

Development and Validation of a Cbct-Based Artificial Intelligence System for Accurate Diagnoses of Dental Diseases

Matvey Ezhov

Diagnocat Inc.

Maxim Gusarev

Diagnocat Inc.

Maria Golitsyna

Diagnocat Inc.

Julian Yates

University of Manchester

Evgeny Kushnerev

University of Manchester

Dania Tamimi

Private Practice

Secil Aksoy

Near East University

Eugene Shumilov

Diagnocat Inc.

Alex Sanders

Diagnocat Inc.

Kaan Orhan (✉ call53@yahoo.com.tr)

Ankara University

Research Article

Keywords: Artificial intelligence, deep learning, tooth detection, CBCT, dentomaxillofacial imaging

DOI: <https://doi.org/10.21203/rs.3.rs-303329/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Cone-beam computed tomography (CBCT) in dental practice is becoming increasingly popular. However, the correct teeth identification, positioning and diagnosis based on CBCT can be tedious and challenging for the untrained eye. This is due to additional training, specific knowledge and time required for analysis and diagnosis. When compared to conventional dental imaging methods. In this study, we introduce a novel artificial intelligence (AI) system that facilitates analysis and diagnosis. This system is based on deep learning approaches that can localize teeth and define pathologies within three-dimensional CBCT scans. The study showed that the diagnostic performance of AI system image interpretation reaches and sometimes exceeds in comparison to clinician's expertise. In this randomized cross-over trial we demonstrated a significant improvement of aided diagnostic accuracy for various dental diseases in comparison to a group of radiologists that made unaided decisions. AI can be used for both stand-alone CBCT interpretation and as a decision support system to improve quality of diagnostics and time efficiency.

Introduction

Radiological examination is an essential part of patient management in dentistry and is frequently used to supplement and aid clinical diagnosis of pathology related to teeth and adjacent structures¹⁻⁴. CBCT was proposed for maxillofacial imaging^{5,6} during the last decade and is now becoming increasingly popular for such diagnoses. It offers distinct advantages including lower radiation doses, compared to medical CT, and the potential of importing and exporting individualized, DICOM and overlap-free reconstructed data to and from other applications⁴⁻⁷. CBCT can supply high-resolution three-dimensional (3D) images without distortion and superimposition of bone and other dental structures that can be seen in conventional radiography⁸⁻⁹.

Several studies have compared diagnostic accuracy of CBCT with conventional and digital panoramic and periapical radiography. CBCT has been shown to significantly increase the detection rate of tooth root canal spaces and periapical areas for the evaluation of dental infection and pathology compared to conventional imaging. This suggests that CBCT enhances recognition of periapical bone lesions and offers improved diagnosis accuracy, treatment planning, and thus prognostic outcomes. These and other possibilities along with increasing access to CBCT imaging for dentists are allowing the transition from 2D to 3D imaging in everyday dental practice¹⁰⁻¹¹.

However, using CBCT for diagnostics requires the dentist to be a highly educated in radiographic diagnosis, which is not always possible. Because of the lack of time or experience of the average dentist, many pathologies can remain unidentified, thus computer-aided systems have been developed to assist in medical and dental imaging diagnosis¹²⁻¹⁵.

Artificial intelligence (AI) provides a promising solution for such medical image interpretation. For object detection and segmentation, Convolutional Neural Networks (CNN) are most commonly used. There are

several studies available with deep-learning methods, including CNNs, to assist clinicians in dentistry. CNNs are used clinically for, apical lesion detection¹⁶, detection of root fractures¹⁷, detection of periodontal disease¹⁸, cystic lesions¹⁹, caries detection²⁰, staging of lower third molar development²¹, tooth detection²²⁻²³, diagnosis of jaw lesion²⁴ and other pathologies detection²⁵.

To optimize their use, AI systems must be applicable to real-world situations and must be designed for clinical evaluation and deployment. Furthermore, an important part of the development and integration of these AI systems is that its functionality (ease of use, speed and accuracy) reaches or exceeds the clinicians' expertise in such situations.

In this study we evaluate a novel AI system called Diagnocat, which is based on deep learning methods, to determine its real-time performance of CBCT imaging diagnosis of anatomical landmarks and pathologies and its clinical effectiveness and safety when used by dentists in a clinical setting.

Results

The output of this process was a set of condition detections, where for every tooth and condition model outputs a probability distribution along with predicted tooth number. This evaluation was calculated only once during this study and was not used for the system training purposes. Moreover, during this process engineers did not have access to the examiners' votes data. To score the results, an independent expert was provided with both ground truth and Diagnocat inference data. The outside expert was responsible for running data analysis and producing performance measurements for primary and secondary endpoints.

The results of AI evaluation are shown in Tables 1 and 2. Table 1 shows the overall sensitivity and specificity for the system and dentomaxillofacial radiology examiners. Outcome counting for Table 3 was summarized over case, tooth, and condition, whilst grouped by the participants. Overall sensitivity values for human examiners ranged between 0.9318–0.9438 while the value for this AI system was 0.9239. Overall specificity values for ground truth examiners were between 0.9899–0.9946 while the value for this AI system was 0.9899. Both sensitivity and specificity were recorded as higher for human examiners. However, since examiners' evaluations were taken as the ground truth votes, this comparison is biased in favor of human examiners. Moreover, the difference in overall performance seems to be rather small, which we interpreted as a proof of Diagnocat standalone capability. Table 2 shows sensitivity and specificity values for the system given per condition. The results of specificity values were high, with the lowest being 0.94 when determining missing tooth. Sensitivity values were condition-dependent, with lowest values being around 0.7 for some difficult or subjective conditions such as evaluating the quality of endodontic treatment (missed canal, short filling, voids in root filling), and signs of dental caries (complex to diagnose using the CBCT). Notably, Diagnocat struggled to detect very rare anatomical configurations of the tooth e.g., 5 canals or 4 roots. Finally, a rare subtype of periapical lesion, periapical radiopacity, did not register in the dataset. Yet, this subtype currently is not claimed as a diagnostic capability of the Diagnocat system.

Table 1

Sensitivity and specificity values of Diagnocat and examiners. DMFR-1,2,3,4 corresponds to the first, second, etc. dental and maxillofacial radiologists that were annotating the dataset.

Participant	Sensitivity	Specificity
Diagnocat	0.9239	0.9899
DMFR-1	0.9411	0.9939
DMFR-2	0.9438	0.9931
DMFR-3	0.9318	0.9913
DMFR-4	0.9337	0.9946

Table 2
Sensitivity and specificity of each evaluated condition for the Diagnocat system predictions.

