Development and evaluation of deep learning for screening dental caries from oral photos

Article *in* Oral Diseases · November 2020

DOI: 10.1111/odi.13735

7 authors, including:

Yuan Liang
University of California, Los Angeles
29 PUBLICATIONS **122** CITATIONS

SEE PROFILE

Wen Li
Macquarie University
16 PUBLICATIONS **190** CITATIONS

SEE PROFILE

Wei-Bin Sun
Nanjing University
67 PUBLICATIONS **975** CITATIONS

SEE PROFILE

Leiying Miao
Nanjing University
67 PUBLICATIONS **1,214** CITATIONS

SEE PROFILE

Received: 17 September 2020 | Revised: 2 November 2020 | Accepted: 17 November 2020

DOI: 10.1111/odi.13735

ORIGINAL ARTICLE

ORAL DISEASES
Leading in Oral, Maxillofacial, Head & Neck Medicine

WILEY

Development and evaluation of deep learning for screening dental caries from oral photographs

Xuan Zhang[1] | **Yuan Liang**[2] | **Wen Li**[3] | **Chao Liu**[4] | **Deao Gu**[4] | **Weibin Sun**[1] | **Leiying Miao**[3]

[1]Department of Periodontology, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China

[2]University of California, Los Angeles, CA, USA

[3]Department of Endodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China

[4]Department of Orthodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, Nanjing, China

Correspondence
Leiying Miao, Department of Endodontics, Nanjing Stomatological Hospital, Medical School of Nanjing University, No.30 Zhongyang Road, Xuanwu District, Nanjing, Jiangsu Province, China.
Email: miaoleiying80@163.com

Funding information
National Natural Science Foundation of China, Grant/Award Number: 51772144; Project of Invigorating Health Care through Science, Technology and Education Jiangsu Provincial Medical Youth Talent, Grant/Award Number: QNRC2016120

Abstract

Objectives: To develop and evaluate the performance of a deep learning system based on convolutional neural network (ConvNet) to detect dental caries from oral photographs.

Methods: 3,932 oral photographs obtained from 625 volunteers with consumer cameras were included for the development and evaluation of the model. A deep ConvNet was developed by adapting from Single Shot MultiBox Detector. The hard negative mining algorithm was applied to automatically train the model. The model was evaluated for: (i) classification accuracy for telling the existence of dental caries from a photograph and (ii) localization accuracy for locations of predicted dental caries.

Results: The system exhibited a classification area under the curve (AUC) of 85.65% (95% confidence interval: 82.48% to 88.71%). The model also achieved an image-wise sensitivity of 81.90%, and a box-wise sensitivity of 64.60% at a high-sensitivity operating point. The hard negative mining algorithm significantly boosted both classification ($p < .001$) and localization ($p < .001$) performance of the model by reducing false-positive predictions.

Conclusions: The deep learning model is promising to detect dental caries on oral photographs captured with consumer cameras. It can be useful for enabling the preliminary and cost-effective screening of dental caries among large populations.

KEYWORDS
artificial intelligence, deep learning, dental caries

1 | INTRODUCTION

Dental caries has been considered to be highly prevalent, with its early lesions present in most of populations. According to a recent study, untreated caries in the permanent dentition has become the most common health condition globally, effecting 34.1% of the population (Peres et al., 2019). It is also affecting 43.1% if the population ages between 2 and 19 in the United States (Fleming & Afful, 2018). Without being properly treated, dental caries can cause pulpitis and periapical diseases (Ali et al., 2018). Since the caries is resulted from a continuous process of many demineralization and remineralization cycles, the early detection of caries with improved hygiene behaviors and corresponding interventions can be promising to arrest or reverse its progress (Featherstone, 2008). Clinically, the detection of dental caries mainly relies on visual-tactile inspection and dental radiography (Gomez, 2015). However, oral health resources are much unbalanced globally, with people from many regions having limited access to dental professionals. Moreover, traditional clinical evaluations can add economic burdens for individuals of low-income,

Xuan Zhang and Yuan Liang contributed equally and were considered as co-first authors.

TABLE 1 Numbers of images with positive and negative findings that are assigned to training, validation, and testing subsets

	Training		Validation		Testing		Total	
	N	*P* (%)	*N*	*P* (%)	*N*	*P* (%)	*N*	*P* (%)
Positive	831	62.53	102	7.67	396	29.8	1,329	100
Negative	1,676	64.39	198	7.61	729	28.01	2,603	100
Total	2,507	63.76	300	7.63	1,125	28.61	3,932	100

Note: Numbers of images and their percentages within the whole dataset are represented as *N* and *P*.

which might also prevent them from regular clinical visits (Petersen et al., 2005). Hence, there has raised a need for systems can detect underlying dental caries with low cost among large populations.

