Detection of dental caries in oral photographs taken by mobile phones based on the YOLOv3 algorithm

Baichen Ding[1,2#], Zhuo Zhang[3#], Yiran Liang[4], Weiwei Wang[3], Siwei Hao[2], Ze Meng[5], Lian Guan[6], Ying Hu[7], Bin Guo[2], Runlian Zhao[3], Yan Lv[7]

[1]Medical School of Chinese People's Liberation Army (PLA), Beijing, China; [2]Department of Stomatology, the First Medical Centre, Chinese PLA General Hospital, Beijing, China; [3]The College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, China; [4]Department of Endodontics, Beijing Stomatological Hospital and School of Stomatology, Capital Medical University, Beijing, China; [5]Department of Stomatology, Handan First Hospital, Handan, China; [6]Department of Stomatology, Beijing Hospital of Integrated Traditional Chinese and Western Medicine, Beijing, China; [7]Beijing Institute of Dental Research, Beijing Stomatological Hospital and School of Stomatology, Capital Medical University, Beijing, China

Contributions: (I) Conception and design: B Ding; (II) Administrative support: None; (III) Provision of study materials or patients: Y Liang; (IV) Collection and assembly of data: S Hao; (V) Data analysis and interpretation: Z Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]The authors contributed equally to this work and are joint first authors.

Correspondence to: Bin Guo. Department of Stomatology, the First Medical Centre, Chinese PLA General Hospital, Fuxing Road, #28, Haidian District, Beijing 100853, China. Email: guobin58@126.com; Runlian Zhao. The College of Information Science and Technology, Beijing University of Chemical Technology, East Beisanhuan Road, #15, Chaoyang District, Beijing 100029, China. Email: rlzhao@mail.buct.edu.cn; Yan Lv. Beijing Institute of Dental Research, Beijing Stomatological Hospital and School of Stomatology, Capital Medical University, West Tiantan Li, #4, Dongcheng District, Beijing 100050, China. Email: lvdentist@126.com.

Background: The aim of this study was to explore the value of using the YOLOv3 algorithm for detection and diagnosis of dental caries in oral photographs taken with mobile phones.

Methods: Oral photographs taken with the mobile phones of 570 patients were used as 3 datasets: the augmented images (n=3,990), the enhanced images (n=3,990), and the combined augmented and enhanced images (n=7,980). Oral photographs taken by mobile phones from another 70 patients were used as an independent test set. We used the YOLOv3 network for migration learning for modelling. Diagnostic precision, recall, F1-score, and mean average precision (mAP) were calculated to obtain the detection and diagnostic performance of the YOLOv3 algorithm.

Results: After 3 independent training, the mAP value of the original group YOLOv3 algorithm was 56.20%, in which the precision for primary caries recognition was 76.92%, recall was 49.59%, and F1-score was 0.60; the precision for secondary caries recognition was 91.67%, recall was 52.38%, and F1-score was 0.67. The mAP value of the enhance group algorithm was 66.69%, in which the precision for primary caries identification was 81.82%, recall was 52.07%, and F1-score was 0.64, and the precision for secondary caries identification was 100%, recall was 33.33%, and F1-score was 0.50. The mAP value of the comprehensive group algorithm was 85.48%, in which the precision for primary caries identification was 93.33%, recall was 69.42%, F1-score was 0.80, and the F1-score for secondary caries identification was 0.50; precision was 100%, recall was 52.38%, and F1-score was 0.69.

Conclusions: The caries detection capability based on the YOLOv3 algorithm highlights the potential utility of deep learning in caries detection and diagnosis. Comparing the 3 experiments, the detection of the model trained after using image augmented and enhancement techniques was significantly improved.

Keywords: Dental caries; oral images; YOLOv3 algorithm; image recognition

Introduction

Dental caries is the most common chronic oral disease in the majority of adolescents and adults all over the world, and it can be divided into primary caries and secondary caries according to previous treatment experience. Primary caries is the most common type of caries in clinical treatment, while the secondary caries is caries that recurs due to the small gap between the edge of the filling and the surrounding dentin for various reasons, or the filling is performed before removal of the original lesion tissue, leaving the plaque in situ. The results of the fourth oral health epidemiological survey in China (1) showed that the caries rate of permanent teeth in China was 89.0% in the age group of 35–44 years old, 95.6% in the age group of 55–64 years old, and 98.0% in the age group of 65–74 years old. With such high prevalence of dental caries, the medical resources of dental professionals are not sufficient, how to improve the efficiency of dental caries treatment with the existing resources and assist doctors in medical diagnosis has become an urgent task. Deep learning techniques have demonstrated excellent capabilities in target recognition in photographic images, and this capability has led us to believe that deep learning may also have the potential to capture the characteristics of lesions in the oral cavity, which could facilitate the early detection of oral diseases.

