BMC Oral Health

RESEARCH

Open Access

Data-driven AI platform for dens evaginatus detection on orthodontic intraoral photographs

Ruiyang Ren[1†], Jialing Liu[1†], Shihao Li[2], Xiaoyue Wu[1], Xingchen Peng[2], Wen Liao[1*] and Zhihe Zhao[1*]

Abstract

Background The aim of our study was to develop and evaluate a deep learning model (BiStageNet) for automatic detection of dens evaginatus (DE) premolars on orthodontic intraoral photographs. Additionally, based on the training results, we developed a DE detection platform for orthodontic clinical applications.

Methods We manually selected the premolar areas for automatic premolar recognition training using a dataset of 1,400 high-quality intraoral photographs. Next, we labeled each premolar for DE detection training using a dataset of 2,128 images. We introduced the Dice coefficient, accuracy, sensitivity, specificity, F1-score, ROC curve as well as areas under the ROC curve to evaluate the learning results of our model. Finally, we constructed an automatic DE detection platform based on our trained model (BiStageNet) using Pytorch.

Results Our DE detection platform achieved a mean Dice coefficient of 0.961 in premolar recognition, with a diagnostic accuracy of 85.0%, sensitivity of 88.0%, specificity of 82.0%, F1 Score of 0.854, and AUC of 0.93. Experimental results revealed that dental interns, when manually identifying DE, showed low specificity. With the tool's assistance, specificity significantly improved for all interns, effectively reducing false positives without sacrificing sensitivity. This led to enhanced diagnostic precision, evidenced by improved PPV, NPV, and F1-Scores.

Conclusion Our BiStageNet was capable of recognizing premolars and detecting DE with high accuracy on intraoral photographs. On top of that, our self-developed DE detection platform was promising for clinical application and promotion.

Keywords Artificial intelligence, Deep learning, Medical image classification, Orthodontic intraoral image

†Ruiyang Ren and Jialing Liu contributed equally to this work.

*Correspondence:
Wen Liao
liaowenssw@126.com
Zhihe Zhao
zhzhao@scu.edu.cn
[1]State Key Laboratory of Oral Diseases & National Clinical Research Center for Oral Diseases, Department of Orthodontics, West China Hospital of Stomatology, Sichuan University, Chengdu, Sichuan 610041, China
[2]Department of Biotherapy, Cancer Center, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China

Introduction

Dens evaginatus (DE) is a developmental dental anomaly characterized by the appearance of an accessory cusp or "tubercle" that protrudes from the occlusal or palatal surfaces of teeth, primarily affecting premolars [1, 2]. Morphologically, DE in premolars averagely possesses a tubercle with 2 mm in width and 3.5 mm in height [1]. This feature often causes occlusal interference and is prone to fracture, abrasion, and subsequent non-carious pulp exposure [3], resulting in a painful therapeutic process and unfavorable prognosis. However, since the patients are asymptomatic until severe abrasion or pulp-involved fracture of the tubercle, overlook is prone to happen.

Early detection of DE premolars is crucial. Conservative treatment methods, such as tubercle preservation, tubercle reduction, and pulp capping, are beneficial to remove occlusal interference and preserve the vitality of dental pulps, significantly decreasing the risk of pulp exposure [4]. Severe complications could also arise from the exposure of pulp tissue: pulp inflammation, pulp necrosis, periapical abscess, and even maxillofacial cellulitis and osteomyelitis of the jaws [5, 6], resulting in long treatment cycles and tortuous therapeutic procedures. It is also noticeable that when DE-caused pulp exposure involves immature permanent teeth, their development will be largely interfered with, specifically resulting in thinner root canal walls, shorter root lengths as well as open root apexes [7].

The timely detection of DE premolars is also vital in orthodontics. Malformed central cusps can lead to malocclusion, uneven distribution of occlusal forces, and other issues that negatively impact a patient's chewing function and oral health [1, 8]. During orthodontic treatment, these anomalies may alter the interproximal contact point and occlusal vertical dimension, necessitating careful consideration by orthodontists when designing treatment plans [8]. Adjustments to the design of orthodontic appliances or selective reduction of the malformed cusps may be necessary to facilitate effective treatment progress. Additionally, premolars are often the first choice for extraction in orthodontic treatment plans, with approximately 34.4% of orthodontic patients opting to have premolars removed [9]. Premolars affected by DE are typically prioritized in extraction plans. However, the excessive number of patients with a relatively small number of dentists in developing countries like China has resulted in a high workload and a heightened risk of misdiagnosis [10]. Therefore, an artificial intelligence (AI)-assisted diagnostic platform is essential to identify patients with DE premolars efficiently.

Nowadays, with the development of AI and its widespread utilization in dentistry, convenience and accuracy have been demonstrated in the field of automatic oral disease detection [10, 11]. Among the algorithms of AI, the convolutional neural network (CNN) embraces strong capacities in medical image feature processing [12], which has predominated the application of AI in medical imaging and won the academic consensus [13].

However, as AI-based studies have proliferated, certain limitations of prior work have also emerged. In terms of inappropriate data selection, poor data quality, selection bias, and limited generalizability due to small sample sizes can significantly impact the performance of AI algorithms [14]. Additionally, regarding the analysis of experimental results, some studies focus solely on the impressive capabilities of AI algorithms, often neglecting comparisons between AI performance and that of dentists. This oversight restricts AI's clinical applications, as it prevents the quantification of clinical effectiveness. Furthermore, result explainability remains contentious since AI's learning outcomes may not fully align with human conceptual understanding. However, completely unexplainable AI algorithms can introduce significant biases [15].

At the application level, while many studies have introduced their algorithms and some have made them publicly available, clinicians still face challenges in understanding and utilizing these AI tools effectively [16]. To minimize data selection bias, a standardized data collection protocol and a well-balanced dataset are essential. Researchers are supposed to add human-algorithm comparison tests to confirm clinical acceptance. Additionally, the internal logic of AI algorithms can be partially explained through visual analyses, such as attention-based visualization in image processing [11]. To enhance clinical accessibility of AI, there is a pressing need for research focused on developing fully functional applications, rather than solely presenting algorithmic code. Moreover, a deeper understanding of human-algorithm interactions should be actively pursued.

