INTERNATIONAL ENDODONTIC JOURNAL

doi:10.1111/iej.13265

Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans

K. Orhan[1], I. S. Bayrakdar[2], M. Ezhov[3], A. Kravtsov[3] & T. Özyürek[4]

[1]Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Ankara University, Ankara; [2]Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Eskisehir Osmangazi University, Eskişehir, Turkey; [3]Diagnocat Inc, San Francisco, CA, USA; and [4]Department of Endodontics, Faculty of Dentistry, Istanbul Medeniyet University, Istanbul, Turkey

Abstract

Orhan K, Bayrakdar IS, Ezhov M, Kravtsov A, Özyürek T. Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. *International Endodontic Journal*.

Aim To verify the diagnostic performance of an artificial intelligence system based on the deep convolutional neural network method to detect periapical pathosis on cone-beam computed tomography (CBCT) images.

Methodology Images of 153 periapical lesions obtained from 109 patients were included. The specific area of the jaw and teeth associated with the periapical lesions were then determined by a human observer. Lesion volumes were calculated using the manual segmentation methods using Fujifilm-Synapse 3D software (Fujifilm Medical Systems, Tokyo, Japan). The neural network was then used to determine (i) whether the lesion could be detected; (ii) if the lesion was detected, where it was localized (maxilla, mandible or specific tooth); and (iii) lesion volume. Manual segmentation and artificial intelligence (AI) (Diagnocat Inc., San Francisco, CA, USA) methods were compared using Wilcoxon signed rank test and Bland–Altman analysis.

Results The deep convolutional neural network system was successful in detecting teeth and numbering specific teeth. Only one tooth was incorrectly identified. The AI system was able to detect 142 of a total of 153 periapical lesions. The reliability of correctly detecting a periapical lesion was 92.8%. The deep convolutional neural network volumetric measurements of the lesions were similar to those with manual segmentation. There was no significant difference between the two measurement methods ($P > 0.05$).

Conclusions Volume measurements performed by humans and by AI systems were comparable to each other. AI systems based on deep learning methods can be useful for detecting periapical pathosis on CBCT images for clinical application.

Keywords: artificial intelligence, cone-beam computed tomography, deep learning, periapical pathology.

Received 2 August 2019; accepted 7 January 2020

Introduction

Periapical diseases are mostly inflammatory lesions with over 90% classified as apical granulomas, apical cysts or abscesses (Koivisto *et al.* 2012). Teeth provide bacteria with pathways into the supporting bone once the dental pulp has become infected. Apical periodontitis is characterized by an immune cell infiltrate and bone destruction and occurs in 34%–61% of individuals and 3%–4% of teeth (Segura-Egea *et al.* 2015, Huumonen *et al.* 2017, Braz-Silva *et al.* 2019). The incidences of cysts and granulomas range from 6 to

Correspondence: I. S. Bayrakdar, Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Eskisehir Osmangazi University, 26240, Eskişehir, Turkey (e-mail: ibrahimsevkibayrakdar@gmail.com).

55% and from 46% to 94%, respectively (Natkin *et al.* 1984, Shrout *et al.* 1993). Periapical pathosis can be seen radiographically as periapical radiolucencies. Frequently, apical periodontitis is an incidental finding detected on periapical radiographs, panoramic radiographs and cone-beam computed tomography scans (CBCT). CBCT creates high-resolution three-dimensional (3D) images without the distortion and superimposition of bone and dental structures seen in conventional radiographs (AAE and AAOMR 2011, ESE 2019). Several studies have compared the diagnostic accuracy of CBCT with panoramic, conventional and digital periapical radiography. CBCT significantly increased detection of canal spaces and periapical areas compared to conventional periapical and panoramic radiographs (Estrela *et al.* 2008, Patel *et al.* 2012, Pope *et al.* 2014, Davies *et al.* 2015). This suggests that CBCT enhances detection of periapical bone lesions and offers improved diagnosis, treatment planning and prognosis (Estrela *et al.* 2008, Patel *et al.* 2012, Pope *et al.* 2014, Davies *et al.* 2015).