Currently, artificial intelligence-aided dental diagnosis is being used initially in dental imaging and images taken with other specialized equipment, but its use is currently limited because patients must first be seen in a medical facility before it can be applied. Tooth recognition and classification in dental images is a prerequisite for disease diagnosis, and many researchers have implemented image processing and segmentation algorithms based on image processing and tooth segmentation for apical slice images (2). Lee *et al.* (3) divided 3,000 apical slice images into 2,400 training and validation datasets and 600 test datasets, using a GoogLeNet-based Inceptionv3 architecture based deep convolutional neural network (CNN) to achieve diagnosis of dental caries. Leite *et al.* evaluated the performance of a novel artificial intelligence (AI) for tooth detection and segmentation on panoramic X-ray films and showed that the AI model had a sensitivity of 98.9% and precision of 99.6% for tooth detection (4). Pierre and Mostafa developed a functional pyramid network (FPN) based on an AI tooth segmentation algorithm for automated tooth segmentation on cone-beam computed tomography (CBCT), which demonstrated excellent capabilities in terms of detection

accuracy and speed (5).

The Beijing University Stomatology Hospital applied deep learning and segmentation algorithms to train and build a dental plaque intelligent interpretation system based on dental image annotation before and after plaque staining (6), allowing people to take pictures at home using an oral endoscope to examine the amount and distribution area of plaque and provide accurate feedback for daily oral cleaning. Casalegno *et al.* (7) used a deep learning model in the instrument output grey-scale images to automatically detect and locate the extent of dental lesions to assist the physician in diagnosis. These studies are limited in practice due to the need for additional specialized equipment, difficulty of operation, and difficulty of obtaining standardized image results. Estai *et al.* (8) compared Mid-level dental providers caries assessments using smartphone-based dental photographs in 2016, arguing that Mid-level dental providers had the potential for remote caries screening using visual inspection by a senior dentist.

In previous studies, special testing instruments generally needed to be achieved or still required a professional doctor to spend time to assess remotely. And these special instruments can show specific characteristics of caries, but fraud is difficult to popularize to the social groups, and telemedicine can solve some problems of lack of medical resources. Still, the overall medical resources of countries or regions can not fundamentally solve the lack of medical help. Based on this situation, obtaining basic images through smartphones can reduce the equipment cost of the examination, and artificial intelligence is used to replace professional doctors to reduce the pressure of diagnosis and treatment in underdeveloped areas. This study used computer vision based on the YOLOv3 algorithm (9,10) instead of the human eye to detect caries images in oral photographs taken by mobile phones, in order to further reduce the difficulty of caries examination, so as to improve the secondary prevention of caries, achieve early detection, diagnosis, and treatment of caries, improve the pre-clinical detection rate, save dental medical resources, and improve the efficiency of clinical diagnosis and treatment. The three datasets used to train in this study are augmented dataset, enhanced dataset and comprehensive dataset. The augmented dataset contains the collected images and the results after the rotation and flipping of these images. Enhanced dataset contains images in augmented dataset are highlighted after local features after color and texture. Comprehensive dataset contains all the images in augmented dataset and enhanced dataset. The effects of
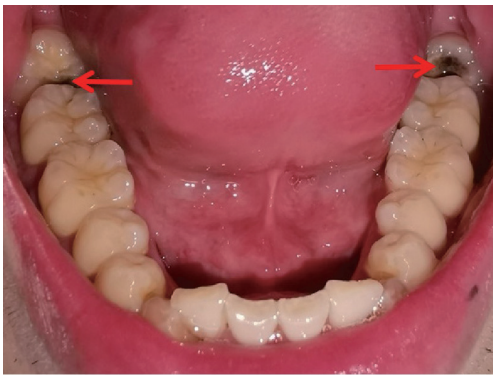
Figure 1 Primary caries collected by mobile phone (arrows).

Figure 2 Secondary caries collected by mobile phone (arrow).

the three training models were tested using an independent training set. The test data and test results of the three models were compared and reported as follows.

Methods

Object of study

Experimental materials

The types of smartphones used for the study included Huawei mate30pro (Huawei Technologies Co. Ltd., Shenzhen, China), Honor 9 (Huawei Technologies Co., Ltd., China), and iPhone 12 (Apple Corp., Cupertino, CA, USA), with no specific or exclusion criteria regarding the images. All data consisted of 640 images of different patients' oral cavities taken with mobile phones from patients attending the stomatology departments of the General Hospital of the Chinese People's Liberation Army (PLA), Beijing Stomatological Hospital, and Handan First Hospital. The primary caries or secondary caries in the

data were diagnosed and labelled by specialist clinicians. The data included images acquired from August 2019 to December 2020. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of the Chinese PLA General Hospital (No. S2021-273-01) and informed consent was taken from all the patients.

Map taking

The frontal occlusal photo required the patient to expose the maximum number of teeth and gums and could be assisted by common objects. The upper teeth photo required the patient to open the mouth as much as possible and take a picture of the maxillary dentition. The patient was required to lie as flat as possible, with the maxillary occlusal surface at 90–120° to the floor after opening the mouth. The handpiece was held as low as possible, depressed or held upside down, and the full maxillary arch was photographed from the bottom up. The patient's head was lowered slightly and their thumb and index finger were held slightly apart at the corner of the mouth while the handpiece was held diagonally upwards towards the bottom. For patients with smaller mouth openings, one photo of each of the right- and left-side teeth could also be taken. No pre-setting of lighting conditions, resolution, exposure, and other jumps in the photographic process were made for this study; all images were obtained by the device in the conventional photographic mode. In order to approximate the image effect in real-life scenarios, all images were captured using a smartphone and without professional equipment such as reflectors (*Figures 1,2*).