Currently, the majority of AI-based dental research relies on radiological images [10, 17], with only a few studies focusing on radiation-free intraoral data [18, 19]. Since intraoral photography only requires relatively inexpensive equipment like a smartphone or camera and has been more widely used even in underdeveloped and remote areas lacking X-ray equipment, an AI-assisted platform based on intraoral photographs has greater potential for widespread adoption. In this study, we aimed to generate a two-stage CNN algorithm model for detecting DE in orthodontic intraoral images. Furthermore, we aimed to develop an interactive and user-friendly platform for DE detection. We evaluated the performance of our platform and conducted a comparative analysis involving three dental interns to further validate its efficacy. We hope that our platform will enable

grassroots dentists and nonspecialists to easily diagnose DE with just a click.

Materials and methods

Dataset

The complete workflow of our research process is shown in Fig. 1. All intraoral images used in this study were collected from the Department of Orthodontics, West China Hospital of Stomatology, Sichuan University. Our investigation was approved by the Research Ethics Committee of West China Hospital of Stomatology (project number: WCHSIRB-D-2021-370). The inclusion criteria of our research are as follows: (1) patients in the permanent dentition stage who initially visited the Department of Orthodontics during the year 2018 to 2022, (2) both patients with and without DE premolars. Images were taken for documentation or educational purposes using professional single-reflex lens cameras (Nikon D300 with a Nikon Micro 105-mm lens) and a macro flash after the teeth were wiped and dried. To ensure image quality, any duplicate images from the same dentition and images with insufficient information (e.g., hazy, distorted, or saliva-contaminated images) were excluded. The original resolution of the included intraoral images, which was 3,008 × 2,008 pixels, was reduced to 450 × 300 pixels in JPEG format (RGB image) to save storage space and improve processing efficiency. To minimize potential bias, images with the following conditions were also excluded: (1) unrecognizable premolar morphology (such as severe abrasion or attrition, and severe dental crowding that obscures occlusal surfaces of premolars), (2) other developmental disorders or defects in premolar areas, including amelogenesis imperfecta, dentinogenesis imperfecta, dens invaginatus, gemination, and tooth fusion, as defined in authoritative articles [20, 21] and differentiated by morphological criteria, and (3) any restorations, resin-fillings, or appliances on premolars.

A total of 1,400 high-quality intraoral images were included in our study. The initial phase was dedicated to the object detection task, where our objective was to identify regions of interest (ROIs) within the intraoral images. To facilitate this, we divided the 1,400 intraoral images into a training set and a testing set, adhering to a 6:1 ratio. Consequently, 1,200 images were allocated to the training set, and 200 images were designated for the testing set. Subsequently, the study progressed to the image classification task, focusing on the differentiation of premolar regions based on the presence or absence of DE. For this purpose, premolar regions were meticulously extracted from the available images, resulting in a curated dataset specifically prepared for this classification challenge. The dataset was then bifurcated into a training subset, consisting of 1,011 premolars diagnosed with DE and 1,017 premolars without DE, and a testing subset, comprising 50 premolars with DE and 50 premolars without DE.

Image annotation

In the first stage, two dentists with more than ten years of clinical experience provided the golden standard of the premolar positions. The tool "labelImg" [22–24] (https://github.com/HumanSignal/labelImg/tree/master, version 1.8.1) was utilized to markup annotations of premolar positions and numbers: experts were required to create a rectangular box around each premolar; and for each box, a tooth number referring to FDI system was labeled. Finally, a total number of 6,312 premolars were labeled.

In the second stage, the same two dentists were asked to independently detect whether each labeled premolar was healthy or with defects of DE. We first conducted image matting according to the labeled rectangular boxes to obtain the images of single labeled premolars. The single premolar images were categorized into two groups, the healthy group and the DE group. All the premolars with inconsistent classification results were discussed and revalued together by a third expert with more than 20 years of clinical experience and these two dentists. If agreements can be reached after discussion, the agreed results were used. However, if agreements were failed to be achieved, we then excluded these premolar images. The manual evaluation results, both the tooth detection and DE determination, served as the golden standard for training our AI models.

BiStageNet model

In this study, we proposed a two-stage deep learning model, namely BiStageNet, for the detection of DE on intraoral photographs. The first stage was an object detection model that used the convolutional layers of VGG16 for feature extraction, followed by a fully connected layer that outputs the coordinates of the center points of four bounding boxes, totaling eight values. Each bounding box measures 90 × 90. The four bounding boxes outputted in the first stage are displayed on a graphical user interface, where dentists can adjust the position of the bounding boxes and modify their width and height to completely cover the area of premolar. The second stage deployed a CNN model (VGG-Lite) for binary classification. The classification model used the 90 × 90 RGB images cropped from the four bounding boxes outputted in the first stage for binary classification prediction. The specific workflow and architectures of our model are shown in Fig. 2.

VGG16 is a deep convolutional neural network architecture known for its effectiveness in image recognition tasks. In the context of our study, we utilized the convolutional layers of VGG16 in the first stage of the BiStageNet model for feature extraction in object detection.