Condition	Sensitivity	Specificity
Artificial crown	0.9546	0.9963
Canals N = 1	0.9864	0.9661
Canals N = 2	0.7759	0.9927
Canals N = 3	0.9531	0.9647
Canals N = 4	0.654	0.9952
Canals N = 5	N/A	0.9998
Caries signs	0.7285	0.9953
Crown defect over 50%	0.8734	0.9975
Endodontically treated tooth	0.9676	0.9953
Filling	0.9721	0.9921
Impaction	0.9137	0.9995
Implant	0.9727	0.9997
Missed canal	0.6695	0.9974
Missing	0.9824	0.9405
Overfilling	0.7831	0.9973
PDL widening along root	0.8794	0.9863
Periapical lesion	0.8383	0.9953
Periapical lesion, PDL widening	0.7587	0.9807
Periapical lesion, poorly circumscribed radiolucency	0.6957	0.9942
Periapical lesion, well circumscribed radiolucency	0.729	0.9974
Periodontal bone loss	0.9489	0.9661
Periodontal bone loss, mild	0.9321	0.9742
Periodontal bone loss, moderate	0.9111	0.9866
Periodontal bone loss, severe	0.9286	0.996
Pontic	0.9101	0.9998
Post and core	0.75	0.9975

Condition	Sensitivity	Specificity
Roots N = 1	0.9593	0.9888
Roots N = 2	0.9181	0.9778
Roots N = 3	0.964	0.9786
Roots N = 4	0.0	1.0
Short filling	0.6981	0.9937
Voids present in the root filling	0.7329	0.9957

Table 3
Description of each annotated Descriptor conditions.

Condition	Description
Artificial crown	A tooth was restored with an artificial crown
Number of canals	Number of root canals in a tooth (1–5)
Caries signs	Signs of dental caries (cases where caries is certain, and there is a low chance of confusion with a metallic artifact or non-contrast filling)
Crown has defect over 50%	A crown is largely destroyed: at least 50% of the crown is missing
Endodontically treated tooth	A tooth displays signs of previous endodontic treatment
Filling	A crown was restored with a filling
Impaction	A tooth is impacted (unerupted)
Implant	There is an implant in place of a tooth
Missed canal	A root canal was missed (not filled) during endodontic treatment of. Should be specified only if a tooth was endodontically treated.
Missing	Absence of tooth, implant, and pontic under specified number including both extracted teeth and teeth that are not visible in the FoV of an image.
Overfilling	Filling material is visualized beyond a radiographic apex. Should be specified only if a tooth was endodontically treated.
Periapical lesion	Presence of inflammatory periapical lesion adjacent to one or more roots of a tooth
Pontic	There is a pontic restoration in place of a tooth (either base or middle part)
Post and core	A tooth was restored with a post and core restoration
Number of roots	Number of roots in the tooth (1–4)
Short filling	Root canal filling is short (ends in 2 mm or more from a radiographic apex). Should be specified only if a tooth was endodontically treated.
Voids present in the root filling	A root canal contains voids (spaces that were not filled during previous endodontic treatment). Should be specified only if a tooth was endodontically treated.

In Table 4 results of the study are presented. Sensitivity and specificity values are shown for aided versus unaided reads for each condition. The lowest sensitivity values for the aided group were 0.1818 and 0.3535, detecting the roots ($n = 4$) and periodontal bone loss. The lowest specificity value was 0.8111 for periodontal bone loss. For the unaided group the lowest sensitivity value was 0.2045 for periapical lesion

and poorly defined radiolucency, the smallest specificity value was 0.7973 for caries. The highest sensitivity and specificity for both groups was 1 for implant detection. The results for the overall sensitivity and specificity for aided and unaided groups, calculated as an aggregate of all conditions. The sensitivity values for aided and unaided groups were 0.8537 and 0.7672 while specificity were 0.9672 and 0.9616 respectively. There was a statistically significant difference between the groups ($p < 0.05$). Statistical tests revealed the group with aided Diagnocat had a superior performance in comparison to ground truth.

Table 4
Sensitivity and specificity of each evaluated condition for aided and unaided annotations.

Condition	Sensitivity		Specificity	
	aided	unaided	aided	unaided
Artificial crown	0.8337	0.6774	0.9804	0.971
Canals N = 1	0.9544	0.9417	0.9655	0.9515
Canals N = 2	0.793	0.7345	0.9713	0.9632
Canals N = 3	0.828	0.7347	0.9719	0.9666
Canals N = 4	0.7345	0.751	0.9807	0.9719
Canals N = 5	N/A	N/A	0.9998	0.9992
Caries signs	0.6693	0.6634	0.8593	0.7973
Crown defect over 50%	0.9088	0.7887	0.9901	0.9812
Endodontically treated tooth	0.9761	0.9476	0.9912	0.9841
Filling	0.9308	0.7771	0.9515	0.9151
Impaction	0.8523	0.5	0.9993	0.9993
Implant	1.0	0.969	1.0	0.9996
Missed canal	0.8233	0.7701	0.9867	0.946
Missing	0.8964	0.8973	0.9403	0.9364
Overfilling	0.7909	0.6199	0.9926	0.9922
PDL widening along root	0.746	0.2779	0.9803	0.9731
Periapical lesion	0.8304	0.683	0.9388	0.9465
Periapical lesion, PDL widening	0.5161	0.3202	0.9332	0.9542
Periapical lesion, poorly circumscribed radiolucency	0.4318	0.2045	0.9897	0.981
Periapical lesion, well circumscribed radiolucency	0.7135	0.6048	0.9912	0.9817
Periodontal bone loss	0.7239	0.4501	0.8111	0.8358
Periodontal bone loss, mild	0.4783	0.2329	0.8739	0.8782
Periodontal bone loss, moderate	0.4357	0.2148	0.9248	0.943

Condition	Sensitivity		Specificity	
	aided	unaided	aided	unaided
Periodontal bone loss, severe	0.3535	0.2727	0.9929	0.9878
Pontic	0.9696	0.9087	0.9988	0.9973
Post and core	0.7877	0.7114	0.9846	0.9715
Roots N = 1	0.9559	0.9635	0.9731	0.9558
Roots N = 2	0.8609	0.8304	0.9728	0.9726
Roots N = 3	0.9376	0.9028	0.9712	0.9785
Roots N = 4	0.1818	0.7045	0.9996	0.9986
Short filling	0.8139	0.6285	0.9846	0.9817
Void present in the root filling	0.7592	0.4938	0.9826	0.9757