Photo information and tagging software

The images were collected in JPG format, with a size of 1–33 Mb. A total of 640 images were collected, and the image annotation software LabelImg (labelImg.1.8.4. November 2020. Tzu Ta Lin. https://pypi.org/project/labelImg/1.8.4/.10/12/2020) as used to annotate the caries in the image data (*Figure 3*).

Methodology

Firstly, all collected oral images were labelled for dental caries using LabelImg software; secondly, 570 of the images were expanded to 3,990 using the data augmentation technique and the data set was divided into a training set and a test set according to a certain ratio. The training set contained 3,990 images from 570 volunteers, while the test
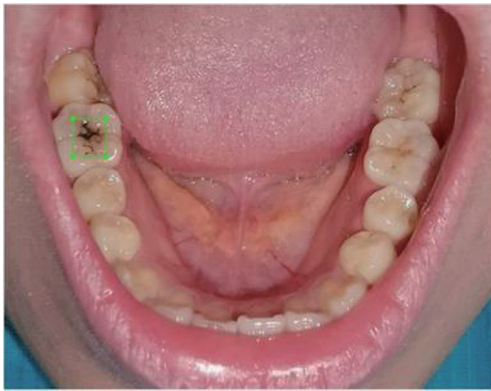
Page 4 of 11

Ding et al. Dental caries was detected by the YOLOv3 algorithm of images

Figure 3 The effect of LabelImg software labelling the caries in a photograph of the mouth, with a green box marking the position of the caries in the image.

Figure 4 Original image.

set contained 70 images from 70 volunteers. We ensured that the 2 subsets, the divided dataset and the test set, were mutually exclusive. Finally, the experiment was divided into 3 parts: the first part used the 3,990 images from the 570 images as the original group, the model was trained using the YOLOv3 algorithm, and test the model was trained using the independent test set; the second part used the image enhancement technique to enhance the images as the original group while keeping the number of images in the original group unchanged. In the third part, the original group data and the enhancement group data were mixed to form the comprehensive group, and the comprehensive group data set consisted of 7,980 images.

YOLOv3-based caries detection (11-13)

Algorithm implementation environment

This study was completed using a desktop computer, the specific environment configuration list was as follows, central processing unit (CPU): Intel® CoreTM i7-6700HQ; memory: 8 G; graphics card: Nvidia GeForce 965M; operating system: Windows 10; Python: Python 3.7.3; CUDA: CUDA 10.1; CUDNN; graphics processing unit (GPU) memory: 8 G. Python: Python 3.7.3; CUDA: CUDA 10.1; CUDNN; graphics processing unit (GPU) video memory: 8 G.

Preparation of data

The proprietary data set for this study were obtained from oral images of patients attending the Department of Stomatology of the General Hospital of the Chinese PLA, Beijing Stomatological Hospital, and the First Hospital of Handan. The inclusion criteria were as follows: age 12 years or older, in good general condition, and able to cooperate in completing intraoral photographs. Those with 1 or more of the following conditions were excluded from the study: undergoing orthodontic treatment; having restricted mouth opening with less than 3 fingers width of mouth opening. The area covered by the images includes mainly the occlusal surfaces of the upper and lower dentition, with some images of the buccal surfaces, labial surfaces, sublingual surfaces, and palatal surfaces of the molars.

The collected data images needed to be manually annotated for each image to complete the pre-production of the dataset. In order for the deep learning network to learn the location features of the target accurately and efficiently, it was necessary to ensure that the labeled locations were reasonable. In this study, LabelImg software was used as the data labelling tool, which not only accurately labels each image with its location, but also converts the label into a Python recognized XML file that stores the location and category of the target frame. After the sample images were labelled using the LabelImg software, the LabelImg software generated the labelling information file in XML format, which became the VOC dataset of the caries images.