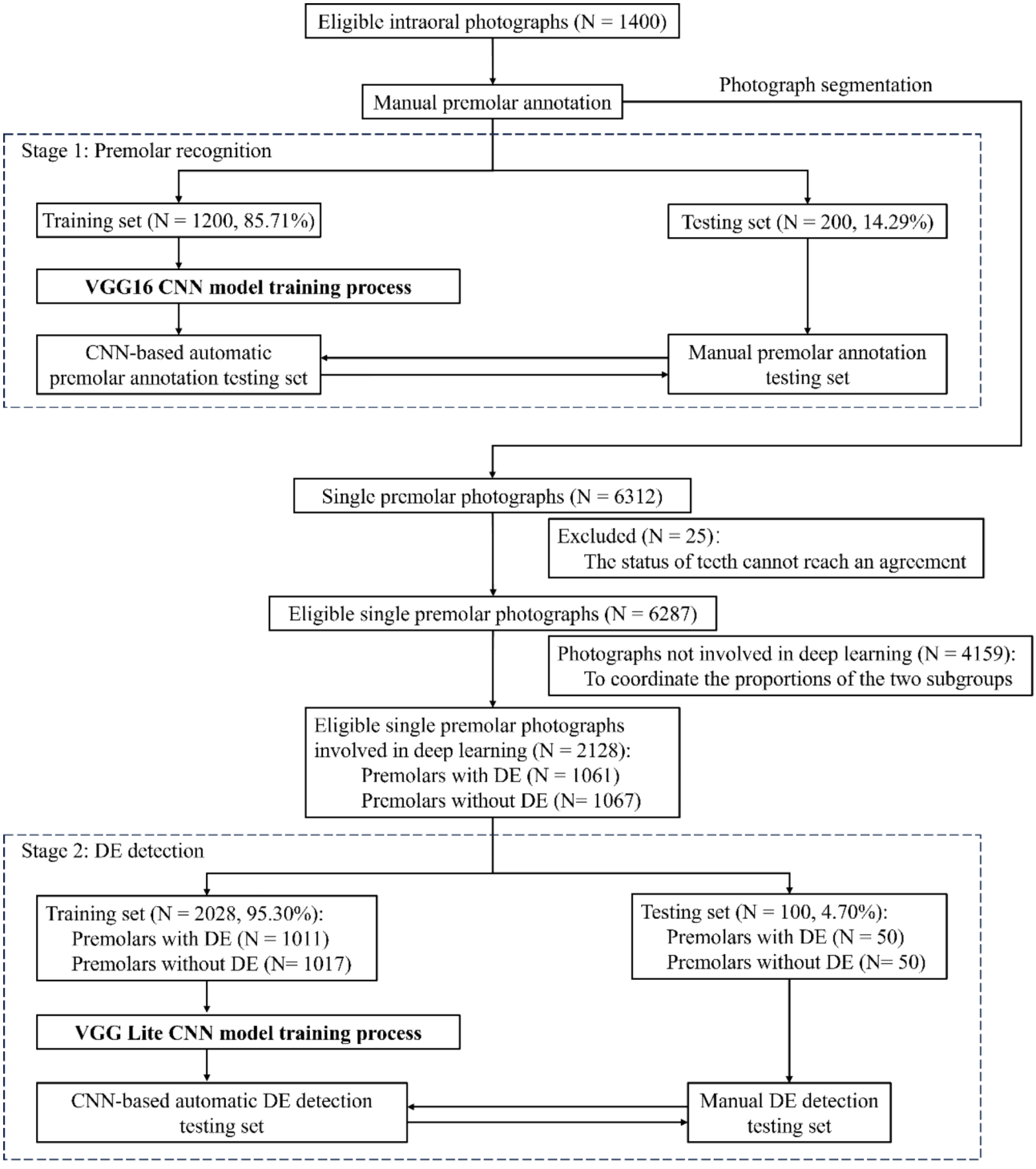

Fig. 1 The workflow of the two-stage research process for CNN model training to conduct automatic DE detection on intraoral photographs. CNN, convolutional neural network; DE, dens evaginatus

These convolutional layers process intraoral photographs with a resolution of 450 × 300 pixels to extract meaningful features that help in identifying the location of premolars. Following the convolutional layers, a fully connected layer outputs the coordinates of the center points of four bounding boxes, yielding a total of eight values. Each bounding box is sized at 90 × 90 pixels and is designed to encompass the premolar regions accurately.

VGG-Lite is a simplified version of the VGG16 model, optimized for computational efficiency without