In recent years, deep learning (DL), an artificial intelligence (AI) method, has been researched to automate decision making process for various clinical dental (Park & Park, 2018). The technique, which consists of multilayer ConvNets, has produced promising accuracy on unforeseen data by automatically learning from datasets with manual annotations from dental experts. Previous works include using DL for segmenting gingival diseases from X-ray scans (Rana et al., 2017), numbering teeth from cone-beam CT (Miki et al., 2017), detecting periodontal bone loss from panoramic dental radiographs (Krois et al., 2019), predicting oral cancer outcomes (Ilhan et al., 2020), and detecting periapical pathosis on cone-beam computed tomography images (Orhan et al., 2020). For automatic dental caries detection, DL has been applied to medical screenings including Near-Infrared-Light Transillumination (NILT) images (Casalegno et al., 2019; Schwendicke et al., 2020) and periapical radiographs (Lee et al., 2018). We observed that the usage of oral images from consumer cameras, which are more convenient and cost-effective to capture, has not been much explored. With consumer cameras including smartphones becoming increasingly ubiquitous and powerful (Lee, 2016), such DL system can be promising to improve the awareness of dental caries among populations.

In this work, we develop a deep ConvNet model for the automatic detection of dental caries from oral photographs. The model classifies the existence of dental caries in a given image and also localizes the findings with bounding boxes. We statistically analyze both the classification and localization performance of the model to illustrate its potential usefulness. To facilitate related studies in the future, we have open-sourced our implementation at https://github.com/liangyuandg/DLCariesScreen.

2 | MATERIALS AND METHODS

2.1 | Dataset collection and annotation

For developing and evaluating the DL model, we built an in-house dataset consisting of 3,932 clinical oral photographs collected from 625 volunteers. All the data were acquired at Department of Periodontics, orthodontics and endodontics, Nanjing Stomatological Hospital, Nanjing University, between January 2018 and December 2019. The project was reviewed by the Institutional Review Board, Nanjing Stomatological Hospital, Medical of Nanjing University (2019NI-065(KS)). The volunteers covered an age range from 14 to 60. To approximate the image quality in the practical scenario, all the images were collected with consumer cameras, which include iPhone 8, iPhone 7, Samsung Galaxy S8, and Canon EOS7D. No specific in- or exclusion criteria about images, for example, lighting and resolution, were applied. All the devices had fully automatic setting for image capture. The images covered regions of oral cavity including occlusal surfaces, buccal surfaces, labial surfaces, lingual surfaces, and palatal surfaces.

We collected reference annotations on dental caries for all the images from three board-certified dentists. In specific, images of a volunteer were independently labeled by one of the three board-certificated dentists based on the results of clinical visual-tactile examination. The labeling applied a single threshold of sound versus carious teeth according to the ICDAS (International Caries Detection and Assessment System) criterial without classifying severity levels of a lesion. Each dental caries was labeled with a bounding box. Since there can be no well-defined boundaries for dental caries in some cases, we followed a common approach of medical labeling by instructing the dentists to focus on the correctness of box centers (Armato et al., 2011).

We divided the dataset into training, validation, and testing subsets by randomly splitting the volunteers into three groups. The training set contains 2,507 (63.76%) images from 389 (62.24%) volunteers, while the testing subset contains 1,125 (28.61%) of images from 187 (29.92%) volunteers. We guaranteed that all image of a same volunteer existed in only one of the three subsets. Note that each image can show none, one or more dental caries findings. Among the training subset, 933 (33.24%) images have positive findings; while among the testing subset, 396 (35.20%) have positive findings. Table 1 shows more details about the split of the dataset. To reduce the model overfitting, we use the validation subset to control the progress of training. We monitored the model performance on the validation subset and detected local minimum of the validation accuracy as the early stopping criteria.

2.2 | ConvNet model

Figure 1 shows the overall architecture of the model. Our model is adapted from Single Shot MultiBox Detector (SSD) (Liu et al., 2016).

Specifically, the input image is firstly resized to 300 × 300 without changing its aspect ratio and processed with VGG-16 (Simonyan & Zisserman, 2014), which is a stack of convolutions with activations and down sampling, for deriving deep feature maps. Then, the feature maps are further encoded into location maps with dimension of 10 × 10, 5 × 5, 3 × 3, and 2 × 2. Five candidates' boxes, or anchor boxes, are assigned at each pixel within the maps, having aspect ratios of 0.25, 0.5, 1, and 2. The aspect ratios are determined by studying the shape distribution of dental caries within the training set, such that they can cover the shapes of most findings. In Figure 1, we show a semantic location map of size 5 × 5 with five anchor boxes at a pixel. The location maps are then regressed for anchor box probabilities, and offsets of sizes and locations. An example output of the model from testing image is also visualized in Figure 1.

2.3 | Training strategies and hard negative mining

The whole model was end-to-end trained. We defined the training object as the smooth L1 loss for bounding box regression (Liu et al., 2016) and cross-entropy loss for box classification (Esteva et al., 2017). In order increase the robustness of our model for in-the-wild application with uncertain factors, we employed intensive augmentations to enhance the diversity of images (Redmon & Farhadi, 2017). Specifically, random image shifting and cropping were performed to counter for possible positions of an oral cavity in the camera view; random image scaling was performed for various camera's distances to an oral cavity; random image rotation was performed to counter for different camera angles; and random changes of image hue, saturation, and exposure were performed to counter for lighting conditions. Each of the aforementioned augmentations was applied on-the-fly during training. To enable the fast training the model, we employed transfer learning by initializing our model from VGG-16 (Simonyan & Zisserman, 2014) that pre-trained on large-scale image recognition tasks for speeding up the training process(Deng et al., 2009).