Training deep neural networks requires a large amount of reasonable data, and too small a data set can lead to overfitting of the neural network. In this study, we used data augmentation techniques to increase the amount of data in the acquired dataset, thereby improving the generalization ability and robustness of the model. The original and augmented images are shown in *Figures 4,5*, respectively, and the annotation files corresponding to each image were transformed simultaneously during this process. The images of 570 participants were expanded to 3,990 for training purposes, and the images of 70 participants were divided
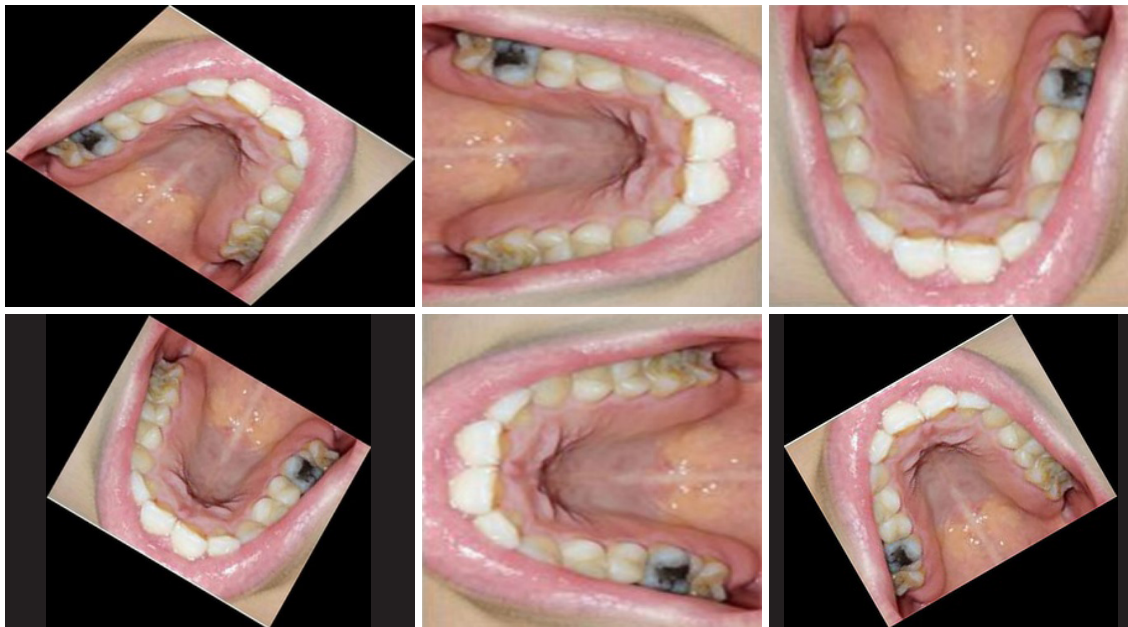
Figure 5 Image after data expansion.

into a separate test set.

Due to the different progression of caries, there may not be enough contrast between the lesion and the surrounding normal areas, making the caries less distinctive. Using the multi scale retinex with color restoration (MSRCR) algorithm, the image parameters could be adjusted so that details such as color and texture of the caries were highlighted after enhancement. The caries could only be located in the enhanced area when the annotation position was the same as the enhanced area, so we reasoned that training the model only on the enhanced image may improve the accuracy of the model (*Figure 6*). Therefore, in this study, the dataset was image enhanced as the enhanced training set.

Experimental procedure

The experiments were divided into 3 groups according to the way the data were processed: the original group used the original data set for model training; the enhance group used the enhanced data set for model training; and the comprehensive group used a combination of the original group data and the enhance group data for model training. The model was trained using a combination of the original group data and the enhancement group data. The 3 experiments were evaluated on the same independent test set.

The experimental steps were as follows: (I) process the dataset and generate 3 datasets by augmentation, enhancement and augmentation plus enhancement; (II) adjust the network parameters according to the dataset requirements to meet the dataset training requirements; (III) load the 3 datasets in turn for model training; (IV) optimize the parameters according to the training results of each dataset again; (V) after the model training, the model effect obtained from the 3 experiments was tested using an independent test set; (VI) the above hypotheses were verified based on the model evaluation metrics such as mean average precision (mAP), recall, and precision.

Detail processing: (I) Dataset enlargement and enhancement meant: using python language to write their own dataset processing tools to get the unified processing results of images and corresponding xml files. (II) Dataset division: original group experiment, 570 images were expanded to 3,990 images as the original group dataset; enhance group experiment, original group data was enhanced with the same number of images as enhance group dataset. Comprehensive group experiments, original group data and enhance group data were mixed as the comprehensive group dataset. (III) Parameter optimization: multiple experiments and statistical results were conducted, and the required parameters were adjusted according to the results of the previous phase of experiments under the current phase of experimental conditions (14). (IV) Model training: to prevent overfitting of the training model, the number of training rounds was reasonably controlled by

Figure 6 Images before and after image enhancement.

invoking the relevant functions. (V) Transfer learning: transferring labeled data or knowledge structures from relevant domains to improve the learning effect of the target domain or task, speeding up model training, and improving model accuracy. (VI) Evaluation metrics uses: all 3 groups were computed under the same conditions for the evaluation metrics.

Statistical analysis

We evaluated the diagnostic precision, recall, F1-score, and mAP values for the test dataset. These results were used to estimate and create better weighting factors for the neural network.

Precision

Precision was the ratio of the actual positive sample to all positive samples in the predicted sample.

$$precision = \frac{TP}{TP + FP} \qquad [1]$$

Recall

Recall was the ratio of actual positive samples to the number of predicted samples in the forecast sample.

$$recall = \frac{TP}{TP + FN} \qquad [2]$$

F1

F1 was the summed average of precision and recall.

$$F1 = \frac{2}{\frac{1}{Precise} + \frac{1}{Recall}} = 2 \cdot \frac{Precise \cdot Recall}{Precise + Recall} \qquad [3]$$

mAP

The mAP was the average AP value for multiple validation sets of individuals, and was used as a measure of detection accuracy in object detection.

AP

AP (average precision) indicating the area under the P-R curve using different values of precision and recall.

TP, FP, FN, AP

TP: true positives, the number of cases that were correctly classified as positive. FP: false positives, the number of cases that were incorrectly classified as positive. FN: false negatives, the number of instances that were incorrectly classified as negative.