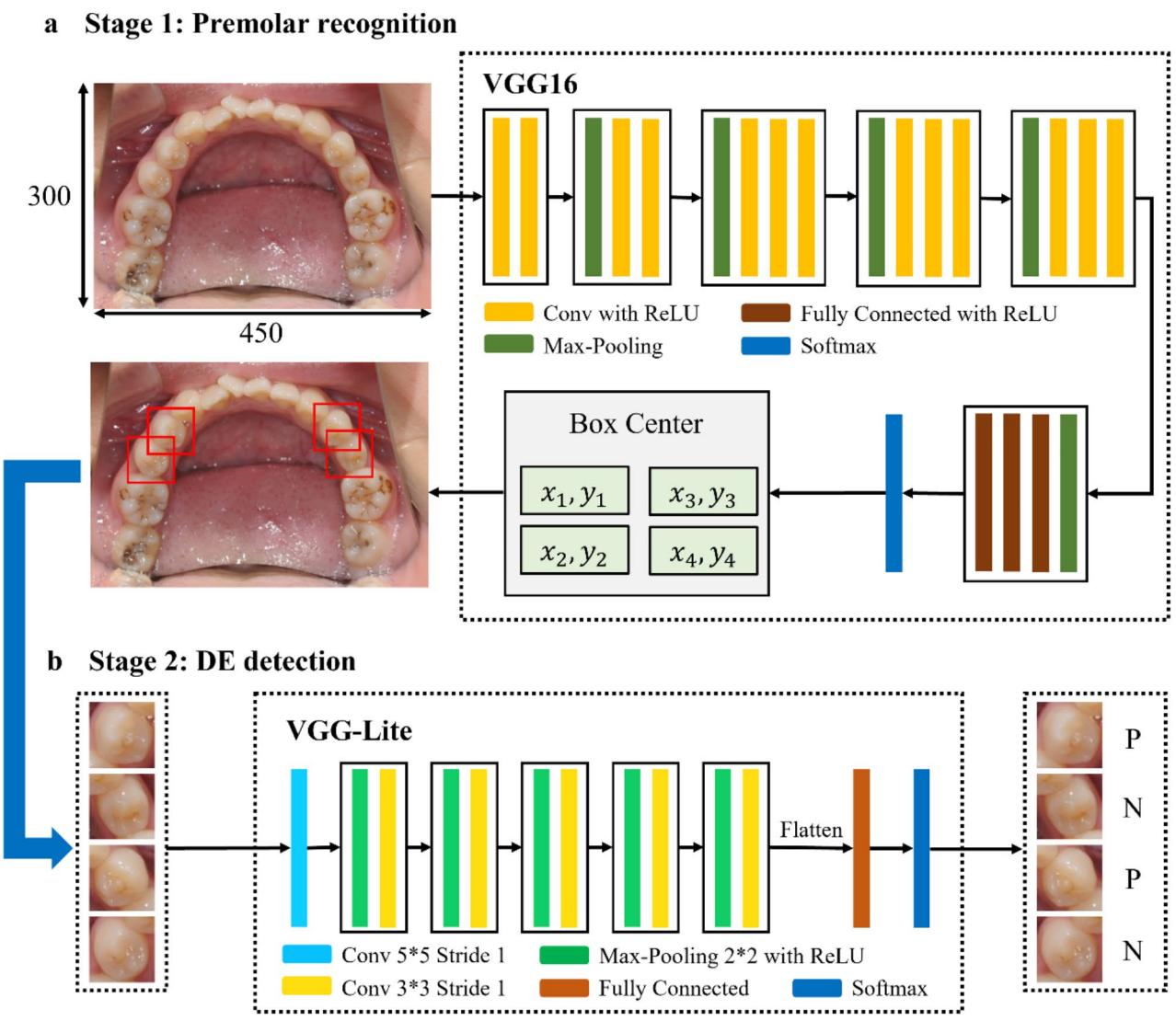

Fig. 2 The workflow and architecture of our BistageNet model for premolar recognition and DE detection. (**a**) The first stage: premolar recognition. Intraoral photographs with 450 × 300 pixels were inputted, and we utilized an object detection model that used the convolutional layers of VGG16, followed by a fully connected layer to output four bounding boxes to recognize the position of the premolars. (**b**) The second stage: DE detection. The images of recognized single premolars were segmented and inputted into our binary classification model, consisting of the VGG-Lite layers and a fully connected layer, to detect the existence of DE. The final results will be presented in two groups where "P" represents "positive" and "N" represents negative. DE, dens evaginatus

significantly compromising performance in classification tasks. In the second stage of the BiStageNet model, VGG-Lite serves as a binary classification model to detect the presence of dens evaginatus in the premolars. The architecture of VGG-Lite consists of one 5 × 5 convolutional layer and five 3 × 3 convolutional layers. Each convolutional layer is followed by a 2 × 2 max pooling layer to reduce spatial dimensions while retaining important features. Unlike the standard VGG16, VGG-Lite concludes with only one fully connected layer, followed by a softmax layer that outputs the probability of the binary classes—positive or negative for dens evaginatus detection.

The experiments for the BiStageNet model were conducted on a Windows 10 workstation equipped with an Intel Core i9-13900 K CPU, 128 GB of DDR4 RAM, and an NVIDIA RTX 4090 GPU to accelerate deep learning computations. The software environment utilized Python 3.9 and PyTorch 1.13 for implementing and training the convolutional neural network models.

Statistical analysis

To evaluate the recognition results of the premolars in intraoral images, we introduced the Dice coefficient, a widely used indicator for evaluating the segmentation

performance of the AI model, which is calculated as follows:

$$\text{Dice coefficient} = \frac{2^* |X \cap Y|}{|X| + |Y|}$$

where X represents the manually labeled area of premolar (the ground truth), and Y represents the auto-detected area conducted by our detect model. $|\mathrm{X} \cap \mathrm{Y}|$ represents the intersected area of X and Y, while $|\mathrm{X}| + |\mathrm{Y}|$ represents the sum area of X and Y. Given that our input intraoral photographs are in RGB formats, input data should be converted into binary format first. Additionally, the area size is determined by the total number of pixels. The Dice coefficient ranges between 0 and 1, with higher values indicating a greater proportion of overlap between X (the manually labeled premolar area) and Y (the auto-detected area by our model). A Dice coefficient closer to 1 signifies improved accuracy in automatic premolar recognition and segmentation.

To figure out the DE detection capacities, we took the following evaluation indicators into consideration: accuracy (ACC), sensitivity (SE), specificity (SP), positive predictive value (PPV), negative predictive value (NPV), as well as F1-score. The calculation formulae are as follows:

$$\mathrm{ACC} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\mathrm{SE} = \frac{TP}{TP + FN}$$

$$\mathrm{SP} = \frac{TN}{TN + FP}$$

$$\mathrm{PPV} = \frac{TP}{TP + FP}$$

$$\mathrm{NVP} = \frac{TN}{TN + FN}$$

$$\text{F1 Score} = \frac{2 * PPV * SE}{PPV + SE}$$

where TP, TN, FP, and FN present true positives, true negatives, false negatives, and false negatives, respectively. Additionally, we also produced the receiver operating characteristic (ROC) curves and calculated the area under the ROC curve (AUC) to evaluate the performance of our model.

Subjective assessment based on BiStageNet

To evaluate the efficacy of our DE detection tool, a comparative analysis was conducted involving the tool and three dental interns, each possessing one year of clinical training experience. This assessment focused on the same testing set of 100 regions, derived from 30 intraoral images encompassing 50 premolars diagnosed with DE and 50 premolars without DE. The interns, who did not receive any additional training specific to this study, initially rendered their diagnostic decisions based solely on their subjective judgment, without the assistance of the DE detection tool.

Two weeks subsequent to the initial assessment, a follow-up evaluation was undertaken, wherein the same three dental interns re-examined the identical 100 regions. During this session, the DE detection tool's predictive results were made available to them. The interns were tasked with determining the concordance between the tool's predictions and their own subjective visual assessments for each region. Discrepancies between the tool's predictions and the interns' judgments prompted a final decision-making process, wherein the interns, leveraging both their clinical insights and the tool's output, adjudicated the ultimate determination regarding the presence of DE.

Additionally, during the comparative analysis, we also randomly selected 100 single premolar images (50 with DE and 50 without DE) and used Cohen's Kappa value to assess the agreement level between the diagnostic outcomes of our DE detection tool and those of three dental interns. Cohen's Kappa value was calculated as follows:

$$\kappa = \frac{p_0 - p_e}{1 - p_e}$$

Where p_0 presents the observed agreement proportion, and p_e presents the expected agreement proportion.

Results

Premolar recognition and DE detection results

In the premolar recognition and segmentation stage, a mean Dice coefficient of 0.961 was demonstrated on the testing set. Further, after CNN training in the second stage, our algorithm could detect DE in different premolars with desirable outcomes: the overall accuracy of DE detection reached 85.0%, with a sensitivity and specificity of 88.0% and 82.0%, respectively. the positive predictive value and negative predictive value are 83.0% and 87.2%, respectively, and the F1-Score is 0.854. The AUC of the overall DE detection was 0.930. Visualizing our two-stage convolutional neural networks in Grad-CAM heatmaps in Fig. 3, we can tell that the attention of the algorithm was mainly paid to the tubercle signatures at the occlusal surfaces of premolars during the automatic detection process.