In recent years, medical imaging has developed at a remarkable pace, including image management via picture archiving systems (PACS), advances in artificial intelligence (AI) and computer-aided diagnostic (CAD) systems. CAD systems can assist physicians and radiologists in the decision-making process for various medical problems (Doi 2007, Tuzoff *et al.* 2019). Deep learning is an AI method used for automated decision-making in various clinical tasks. These techniques, which are constructed through the development of artificial neural networks, allow data sets to be categorized automatically and promote learning features contained within data via multilayer convolutional neural networks (CNNs). CNNs aim to simulate the architecture of the human brain, processing data using a series of interconnected 'neurons'. This AI method can learn adaptive image characteristics and simultaneously make image classifications (LeCun *et al.* 2015, Shin *et al.* 2016, Tajbakhsh *et al.* 2016, Kim & MacKinnon 2018). CNNs have been successfully used for automatic assessment of various medical and dental problems, including image-based automated diagnosis to detect lung and brain lesions (Akkus *et al.* 2017, Song *et al.* 2017, Wang *et al.* 2017a, Blanc-Durand *et al.* 2018), breast cancer in mammography images (Becker *et al.* 2017), colorectal polyps and prostate cancer (Wang *et al.* 2017b, Byrne *et al.* 2019), skin cancer (Esteva *et al.* 2017), diabetic retinopathy in retinal fundus photographs (Gulshan *et al.* 2016), hip osteoarthritis (Xue *et al.* 2017) and bone age assessment (Lee *et al.* 2017). In dentistry, CNNs have been applied to detect carious lesions, periapical lesions, tooth eruption and numbering, vertical root fractures, assess root morphology or periodontal bone loss, dental and jaw pathosis, osteoporosis, and maxillary sinusitis on dental radiographs (Kositbowornchai *et al.* 2013, Miki *et al.* 2017, Ezhov *et al.* 2018, Murata *et al.* 2018, Poedjiastoeti & Suebnukarn 2018, Lee *et al.* 2018a,b, Zakirov *et al.* 2018a, Zakirov *et al.* 2018b, Chen *et al.* 2019, Ekert *et al.* 2019, Hiraiwa *et al.* 2019, Hwang *et al.* 2019, Krois *et al.* 2019, Tuzoff *et al.* 2019).

The purpose of the present study was to verify the diagnostic performance of an artificial intelligence system based on the deep convolutional neural network method to detect periapical pathosis in CBCT images.

Materials and methods

Use of deep convolutional neural network

A two-step process was used to create, train and validate the deep convolutional neural network. To train the neural net, a data set of depersonalized 3D CBCT scans was used. A set of 2800 scans with periapical lesions around teeth were annotated using per-voxel label assignment (i.e. each voxel was labelled as background or pathology) by maxillofacial radiologists. To obtain precise segmentation results, specialists used ITK-SNAP software that allows users to navigate 3D images in all three planes. Once annotated, each mask was automatically examined to eliminate human factors, for example, misalignment of tooth volume and resulting mask.

To provide a negative control, 1100 examples without periapical pathosis were obtained during the data set collection stage. In the present study, these examples were considered as 'hard negatives'. Identification of periapical pathosis was complicated due to anatomical variation, metal-induced artefacts and CBCT image quality. Two specialists validated the aforementioned samples in a 2D manner with a cross-section of teeth in each dimension. The model was normalized using a separate deep convolutional neural network examining so-called 'soft negatives' or teeth for which there was a high level of confidence in the absence of periapical lesions. This type of training formed half of the resulting data set.

Architecture of the deep convolutional neural network

The deep learning process was performed using U-net-like architecture. U-Net is an encoder–decoder style neural network that solves semantic segmentation tasks end to end; it extends the fully convolutional network found in the work of Long *et al.* (2015).

The problem formulation in terms of the machine learning task is semantic segmentation, including segmenting background and periapical pathology. For this purpose, specificity and sensitivity metrics were used to measure the number of incorrectly identified positive and negative conditions and to evaluate diagnostic performance. To measure the pathosis localization capabilities of the model, binary voxel-wise intersection over union (IoU) of the ground truth mask and prediction were used.

Patient selection

CBCT scans for test data sets taken for various diagnostic purposes were obtained from the CBCT archive of Eskisehir Osmangazi University Faculty of Dentistry. A total of 153 periapical lesion images obtained from 109 patients were included in this study. The research protocol was approved by the Non-interventional Clinical Research Ethical Committee of Eskisehir Osmangazi University (decision date and number: 28.05.2019/48) and was performed in accordance with the principles of the Declaration of Helsinki.

Imaging

The same CBCT scanner (ProMax 3D Mid; Planmeca, Helsinki, Finland) was used for all patients, who were in a standing position during imaging. Diagnostic settings were as follows: 94 kVp, 14 mA, 360° rotation, 27 s. The scanner offers multiple fields of view (FOVs) allowing the dentist to select the optimum scan on a case-by-case basis. Images were obtained using a 5 × 5.5 FOV (0.075 mm^3 and 0.100 mm^3 voxel size), a 5 × 5.5 FOV (0.150 mm^3 voxel size) and a 10 × 5.5 FOV (0.200 mm^3 voxel size) with isotropic voxels.

Evaluation

Tomography data with periapical lesions were anonymized in DICOM format, and images were evaluated by an oral and maxillofacial radiologist (İ.Ş.B.) with 8 years of professional experience. Jaw location and lesion-associated teeth were recorded, and lesion volume was calculated by manual segmentation methods using Fujifilm-Synapse 3D software (Fujifilm Medical Systems, Tokyo, Japan). Then, the files were randomly uploaded to the deep convolutional neural network (Diagnocat, Inc., San Francisco, CA, USA) for calculation of lesion volumes. Volumes calculated by the manual segmentation and artificial intelligence (AI) methods were compared (Figures 1 and 2).