Discussion

The integration of AI into the medical field has dramatically accelerated in the past decade. The use of deep learning advanced almost synchronously in both medical and dental fields⁹. Previous studies in dentistry focused on image-processing algorithms to achieve high-accuracy classification and segmentation in dental radiographs. They used mathematical morphology, active contour models, level-set methods, Fourier descriptors, textures, Bayesian techniques, linear models, or binary support vector machines^{15,26}. However, image components are usually obtained manually using these image-enhancement algorithms. The deep learning method used in this AI system (Diagnocat) yielded fairer outcomes by automatically obtaining image features. The objects detected in an image are classified into a pretrained network without preliminary diagnostics, as a result of processes such as various filtering and subdivision. With its direct problem-solving ability, deep learning is used extensively in the medical field. Deep learning methods using CNNs are a cornerstone of medical image analysis²⁷. Such methods have also been preferred in AI studies in dental radiology as well. Tooth detection, identification and numeration are the first diagnostic steps in dental radiography. Image-processing algorithms have been developed with classification and segmentation in dental radiographs using mathematical morphology, active contour, or level-set method. Mahoor et al.¹³ presented an automated dental identification system to classify and identify teeth in bitewing radiographs using Bayesian classification. Lin et al.²⁸ recommended a tooth classification and numbering system to efficiently segment, classify, and number teeth using an image enhancement technique in bitewing radiographs. Tooth detection and numbering have been researched intensively during the last few decades mainly using threshold and region-based techniques. CNN as a popular deep learning method has been used to detect and number teeth as well. Eun et al.²⁹ emphasized that localization of teeth is important for dental image applications in their study. They suggested an original teeth localization technique for periapical radiographs by means of

oriented tooth detection using a CNN. The results of their study showed that the proposed method is effective to localize teeth successfully.

Similar studies were done and reported in the literature for CBCT as well. Miki et al.²² considered automatic teeth classifying system into 7 types from an axial slice on CBCT images using a CNN. They concluded that a 7-tooth type classification system can be used efficiently for the automatic charting of tooth lists. Another study performed by Oktay²³ described a CNN model modified with AlexNet architecture for tooth detection in panoramic radiographs. This study defined mouth gap detection that showed possible placement of teeth for pre-processing step. It was concluded that this model could be efficiently used for the detection of teeth. Jader et al.³⁰ presented a study that used segmentation of teeth using mask region-based CNN method with transfer learning strategies. In a similar study Lee et al.³¹ used a fully deep learning mask region-based convolutional neural network (R-CNN) method implemented through a fine-tuning process for automated tooth segmentation. This technique showed high performance for automatic teeth segmentation on panoramic radiographs. Recently published paper by Chen et al.¹⁵ proposed a deep learning CNN model with a VGG16 network structure for the teeth detection and classification of periapical radiographs. They offered to distinguish teeth from a different position and experimental findings. Another study conducted by Tuzoff et al.³² also used state-of-the-art Faster R-CNN model of tooth detection and numbering. VGG-16 Net that is a 16-layer CNN architecture was used as a base CNN. They concluded that the AI model had a close performance to the experts' level. In 2018 we published our AI algorithm (later called Diagnocat)³³ that presented coarse-to-fine volumetric segmentation of teeth in CBCT images which was efficient for handling large volumetric images for tooth segmentation. Diagnocat's approach in diagnosing lesions is based on a deep convolutional neural network using a U-net-like architecture³⁴. The problem formulation in the study in terms of machine learning tasks was semantic segmentation, including segmenting background and periapical pathology. Specificity and sensitivity metrics were used to evaluate diagnostic performance and to measure the localization capabilities of our model, binary voxel-wise intersection over union (IoU) of the ground truth mask and prediction were used. In the current study, we tested the diagnostic performance of this AI system compared to the experienced dentomaxillofacial examiners. Secondly, we compared the diagnostic performance and required diagnostic time within aided and unaided groups in real time clinical environment. To the best of our knowledge, no study was performed for a real time clinical performance trial of CBCT imaging diagnosis to demonstrate clinical safety and effectiveness of its use by dentists in clinical settings. In the first study, ground truth set-up results showed the sensitivity and specificity values for human examiners in between 0.9318–0.9438 while the value for this AI system was 0.9239. There was no statistically significant difference between Diagnocat and the experienced dentomaxillofacial radiologists ($P > 0.05$). The results of sensitivity and specificity values were also similar between this AI system and the examiners. For the second study it was shown that the AI-aided group had a superior performance compared to the unaided group. The group with AI support had 0.85 averaged by condition sensitivity value and 0.97 specificity value opposed to 0.77 and 0.96 for the unaided group, correspondingly. The average time for the aided group was 17.55 minutes while it took 18.74 minutes for the unaided respectively. There was a significant difference between two groups ($p <$

0.05). Statistical tests revealed that the AI-aided group had lower evaluation period in comparison with unaided group ($p < 0.05$).

The results showed that the evaluation process was improved with AI-aid and required diagnosis time was reduced with a better specificity and sensitivity. In conclusion, we presented a novel framework that analyses clinical CBCT scans and makes referral suggestions to a standard that is comparable to dental professional expertise. This study showed that the proposed AI system (Diagnocat) significantly improved diagnostic capabilities of radiologist. Moreover, AI can save time by enabling automatic preparation of dental charts and electronic dental records together with automated pathology detection.

Methods

Ethics and information governance. A written informed consent was obtained from all patients before CBCT examinations. The research protocol was performed in accordance with the principles of the Declaration of Helsinki and was approved by the non-interventional Institutional Review Board (IRB) within the research project of the YDU 82-1147 of Near East University, Faculty of Medicine, Health Sciences Ethics Committee (Nicosia, Cyprus). Deidentification was performed in line with the Information Commissioner's Anonymization: managing data protection risk code of practice (https://ico.org.uk/media/1061/anonymisation_code.pdf), and validated by the aforementioned institution. Only de-identified anonymized retrospective data were used for research, without the active involvement of patients.

Testing the system. The primary goal of this study is to evaluate the ability of this AI system (Diagnocat) to enhance the diagnostic capabilities of the dentist and radiologist. In order to test this, a few steps had to be taken to prepare the dataset for viewing and analysis. These steps are necessary due to the inherent variability of CBCT datasets coming from CBCT machines as well as the variability in clinical experience on the part of the examiners. Thus, this study has two distinct parts. The first was preparing the dataset for evaluation and the second was evaluating the usefulness of the system for enhancing diagnostic capabilities.

Part (A): Preparing the dataset for evaluation:

1. **Image processing.** Due to high variety of CBCT scanning devices and different calibration settings, CBCT images need to be normalized for both manual and automatic diagnostics. This is usually done with the help of window level and width DICOM properties extracted from scan metadata. Unfortunately, the radiodensity of bone and tissue of scans from the same scanning device manufacturer differ when the extracted window is applied. The difference is significantly higher when corresponding windows are applied to images from different devices. We apply the normalization process based on voxel radiodensity measured in Hounsfield Units (HU):
2. HU values below -1000 (air radiodensity) are clipped.
3. HU values below 5th and above 95th percentiles of an image are clipped.

4. HU values are standardized by subtracting the mean value and dividing the difference by the standard deviation.

The last step may vary depending on the task. When there is no need to preserve difference between a dense bone (2000 HU) and metal (3000 HU) radiodensities, HU values above 2000 can be clipped, and resulting values can be rescaled to [0, 1] range.