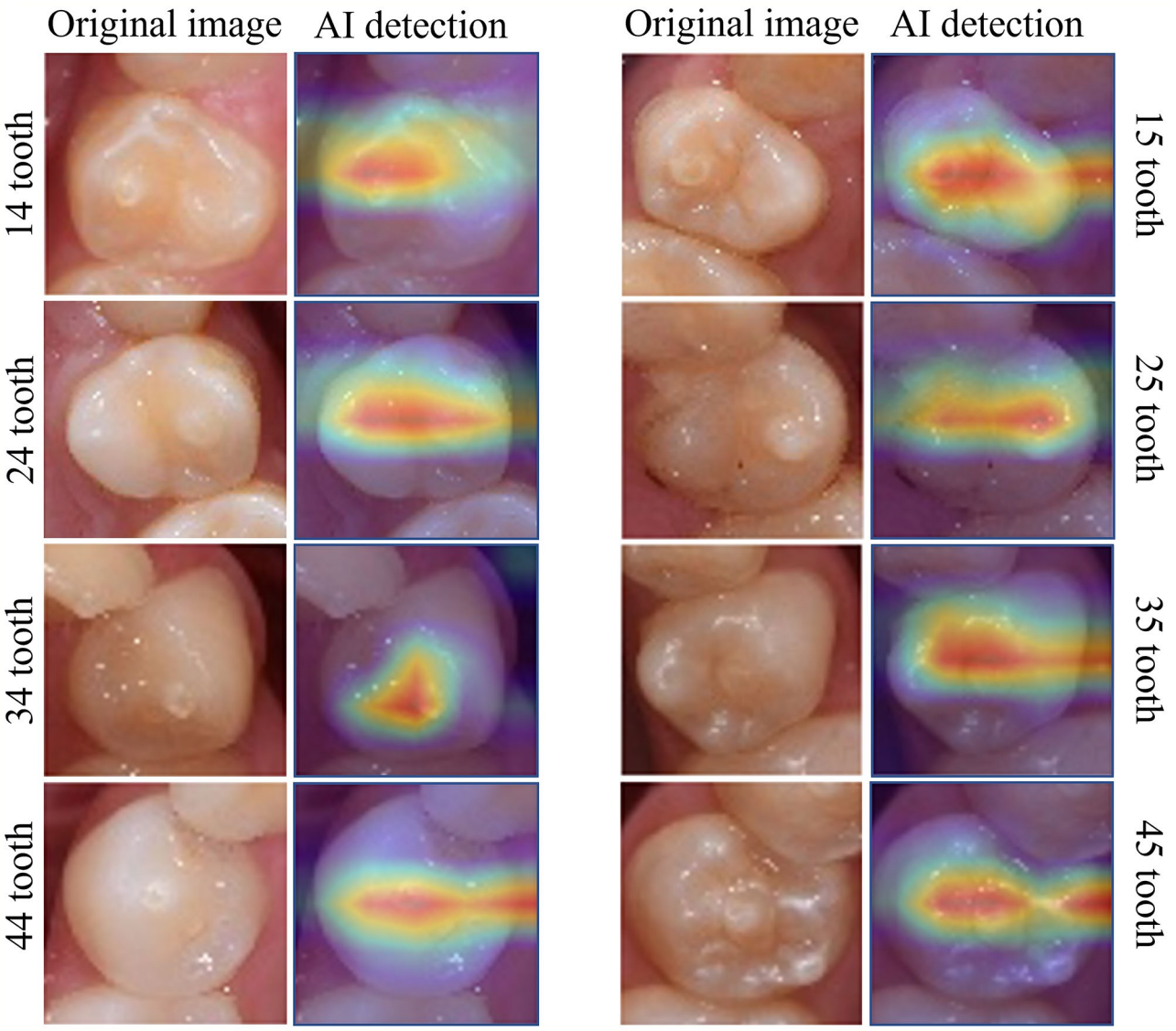

Fig. 3 Heatmaps of the attentions of our BiStageNet. The heatmap illustrates examples of BiStageNet's attention during single-tooth DE detection. The color intensity reflects the degree of attention BiStageNet assigns to various regions of the tooth image, with red indicating the highest attention and purple indicating the lowest

The test of DE detection tool

Based on our two-stage CNN algorithm and training results, we have eventually constructed an application tool that can automatically recognize premolars and detect the existence of DE on premolars using Pytorch (Fig. 4). To start with, by clicking the "Open Image" button we can import one intraoral image with any size and resolution, and then click the "Gen Boxes" button and four rectangular boxes appear to automatically select four premolars. Manual exact adjustment of these boxes can be achieved by dragging them with your mouse (changing their positions) or clicking the "height+", "height-", "width+", and "width-" buttons (changing their sizes). After selecting premolars properly, clicking the "Central Cusp Deformity" button will generate the automatic decision-making results of whether DE exists in these areas, where "P" presents the positive result and "N" for the negative outcome.

Our DE detection tool exhibited high consistency with dental interns during comparative analysis, where The Kappa values obtained were 0.859 for the model versus Dental Intern 1, 0.839 for the model versus Dental Intern 2, and 0.818 for the model versus Dental Intern 3. Table 1 presents the outcomes of subjective assessments conducted by three dental interns, both with and without the assistance of the DE detection tool, alongside the performance metrics of the tool operating autonomously. The data reveal that, when operating manually (without tool assistance), the dental interns generally exhibited high sensitivity, indicating a strong ability to correctly

Fig. 4 The operation procedures and diagnostic examples of our self-developed application tool for automatic premolar recognition and DE detection. (**a**) The instructions for our DE detection tool: (1) Click the "Open Image" button to import one intraoral image with any size and resolution; (2) Click the "Gen Boxes" button to generate four rectangular boxes which can automatically select four premolars. Fine-tuning these boxes can be implemented by dragging them with your mouse or clicking the "height+", "height-", "width+", and "width-" buttons; (3) Click the "Central Cusp Deformity" button to automatically determine whether DE exists in these areas, with "P" indicating a positive result and "N" indicating a negative outcome. (**b**) Three examples of DE detection outcomes by our software. DE, dens evaginatus

Table 1 The results of subjective assessment with and without DE detection tool

Operator	Method	TP	FN	TN	FP	SE	SP	PPV	NPV	F1-Score
Our tool	Automatic	44	6	41	9	0.88	0.82	0.83	0.87	0.854
Dental intern 1	Manual	48	2	42	8	0.96	0.84	0.86	0.95	0.906
	Tool-aid	48	2	45	5	0.96	0.90	0.91	0.96	0.932
Dental intern 2	Manual	47	3	40	10	0.94	0.80	0.82	0.93	0.879
	Tool-aid	48	2	44	6	0.96	0.88	0.89	0.96	0.923
Dental intern 3	Manual	49	1	39	11	0.98	0.78	0.82	0.98	0.891
	Tool-aid	49	1	45	5	0.98	0.90	0.91	0.98	0.942

identify true cases of DE. Upon integration of the DE detection tool, a noticeable enhancement in specificity was observed for all three interns, without a significant compromise in sensitivity. This improvement in specificity underscores the tool's utility in reducing false positive rates, thereby refining the accuracy of DE detection. Notably, the tool's influence also led to improvements in the PPV, NPV, and F1-Scores, further evidencing its positive impact on diagnostic precision.