Moreover, according to our observation, some conditions having similar appearances as dental caries, for example, dental stains, can be mispredicted as positive findings. To improve the model sensitivity, we further trained the model with hard negative mining (Felzenszwalb et al., 2010; Liu et al., 2016): The model was fine-tuned with true positives against only false positives that have highest predicted probabilities. To be specific, we controlled the ratio between positive boxes and hard negative boxes that included in the loss calculation. We applied two times of hard negative mining with the ratios set to be 1:5 and 1:1. For each of the hard negative mining, we followed the same early stopping criterial as aforementioned. We show that this strategy can significantly improve the model by reducing false-positive predictions. Note that a larger ratio than 1:1 does not further improved the model performance according to our experiments, mainly because: (i) the total number of boxes involved in the loss is reduced, such that the training becomes less efficient, and (ii) the model specificity is reduced since boxes are more likely to be classified as positive.

2.4 | Statistical analyses

We evaluated the performance of the models for: (i) classification for telling the existence of dental caries in an input image and (ii) localization for telling the regions of dental caries. In terms of classification performance, we measured the receiver operating characteristic (ROC) curve that shows the true-positive rate (TPR), or sensitivity, against its false-positive rate (FPR), or 1—specificity, under varying discrimination thresholds. We also reported the area under curve (AUC) with 95% confidence intervals (95% CIs) as the summarized indicator of model performance. A larger AUC implies better classification performance, while a perfect model and random guessing will have AUC of 1.0 and 0.5, respectively.

In terms of localization performance, we measured the free-response ROC (FROC) curve, which shows the bounding box-wise TPR, or sensitivity, against the average number of false-positive (FP) boxes per image, under varying thresholds for box probabilities. To determine if a predicted box was a hit, we followed the practice of Van et al[17] and calculate if the center of the box falls into the range of a ground-truth box. A perfect model will have a TPR of 1.0 at a FP of 0.0, meaning the model can detect all the dental caries regions without predicting any false-positive bounding box.

We also reported the sensitivities and specificities for image-wise classification and box-wise localization under the two selected operating points of our model.

3 | RESULTS ANALYSIS

The ROC and FROC curves with 95% CI for our models with and without hard negative mining are detailed in Figure 2. It indicates that the model with two times of hard negative mining (M3)

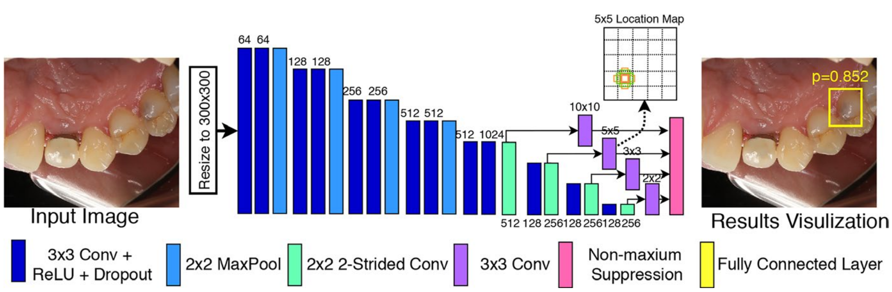

FIGURE 1 Overview of the model architecture. The model is adapted from Single Shot MultiBox Detector (SSD). The anchor boxes are designed by considering the shapes of dental caries. The outputs of the models are bounding boxes with predicted probabilities

achieved the highest accuracy for both classification and localization. Figure 3 quantitatively compares the classification and localization performance. Our model M3 achieved the mean AUC (95% CI) of 85.65% (82.48% to 88.71%), which is significantly higher than both the model (M2) with one-time hard negative mining ($p < .001$) and the baseline model (M1) without hard negative mining ($p < .001$). The results are also consistent for the localization performance: Our model with two times of hard negative mining (M3) achieved highest bounding box detection performance for dental caries. To be specific, the model reached box-wise TPR (95% CI) of 46.74% (42.75% to 55.02%), 61.70% (56.77% to 66.36%), and 82.10% (78.00% to 86.17), with a trade-off of having 1/2, 1, and 2 false-positive boxes per image in average, respectively. The results prove that hard negative mining can be effective for improving model performance in detecting dental caries from oral cavity images. The performance gain can be explained as the loss function can better reflect model improvement without it being averaged out by large number of negative findings that are easy to distinguish.