Results

The mAP value of the original group training model for caries image recognition in the test set was 56.20%, in

Table 1 Precision, recall, F1-score, and AP values for the three training set detection results

Group	Primary caries				Secondary caries			
	Precision (%)	Recall (%)	F1-score	AP (%)	Precision (%)	Recall (%)	F1-score	AP (%)
Original group	76.92	49.59	0.60	55.63	91.67	52.38	0.67	56.78
Enhance group	81.82	52.07	0.64	68.21	100.00	33.33	0.50	65.17
Comprehensive group	93.33	69.42	0.80	85.09	100.00	52.38	0.69	85.87

which the precision of primary caries was 76.92%, recall was 49.59%, F1-score was 0.60, and AP value was 55.63%; the secondary caries algorithm recognized 91.67% of precision, 52.38% of recall, F1-score was 0.67, and AP value was 56.78%. The mAP value of the enhance group training model for caries image recognition in the test set was 66.69%, and the primary caries algorithm recognized a precision of 81.82%, recall was 52.07%, F1-score was 0.64, and AP value was 68.21%; the secondary caries algorithm recognized 100% of precision, 33.33% of recall, F1-score was 0.50, and AP value was 65.17%. The mAP value of the comprehensive group training model for caries image recognition in the test set was 85.48%, with 93.33% for primary caries, 69.42% for recall, 0.80 for F1-score, and 85.09% for AP. The secondary caries algorithm had a precision of 100%, a recall of 52.38%, an F1-score of 0.69, and an AP of 85.87% (*Table 1*).

The results showed good caries recognition by the algorithmic model after learning on the combined training set *Figure 7*, comparing the test results of the 3 models, training the enhanced images together with the unenhanced images enhanced the model's ability to recognize caries (*Figure 8*).

Discussion

Regular oral check-ups enable patients to understand their oral health, gain oral health care knowledge, detect oral conditions, and treat dental lesions as early as possible through check-ups in the early stages of the oral disease process. On the one hand, most common diseases in the oral cavity such as caries and periodontal diseases are chronic diseases of which the early symptoms are not obvious, although regular check-ups can achieve early detection, early diagnosis, and early treatment. On the other hand, some systemic diseases have special manifestations in the oral cavity, and stomatology is sometimes the first department to see them, such as scurvy, acute leukemia, and measles. By intervening in the early stages of oral

diseases, patients will achieve better treatment outcomes, shorter treatment times, and lower treatment costs. Regular oral check-ups are one of the most important measures to maintain oral health and promote general health. However, the results of the Fourth National Oral Health Epidemiological Survey show that oral health problems are still serious in the Chinese population. The need for early screening and diagnosis of oral lesions is in stark contrast to the limited medical resources currently available, and the development of AI technology holds promise for resolving this conflict.

The YOLOv3 algorithm is a type of target detection algorithm, and its basic idea is as follows: a series of candidate regions is generated on the image according to certain rules, and then the candidate regions are labelled according to the position relationship between these candidate regions and the real frame of the object on the image; a CNN is used to extract image features and predict the position and class of the candidate regions (*Figure 9*). The YOLOv3 algorithm generates candidate frames much less than Faster-RCNN and does not require 2 stages to complete the detection task. Therefore, these characteristics make the YOLOv3 algorithm faster, capable of achieving real-time response levels, and more suitable for the application scenario of AI-assisted diagnosis of dental caries (15).