Model pipeline

Diagnocat's approach to diagnosing periapical lesions is based on a deep convolutional neural network using a U-net-like architecture. At this point, the exact type of periapical pathosis must be specified using a separate deep CNN. Once trained, the model is utilized to identify the presence of apical periodontitis in the following manner: (1) the whole CBCT scan volumetric image is fed into the model, followed by (2) pre-processing of the incoming image, (3) localization of each present tooth in the 3-D volume, (4) extraction of the tooth of interest with its surrounding context, (5) rescaling of the tooth image to establish a 0.25mm isotropic voxel resolution using linear interpolation, (6) prediction of the condition and pathosis mask (semantic segmentation), (7) post-processing of the mask, which includes thresholding and splitting the pathosis mask into components and (8) measuring the pathosis volume of each component to be localized. Steps 3 and 6 are implemented as separate deep convolutional neural networks, whereas steps 2, 4, 5, 7 and 8 are algorithmic procedures. The neural network was used to determine (1) whether the lesion

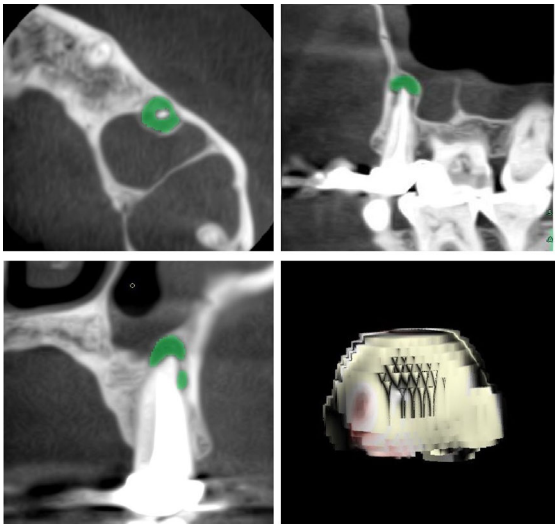

Figure 1 Volume measurement using the manual segmentation method.

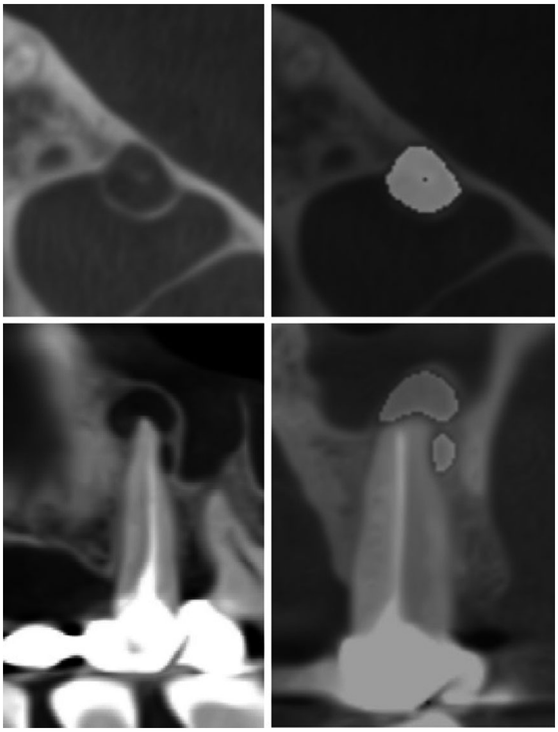

Figure 2 Volume measurement using the artificial intelligence (AI) method.

could be detected and, if so, (2) where it was localized (maxilla, mandible or specific tooth).

Examiner consistency

Intraclass correlation coefficients were used to evaluate intra-examiner agreement and reliability. Some (20%) measurements were repeated to determine intraobserver compliance; the intraobserver compliance coefficient was calculated by the same investigator 2 weeks after the first round of observations. The intraclass correlation coefficient (95% confidence interval) was 0.998 (0.996–0.999) for these measurements.

Statistical analyses

The SPSS 21.0 Package Data Program (SPSS 21.0 Software Package Program, Inc., Chicago, IL, USA) was used to evaluate all data. Data were assessed for normality using the Kolmogorov–Smirnov test. For comparisons of the volumes calculated by the manual segmentation and artificial intelligence (AI) methods, the Wilcoxon signed rank test was used. A value of $P < 0.05$ was considered statistically significant. Volumetric agreement between manual segmentation method and deep CNN system was evaluated using the Bland–Altman analysis.

Diagnostic performance was evaluated according to Özdemir *et al.* (2010) with recall, precision and F-measure values defined as follows:

Recall = number of correctly detected periapical lesions/number of all periapical lesions.