1. **Localization datasets.** To obtain precise segmentation results for training purposes, dental and radiology specialists used ITK-SNAP software³⁵ that allows users to navigate 3D images in three planes. Once being annotated, each segmentation mask was automatically examined to eliminate human factor, e.g. misalignment of a tooth volume and a resulting mask. Each dataset split mimics the distribution of the original one. These modules are:

- *ROI localization module.* The dataset consists of 562 CBCT scans with segmented teeth and jaws. The scans are equally distributed among 19 scanner models of 12 scanner manufacturers.

- *Tooth localization and numeration module.* The dataset consists of 684 CBCT scans with segmented and numerated teeth. The scans are equally distributed among 24 scanner models of 15 scanner manufacturers.

- *Periodontitis module.* The dataset consists of 99 CBCT scans with precisely segmented alveolar bone area and 120 CBCT scans with precisely segmented enamel area of teeth. The scans are equally distributed among 11 scanner models of 8 scanner manufacturers. Each side of a tooth (mesial, distal, oral, and vestibular) has a group of three periodontium landmark points: a point of cemento-enamel junction, a point of bone attachment, and a point of bone peak within 2 mm tooth vicinity. The dataset with segmented enamel area is used to obtain the first point, while the dataset with segmented alveolar bone is used to obtain two latter points.

- *Caries localization module.* The dataset consists of 4398 tooth volumes with a context area. The class labels are: background (no pathology), caries sign, metallic artefact, and non-contrast filling. One instance can have multiple conditions. The dataset was additionally validated by a lead radiologist.

- *Periapical lesion localization module.* The dataset consists of 2800 tooth volumes with a context area. The class labels are: background (no pathology), periodontal ligament (PDL) widening, poorly circumscribed radiolucency, well circumscribed radiolucency, and radiopacity. One instance can have multiple conditions. The dataset was additionally validated by a lead radiologist.

1. **Classification (Descriptor) datasets.** Descriptor, the main diagnostic module, is a complex model that, besides accurate data collection, requires several iterations of dataset formation and annotation regulations. We provided detailed description of the annotation process and insights on managing class imbalance and high model uncertainty.

- *Annotation protocol.* Every radiologist was provided with an instruction describing annotation including a list of required pathologies, access to the internal web-based application that provided a data collection form, and an option to download study DICOM for a standalone viewing. Additionally, every radiologist reviewed and described 3 sample CBCTs containing all target pathologies, which were then reviewed by the study supervisor, highly experienced oral and maxillofacial radiologist. Then, the study supervisor provided feedback to the radiologist. Each radiologist independently studied a CBCT image in a clinical viewer software and noted the presence or absence of each condition for each tooth in the target list. Radiologists were required to answer either “applicable”, or “not applicable” for every condition in table 4.

- *Initial annotation.* During the first stage of the annotation process, a group of experienced radiologists annotated a large set of images following the annotation protocol. Images were randomly sampled, filtered by the study coordinator according to the inclusion and exclusion criteria, and then passed to radiologists. Before the main annotation process, annotators were trained and evaluated by the study coordinator:

- Participant studied annotation instruction and protocol
- Participant annotated a small set of exemplary images, the study coordinator evaluated the results and provided feedback to the participant

During this stage each sample (distinct patient-tooth) received 1 diagnostic vote for every condition in consideration.

- *Test set separation.* Following the completion of the first stage of annotation, a test was separated from the annotated data pool and excluded from all following development activities. Test images were sampled in a way to have at least N positives and N negatives for every condition. The choice of N = 300 was motivated by the available number of positive samples for rare conditions. The sampling procedure was as follows.

1. Randomly sample a condition.
2. If the test set contained less than N positives of the condition, sample a random positive example from the data pool and allocate it to the test set. Additionally, allocate all other samples from the same image.
3. Repeat until the test set contains at least N positives and N negatives for each condition.

Each sampled example contains annotation for all target conditions, so the resulting test set contains more than N positives/negatives for the majority of conditions. Additionally, the test set contains a different number of positives and negatives for each condition, typically, negatives outnumbering the positives (class imbalance). This influenced our decision to choose the AUPRC metric for evaluation as it is robust to significant class imbalance.

- *Test set additional annotation.* An additional vote from a second radiologist was obtained for each tooth-condition (sample). Then, for the sample where the first two radiologists disagreed, another vote

from the third radiologist was obtained. Ground truth was decided by the majority vote (2-vote agreement).

- *Model development dataset.* A set for model development purposes formed from remaining annotated data pool (i.e. not included in the test set) was split into training and validation subsets as it was fit for the task. As the majority of examples in the train set had only 1 vote, it was expected that some labels would be incorrect. However, deep learning is known to have some level of robustness against noisy training labels, and we hypothesized that the models would be able to learn the correct labels and achieve satisfactory scores. Additionally, the partially trained model could be used to find and correct the erroneous votes by measuring disagreement between votes and model predictions. In the course of this project, this hypothesis was confirmed. While samples with 1 vote were widely used in the train set, model validation was performed using standard 2-vote agreement protocol.

- *Rare case mining.* Following the separation of the train set, a series of models was trained. Then, the best model was used to enrich the train set by mining rare cases and finding potentially erroneous votes in the train set. Initially rare conditions did not contain enough positive examples in the train set. To rectify this, following mining procedure was implemented:

1. Define a set of rare condition list where additional data is required.
2. Perform inference of the best model available at a time on studies from the non-annotated data pool.
3. Calculate information entropy for every condition in the rare condition list.
4. Sample teeth with high information entropy.
5. Run images containing sampled teeth through the annotation process.

Information entropy is defined as

$$S = -P_i \log P_i$$

where P_i is the probability of i th outcome of a set of all possible outcomes. For binary task, such as our formulation, i iterates over “present” and “not present”. Information entropy is highest when probabilities of “present” and “not present” are equal to 0.5. Intuitively, information entropy is a measure of uncertainty in the probability distribution. High uncertainty on an example excluded from the training and validation set means that the training process is likely to improve if the example is annotated and added to the training set.

- *Incorrect vote mining and rectification.* Since we collected only 1 vote for a large number of images allocated to the train set, some of these votes were submitted incorrectly. To rectify this, we implemented the following procedure:

1. Perform K -fold inference on all images in the train set using the best model available at a time. K -fold inference procedure:

- (a) Split train images into K disjoint sets
- (b) Pick a subset i
- (c) Train model on all subsets except i
- (d) Perform inference on set i and record resulting scores
- (e) Repeat for every i in K
 1. Calculate radiologist-model disagreement.
 2. Sample from cases where radiologist-model disagreement was high.
 3. Collect additional votes for sampled cases using the annotation process.

- *Dataset statistics.* We evaluated the performance of our models on a dataset with 705017 samples consisting of 28745 teeth and 25 conditions across 1135 CBCT scans. The scans are distributed among 31 scanner models of 17 scanner manufacturers.