Due to the excellent detection speed and capability of the YOLOv3 target detection algorithm for targets, related research has been conducted in many fields (16-20). This study attempted to address the problem of low prevalence of dental caries examination and insufficient oral health care resources through the YOLOv3 algorithm, which in turn provides support for the early diagnosis of oral diseases. As the training model of target detection needs a certain number of images to support and also needs to enhance the local features, the mAP of the training results of original group and enhance group was low, and the mAP of the training results of comprehensive group on caries detection was greatly improved. By mixing the original and enhanced images, the model achieved 69.42% of recall
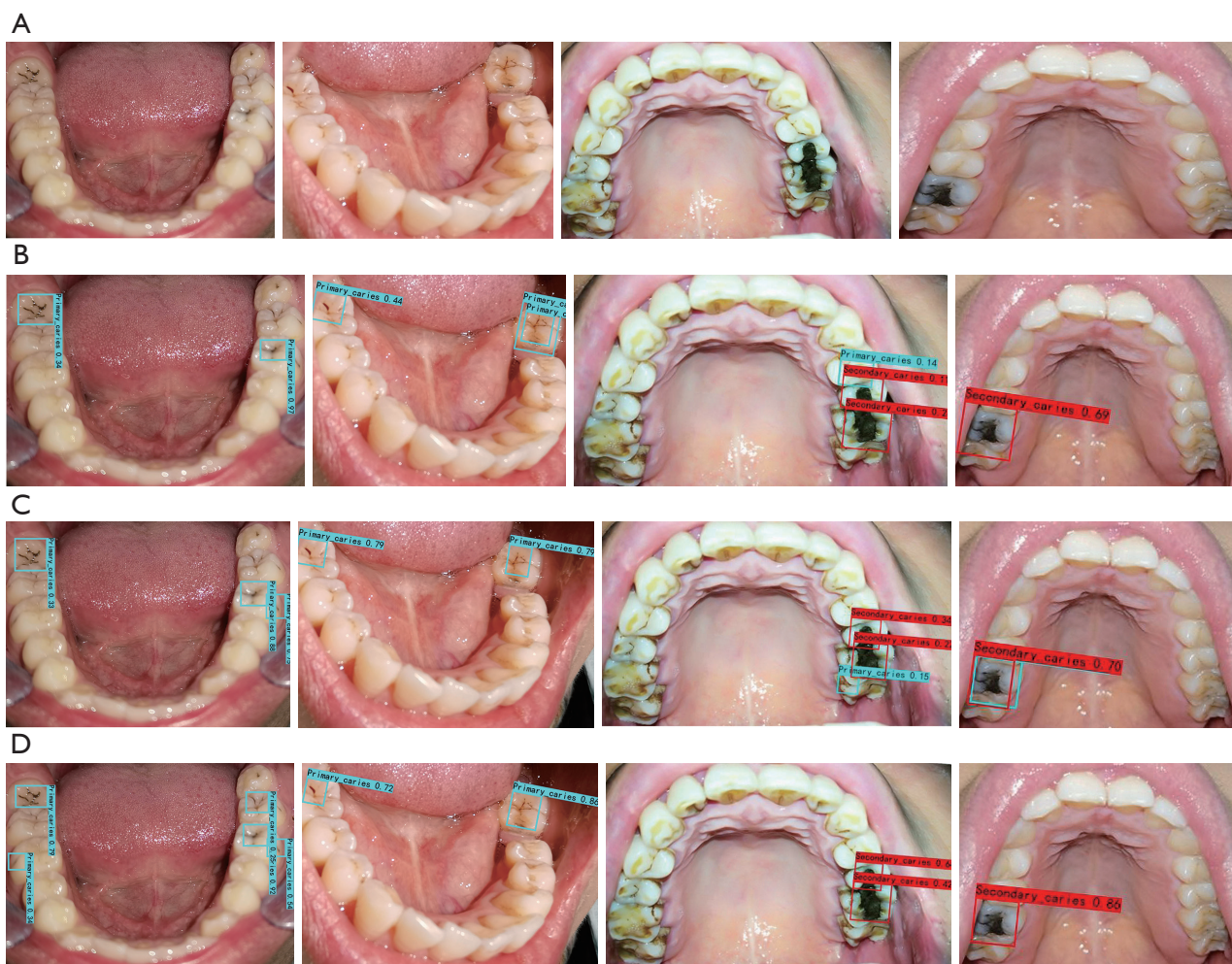
Page 8 of 11

Ding et al. Dental caries was detected by the YOLOv3 algorithm of images

Figure 7 Comparison of the detection results of the 3 training models in the test set: (A) original; (B) original group; (C) enhancement group; (D) comprehensive group. The detection and localization of primary caries are marked by the blue box and secondary caries is marked by the red box, and the values are shown as confidence levels.
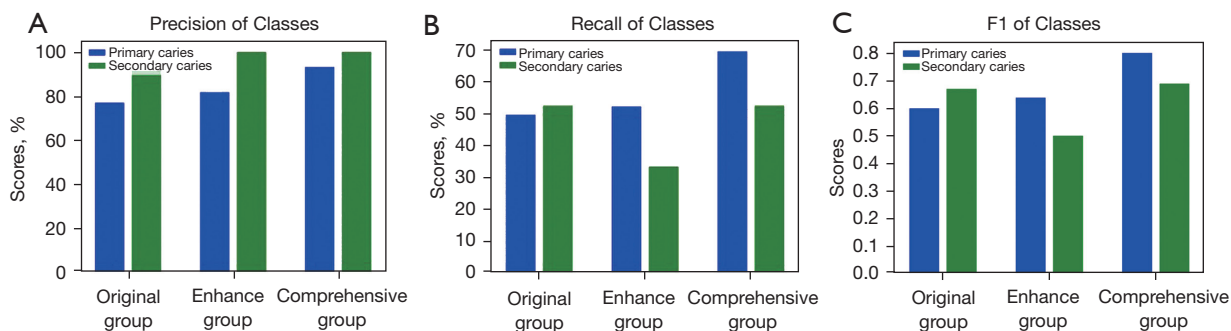

Figure 8 The YOLOv3 network was trained by the original group, enhance group, and comprehensive group models to detect primary caries (green) and secondary caries (blue) in the test set. The detection ability of each group is shown by the precision (A), recall (B) and F1 values (C) of the detection results.
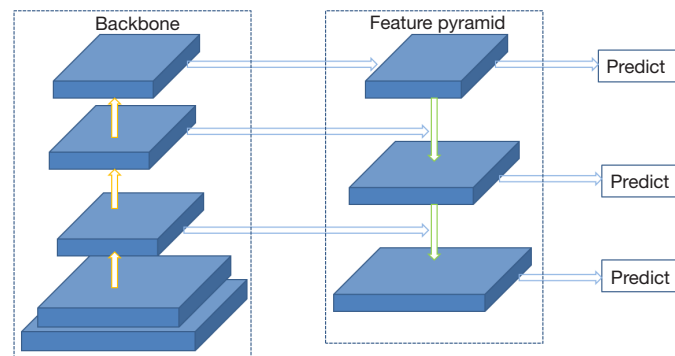
Figure 9 YOLOv3 network structure.