Precision = number of correctly detected periapical lesions/(number of correctly detected periapical lesions + number of falsely detected periapical lesions).

F-measure = 2 × (recall + precision)/(recall + precision).

where the ‘number of all periapical lesions’ is the number of teeth that had a periapical lesion ($n = 153$), and the F-measure indicates the harmonic mean of the recall and precision values. Diagnostic performance was defined as the mean of the CNN results. Random distribution of the data around the zero is provided by Bland–Altman method using logarithmic transformation, and real limits are also presented by antilog transformation.

Results

The deep convolutional neural network system was successful in detecting tooth presence and specific tooth numbers. There was only one incorrect numbering, which was associated with a congenitally missing tooth being misidentified as a retained primary tooth. The system was able to detect 142 periapical lesions from 153 periapical lesions, a reliability of 92.8% in correctly detecting a periapical lesion. The recall rates were high, whilst the precision rates were also high for detecting the periapical lesions. Consequently, the estimated recall, precision and F-measure values were 0.89, 0.95 and 0.93 respectively.

Comparisons of volume measurements were performed by both the manual segmentation method (Fujifilm-Synapse 3D software, Fujifilm Medical Systems) and the deep CNN (Diagnocat Inc.) system. There was no significant difference between the two measurement methods ($P > 0.05$) (Table 1), although the values were higher for the manual segmentation method (Figure 3). The confidence intervals for the differences between the two techniques are presented in Table 2. The degree of agreement between two methods was assessed using

Table 1 Comparison of volume slices calculated by manual segmentation and artificial intelligence (AI)

Methods	*n*	Mean	Median	Minimum	Maximum	ss	Wilcoxon sign test *z*	Wilcoxon sign test *P*
Synapse-Manuel/mm^3	142	191.41	72.45	0.91	2740.00	341.80	−1.9	0.051
Diagnocat-mm^3	142	143.84	52.61	0.38	1210.88	214.78		

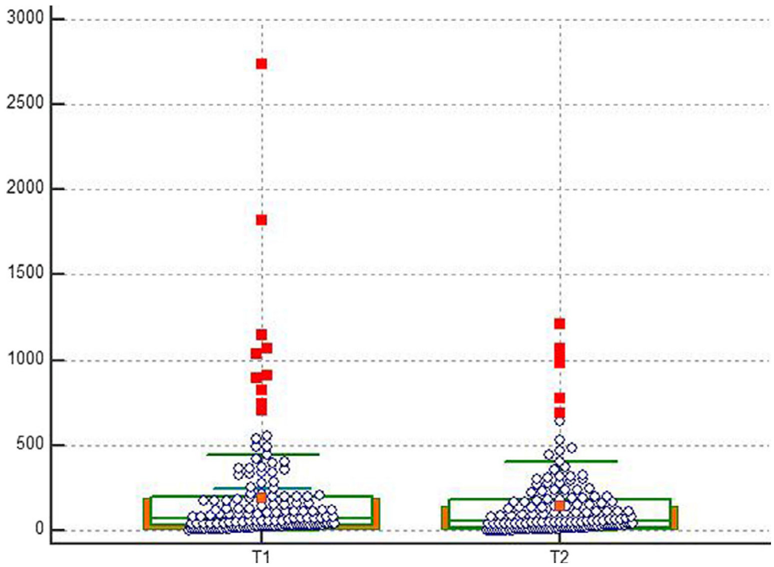

Figure 3 Box graph of the measured values, T1: Synapse-Manuel/mm3, T2-Diagnocat/mm^3.

Table 2 Confidence interval results for differences between the two techniques

	Mean	Standard deviation (SD)	%95 CI*
Differences (Logarithmic)	1.71	0.723	3.127-0.292
Differences (Original)	5.535	2.062	22.852-1.341

*Mean ± 1.96*SD.

Bland–Altman analysis (Figure 4). The Bland–Altman plot showed the mean logarithmic transformations of manual and CNN measurements as 0.15, and their corresponding 95% limits of agreement were −1.21 to 1.52, respectively. Similarly, high degree of agreement was found for all the parameters derived by two methods of periapical lesion volumetric measurements.

Discussion

The integration of AI into the medical field has accelerated with the development of deep learning and neural methods, with AI being used to solve multiple clinical problems. Recently, its use in dentistry has grown in parallel with the use of deep learning methods in the medical field. Previously, most studies have aimed to assess the impact of AI in the dental field. Miki *et al.* (2017) investigated an automated method for classifying tooth types on dental cone-beam CT images using a deep convolutional neural network (DCNN) as a component of automated dental charting. They found high accuracy (up to 91.0%) in their DCNN for differentiating teeth and concluded that AI could be efficiently used for automatic dental charting, which may be valuable in forensic identification (Miki *et al.* 2017). Tuzoff *et al.* (2019) focused on analysing panoramic radiographs using CNN-based