Discussion

The early-stage DE causes no symptoms, thus can be easily neglected during dental visits. However, its high possibility of abrasion and fracture is prone to bring underlying severe problems. Pulp-unaffected dental abrasion or fracture can be restored with satisfactory results, but irreversible pulp lesions usually occur early, which largely prevents young permanent premolars from development and leads to unfavorable prognosis [25]. Also, of clinical imperative to the orthodontist is that premolar extraction cases should be planned to include extraction of the DE premolars instead of the normal ones [26]. Therefore, it is of necessity and significance for dentists to conduct early DE detection to prevent its progression in time and to design optimal treatment schemes. To our best knowledge, our study is particularly novel in first presenting an effective AI-based approach to detect DE on orthodontic intraoral images. The results suggest that our AI model can be a useful tool for assisting DE detection with high accuracy and inexperienced dental interns can expediently screen for DE in premolars using our tool. Our findings demonstrate that with the assistance of our AI model, grassroots dentists or nonspecialists could achieve an average accuracy of 93.5% in detecting DE premolars.

Our model demonstrated an overall diagnostic accuracy of 85.0%, a sensitivity of 0.880, an F1-Score of 0.854, and an AUC of 0.930 when recognizing all the premolars with DE in the testing dataset, indicating superior diagnostic outcomes were obtained. However, a shortcoming of our tool is that its specificity (0.820) is relatively lower than its sensitivity, indicating that while our tool effectively screens for potential DE cases, it may mistakenly diagnose healthy premolars as having DE. The tool's assistance to dental interns significantly improved specificity, suggesting that in clinical applications, DE misdiagnosis rates may be reduced. PPV and NPV are also critical indicators in clinical diagnosis. Given that DE is a dental deformity with relatively low prevalence, NPV is expected to be high while PPV remains relatively lower, consistent with our findings. With the aid of our tool, the PPV values for dental interns also improved markedly, suggesting that our tool offers substantial clinical application value. In addition, to simply visualize how our tool's decisions were made, we provided heatmaps of the attentions of our BiStageNet in Fig. 3. The heatmap examples of automatic DE detection demonstrated that our unsupervised learning algorithm has successfully paid its attention to the tubercle sites on the occlusal surfaces of premolars, the characteristic structure of DE deformity, which confirmed our detection results as reliable. And finally, we have made a DE detection application program on the basis of our CNN model, which can automatically select premolars and judge the existence of DE.

In orthodontic clinics, DE is a relatively uncommon dental anomaly, making it crucial to prioritize diagnostic sensitivity to capture as many DE cases as possible. On the basis of high sensitivity, specificity must also be carefully considered to minimize misdiagnosis. Our analysis of three dental interns, which mimicked the diagnostic behaviors of less experienced dentists, exhibited high sensitivity but clinically unacceptable specificity, with intern 3's specificity dropping below 80%. With the assistance of our DE diagnostic program, they succeeded in maintaining high sensitivity while significantly enhancing specificity by 6–12%, significantly reducing the risk of clinical misdiagnosis and subsequent incorrect orthodontic decisions. The human-algorithm comparison test showed that our detection program is a valuable aid for DE preliminary screening in clinical practice.

In addition to clinical reliability, our DE diagnostic tool is also highly user-friendly. As previously mentioned, dentists and patients are required only to input intraoral images in the program and fine-tune the positions and sizes of automatic-generated premolar-targeted boxes, and our decision-making algorithm will evaluate the involved scope of DE simultaneously, which greatly facilitates the DE clinical diagnosis of dentists. Moreover, such

a simple-to-use tool also minimizes obstacles for non-specialists to screen for premolar DE, which brings much convenience to improve the oral health of the public.

Our model demonstrated a satisfactory premolar detection outcome with the mean Dice coefficient of 0.961, indicating that the auto-segmented area and the manual premolar recognition area were almost completely overlapped. The AI-based tooth recognition has previously been implemented in multiple studies, most of which were conducted using radiographs [27, 28]. Tuzoff et al. [28] utilized Faster R-CNN to detect permanent tooth position, by training with 1,352 panoramic radiographs, a sensitivity of 0.9941 and a precision of 0.9945 were achieved. Another study possessed a sensitivity and precision of 0.9804 and 0.9571, respectively, when implementing deciduous teeth recognition using the same type of algorithm [27]. Only one study conducted landmark detection on intraoral occlusal images [29], where the VGG19 model demonstrated the best landmark detection capacity with a mean error of 0.84 mm in the maxilla and 1.06 mm in the mandible. Our results also verified the outstanding tooth recognition capacity of VGGnet. Therefore, Faster R-CNN and VGGnet are promising automatic tooth recognition tools that can be utilized in future research.

There have been several studies developing AI models concerning intraoral data, most of them aimed at distinguishing the multiple status of teeth, including the existence of dental caries [30], the existence of various restoration materials [31], and the existence of malocclusion [19]. And only one study focused on the status of soft tissue (automatic gingivitis detection) [32]. The limited number of studies suggests that intraoral images still need further exploration with AI approaches. As reported in other studies, the detection accuracies of dental diseases on photographs carried out by AI-based models ranged widely from 64% (Angle malocclusion classification on intraoral dentition photographs) [19] to 99.4% (gold restoration detection on intraoral single tooth photographs) [18] based on various detection difficulties and the complexity of the information provided by photographs. The morphology and color of the accessory dental cusps are similar to those of the normal ones since the histological composition of them is identical. And clinically, the accessory cusps were usually fractured or worn, with a statistical rate of appropriate 75% [26], so that the teeth with DE are hard to differentiate from the ones with deep occlusal pits and fissures, or with occlusal dental caries in photographs. In consideration of the above difficulties, therefore, our diagnostic tools possess very high accuracy for the distinguishment of DE teeth from patients' image data, and our independently developed DE diagnostic tool can help dentists to diagnose DE more preciously and develop treatment plans more intensively.