and 93.33% of precision for primary caries and 52.38% of recall and 100% of precision for secondary caries, and both F1-scores reached 0.6 or above, indicating that the enhancement groups' secondary caries showed a decrease in recall; but in the comprehensive group, the accuracy of the secondary caries remained the same while recall increased in the comprehensive group. We analyzed the following reasons for the low recall of enhance group secondary caries and the lack of significant improvement in recall in the comprehensive group: firstly, the number of secondary caries labelled in the dataset was small and the true recall of the model may not have been fully revealed during the testing process. Secondly, the clinical performance of the secondary caries was very different, for example, there were some differences in the images of the secondary caries with silver amalgam fillings and those with oral composite resin fillings. Finally, the dataset used in the comprehensive group experiment was a combination of the data from the original group and enhance group experiments, which had a larger sample size. Although the recall and original group results of the comprehensive group experiment were equal, the number of positive samples predicted by the comprehensive group experiment was higher than that of the original group experiment, and the number of positive samples predicted by the comprehensive group was higher than that of the original group experiment. The number of positive samples predicted in the comprehensive group was higher than that in the original group, and the number of positive samples predicted in the comprehensive group was higher than that in the enhance group, although the number of positive samples predicted was the same.

With their popularity and rapidly increasing performance, smartphones are gradually being used in all scenarios of human life. Scientists are also exploring how to combine smartphones with healthcare. From 2017 to 2018, there were already studies conducted in which scientists attempted to take smartphone images of teeth in the oral cavity and transmit them via the Internet to a remote location for oral examination by a dentist, and concluded that photographic diagnosis using smartphone images was feasible and accurate for tooth surface and extensive caries lesions, but the conditions of smart devices at that time still had some limitations for detecting mild and moderate lesions of caries (21,22). In 2019, Moutselos et al. used a deep learning model to classify and detect caries in 88 oral images taken by intraoral cameras with good accuracy (23). Askar et al. in Germany applied deep learning techniques to detect white spot lesions in dental photographs taken with a digital camera and showed satisfactory accuracy in detecting white spot lesions, especially fluorosis (24). Research on the acquisition of images of the oral cavity via smartphones and AI detection began to increase in 2020, with a sensitivity of 94.9% from Fu et al. who used a smartphone app to take images and train AI to identify lesion patterns in oral squamous cell carcinoma (25). While previous studies have typically pre-processed images to split them into multiple images and classify each image, in our study the detection, localization, and classification of lesions were performed simultaneously in the whole image. Most relevant to our study is the study by Zhang et al. (26), who examined the performance of a deep learning system for identifying dental caries in smartphone photographs of the oral cavity and significantly improved the performance of the model using the hard negative-mining algorithm. Although the same study investigated the detection of dental caries in oral photographs taken by smartphones by a deep learning system, our research was conducted differently to theirs. (I) Our study got closer to the final application scenario by using a smartphone to acquire

Page 10 of 11

Ding et al. Dental caries was detected by the YOLOv3 algorithm of images

the photos and not using professional dental equipment such as shades to assist in the photography, we choose 3 common oral actions (upper teeth photo, lower teeth photo, frontal occlusal photo) to show the maximum number of teeth images in the mouth, as a way to create a more realistic oral image dataset. (II) We used the YOLOv3 algorithm with migration learning to obtain a lighter training model, which can detect 15 images in 12 s. The YoloV3 algorithm shows excellent performance in detecting dental caries, shows good classification ability and positioning ability under the condition that the training data are not extensive, and shows superior performance in the detection speed of targets. Thus, detection of dental caries using the YOLOv3 algorithm can achieve very rapid detection results without sacrificing too much accuracy. (III) We tested the application of image enhancement techniques in this study and compared the data before and after enhancement to find a training method that could achieve better detection performance.

Our work still had some shortcomings. Firstly, our dataset was relatively limited as the acquisition and annotation of images requires significant human cost. Resource sharing and cooperation among multiple centers is important to support and facilitate the progress of the algorithm. Secondly, although the results of the study were highly accurate for the detection of dental caries in oral photographs, it was not possible to collect image information of all orientations of the teeth due to the specificity of the oral structure. Considering the fast detection capability of YOLOv3, image recognition of smartphone video content for oral structures may have the potential to solve this problem in order to obtain more image data and increase the number of tooth structures covered by the detection. It is hoped that this problem can overcome by obtaining larger image data and reducing the difficulty in the data collection process. Thirdly, the clinical manifestations of caries can present with a number of clinical symptoms in addition to the chromatic and morphological changes in the teeth, and a number of poor lifestyle habits can contribute to the diagnosis of the disease. Coding and weighting these factors and incorporating them into the algorithm will help to improve its accuracy. Finally, a limitation of our study was that the algorithm cannot make definitive predictions of occult caries, mainly because the photographs used to train the deep neural network may not fully represent the diversity and heterogeneity of occult caries lesions.