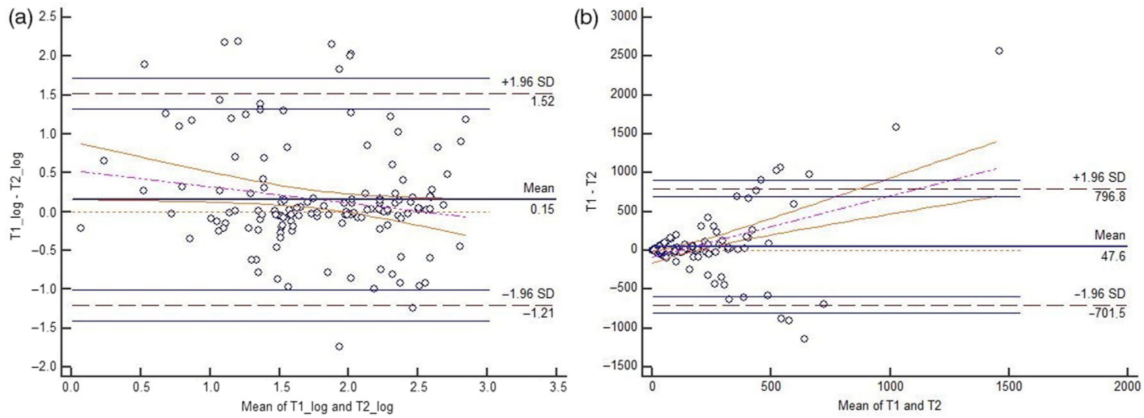

Figure 4 (a) Bland–Altman graphic of logarithmic transformations (b) Bland–Altman graphic of original values, T1: Synapse-Manuel/mm3, T2-Diagnocat/mm^3.

models. They sought to describe the upper and lower jaws within a single image, using a CNN-based deep learning model trained to detect and number teeth during automated dental charting. Their results demonstrated that AI deep learning algorithms have potential for practical application within a clinical setting (Tuzoff *et al.* 2019). Chen *et al.* (2019) presented a study to detect and number teeth in dental periapical films using faster regions with convolutional neural network features (faster R-CNN) in the TensorFlow library. They used three post-processing methods to integrate the basis of faster R-CNN to improve detection predictions. Their work revealed that both predictions and recalls were 90% accurate using the faster R-CNN. To demonstrate the system's robust performance, the algorithms were compared to the responses of three dentists, who reviewed the data set independently. Chen *et al.* (2019) concluded the AI machines performed at a success rate close to that of a junior dentist.

Devito *et al.* (2008) reported a study utilizing an artificial multilayer perceptron neural network models to diagnose interproximal dental caries. They reported a 39.4% improvement using the neural network. Valizadeh *et al.* (2015) also designed computer software for detection of interproximal caries in posterior teeth. Their software diagnosed 60% of enamel caries and 97% of dentine caries presented but struggled to analyse enamel caries effectively (Valizadeh *et al.* 2015). Lee *et al.* (2018a) evaluated the efficacy of deep CNN algorithms for detection and diagnosis of dental caries on periapical radiographs. A GoogLeNet Inception v3 CNN network was used for pre-processing and transfer learning. The diagnostic accuracies for premolar, molar and both combined-teeth models were 89.0%, 88.0%, and 82.0% respectively. The premolar model was more successful than other models. The authors concluded that a deep learning-based CNN algorithm can detect dental caries in periapical radiographs (Lee *et al.* 2018a).

Krois *et al.* (2019) designed deep CNNs to detect periodontal bone loss on panoramic radiographs. A CNN trained on a limited amount of image segments showed discrimination ability similar to that of live dentists assessing periodontal bone loss with panoramic radiographs (Krois *et al.* 2019). Johari *et al.* (2017) modelled a probabilistic neural network (PNN) to detect vertical root fractures in vital and endodontically treated teeth using periapical and CBCT radiographs. They confirmed that the neural network diagnosed fractures more effectively using CBCT images than periapical radiographs, suggesting that this model may benefit endodontic assessments (Johari *et al.* 2017). Kositbowornchai *et al.* (2013) also created a neural network to diagnose vertical root fractures using intraoral digital radiographs and evaluated the diagnostic performance of their neural network. Their study indicated that an artificial neural network could successfully be trained to make correct interpretations of root fractures and surrounding bone (Kositbowornchai *et al.* 2013).

Hiraiwa *et al.* (2019) evaluated the diagnostic performance of a deep learning system viewing panoramic radiographs to assess the number of distal roots present on mandibular first molars-based training through CBCT findings. Their system was capable of detecting additional roots at a consistent performance level (Hiraiwa *et al.* 2019). Poedjiastoeti & Suebnukarn (2018) created a CNN to detect ameloblastomas and keratocystic odontogenic tumours, two of the most common dental tumours seen in the mandible. Whilst the sensitivity, specificity, accuracy and diagnostic time were 81.8%, 83.3%, 83.0% and 38 seconds, respectively, for the CNN, the oral and maxillofacial specialist matched the AI in all these parameters, except diagnostic time, which took 23.1 minutes. The authors concluded that ameloblastomas and keratocystic odontogenic tumours could be detected based on digital panoramic images using CNN, leading to a substantially shorter time to diagnosis (Poedjiastoeti & Suebnukarn 2018).