Each step analyzes data in progressively higher spatial resolution, from a coarse voxel size of 3 mm^3 at the initial stage of region of interest (ROI) localization to a fine voxel size up to 0.15 mm^3 for final per-tooth diagnosis. This multi-step pathway was required due to the large memory size of CBCT images at the original resolution. The flow of the system pipeline is shown in figure 1.

The first step is ROI localization module (fig. 1). Reduction of field of view (FoV) to ROI sufficient for analysis of dental diseases allows completion of the diagnostics without any information loss. The ROI localizer identifies specific regions of jaws and teeth with some extended context and excludes other anatomical regions. The localization module is based on the volumetric modification of U-Net architecture³⁶ performing 3-class semantic segmentation: background, teeth, and jaw bone. To fit large FoV volumes, the module operates at 3 mm^3 per voxel resolution.

The cropped image is further passed to the tooth localization and numeration module (fig. 1) that plays a crucial role in diagnostic pipeline. Tooth localization allows further analysis of different conditions inside and around a tooth, while tooth numeration helps with determining number-specific attributes and inter-tooth relations. The localization and numeration module is implemented as a volumetric U-Net network performing semantic segmentation on 54 classes (the background, 52 possible teeth, and an additional class for supernumerary teeth). It operates at 1 mm^3 per voxel resolution.

At the next step each localized tooth area is extended with some context and passed to Descriptor (fig. 1) that defines the probabilities of a tooth being affected by a set of conditions (table 1). Descriptor is a principal classification module and is implemented as an ensemble of a ResNeXt³⁷ with integrated squeeze-and-excitation blocks³⁸ and a DenseNet³⁹ architectures performing multiple binary classification on 25 classes.

Each tooth volume is further examined by three modules for auxiliary classification purpose. (1) The periodontitis module detects and evaluates alveolar bone loss in close vicinity to a tooth. It allows classification of 3 bone loss types of different severity by calculating distances between pairs of periodontium landmarks segmented by a separate landmark localizer. (2) The caries localization module defines signs of caries probability using segmentation of carious lesions found inside a tooth area. (3) The periapical lesion localization module detects periapical lesion presence and allows classification of 4 lesion types found around a tooth. The embedded localizers of three classification modules are implemented as volumetric U-Nets performing semantic segmentation.

Part (B) Evaluating the ability of the AI system (Diagnocat) to enhance the diagnostic capabilities of the dentist and radiologist:

1. **Evaluating diagnostic capabilities of the Diagnocat AI system.** The primary end-point was to test end-to-end performance of this AI system, measuring tooth localization, numeration, and diagnostic sub-modules as a single system. It allowed to estimate overall safety and performance of the proposed system.

The Diagnocat AI software was used to obtain a binary condition prediction made on 3D CBCT scans using its predefined operating point (checkpoints of the trained models), which was then compared to ground truth to calculate sensitivity (proportion of correctly defined conditions) and specificity (proportion of correctly defined teeth not having conditions).

Secondary endpoint was to evaluate examiners performance and compare it to the AI results. Although examiners were tested on the data that was beforehand annotated by each of them the results showed comparable diagnostic quality of Diagnocat and the examiners. For the performance evaluation a set of 300 CBCT maxillofacial images in DICOM format was sourced consecutively from three clinics (100 images from each site) and anonymized by replacing "PatientName" with an empty string and truncated "PatientBirthDate" to the first day of the nearest year. Subsequently, images were screened against the inclusion and exclusion criteria.

The inclusion criteria were:

- a patient with the ability to consent to participate in the project
- a patient of 21 years or older
- anonymized CBCT image of maxillofacial region, and
- both model and manufacturer of imaging device are not present in the training dataset of the system (allows testing generalizability to new imaging devices).

The exclusion criteria were:

- images containing significant motion artifacts (as judged by radiologist coordinating the study)

- images containing severe artifacts such as streak artifacts, beam hardening (low and medium artifact remover was applied using device-specific software when available for standardization of the images) and;
- images of patients with cleft lip and palate, trauma, bone lesions, and severe bone erosions.

Final set of images was then reviewed by a scientific coordinator (an internationally recognized dentomaxillofacial radiologist with at least 18 years of experience) and 10 images were rejected due to significant motion artifacts. To establish the ground truth, examiners were recruited from experienced dentomaxillofacial radiologists. In total, the data was evaluated by four of them with a mean of 10 years professional experience.

Each examiner was responsible for annotation of CBCT anatomy on their own. Moreover, the examiners were unaware of conditions of the patients. Each examiner was then trained by the study coordinator to annotate 3D CBCT scans and fill the provided form correctly. After the study coordinator evaluated the examiners and approved them as sufficiently trained, the study proceeded to actual data collection. Each radiologist received a random, non-overlapping portion of the dataset via electronic means (shared folder). They evaluated the cases in their clinical environment and filled the spreadsheet, then saved them to the separate shared folders. The examiners could not access each other's forms. After they evaluated cumulatively and annotated the full CBCT dataset, a second round of annotation started, where the examiners were assigned a different random subset of the dataset. After the second round was finished, the third commenced. At the end of the third round, the scientific coordinator collected the examination from 3 radiologists for every sample. Evaluations took place between December 2019 and April 2020. Data was extracted on an individual and group comparison level. To establish the true values of conditions, a consensus process was performed, where the ground truth was taken as a majority (at least 2 of 3 votes) per each case, tooth, and condition. The whole process was then reviewed again by the study coordinator for last adjustments and establish final ground truth evaluations of each patient and teeth as well as for each condition. Inference of Diagnocat system was performed once for the full dataset: an engineer performed inference using the production version of the system.

1. **Evaluating the clinical performance.** After evaluation of diagnostic capability of the Diagnocat AI system, the next step is to evaluate the clinical performance of the system which can be achieved by comparing the accuracy of the diagnosis and time required for the reading for aided and unaided cases. Evaluation duration was compared between aided and unaided to determine if the addition of Diagnocat suggestions changes the time required to review the case. It was estimated that approximately eight weeks were required to conduct this study. This was addressed at each stage, from recruitment to analysis: recruitment and consent - 1 week, training and randomization - 1 week, investigation - 1 week, washout - at least 1 month, investigation - 1 week and analysis - 2 weeks. The washout period was at least 1 month. This was to minimize memory bias and confounding factors. Crossover design reduced confounding factors as well. To identify the number of required examiners a power analysis was performed⁴⁰. To hold the study at least 20 examiners were required in total.