Conclusions

In conclusion, deep learning algorithms can improve the uneven distribution of public health resources by using photographs taken by smartphones for caries detection. The deep learning models we have developed show satisfactory results for the accurate detection of caries, especially primary caries. The results also show that the image enhancement approach can significantly improve the performance of the model.

Acknowledgments

Footnote

Data Sharing Statement: Available at https://dx.doi.org/10.21037/atm-21-4805

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-4805). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of the Chinese PLA General Hospital (No. S2021-273-01) and informed consent was taken from all the patients.

References

1. The Fourth National Oral Health Epidemiological Survey Technology Group, Chinese Stomatological Association. Work manual of the fourth national Oral health

epidemiological survey (2015-2018). Beijing: Chinese Stomatological Association, 2015:27-9.

2. Chen H, Zhang K, Lyu P, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. Sci Rep 2019;9:3840.

3. Lee JH, Kim DH, Jeong SN, et al. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent 2018;77:106-11.

4. Leite AF, Gerven AV, Willems H, et al. Artificial intelligence-driven novel tool for tooth detection and segmentation on panoramic radiographs. Clin Oral Investig 2021;25:2257-67.

5. Lahoud P, EzEldeen M, Beznik T, et al. Artificial Intelligence for Fast and Accurate 3-Dimensional Tooth Segmentation on Cone-beam Computed Tomography. J Endod 2021;47:827-35.

6. Huang Z, Luan QX. Evaluation of dental plaque by quantitative digital image analysis system. Beijing Da Xue Xue Bao Yi Xue Ban 2016;48:320-3.

7. Casalegno F, Newton T, Daher R, et al. Caries Detection with Near-Infrared Transillumination Using Deep Learning. J Dent Res 2019;98:1227-33.

8. Estai M, Kanagasingam Y, Huang B, et al. The efficacy of remote screening for dental caries by mid-level dental providers using a mobile teledentistry model. J Community Dentistry & Oral Epidemiology, 2016.

9. Zhao L, Li S. Object Detection Algorithm Based on Improved YOLOv3. Electronics 2020;9:537.

10. Redmon J, Divval A, Girshick R, et al. You only look once: Unified, real-time object detection. Available online: cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf

11. Pang L, Liu H, Chen Y, et al. Real-time Concealed Object Detection from Passive Millimeter Wave Images Based on the YOLOv3 Algorithm. Sensors (Basel) 2020;20:1678.

12. Ouyang L, Wang H. Aerial Target Detection Based on the Improved YOLOv3 Algorithm. 2019 6th International Conference on Systems and Informatics (ICSAI), 2019.

13. Liu X, Gan H, Yan Y. Study on Improvement of YOLOv3 Algorithm. J Phys Conf Ser 2021;1884:012031.

14. Larsen J, Hansen LK, Svarer C, et al. Design and regularization of neural networks: the optimal use of a validation set. Neural Networks for Signal Processing VI. Proceedings of the 1996 IEEE Signal Processing Society Workshop, 2002.

15. Girshick R. Fast R-CNN. IEEE International Conference on Computer Vision. New York: IEEE Computer Society, 2015:1440-8.

16. Shakarami A, Menhaj MB, Mahdavi-Hormat A, et al. A fast and yet efficient YOLOv3 for blood cell detection. Biomed Signal Process Control 2021;66:102495.

17. Atila Ü, Uçarb M, Akyol K, et al. Plant leaf disease classification using EfficientNet deep learning model. Ecol Inform 2021;61:101182.

18. Chen J, Wang Z, Wu J, et al. An improved Yolov3 based on dual path network for cherry tomatoes detection. J Food Process Eng 2021. doi: 10.1111/jfpe.13803.

19. Deepan P, Sudha LR. Effective utilization of YOLOv3 model for aircraft detection in Remotely Sensed Images. Mater Today Proc 2021. [Epub ahead of print]. doi: 10.1016/j.matpr.2021.02.831.

20. Jiang N, Wang J, Kong L, et al. Optimization of Underwater Marker Detection Based on YOLOv3. Procedia Comput Sci 2021;187:52-9.

21. Estai M, Kanagasingam Y, Huang B, et al. Comparison of a Smartphone-Based Photographic Method with Face-to-Face Caries Assessment: A Mobile Teledentistry Model. Telemed J E Health 2017;23:435-40.

22. Kohara EK, Abdala CG, Novaes TF, et al. Is it feasible to use smartphone images to perform telediagnosis of different stages of occlusal caries lesions? PLoS One 2018;13:e0202116.

23. Moutselos K, Berdouses E, Oulis C, et al. Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:1617-20.

24. Askar H, Krois J, Rohrer C, et al. Detecting white spot lesions on dental photography using deep learning: A pilot study. J Dent 2021;107:103615.

25. Fu Q, Chen Y, Li Z, et al. A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study. EClinicalMedicine 2020;27:100558.

26. Zhang X, Liang Y, Li W, et al. Development and evaluation of deep learning for screening dental caries from oral photographs. Oral Dis 2020. [Epub ahead of print]. doi: 10.1111/odi.13735.