In Figure 2, we also show the operating points for our best version of model (M3). In order to convey confidence information to users, we designed two operating points for the model, and detection boxes of different operations points can be thus visualized differently. By following previous studies on DL for medical diagnosis (De Vrijer et al., 2009; Gulshan et al., 2016), we set a high-specificity operating point with a higher discrimination threshold that aims for reducing false positives, and the high-sensitivity operating point with a lower discrimination threshold for keeping the missing rate low. For the classification task of telling the existence of dental caries within an image, our model achieved the specificities of 81.90% at a sensitivities of 68.74% under the high-specificity operating point, meanwhile the mean sensitivities of 84.53% at a specificities of 63.82% under the high-sensitivity operating point. Accordingly, the model can accurately point out dental caries using bounding boxes with a sensitivity of 41.39% at the high-specificity operating point and a sensitivity of 64.60% at the high-sensitivity operating point.

Figure 3 depicts selected results obtained on the testing images for a qualitative evaluation. For better comparison between different models, we visualize all bounding boxes that have a predicted probability higher than the aforementioned high-sensitivity threshold. The results also confirm that hard negative mining can effectively reduce false-positive predictions.

4 | DISCUSSION

Dental caries as the common disease that influence health and the quality of live (Frencken et al., 2017). When dental caries is not properly treated, the lesion may influence the enamel, dentin, and even pulp tissue, and ultimately leading to severe pain even dental loss. Different from the existing DL systems based on medical imaging for computer-assisted diagnosis, for example, panoramic radiographs, cone-beam CT images, and histological images, our work utilized clinical oral photographs for screening dental caries. With the advent of low-cost, sensor-equipped smartphones, such method can be promising to benefit large populations for preliminary assessment of dental health. Enabled by the method, systems can be developed to instruct hygiene behaviors (Matthews, 2008), health-enhancing activities (Philip et al., 2018), or recommend clinical treatments based on the screening results. Such systems can be effective for areas with limited dental care resources and individuals that cannot afford regular dental visits.

Previous studies have explored dentists using teeth photographs for the detection of dental caries in the purpose of epidemiology. Boye et al. compared diagnostic performance for the telling the existence of caries using photographs with the established visual examination method (Boye et al., 2012). Gomez et al verified the detection performance of non-cavitated sessions from photographs (Gomez et al., 2013). Bottenberg et al further compared the difference of ICDAS scores that drawn from visual examination and photographs (Bottenberg et al., 2016). All the studies show teeth photographs were adequate for dentists to draw statistically similar results as directly from teeth. However, those studies were based on

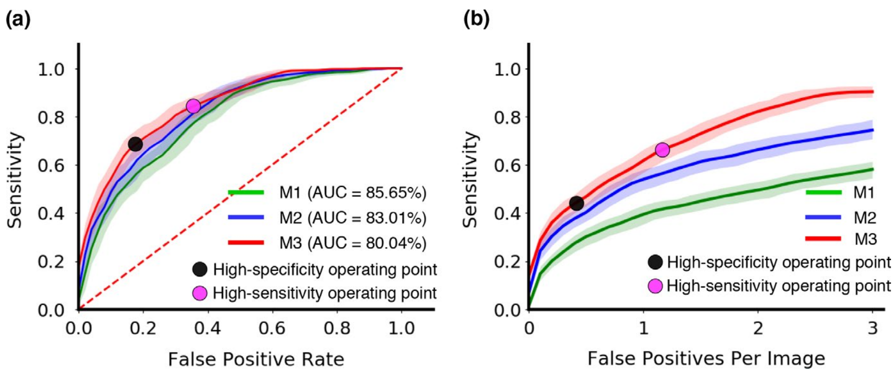

FIGURE 2 Quantitative analysis of the performance for the classification task (a) and the bounding box localization task (b) achieved by: (i) the baseline model without hard negative mining (M1), (ii) the model with one time hard negative mining (M2), and (iii) the model with two times hard negative mining (M3). High-specificity and high sensitivity operating point are marked on the curves