Thus, 24 dentists were enrolled in the study as the examiners and divided into two groups at a 1:1 ratio: (1) Group 1 examined the CBCTs with AI system-aided; (2) Group 2 examined the CBCTs unaided. The confidence interval is 0.80, and 5% was used for error. Enrolled examiners were qualified general dental practitioners of various experience with no defined specialty interest. Following inclusion and exclusion criteria were applied:

- Inclusion criteria

1. Qualified dentist - General Dental Practitioner
2. At least five years of experience in dentistry
3. Ability to interpret dental CBCT
4. Access to CBCT software at the workplace

- Exclusion criteria

1. Unable to commit to the study
2. Employees of Diagnocat - and their relatives were excluded from participation

The scientific advisor for CBCT scan reading conducted training sessions for examiners for one hour including the use of the Diagnocat AI system. 10 training CBCT scans were used for training and practice purposes; those encompassed the full spectrum of required diagnostics. A list of all possible diagnoses was given as well to ensure that the scope of diagnosis was calibrated and participants were aware of that. There was also remote support available to guide through the training process. For this study overall dataset contained 40 CBCT images, including 30 study images and 10 images for examiners training. These scans were sampled randomly from the dataset of the standalone performance test. 30 study images were sequentially numbered after randomization was performed. Thus each participant had a different sequence of clinical cases. Each CBCT scan required all 32 teeth to be diagnosed with none, one or more pathologies. Thus, 32 units in each CBCT, multiplied by the number of pathologies identified in each unit, with a total number of 30 CBCT scans per group. In this way, 960 (30x32=960) diagnostic activities were carried out in each investigation by each participant. The crossover nature of this study ensured that this was performed twice by each participant. Table 3 shows the conditions that were asked to diagnose by the examiners. Once investigations were completed, the raw data from forms filled by unaided group and Diagnocat was transformed into the same format using automated scripts written before the study and then sent to an independent blinded assessor. This assessor analyzed the data and compared it with ground truth (same as in the standalone Diagnocat performance test). Raw data was compared to ground truth electronically. Scoring was performed via electronic means and data stored securely. Once this was completed, groups were decoded and results compared.

Declarations

Acknowledgement

Since the present study have been conducted by the retrospective radiologic images, it is not subject to the "registration" and "clinical trial number" procedures required for clinical trials (Clinical Trials or Observational Studies)

Author Contributions

M.E., M. Gusarev, M. Golitsyna and E.S. collected and analyzed the data. M.E., A.S., E.S. and K.O. designed the study. J.M.Y., E.K., D.T. and S.A. participated writing the manuscript. All authors discussed the results and reviewed the manuscript

Competing Interests

Financial support was received by Diagnocat Co. Ltd., San Francisco CA. Matvey Ezhov, Maxim Gusarev, Maria Golitsyna, Eugene Shumilov, and Alex Sanders are employees of Diagnocat Co. Ltd. Julian M Yates, Evgeny Kushnerev, Dania Tamimi, Secil Aksoy, and Kaan Orhan declare no potential conflict of interest.

References

1. Chan, M., Dadul, T., Langlais, R., Russell, D. & Ahmad, M. Accuracy of extraoral bite-wing radiography in detecting proximal caries and crestal bone loss. *J. Am. Dent. Assoc.***149(1)**, 51-58 (2018).
2. Vandenberghe, B., Jacobs, R. & Bosmans, H. Modern dental imaging: a review of the current technology and clinical applications in dental practice. *Eur. Radiol.***20(11)**, 2637-55 (2010).
3. White, S.C. & Pharoah, M.J. Oral radiology-E-Book: Principles and interpretation. Elsevier Health Sciences (2014).
4. Oz, U., Orhan, K. & Abe, N. Comparison of linear and angular measurements using two-dimensional conventional methods and three-dimensional cone beam CT images reconstructed from a volumetric rendering program in vivo. *Dentomaxillofac. Radiol.***40(8)**, 492-500 (2011).
5. Farman, A.G. & Scarfe, W.C. Development of imaging selection criteria and procedures should precede cephalometric assessment with cone-beam computed tomography. *Am. J. Orthod. Dentofacial Orthop.***130 (2)**, 257-265 (2006).
6. Scarfe, W.C., Farman, A.G. & Sukovic, P. Clinical applications of cone-beam computed tomography in dental practice. *J. Can. Dent. Assoc.***72(1)**, 75-80 (2006).
7. Korbmacher, H., Kahl-Nieke, B., Schöllchen, M. & Heiland, M. Value of two cone-beam computed tomography systems from an orthodontic point of view. *J. Orofac. Orthop.***68 (4)**, 278-289 (2007).
8. Bayrakdar, I.S. et al. Cone beam computed tomography and ultrasonography imaging of benign intraosseous jaw lesion: a prospective radiopathological study. *Clin. Oral Investig.***22(3)**, 1531-1539 (2018).
9. Orhan, K., Bayrakdar, I.S., Ezhov, M., Kravtsov, A. & Ozyurek, T. Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. *Int. Endod. J.***53(5)**, 680-

- 689 (2020).
10. Estrela, C., Bueno, M.R., Leles, C.R., Azevedo, B. & Azevedo, J.R. Accuracy of cone beam computed tomography and panoramic and periapical radiography for detection of apical periodontitis. *J. Endod.***34(3)**, 273-279 (2018).
 11. Davies, A., Mannocci, F., Mitchell, P., Andiappan, M. & Patel, S. The detection of periapical pathoses in root filled teeth using single and parallax periapical radiographs versus cone beam computed tomography-a clinical study. *Int. Endod. J.***48(6)**, 582-592 (2015).
 12. Hung, K., Montalvao, C., Tanaka, R., Kawai, T. & Bornstein, M.M. The use and performance of artificial intelligence applications in dental and maxillofacial radiology: A systematic review. *Dentomaxillofac. Radiol.***49(1)**, 20190107 (2020). DOI: 10.1259/dmfr.20190107.
 13. Mahoor, M.H. & Abdel-Mottaleb, M. Classification and numbering of teeth in dental bitewing images. *Pattern Recognition***38 (4)**, 577-586 (2005).
 14. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H. & Aerts, H.J. Artificial intelligence in radiology. *Nat. Rev. Cancer***18(8)**, 500-510 (2018).
 15. Chen, H. et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci. Rep.***9(1)**, 1-11 (2019).
 16. Ekert, T. et al. Deep learning for the radiographic detection of apical lesions. *J. Endod.***45(7)**, 917-922 (2019).
 17. Fukuda, M. et al. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. *Oral Radiol.***36(4)**, 337-343 (2019).
 18. Krois, J. et al. Deep learning for the radiographic detection of periodontal bone loss. *Sci. Rep.***9(1)**, 8495 (2019).
 19. Lee, J.H., Kim, D.H. & Jeong, S.N. Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. *Oral Dis.***26(1)**, 152-158 (2020).
 20. Lee, J.H., Kim, D.H., Jeong, S.N. & Choi, S.H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J. Dent.***77**, 106-111 (2018).
 21. Merdietio Boedi, R. et al. Effect of lower third molar segmentations on automated tooth development staging using a convolutional neural network. *J. Forensic Sci.***65(2)**, 481-486 (2020).
 22. Miki, Y. et al. Classification of teeth in cone- beam CT using deep convolutional neural network. *Comput. Biol. Med.***80**, 24-29 (2017).
 23. Oktay, A.B. Tooth detection with convolutional neural networks. 2017 Medical Technologies National Congress (TIPTEKNO): IEEE, (2017) DOI:10.1109/TIPTEKNO.2017.8238075.
 24. Poedjiastoeti, W. & Suebnukarn, S. Application of convolutional neural network in the diagnosis of jaw tumors. *Healthc. Inform. Res.***24(3)**, 236-241 (2018).
 25. Schwendicke, F., Golla, T., Dreher, M. & Krois, J. Convolutional neural networks for dental image diagnostics: A scoping review. *J. Dent.***91**, 103226 (2019).

