8. https://doi.org/10.1177/154405910408301s27
3. Kamburoglu K, Kolsuz E, Murat S, Yuksel S, Ozen T (2012) Proximal caries detection accuracy using intraoral bitewing radiography, extraoral bitewing radiography and panoramic radiography. Dentomaxillofac Radiol 41:450–459. https://doi.org/10.1259/dmfr/30526171
4. Schwendicke F, Tzschoppe M, Paris S (2015) Radiographic caries detection: a systematic review and meta-analysis. J Dent 43:924–933. https://doi.org/10.1016/j.jdent.2015.02.009

5. Baelum V (2010) What is an appropriate caries diagnosis? Acta Odontol Scand 68:65–79. https://doi.org/10.3109/00016350903530786

6. Berkhout WE, Beuger DA, Sanderink GC, van der Stelt PF (2004) The dynamic range of digital radiographic systems: dose reduction or risk of overexposure? Dentomaxillofac Radiol 33:1–5. https://doi.org/10.1259/dmfr/40677472

7. Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol 160:106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

8. Fukushima K (1980) Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36:193–202. https://doi.org/10.1007/BF00344251

9. Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez AJN (2020) A comprehensive survey on support vector machine classification: applications, challenges and trends. Neurocomputing 408:189–215

10. Altman NS (1992) An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 46:175–185

11. Quinlan JR (1990) Decision trees and decision-making. IEEE Trans Syst Man Cybern 20:339–346

12. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88

13. Al-masni MA, Al-antari MA, Choi M-T, Han S-M, Kim T-S (2018) Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. Comput Methods Programs Biomed 162:221–231. https://doi.org/10.1016/j.cmpb.2018.05.027

14. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:115–118. https://doi.org/10.1038/nature21056

15. Wang J, Yang X, Cai H, Tan W, Jin C, Li L (2016) Discrimination of breast cancer with microcalcifications on mammography by deep learning. Sci Rep 6:27327. https://doi.org/10.1038/srep27327

16. Puente-Castro A, Fernandez-Blanco E, Pazos A, Munteanu CR (2020) Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. Comput Biol Med 120:103764. https://doi.org/10.1016/j.compbiomed.2020.103764

17. Zhang HT, Zhang JS, Zhang HH, Nan YD, Zhao Y, Fu EQ, Xie YH, Liu W, Li WP, Zhang HJ, Jiang H, Li CM, Li YY, Ma RN, Dang SK, Gao BB, Zhang XJ, Zhang T (2020) Automated detection and quantification of COVID-19 pneumonia: CT imaging analysis by a deep learning-based software. Eur J Nucl Med Mol Imaging 47:2525–2532. https://doi.org/10.1007/s00259-020-04953-1

18. Schwendicke F, Golla T, Dreher M, Krois J (2019) Convolutional neural networks for dental image diagnostics: a scoping review. J Dent 91:103226. https://doi.org/10.1016/j.jdent.2019.103226

19. Lee JH, Kim DH, Jeong SN, Choi SH (2018) Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci 48:114–123. https://doi.org/10.5051/jpis.2018.48.2.114

20. Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, Gaudin R, Elhennawy K, Schwendicke F (2020) Detecting caries lesions of different radiographic extension on bitewings using deep learning. J Dent 100:103425. https://doi.org/10.1016/j.jdent.2020.103425

21. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, Schwendicke F (2019) Deep learning for the radiographic detection of apical lesions. J Endod 45:917-922 e5. https://doi.org/10.1016/j.joen.2019.03.016

22. Lee JH, Kim DH, Jeong SN, Choi SH (2018) Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent 77:106–111. https://doi.org/10.1016/j.jdent.2018.07.015

23. Schwendicke F, Elhennawy K, Paris S, Friebertshauser P, Krois J (2020) Deep learning for caries lesion detection in near-infrared light transillumination images: a pilot study. J Dent 92:103260. https://doi.org/10.1016/j.jdent.2019.103260

24. Holtkamp A, Elhennawy K, Cejudo Grano de Oro JE, Krois J, Paris S and Schwendicke F (2021) Generalizability of deep learning models for caries detection in near-infrared light transillumination images. J Clin Med 10. https://doi.org/10.3390/jcm10050961

25. Casalegno F, Newton T, Daher R, Abdelaziz M, Lodi-Rizzini A, Schurmann F, Krejci I, Markram H (2019) Caries detection with near-infrared transillumination using deep learning. J Dent Res 98:1227–1233. https://doi.org/10.1177/0022034519871884

26. Geetha V, Aprameya KS, Hinduja DM (2020) Dental caries diagnosis in digital radiographs using back-propagation neural network. Health Inf Sci Syst 8:8. https://doi.org/10.1007/s13755-019-0096-y

27. Askar H, Krois J, Rohrer C, Mertens S, Elhennawy K, Ottolenghi L, Mazur M, Paris S, Schwendicke F (2021) Detecting white spot lesions on dental photography using deep learning: a pilot study. J Dent 107:103615. https://doi.org/10.1016/j.jdent.2021.103615

28. Zhang X, Liang Y, Li W, Liu C, Gu D, Sun W, Miao L (2020) Development and evaluation of deep learning for screening dental caries from oral photographs. Oral Dis. https://doi.org/10.1111/odi.13735

29. Firestone AR, Lussi A, Weems RA, Heaven TJ (1994) The effect of experience and training on the diagnosis of approximal coronal caries from bitewing radiographs. A Swiss-American comparison. Schweiz Monatsschr Zahnmed 104:719–723

30. Redmon J, Divvala S, Girshick R and Farhadi A (2016) You only look once: unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition. 779–788

31. Girshick R, Donahue J, Darrell T and Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition. 580–587

32. Redmon, J. and Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition 7263–7271

33. Redmon J and Farhadi A (2018) Yolov3: an incremental improvement. arXiv preprint arXiv 180402767

34. Bochkovskiy A, Wang C-Y and Liao H-YML (2020) YOLOv4: optimal speed and accuracy of object detection. arXiv preprint arXiv:10934

35. Yosinski J, Clune J, Bengio Y and Lipson H (2014) How transferable are features in deep neural networks? arXiv preprint arXiv:1411.1792

36. Bozdemir E, Aktan AM, Ozsevik A, Sirin Kararslan E, Ciftci ME, Cebe MA (2016) Comparison of different caries detectors for approximal caries detection. J Dent Sci 11:293–298. https://doi.org/10.1016/j.jds.2016.03.005

37. Abdinian M, Razavi SM, Faghihian R, Samety AA, Faghihian E (2015) Accuracy of digital bitewing radiography versus different views of digital panoramic radiography for detection of proximal caries. J Dent (Tehran) 12:290–297

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

39. Walsh T (2018) Fuzzy gold standards: approaches to handling an imperfect reference standard. J Dent 74(Suppl 1):S47–S49. https://doi.org/10.1016/j.jdent.2018.04.022

40. Koppanyi Z, Iwaszczuk D, Zha B, Saul CJ, Toth CK and Yilmaz A (2019) Chapter 3 - multimodal semantic segmentation: fusion of RGB and depth data in convolutional neural networks. Multimodal scene understanding 1st edition. Academic Press. 41-64

41. Shorten C, Khoshgoftaar T (2019) A survey on image data augmentation for deep learning. J Big Data 6:1–48