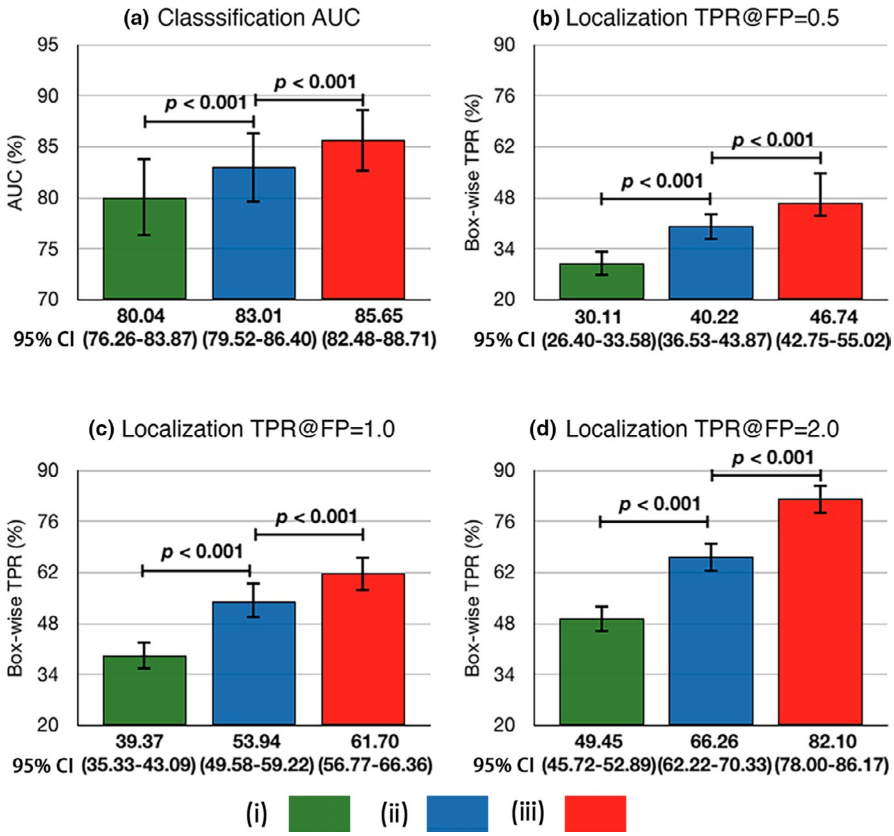

FIGURE 3 (a): Comparison of classification performance between different models: (i) without hard negative mining (M1), (ii) with one hard negative mining (M2), and with two hard negative mining (M3). The box-wise TPR is measured at average FP = 0.5 (b), 1.0 (c) and 2.0 (d). The error bars represent the 95% confidence interval, and the numbers on the bridges indicate the level of significance measured by *p*-Value. The numbers below the x-axes show detailed values

extracted teeth, which can be different from the application scenarios of large-scale epidemiological survey. Instead, Boye et al studied the dentists' detection accuracy of dental caries from intra-oral photographs captured from an intra-oral camera and showed its equivalent to visual examination (Boye et al., 2013).

Since the smartphones with powerful cameras are becoming ubiquitous, multiple studies have also discussed the possibility of teledentistry based on oral photographs from smartphones. Estai M et al evaluated the efficacy of detecting dental caries from smartphone photographs, where the sensitivity scores of dentists ranged from 60% to 63%, which is lower than the benchmark visual examination (Estai et al., 2017). However, only labial and occlusal views of teeth were captured in the study, which was reported to cause the missing of dental caries detections. Kohara et al further studied the performance of dentist's detection of different stages of caries lesions from smartphone photographs (Kohara et al., 2018). Their results show that the sound teeth and extensive lesions can be detected with a high sensitivity from 75% to 100%, while the initial and moderate lesions have a detection sensitivity lower than 60%. Moreover, the detection specificities for lesions of all stages were always above 83.3%.

Different from the aforementioned works, we propose to automate the detection progress from oral photographs with a deep learning model. In a review of the literature, the most related work is from Moutselos et al (Moutselos et al., 2019). They assessed the dental caries detection with deep learning on dental intra-oral images that captured from intraoral cameras. 88 posterior permanent molars images were collected from patients, and a mask R-CNN model was used with pre-processing and transfer learning. Their best segmentation accuracy was 0.684 in F-score. Different from their work, we: (i) modeled with oral photographs from consumer cameras, (ii) collected a larger annotated dataset of 3,932 images for effectively model training, and (iii) formulated a localization task rather than a segmentation one. We argue that bounding boxes labeling has been efficient for indicating the locations of positive findings and also saves labor cost for annotation.

Besides the dental researches, DL has enabled the automatic screening of several diseases using photographs captured from regular cameras. For example, Mariakakis et al. automatic detected live disorders by capturing Jaundice color changes from photographs (Mariakakis, Banks, et al., 2017); Vardell et al introduced a skin disease diagnosis system with skin photographs as input (Vardell & Boucrick, 2012). Mariakakis et al. evaluated using pupil photographs in response to stimulations for the fast detection of the traumatic brain injury (Mariakakis, Baudin, et al., 2017). All those works show that deep learning is promising to detect health issues for everyday users as a supplement to clinical visits. In the case of dental caries detection, Devito et al evaluated a neural network could improve the performance of diagnosing proximal caries (Devito et al., 2008). Valizadeh S et al designed a computer software for detection of approximal caries in posterior teeth. They applied data-clustering algorithm named fuzzy c-means in computer software and diagnosed 60% of enamel caries (Valizadeh et al., 2015). Hung M et al used public data from the National Health and Nutrition Examination Survey, evaluated machine learning for diagnostic prediction of root caries, the algorithms performed well and allow for clinical implementation

FIGURE 4 Examples of dental caries detection on the testing images. We show ground-truth bounding boxes from manual annotations (Reference), predictions from the baseline model without hard negative mining (M1), and predications from the model with two times of hard negative mining (M3)

(Hung et al., 2019). Casalegno F et al evaluated the efficacy of deep CNN algorithms for detection and diagnosis of dental caries on near-infrared transillumination. Their model achieved an overall mean intersection-over-union (IOU) score of 72.7% on a 5-class segmentation task, suggested that a deep learning approach for the analysis of dental images holds promise for increasing the speed and accuracy of caries detection (Casalegno et al., 2019). However, none of the works were based on oral photographs from consumer cameras.