Various studies have also evaluated the performance of deep learning image classification for the diagnosis of lymph node metastases. The performance of these AI systems was not significantly different from that of radiologists, suggesting that these systems could be a useful method for diagnostic support (Ariji *et al.* 2019). Deep learning systems are also capable of detecting the impact of systemic diseases on oral tissues. A DCNN-based CAD system showed strong agreement with experienced oral and maxillofacial radiologists in detecting osteoporosis; this system could provide information to dentists for early detection of osteoporosis, allowing asymptomatic patients to be referred to the appropriate medical professionals for preventive care (Lee *et al.* 2018b). A recent study by Murata *et al.* (2018) reported that deep learning systems could diagnose maxillary sinusitis on a panoramic radiograph at a rate comparable to that of radiologists and that their performance was superior to that of dental residents (Murata *et al.* 2018).

In a review of the literature, only one study was found assessing apical pathosis detection with an AI system. Ekert *et al.* (2019) assessed the ability of deep CNNs to detect apical lesions on panoramic radiographs. Although their study only included a limited number of panoramic radiographs, the CNN was capable of detecting lesions. However, the authors cautioned that the sensitivity of their system should be improved before clinical use (Ekert *et al.* 2019).

Patients with root canals that have periapical lesions and associated symptomatology can pose a serious challenge in terms of diagnosis and treatment planning. The exact problem is often hard to discern precisely and a patient may continue to experience symptoms without any radiographic signs of further periapical disease (Patel *et al.* 2009a). For endodontic procedures, two-dimensional (2D) periapical radiographs are the standard for diagnosis and follow-up. However, 2D radiographs have several drawbacks, including errors that are classified as either 'errors of projection' or 'errors of identification'. Like all conventional radiographic techniques, periapical radiographs collapse a 3D structure onto a 2D plane. The resulting superimposition of anatomical structures complicates image interpretation and landmark identification, where distortion and magnification may lead to a reduction in measurement accuracy (Oz *et al.* 2011). One of the major problems encountered in the diagnosis and management of periapical lesions is that intraoral radiographs provide limited diagnostic information. The information gained from conventional and digital periapical radiographs is incomplete due to the fact that the 3D anatomy of the area being radiographed is compressed into a 2D image or shadowgraph (Patel *et al.* 2009b). Although intraoral radiography is reasonably accurate in diagnosing endodontic pathosis, CBCT has proven to be beneficial for diagnosing periapical lesions not identified by periapical radiographs (Scarfe *et al.* 2009). In the clinical setting, patients with endodontic problems can pose a serious challenge in terms of diagnosis and treatment planning. The problem is often hard to precisely discern when a patient has symptoms without any radiographic signs of further periapical disease. The high accuracy afforded by CBCT makes it a valuable tool for the analysis of both tooth structure and the adjacent anatomy (Estrela *et al.* 2009). CBCT is highly accurate for detecting periapical lesions at early stages and establishing a differential diagnosis (Leonardi *et al.* 2016).

In the current study, the AI system detected 142 of the 153 (92.8 %) periapical lesions examined. A significant positive correlation was found between the volumetric measurements taken by the radiologists and those dome by the machine. In the present study, differences between manual volume segmentation and AI measurements were observed in a few cases. The causes of these differences were considered due to several factors. Separate lesions in the neighbouring teeth could be segmented together, and buccal–palatal and lingual cortical bone perforations could affect the AI system's ability to distinguish between lesion area and soft tissue. CBCT images do not provide superior soft tissue resolution, making soft tissue and lesion density similar. Also, the presence of endo-perio lesions, periodontal defects and alveolar bone loss may have altered the AI system's measurement. Neighbouring normal anatomic structures including the maxillary incisive canal, inferior alveolar canal, mental foramen, maxillary sinus and nasal fossa can be segmented, impacting AI analysis. Large lesions associated with multiple teeth and root canal-treated teeth associated with lesions may also change the AI system's measurements. Dental anomalies such as dens in dente, incomplete apex development, an open apex or a larger than normal root canal might also influence the analysis. Further programming with variations on normal anatomy will be needed to address some of these concerns.

Conclusion

Volumetric measurements created by a manual segmentation method and by an AI system were comparable to each other. There was no significant difference between the two measurement methods. AI systems based on deep learning methods can be useful in detecting periapical pathosis in CBCT images for clinical application.

Conflict of interest

The authors have stated explicitly that there are no conflicts of interest in connection with this article.