Existing methods for disease detection using intraoral data often rely on segmentation tasks driven by deep learning, as demonstrated by [33] who developed a segmentation-based approach to identify dental calculus, gingivitis, and dental caries from intraoral photographic images. While effective, these methods require pixel-level annotation, which is labor-intensive, time-consuming, and resource-intensive, posing challenges for scalability in practical applications. In contrast, our proposed method addresses these limitations by utilizing bounding box annotations combined with classification-based approaches, eliminating the need for pixel-level detail. This reduces the annotation workload and associated costs, providing a more efficient and cost-effective solution for disease detection in intraoral images while maintaining high levels of accuracy and applicability in large-scale screening scenarios.

However, our study also has some limitations. Firstly, the diagnostic accuracy of DE teeth needs to be further improved. More intraoral images are needed for model training, and transfer learning can also be an effective approach [34]. Secondly, our model focused not only on the accessory dental cusps in DE premolars, but in some cases the normal cusps were also taken into consideration, as shown in Fig. 3 (e.g., the 44 tooth and the 45 tooth). To further optimize our model, we can conduct supervised learning by marking annotations on the specific accessory cusps to narrow its range of attention.

The detection of DE teeth is the starting point of a novel research direction of AI-based disease diagnosis, and we envisage that future studies can dig deeper into the automatic detection of DE by involving more tooth types rather than only premolars, and by classifying DE into several categories according to the position of accessary cusps or their fracture and abrasion degree, to better guide clinical decision making. Our DE diagnostic tool is the first step of the automated diagnosis of multiple dental diseases, further researchers are awaited to integrate the algorithms of various dental disease detection based on intraoral data and form a high-accuracy mode for clinical usage, achieving the simultaneous and comprehensive reflection of both the status of soft and hard tissues in one tool, which will largely improve the diagnostic efficiency of dentists. AI is a powerful and promising tool for advancing public health in the future. When applied to public health, information security becomes essential. Clinical data contain significant amounts of patients' privacy information; therefore, developers of future diagnostic tools must focus not only on algorithm optimization but also on approaches to face legal, ethical, and cybersecurity challenges. To make AI algorithms truly beneficial for public health, robust laws against

AI-based privacy invasion must be established, along with user informed consent protocols and measures to prevent data leakage and misuse [35]. It is hoped that, by simply pressing the camera shutter and uploading intraoral photographs, one can have a preliminary view of his or her own oral health status, which is of great use in the promotion of public oral hygiene.

Conclusion

In this study, we have successfully constructed a BiStageNet model and demonstrated that deep learning methods are capable of achieving automatic premolar recognition and DE detection with high accuracy in intraoral photographs. Based on our custom-made CNN algorithms, we have also developed an automatic DE detection platform, which was applicable to both dentists and nonspecialists with promising diagnostic results. CNNs are powerful tools to improve DE early diagnosis rate, diagnostic accuracy as well as clinical work efficiency, and further improve the public oral health status.

Abbreviations

DE	Dens evaginatus
CNN	Convolutional neural network
ROIs	Regions of interest
SEN	Sensitivity
SP	Specificity
PPV	Positive predictive value
NPV	Negative predictive value

Acknowledgements

We are grateful to Dr. Fei Yu for her assistance and support in funding: Natural Science Foundation of Sichuan Province (2023NSFSC0562).

Author contributions

R.Y. Ren: Conceptualization, Methodology, Formal analysis, Investigation, Writing-Original Draft, Writing-Review & Editing, Data curation, Visualiation; J.L. Liu: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing-Original Draft, Writing-Review & Editing; S.H. Li: Methodology, Software, Validation, Formal analysis, Writing-Original Draft, Visualization; X.Y. Wu: Investigation, Data Curation, Writing-Original Draft; X.C. Peng: Software, Validation, Resources, Writing-Review & Editing; W. Liao: Conceptualization, Validation, Writing-Review & Editing, Supervision, Funding acquisition; Z.H. Zhao: Conceptualization, Resources, Supervision, Project administration, Funding acquisition.

Funding

This work was supported by Major project of Sichuan Science and Technology Department (2022ZDZX0031), Chengdu Artificial Intelligence Applications and Development Industrial Technology Basic Public Service Platform (2021-0166-1-2), International Orthodontics Foundation Young Research Grants Award (IOF2022Y08), the central government guides the construction of local S&T innovation funds (2021ZYD104), Natural Science Foundation of Sichuan Province (2023NSFSC0562), Sichuan Science and Technology Program (2025ZNSFSC1592), and the College Students' Innovative Entrepreneurial Training Plan Program (S202310610252).

Data availability

The source code for our deep learning model, along with the graphical user interface of our tool, is publicly available on GitHub: https://github.com/che006/dens-evaginatus-detection.

Declarations

Ethics approval and consent to participate

This study was approved by the Research Ethics Committee of West China Hospital of Stomatology (project number: WCHSIRB-D-2021-370).

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Received: 25 September 2024 / Accepted: 19 November 2024
Published online: 01 March 2025