In this study, we developed a trained DL model to automatically classify the existence of dental caries from an oral photograph and also localize positive findings on the photograph. By comparing with reference annotations from experts, the proposed model achieved a classification AUC of 85.65%, and localization sensitivities of 64.60% at the high-sensitivity operating point. To make the model better distinguish dental caries from dental stains of similar appearance, we employed the negative mining strategy during training: Only positive bounding boxes and negative bounding boxes of high uncertainty are included in the loss term. We show that employing the strategy can significantly ($p < .001$) improve both classification and localization performance.

Our work still has serval shortcomings. Here we discuss them and propose possible solutions for future researches. First, our dataset is limited in the sense that the images were collected from a single organization. We believe enriching the dataset with diverse data from multiple sites globally can help better validate the method and also possibly improve the generalization of the method. Second, despite the high accuracy of classifying an oral photograph for the existence of dental caries, it is still challenging for the model to accurately localize the findings (bounding box-wise sensitivities of 41.39%/64.60% at high-specificity/high-sensitivity operating points). As the examples show in Figure 4, it is mainly due to the reduced model specificity with some dental stains can be misclassified as dental caries. These black stains are present as dots or small areas of dark coloration that may appear as diffuse stains, covering a part of the tooth crown. Such extrinsic dark dental pigmentation can also be a disruptive factor for dentists when doing visual assessment of dental caries (Koch et al., 2001). In order to reduce such false-positive predictions, a separate DL classifier that cascaded with our model can be promising, which takes all the positive predictions from our model and filters out dental stains (Ding et al., 2017). Moreover, self-reported symptoms, for example, pains, tooth sensitivity, and eating habits, have been proven to be highly related to the existence of dental caries (Barber & Wilkins, 2002). Such factors can also be fused with diagnostic model to improve the accuracy by feature encoding (Choi et al., 2018).

5 | CONCLUSION

Deep learning algorithms could improve public health by enabling automatic screening of diseases using photographs captured from well-accessed cameras. In this work, we demonstrated the use of a deep learning model for the detection of dental caries from clinical oral photographs and discussed its possibility of being used for public self-examination. Our model was assessed for handling both image classification and finding localization tasks. Our results also showed the hard negative mining strategy can significantly improve the performance of model.

ACKNOWLEDGEMENTS

This research was supported by the National Natural Science Foundation of China (51772144); Project of Invigorating Health Care through Science, Technology and Education Jiangsu Provincial Medical Youth Talent (QNRC2016120). The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

CONFLICT OF INTEREST

None to declare.

AUTHOR CONTRIBUTIONS

Xuan Zhang: Data curation; Methodology; Writing-original draft. **Yuan Liang:** Methodology; Software; Writing-original draft. **Wen Li:** Data curation. **Chao Liu:** Data curation. **Deao Gu:** Data curation. **Weibin Sun:** Conceptualization; Writing-review & editing. **Leiying Miao:** Funding acquisition; Writing-review & editing.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/odi.13735.

ORCID

Xuan Zhang https://orcid.org/0000-0003-4523-3687
Leiying Miao https://orcid.org/0000-0001-6915-8218

REFERENCES

Ali, A. H., Koller, G., Foschi, F., Andiappan, M., Bruce, K. D., Banerjee, A., & Mannocci, F. (2018). Self-limiting versus conventional caries removal: A randomized clinical trial. *Journal of Dental Research*, *97*(11), 1207–1213. https://doi.org/10.1177/0022034518769255

Armato, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., Zhao, B., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., MacMahon, H., van Beek, E. J. R., Yankelevitz, D., Biancardi, A. M., Bland, P. H., Brown, M. S., Engelmann, R. M., Laderach, G. E., … Clarke, L. P. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference Database of lung nodules on CT scans. *Medical Physics*, *38*(2), 915–931. https://doi.org/10.1118/1.3528204

Barber, L. R., & Wilkins, E. M. (2002). Evidence-based prevention, management, and monitoring of dental caries. *Journal of Dental Hygiene*, *76*(4), 270–275. PMID: 12592918.