26. Rad, A.E., Rahim, M.S.M. & Norouzi, A. Digital dental X-ray image segmentation and feature extraction. *TELKOMNIKA Indonesian Journal of Electrical Engineering***11(6)**, 3109-3114 (2013).
27. Yasaka, K., Akai, H., Kunimatsu, A., Kiryu, S. & Abe, O. Deep learning with convolutional neural network in radiology. *Jpn. J. Radiol.***36(4)**, 257-272 (2018).
28. Lin, P., Lai, Y. & Huang, P. An effective classification and numbering system for dental bitewing radiographs using teeth region and contour information. *Pattern Recognition***43(4)**, 1380-1392 (2010).
29. Eun, H. & Kim, C. Oriented tooth localization for periapical dental X-ray images via convolutional neural network. 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), IEEE, 2016 DOI:10.1109/APSIPA.2016.7820720
30. Jader, G. et al. Deep instance segmentation of teeth in panoramic X-ray images. 2018 31st SIBGRAP Conference on Graphics, Patterns and Images (SIBGRAP): IEEE, (2018) DOI: 10.1109/SIBGRAP.2018.00058.
31. Lee, J.H., Han, S.S., Kim, Y.H., Lee, C. & Kim, I. Application of a fully deep convolutional neural network to the automation of tooth segmentation on panoramic radiographs. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.***129(6)**, 635-642 (2019).
32. Tuzoff, D.V. et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac. Radiol.***48(4)**, 20180051 (2019). DOI: 10.1259/dmfr.20180051.
33. Zakirov, A., Ezhov, M., Gusarev, M., Alexandrovsky, V. & Shumilov, E. End- to-end dental pathology detection in 3D cone-beam computed tomography images. Proc. 1st Conference on Medical Imaging with Deep Learning, 1–9 (2018).
34. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. ArXiv e-prints (2015). arXiv:1411.4038, DOI:10.1109/CVPR.2015.729896
35. Yushkevich, P.A. et al. User-guided 3-D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage***31(3)**, 1116-1128, (2006).
36. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. ArXiv e-prints (2015). arXiv:1505.04597, DOI: 10.1007/978-3-319-24574-4 28.
37. Xie, S., Girshick, R., Dollar, P., Tu, Z. & He, K. Aggregated residual transformations for deep neural networks. ArXiv e-prints (2017). arXiv:1611.05431, DOI: 10.1109/CVPR.2017.634.
38. Hu, J., Shen, L. & Sun, G. Squeeze-and-excitation networks. ArXiv e-prints (2018). arXiv:1709.01507, DOI:10.1109/CVPR.2018.00745.
39. Huang, G., Liu, Z., Weinberger, K.Q. & Maaten, L. Densely connected convolutional networks. ArXiv e-prints (2017). arXiv:1608.06993, DOI: 10.1109/CVPR.2017.243.
40. Lin, T.Y., Dollar P., Girshic, He, R.K., Hariharan, B. & Belongie, S. Feature pyramid networks for object detection. ArXiv e-prints (2017). arXiv:1612.03144, DOI:10.1109/CVPR.2017.106.

Tables

Table 1. Sensitivity and specificity values of Diagnocat and examiners. DMFR-1,2,3,4 corresponds to the first, second, etc. dental and maxillofacial radiologists that were annotating the dataset.

Participant	Sensitivity	Specificity
Diagnocat	0.9239	0.9899
DMFR-1	0.9411	0.9939
DMFR-2	0.9438	0.9931
DMFR-3	0.9318	0.9913
DMFR-4	0.9337	0.9946

Table 2. Sensitivity and specificity of each evaluated condition for the Diagnocat system predictions.

Condition	Sensitivity	Specificity
Artificial crown	0.9546	0.9963
Canals N=1	0.9864	0.9661
Canals N=2	0.7759	0.9927
Canals N=3	0.9531	0.9647
Canals N=4	0.654	0.9952
Canals N=5	N/A	0.9998
Caries signs	0.7285	0.9953
Crown defect over 50%	0.8734	0.9975
Endodontically treated tooth	0.9676	0.9953
Filling	0.9721	0.9921
Impaction	0.9137	0.9995
Implant	0.9727	0.9997
Missed canal	0.6695	0.9974
Missing	0.9824	0.9405
Overfilling	0.7831	0.9973
PDL widening along root	0.8794	0.9863
Periapical lesion	0.8383	0.9953
Periapical lesion, PDL widening	0.7587	0.9807
Periapical lesion, poorly circumscribed radiolucency	0.6957	0.9942
Periapical lesion, well circumscribed radiolucency	0.729	0.9974
Periodontal bone loss	0.9489	0.9661
Periodontal bone loss, mild	0.9321	0.9742
Periodontal bone loss, moderate	0.9111	0.9866
Periodontal bone loss, severe	0.9286	0.996
Pontic	0.9101	0.9998
Post and core	0.75	0.9975
Roots N=1	0.9593	0.9888
Roots N=2	0.9181	0.9778
Roots N=3	0.964	0.9786
Roots N=4	0.0	1.0
Short filling	0.6981	0.9937
Voids present in the root filling	0.7329	0.9957

Table 3. Description of each annotated Descriptor conditions.

Condition	Description
Artificial crown	A tooth was restored with an artificial crown
Number of canals	Number of root canals in a tooth (1-5)
Caries signs	Signs of dental caries (cases where caries is certain, and there is a low chance of confusion with a metallic artifact or non-contrast filling)
Crown has defect over 50%	A crown is largely destroyed: at least 50% of the crown is missing
Endodontically treated tooth	A tooth displays signs of previous endodontic treatment
Filling	A crown was restored with a filling
Impaction	A tooth is impacted (unerupted)
Implant	There is an implant in place of a tooth
Missed canal	A root canal was missed (not filled) during endodontic treatment of. Should be specified only if a tooth was endodontically treated.
Missing	Absence of tooth, implant, and pontic under specified number including both extracted teeth and teeth that are not visible in the FoV of an image.
Overfilling	Filling material is visualized beyond a radiographic apex. Should be specified only if a tooth was endodontically treated.
Periapical lesion	Presence of inflammatory periapical lesion adjacent to one or more roots of a tooth
Pontic	There is a pontic restoration in place of a tooth (either base or middle part)
Post and core	A tooth was restored with a post and core restoration
Number of roots	Number of roots in the tooth (1-4)
Short filling	Root canal filling is short (ends in 2 mm or more from a radiographic apex). Should be specified only if a tooth was endodontically treated.
VOIDS present in the root filling	A root canal contains voids (spaces that were not filled during previous endodontic treatment). Should be specified only if a tooth was endodontically treated.