References

Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ (2017) Deep learning for brain MRI segmentation: state of the art and future directions. *Journal of Digital Imaging* **30**, 449–59.

AAE and AAOMR. (2011) Use of cone-beam computed tomography in endodontics Joint Position Statement of the American Association of Endodontists and the American Academy of Oral and Maxillofacial Radiology. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology* **111**, 234–7.

Ariji Y, Fukuda M, Kise Y *et al.* (2019) Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of artificial intelligence. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology* **127**, 458–63.

Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T, Boss A (2017) Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. *Investigative Radiology* **52**, 434–40.

Blanc-Durand P, Van Der Gucht A, Schaefer N, Itti E, Prior JO (2018) Automatic lesion detection and segmentation of 18F-FET PET in gliomas: a full 3D U-Net convolutional neural network study. *PLoS ONE* **13**, e0195798.

Braz-Silva PH, Bergamini ML, Mardegan AP, De Rosa CS, Hasseus B, Jonasson P (2019) Inflammatory profile of chronic apical periodontitis: a literature review. *Acta Odontologica Scandavica* **77**, 173–80.

Byrne MF, Chapados N, Soudan F *et al.* (2019) Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* **68**, 94–100.

Chen H, Zhang K, Lyu P *et al.* (2019) A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Scientific Reports* **9**, 3840.

Davies A, Mannocci F, Mitchell P, Andiappan M, Patel S (2015) The detection of periapical pathoses in root filled teeth using single and parallax periapical radiographs versus cone beam computed tomography—a clinical study. *International Endodontic Journal* **48**, 582–92.

Devito KL, de Souza Barbosa F, Felippe Filho WN (2008) An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology* **106**, 879–84.

Doi K (2007) Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics* **31**, 198-211.

Ekert T, Krois J, Meinhold L *et al.* (2019) Deep learning for the radiographic detection of apical lesions. *Journal of Endodontics* **45**, 917–22.

ESE (2019) European Society of Endodontology position statement: Use of cone beam computed tomography in Endodontics: European Society of Endodontology (ESE). *International Endodontic Journal* **52**, 1675–8.

Esteva A, Kuprel B, Novoa RA *et al.* (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–8.

Estrela C, Bueno MR, Leles CR, Azevedo B, Azevedo JR (2008) Accuracy of cone beam computed tomography and panoramic and periapical radiography for detection of apical periodontitis. *Journal of Endodontics* **34**, 273–79.

Estrela C, Bueno MR, Alencar AH *et al.* (2009) Method to evaluate inflammatory root resorption by using cone beam computed tomography. *Journal of Endodontics* **35**, 1471–7.

Ezhov M, Zakirov A, Gusarev M (2018) Coarse-to-fine volumetric segmentation of teeth in vone-beam CT. arXiv:1810.10293 [cs.CV] URL https://arxiv.org/abs/1810.10293[accessed on 24 Oct 20187]

Gulshan V, Peng L, Coram M *et al.* (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Journal of the American Medical Association* **316**, 2402–10.

Hiraiwa T, Ariji Y, Fukuda M *et al.* (2019) A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. *Dentomaxillofacial Radiology* **48**, 20180218.

Huumonen S, Suominen AL, Vehkalahti MM (2017) Prevalence of apical periodontitis in root filled teeth: findings from a nationwide survey in Finland. *International Endodontic Journal* **50**, 229–36.

Hwang JJ, Jung YH, Cho BH, Heo MS (2019) An overview of deep learning in the field of dentistry. *Imaging Science in Dentistry* **49**, 1–7.

Johari M, Esmaeili F, Andalib A, Garjani S, Saberkari H (2017) Detection of vertical root fractures in intact and endodontically treated premolar teeth by designing a probabilistic neural network: an ex vivo study. *Dentomaxillofacial Radiology* **46**, 20160107.

Kim DH, MacKinnon T (2018) Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clinical Radiology* **73**, 439–45.

Koivisto T, Bowles WR, Rohrer M (2012) Frequency and distribution of radiolucent jaw lesions: a retrospective analysis of 9,723 cases. *Journal of Endodontics* **38**, 729–32.

Kositbowornchai S, Plermkamon S, Tangkosol T (2013) Performance of an artificial neural network for vertical root fracture detection: an ex vivo study. *Dental Traumatology* **29**, 151–5.

Krois J, Ekert T, Meinhold L *et al.* (2019) Deep learning for the radiographic detection of periodontal bone loss. *Scientific Reports* **9**, 8495.

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* **521**, 436–44.

Lee H, Tajmir S, Lee J *et al.* (2017) Fully Automated deep learning system for bone age assessment. *Journal of Digital Imaging* **30**, 427–41.

Lee JH, Kim DH, Jeong SN, Choi SH (2018a) Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *Journal of Dentistry* **77**, 106–11.