References

1. Levitan ME, Himel VT. Dens evaginatus: literature review, pathophysiology, and comprehensive treatment regimen. J Endod. 2006;32(1):1–9.
2. Dankner E, Harari D, Rotstein I. Dens evaginatus of anterior teeth. Literature review and radiographic survey of 15,000 teeth. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 1996;81(4):472–5.
3. Merrill R.G. Occlusal anomalous tubercles on premolars of alaskan eskimos and Indians. Oral Surg Oral Med Oral Pathol. 1964;17:484–96.
4. Chen JW, Huang GT, Bakland LK. Dens evaginatus: Current treatment options, Journal of the American Dental Association (1939) 151(5) (2020) 358–367.
5. Stecker S, DiAngelis AJ. Dens evaginatus: a diagnostic and treatment challenge, Journal of the American Dental Association (1939) 133(2) (2002) 190-3.
6. Rao YG, Guo LY, Tao HT. Multiple dens evaginatus of premolars and molars in Chinese dentition: a case report and literature review. Int J Oral Sci. 2010;2(3):177–80.
7. Shah A, Peacock R, Eliyas S. Pulp therapy and root canal treatment techniques in immature permanent teeth: an update. Br Dent J. 2022;232(8):524–30.
8. Yordanova-Kostova GR, Grancharov MV, Gurgurova GD. Abnormality in the morphogenesis of Tooth Development and relationship with Orthodontic deformities and treatment approaches. Case Rep Dent. 2021;2021:1183504.
9. Mahtani A, Jain RK. Frequency of premolar teeth extractions for orthodontic treatment. Bioinformation. 2020;16(12):1080–7.
10. Liu JL, Li SH, Cai YM, Lan DP, Lu YF, Liao W, Ying SC, Zhao ZH. Automated Radiographic evaluation of adenoid hypertrophy based on VGG-Lite. J Dent Res. 2021;100(12):1337–43.
11. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: chances and challenges. J Dent Res. 2020;99(7):769–74.
12. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N. Deep Learn Med Imaging: Gen Overv Korean J Radiol. 2017;18(4):570–84.
13. Ren R, Luo H, Su C, Yao Y, Liao W. Machine learning in dental, oral and craniofacial imaging: a review of recent progress. PeerJ. 2021;9:e11451.
14. Nordblom NF, Büttner M, Schwendicke F. Artificial Intelligence in Orthodontics: critical review. J Dent Res. 2024;103(6):577–84.
15. Chaddad A, Peng J, Xu J, Bouridane A. Survey of explainable AI techniques in Healthcare. Sensors. 2023;23(2).
16. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17(1):195.
17. Li S, Liu J, Zhou Z, Zhou Z, Wu X, Li Y, Wang S, Liao W, Ying S, Zhao Z. Artificial intelligence for caries and periapical periodontitis detection. J Dent. 2022;122:104107.
18. Engels P, Meyer O, Schönewolf J, Schlickenrieder A, Hickel R, Hesenius M, Gruhn V, Kühnisch J. Automated detection of posterior restorations in permanent teeth using artificial intelligence on intraoral photographs. J Dent. 2022;121:104124.
19. Cejudo Grano de Oro JE, Koch PJ, Krois J, Garcia Cantu Ros A, Patel J, Meyer-Lueckel H, Schwendicke F. Hyperparameter tuning and Automatic Image Augmentation for Deep Learning-Based Angle classification on intraoral Photographs-A Retrospective Study, Diagnostics (Basel. Switzerland). 2022;12(7).
20. Ahmed HMA, Dummer PMH. A new system for classifying tooth, root and canal anomalies. Int Endod J. 2018;51(4):389–404.
21. American Association of Endodontists. Glossary of Endodontic Terms, 2020.
22. Zheng L, Zhang X, Hu J, Gao Y, Zhang X, Zhang M, Li S, Zhou X, Niu T, Lu Y, Wang D. Establishment and applicability of a Diagnostic System for

Advanced Gastric Cancer T staging based on a faster region-based convolutional neural network. Front Oncol. 2020;10:1238.
23. Kim J, Kwak CW, Uhmn S, Lee J, Yoo S, Cho MC, Son H, Jeong H, Choo MS. A Novel Deep Learning-based Artificial Intelligence System for Interpreting Urolithiasis in Computed Tomography, European urology focus (2024).
24. Quispe MD, Quispe CC, Serrano-Arriezu L, Trigo JD, Bengoechea JJ, Quispe EC. Development and validation of a smart system for medullation and diameter assessment of alpaca, llama and mohair fibres. Animal: Int J Anim Bioscience. 2023;17(5):100800.
25. Almutairi W, Yassen GH, Aminoshariae A, Williams KA, Mickel A. Regenerative endodontics: a systematic analysis of the failed cases. J Endod. 2019;45(5):567–77.
26. Turner JW, Kluemper GT, Chance K, Long LS. Dens evaginatus: the hornet's nest of adolescent orthodontics, American journal of orthodontics and dentofacial orthopedics: official publication of the American Association of Orthodontists, its constituent societies. Am Board Orthod. 2013;143(4):570–3.
27. Kılıc MC, Bayrakdar IS, Çelik Ö, Bilgir E, Orhan K, Aydın OB, Kaplan FA, Sağlam H, Odabaş A, Aslan AF, Yılmaz AB. Artificial intelligence system for automatic deciduous tooth detection and numbering in panoramic radiographs. Dento Maxillo Fac Radiol. 2021;50(6):20200172.
28. Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, Sveshnikov MM, Bednenko GB. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dento Maxillo Fac Radiol. 2019;48(4):20180051.
29. Ryu J, Kim YH, Kim TW, Jung SK. Evaluation of artificial intelligence model for crowding categorization and extraction diagnosis using intraoral photographs. Sci Rep. 2023;13(1):5177.
30. Kühnisch J, Meyer O, Hesenius M, Hickel R, Gruhn V. Caries Detection on Intraoral images using Artificial Intelligence. J Dent Res. 2022;101(2):158–65.
31. Schlickenrieder A, Meyer O, Schönewolf J, Engels P, Hickel R, Gruhn V, Hesenius M, Kühnisch J. Automatized Detection and Categorization of Fissure Sealants from Intraoral Digital Photographs Using Artificial Intelligence, Diagnostics (Basel, Switzerland). 2021;11(9).
32. Alalharith DM, Alharthi HM, Alghamdi WM, Alsenbel YM, Aslam N, Khan IU, Shahin SY, Dianišková S, Alhareky MS, Barouch KK. A deep learning-based Approach for the Detection of Early Signs of Gingivitis in Orthodontic patients using faster region-based convolutional neural networks. Int J Environ Res Public Health. 2020;17(22).
33. Liu Y, Cheng Y, Song Y, Cai D, Zhang N. Oral screening of dental calculus, gingivitis and dental caries through segmentation on intraoral photographic images using deep learning. BMC Oral Health. 2024;24(1):1287.
34. Kotaki S, Nishiguchi T, Araragi M, Akiyama H, Fukuda M, Ariji E, Ariji Y. Transfer learning in diagnosis of maxillary sinusitis using panoramic radiography and conventional radiography. Oral Radiol. 2023;39(3):467–74.
35. Kováč P, Jackuliak P, Bražinová A, Varga I, Aláč M, Smatana M, Lovich D, Thurzo A. Artificial Intelligence-Driven Facial Image Analysis for the Early Detection of Rare Diseases: Legal, Ethical, Forensic, and Cybersecurity Considerations. 2024;5(3):990–1010.

Publisher's note