Table 4. Sensitivity and specificity of each evaluated condition for aided and unaided annotations.

Condition	Sensitivity		Specificity	
	aided	unaided	aided	unaided
Artificial crown	0.8337	0.6774	0.9804	0.971
Canals N=1	0.9544	0.9417	0.9655	0.9515
Canals N=2	0.793	0.7345	0.9713	0.9632
Canals N=3	0.828	0.7347	0.9719	0.9666
Canals N=4	0.7345	0.751	0.9807	0.9719
Canals N=5	N/A	N/A	0.9998	0.9992
Caries signs	0.6693	0.6634	0.8593	0.7973
Crown defect over 50%	0.9088	0.7887	0.9901	0.9812
Endodontically treated tooth	0.9761	0.9476	0.9912	0.9841
Filling	0.9308	0.7771	0.9515	0.9151
Impaction	0.8523	0.5	0.9993	0.9993
Implant	1.0	0.969	1.0	0.9996
Missed canal	0.8233	0.7701	0.9867	0.946
Missing	0.8964	0.8973	0.9403	0.9364
Overfilling	0.7909	0.6199	0.9926	0.9922
PDL widening along root	0.746	0.2779	0.9803	0.9731
Periapical lesion	0.8304	0.683	0.9388	0.9465
Periapical lesion, PDL widening	0.5161	0.3202	0.9332	0.9542
Periapical lesion, poorly circumscribed radiolucency	0.4318	0.2045	0.9897	0.981
Periapical lesion, well circumscribed radiolucency	0.7135	0.6048	0.9912	0.9817
Periodontal bone loss	0.7239	0.4501	0.8111	0.8358
Periodontal bone loss, mild	0.4783	0.2329	0.8739	0.8782
Periodontal bone loss, moderate	0.4357	0.2148	0.9248	0.943
Periodontal bone loss, severe	0.3535	0.2727	0.9929	0.9878
Pontic	0.9696	0.9087	0.9988	0.9973
Post and core	0.7877	0.7114	0.9846	0.9715
Roots N=1	0.9559	0.9635	0.9731	0.9558
Roots N=2	0.8609	0.8304	0.9728	0.9726
Roots N=3	0.9376	0.9028	0.9712	0.9785
Roots N=4	0.1818	0.7045	0.9996	0.9986
Short filling	0.8139	0.6285	0.9846	0.9817
Voids present in the root filling	0.7592	0.4938	0.9826	0.9757

Figures

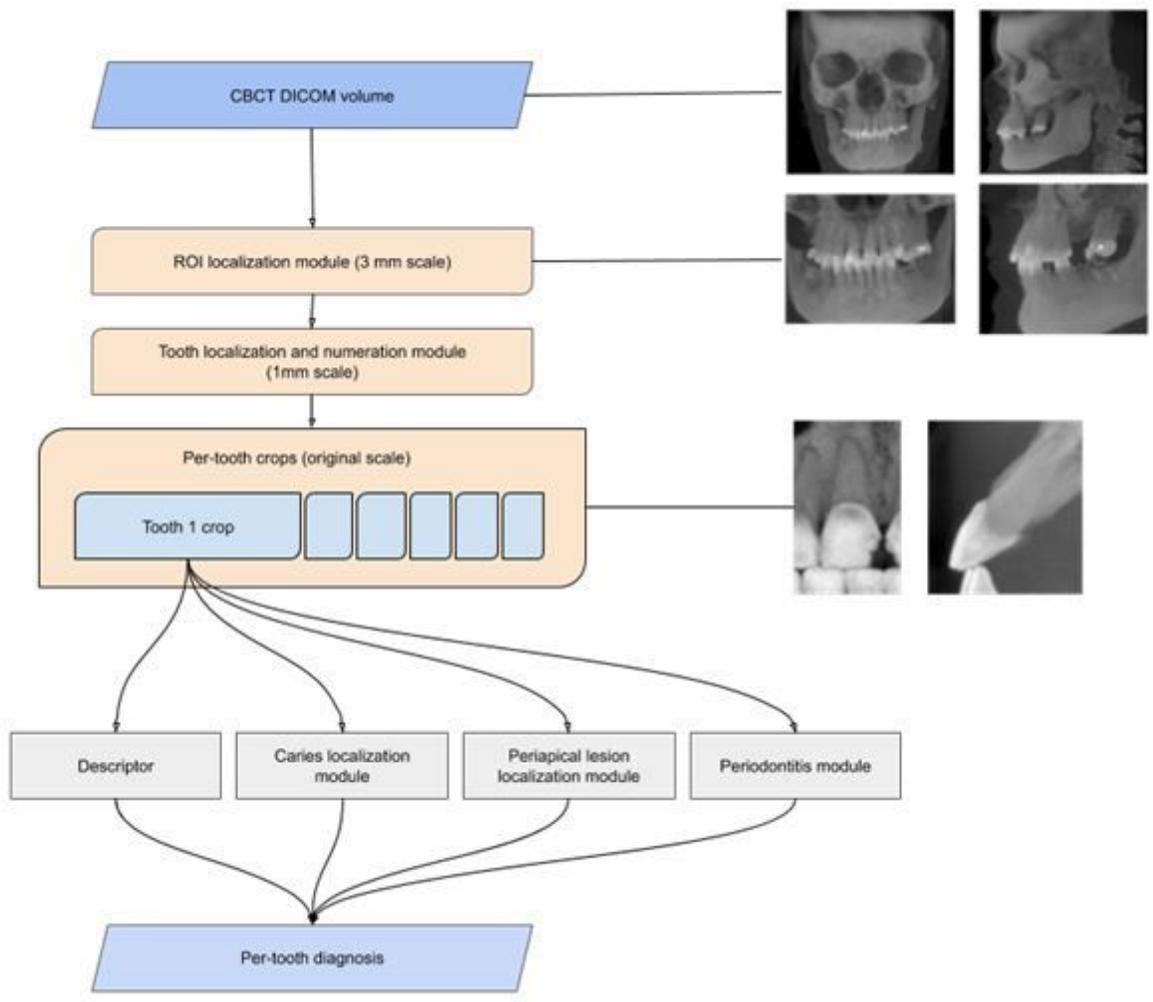


Figure 1

Flow diagram of CBCT processing pipeline

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplements.docx](#)
- [Supplements.docx](#)