Lee JS, Adhikari S, Liu L, Jeong HG, Kim H, Yoon SJ. (2018b) Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: a preliminary study. *Dentomaxillofacial Radiology* 20170344.

Leonardi DK, Haas L, Porporatti AL *et al.* (2016) Diagnostic accuracy of cone-beam computed tomography and conventional radiography on apical periodontitis: a systematic review and meta-analysis. *Journal of Endodontics* **42**, 356–64.

Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. IEEE Conference Computer Vision Pattern Recognition (CVPR) Piscataway, NJ, USA, pp 3431–40.

Miki Y, Muramatsu C, Hayashi T *et al.* (2017) Classification of teeth in cone-beam CT using deep convolutional neural network. *Computers in Biology and Medicine* **80**, 24–9.

Murata M, Ariji Y, Ohashi Y *et al.* (2018) Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. *Oral Radiology* **35**, 301–7.

Natkin E, Oswald RJ, Carnes LI (1984) The relationship of lesion size to diagnosis, incidence, and treatment of periapical cysts and granulomas. *Oral Surgery, Oral Medicine, Oral Pathology* **57**, 82–94.

Oz U, Orhan K, Abe N (2011) 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. *Dentomaxillofacial Radiology* **40**(8), 492–500.

Özdemir B, Aksoy D, Eckert D, Pesaresi M, Ehrlich D (2010) Performance measures for object detection evaluation. *Pattern Recognition Letters* **31**, 1128–37.

Patel S, Dawood A, Wilson R, Horner K, Mannocci F (2009a) The detection and management of root resorption lesions using intraoral radiography and cone beam computed tomography–an in vivo investigation. *International Endodontic Journal* **42**, 831–8.

Patel S, Dawood A, Whaites E, Pitt Ford T. (2009b) New dimensions in endodontic imaging: part 1. Conventional and alternative radiographic systems. *International Endodontic Journal* **42**(6), 447–462.

Patel S, Wilson R, Dawood A, Foschi F, Mannocci F (2012) The detection of periapical pathosis using digital periapical radiography and cone beam computed tomography—part 2: a 1-year post-treatment follow-up. *International Endodontic Journal* **45**, 711–23.

Poedjiastoeti W, Suebnukarn S (2018) Application of convolutional neural network in the diagnosis of jaw tumors. *Healthcare Informatics Research* **24**, 236–41.

Pope O, Sathorn C, Parashos P (2014) A comparative investigation of cone-beam computed tomography and periapical radiography in the diagnosis of a healthy periapex. *Journal of Endodontics* **40**, 360–5.

Scarfe WC, Levin MD, Gane D, Farman AG (2009) Use of cone beam computed tomography in endodontics. *International Journal of Dentistry* **2009**, 634567.

Segura-Egea JJ, Martin-Gonzalez J, Castellanos-Cosano L (2015) Endodontic medicine: connections between apical periodontitis and systemic diseases. *International Endodontic Journal* **48**, 933–51.

Shin HC, Roth HR, Gao M *et al.* (2016) Deep convolutional neural networks for computer-aided detection: cnn architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging* **35**, 1285–98.

Shrout MK, Hall JM, Hildebolt CE (1993) Differentiation of periapical granulomas and radicular cysts by digital radiometric analysis. *Oral Surgery, Oral Medicine, Oral Pathology* **76**, 356–61.

Song Q, Zhao L, Luo X, Dou X (2017) Using deep learning for classification of lung nodules on computed tomography images. *Journal of Healthcare Engineering* **2017**, 8314740.

Tajbakhsh N, Shin JY, Gurudu SR *et al.* (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Transactions on Medical Imaging* **35**, 1299–312.

Tuzoff DV, Tuzova LN, Bornstein MM *et al*. (2019) Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofacial Radiology* **48**, 20180051.

Valizadeh S, Goodini M, Ehsani S, Mohseni H, Azimi F, Bakhshandeh H (2015) Designing of a computer software for detection of approximal caries in posterior teeth. *Iranian Journal of Radiology* **12**, e16242.

Wang H, Zhou Z, Li Y *et al*. (2017a) Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from (18)F-FDG PET/CT images. *European Journal of Nuclear Medicine and Molecular Imaging* **7**, 11.

Wang X, Yang W, Weinreb J *et al*. (2017b) Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning. *Scientific Reports* **7**, 15415.

Xue Y, Zhang R, Deng Y, Chen K, Jiang T (2017) A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis. *PLoS ONE* **12**, e0178992.

Zakirov A, Ezhov M, Gusarev M, Alexandrovsky V,Shumilov E (2018a) End-to-end dental pathology detection in 3D cone-beam computed tomography images. 1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

Zakirov A, Ezhov M, Gusarev M, Alexandrovsky V,Shumilov E. (2018b) Dental pathology detection in 3D cone-beam CT. arXiv:1810.10309 [cs.CV] https://arxiv.org/abs/1810.10309 [accessed on 24 Oct 2018].