Bottenberg, P., Jacquet, W., Behrens, C., Stachniss, V., & Jablonskimomeni, A. J. (2016). Comparison of occlusal caries detection using the ICDAS criteria on extracted teeth or their photographs. *BMC Oral Health*, *16*(1), 93. https://doi.org/10.1186/s12903-016-0291-z

Boye, U., Walsh, T., Pretty, I. A., & Tickle, M. J. (2012). Comparison of photographic and visual assessment of occlusal caries with histology as the reference standard. *BMC Oral Health*, *12*(1), 10. https://doi.org/10.1186/1472-6831-12-10

Boye, U., Willasey, A., Walsh, T., Tickle, M., & Pretty, I. A. (2013). Comparison of an intra-oral photographic caries assessment with an established visual caries assessment method for use in dental epidemiological studies of children. *Community Dentistry and Oral Epidemiology*, *41*(6), 526–533. https://doi.org/10.1111/cdoe.12049

Casalegno, F., Newton, T. H., Daher, R., Abdelaziz, M., Lodirizzini, A., Schurmann, F., & Markram, H. (2019). Caries detection with near-infrared transillumination using deep learning. *Journal of Dental Research*, *98*(11), 1227–1233. https://doi.org/10.1016/j.jdent.2019.103260

Choi, E., Xiao, C., Stewart, W. F., & Sun, J. (2018). MiME: Multilevel medical embedding of electronic health records for predictive healthcare. *Neural Information Processing Systems*, *10*, 4552–4562.

De Vrijer, M., Medendorp, W. P., & Van Gisbergen, J. A. (2009). Accuracy-precision trade-off in visual orientation constancy. *Journal of Vision*, *9*(2), 9. https://doi.org/10.1167/9.2.9

Deng, J., Dong, W., Socher, R., Li, L. J., & Li, F. F. (2009). ImageNet: A large-scale hierarchical image database. *Computer Vision and Pattern Recognition*, *06*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Devito, K. L., Barbosa, F. D., & Filho, W. N. (2008). An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. *Oral Surgery Oral Medicine Oral Pathology Oral Radiology and Endodontology*, *106*(6), 879–884. https://doi.org/10.1016/j.tripleo.2008.03.002

Ding, J., Li, A., Hu, Z., & Wang, L. (2017). Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. *Medical Image Computing and Computer Assisted Intervention*, *06*, 559–567. arXiv:1706.04303.

Estai, M., Kanagasingam, Y., Huang, B., Shiikha, J., Kruger, E., Bunt, S., & Tennant, M. (2017). Comparison of a smartphone-based photographic method with face-to-face caries assessment: A Mobile Teledentistry Model. *Telemedicine Journal and E-health*, *23*(5), 435–440. https://doi.org/10.1089/tmj.2016.0122

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J. M., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Featherstone, J. D. (2008). Dental caries: A dynamic disease process. *Australian Dental Journal*, *53*(3), 286–291. https://doi.org/10.1111/j.1834-7819.2008.00064.x

Felzenszwalb, P. F., Girshick, R., Mcallester, D., & Ramanan, D. (2010). Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(9), 1627–1645. https://doi.org/10.1109/TPAMI.2009.167

Fleming, E., & Afful, J. (2018). Prevalence of total and untreated dental caries among youth: United States, 2015–2016. *NCHS Data Brief*, *2018*, 1.

Frencken, J. E., Sharma, P., Stenhouse, L., Green, D., Laverty, D., & Dietrich, T. (2017). Global epidemiology of dental caries and severe periodontitis – A comprehensive review. *Journal of Clinical Periodontology*, *44*(Suppl 18), S94–S105. https://doi.org/10.1111/jcpe.12677

Gomez, J. (2015). Detection and diagnosis of the early caries lesion. *BMC Oral Health*, *15*(1), 1–7. https://doi.org/10.1186/1472-6831-15-S1-S3

Gomez, J., Zakian, C., Salsone, S., Pinto, S. C. S., Taylor, A., Pretty, I. A., & Ellwood, R. P. (2013). In vitro performance of different methods in detecting occlusal caries lesions. *Journal of Dentistry*, *41*(2), 180–186. https://doi.org/10.1016/j.jdent.2012.11.003

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, *316*(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216

Hung, M., Voss, M. W., Rosales, M. N., Li, W., Su, W., Xu, J., Bounsanga, J., Ruiz-Negrón, B., Lauren, E., & Licari, F. W. (2019). Application of machine learning for diagnostic prediction of root caries. *Gerodontology*, *36*(4), 395–404. https://doi.org/10.1111/ger.12432

Ilhan, B., Lin, K., Guneri, P., & Wildersmith, P. J. (2020). Improving oral cancer outcomes with imaging and artificial intelligence. *Journal of Dental Research*, *99*(3), 241–248. https://doi.org/10.1177/0022034520902128

Koch, M. J., Bove, M., Schroff, J., Perlea, P., Garciagodoy, F., & Staehle, H. J. (2001). Black stain and dental caries in schoolchildren in Potenza, Italy. *ASDC Journal of Dentistry for Children*, *68*, 353–355. PMID: 11985199.

Kohara, E. K., Abdala, C. G., Novaes, T. F., Braga, M. M., Haddad, A. E., & Mendes, F. M. (2018). Is it feasible to use smartphone images to perform telediagnosis of different stages of occlusal caries lesions. *PLoS One*, *13*(9), e0202116. https://doi.org/10.1371/journal.pone.0202116

Krois, J., Ekert, T., Meinhold, L., Golla, T., Kharbot, B., Wittemeier, A., Dörfer, C., & Schwendicke, F. (2019). Deep learning for the radiographic detection of periodontal bone loss. *Scientific Reports*, *9*(1), 8495. https://doi.org/10.1038/s41598-019-44839-3

Lee, J. J. (2016). Future of the smartphone for patients and healthcare providers. *Healthcare Informatics Research*, *22*(1), 1–2. https://doi.org/10.4258/hir.2016.22.1.1

Lee, J., Kim, D., Jeong, S. N., & Choi, S. J. (2018). Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *Journal of Dentistry*, *77*, 106–111. https://doi.org/10.1016/j.jdent.2018.07.015

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., & Berg, A. C. (2016). *SSD: Single shot multibox detector*. European Conference on Computer Vision. 21–37. arXiv:1512.02325.

Mariakakis, A., Banks, M. A., Phillipi, L., Yu, L., Taylor, J., & Patel, S. N. (2017). BiliScreen. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(2), 1–26. https://doi.org/10.1145/3090085

Mariakakis, A., Baudin, J., Whitmire, E., Mehta, V., Banks, M. A., Law, A., Patel, S. N. (2017). PupilScreen. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(3), 1–27. https://doi.org/10.1145/3131896

Matthews, D. C. (2008). Powered toothbrush plus triclosan only as effective as manual brush and fluoride toothpaste for periodontal maintenance patients. *Evidence-based Dentistry*, *9*(3), 74–75. https://doi.org/10.1038/sj.ebd.6400594

Miki, Y., Muramatsu, C., Hayashi, T., Zhou, X., Hara, T., Katsumata, A., & Fujita, H. J. (2017). Classification of teeth in cone-beam CT using deep convolutional neural network. *Computers in Biology and Medicine*, *80*, 24–29. https://doi.org/10.1016/j.compbiomed.2016.11.003

Moutselos, K., Berdouses, E. D., Oulis, C., & Maglogiannis, I. (2019). *Recognizing occlusal caries in dental intraoral images using deep learning*. International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 1617–1620. https://doi.org/10.1109/EMBC.2019.8856553

Orhan, K., Bayrakdar, I. S., Ezhov, M., Kravtsov, A., & Ozyurek, T. J. (2020). Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. *International Endodontic Journal*, *53*(5), 680–689. https://doi.org/10.1111/iej.13265

Park, W., & Park, J. (2018). History and application of artificial neural networks in dentistry. *European Journal of Dentistry*, *12*(4), 594–601. https://doi.org/10.4103/ejd.ejd_325_18

Peres, M. A., Macpherson, L. M. D., Weyant, R. J., Daly, B., Venturelli, R., Mathur, M. R., Listl, S., Celeste, R. K., Guarnizo-Herreño, C. C., Kearns, C., Benzian, H., Allison, P., & Watt, R. G. (2019). Oral diseases: A global public health challenge. *The Lancet*, *394*, 249–260. https://doi.org/10.1016/S0140-6736(19)31146-8

Petersen, P. E., Bourgeois, D., Ogawa, H., Estupinanday, S., & Ndiaye, C. J. (2005). The global burden of oral diseases and risks to oral health. *The Lancet*, *83*(9), 661–669.

Philip, N., Suneja, B., & Walsh, L. J. (2018). Ecological approaches to dental caries prevention: Paradigm Shift or Shibboleth? *Caries Research*, *52*, 153–165. https://doi.org/10.1159/000484985

Rana, A., Yauney, G., Wong, L. C., Gupta, O., & Shah, P. (2017). *Automated segmentation of gingival diseases from oral images*. 2017 IEEE Healthcare Innovation Point-of-Care Technologies (HI-POCT). 144–147. https://doi.org/10.1109/HIC.2017.8227605

Redmon, J., & Farhadi, A. (2017). *YOLO9000: Better, faster, stronger*. Computer Vision and Pattern Recognition, 6517–6525. https://doi.org/10.1109/CVPR.2017.690

Schwendicke, F., Elhennawy, K., Paris, S., Friebertshauser, P., & Krois, J. J. (2020). Deep learning for caries lesion detection in near-infrared light transillumination images: A pilot study. *Journal of Dentistry*, *92*, e103260. https://doi.org/10.1016/j.jdent.2019.103260

Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition. computer vision and pattern recognition*. arXiv 1409.1556.

Valizadeh, S., Goodini, M., Ehsani, S., Mohseni, H., Azimi, F., & Bakhshandeh, H. J. (2015). Designing of a computer software for detection of approximal caries in posterior teeth. *Iranian Journal of Radiology*, *12*(4), e16242. https://doi.org/10.5812/iranjradiol.12(2)2015.16242

Vardell, E., & Boucrick, C. J. (2012). VisualDx: A visual diagnostic decision support tool. *Medical Reference Services Quarterly*, *31*(4), 414–424. https://doi.org/10.1080/02763869.2012.724287

How to cite this article: Zhang X, Liang Y, Li W, et al. Development and evaluation of deep learning for screening dental caries from oral photographs. *Oral Dis*. 2020;00:1–9. https://doi.org/10.1111/odi.13735