Dentomaxillofacial Radiology (2020) 49, 20190107

birpublications.org/dmfr

SYSTEMATIC REVIEW

The use and performance of artificial intelligence applications in dental and maxillofacial radiology: A systematic review

[1]Kuofeng Hung, [1]Carla Montalvao, [1]Ray Tanaka, [2]Taisuke Kawai and [1]Michael M. Bornstein

[1]Oral and Maxillofacial Radiology, Applied Oral Sciences and Community Dental Care, Faculty of Dentistry, The University of Hong Kong, Hong Kong SAR, China; [2]Department of Oral and Maxillofacial Radiology, School of Life Dentistry at Tokyo, Nippon Dental University, Tokyo, Japan

Objectives: To investigate the current clinical applications and diagnostic performance of artificial intelligence (AI) in dental and maxillofacial radiology (DMFR).
Methods: Studies using applications related to DMFR to develop or implement AI models were sought by searching five electronic databases and four selected core journals in the field of DMFR. The customized assessment criteria based on QUADAS-2 were adapted for quality analysis of the studies included.
Results: The initial electronic search yielded 1862 titles, and 50 studies were eventually included. Most studies focused on AI applications for an automated localization of cephalometric landmarks, diagnosis of osteoporosis, classification/segmentation of maxillofacial cysts and/or tumors, and identification of periodontitis/periapical disease. The performance of AI models varies among different algorithms.
Conclusion: The AI models proposed in the studies included exhibited wide clinical applications in DMFR. Nevertheless, it is still necessary to further verify the reliability and applicability of the AI models prior to transferring these models into clinical practice.
Dentomaxillofacial Radiology (2020) **49**, 20190107. doi: 10.1259/dmfr.20190107

Cite this article as: Hung K, Montalvao C, Tanaka R, Kawai T, Bornstein MM. The use and performance of artificial intelligence applications in dental and maxillofacial radiology: A systematic review. *Dentomaxillofac Radiol* 2020; **49**: 20190107.

Keywords: artificial intelligence; diagnostic imaging; radiography; dentistry; computer-assisted

Introduction

Artificial intelligence (AI) is defined as the capability of a machine to imitate intelligent human behavior to perform complex tasks, such as problem solving, object and word recognition, and decision-making.[1,2] AI technologies have achieved remarkable success and have also influenced daily life in the form of search engines (such as Google search), online assistants (such as Siri) and games (such as AlphaGo), and are developing quite broadly into various other fields,[3] including medicine. In the field of clinical medicine, a large number of AI models are being developed for automatic prediction of disease risk, detection of abnormalities/pathologies, diagnosis of disease, and evaluation of prognosis.[4–6] Radiology is seen to offer more straightforward access for AI into medicine due to its nature of producing digitally coded images that can be more easily translated into computer language.[7]

Machine learning is one of the core subfields of AI that enables a computer model to learn and make predictions by recognizing patterns.[2] Just as radiologists are trained by repeatedly evaluating medical images, the main advantage of machine learning is that the respectively designed AI model is able to improve and learn with experience through increased training based on large and novel image data sets.[8] A large number of studies have reported the applications of AI diagnostic models, for example, to automatically detect pulmonary nodules,[9,10] colon polyps,[11] cerebral aneurysms,[12] prostate cancer,[13,14]coronary artery calcification,[15] to differentiate skin lesions[16] as well as lung nodules into benign or malignant,[17,18] and to assess bone age.[19] With

Correspondence to: Michael M. Bornstein, E-mail: bornst@hku.hk
Received 21 March 2019; revised 14 June 2019; accepted 31 July 2019

the assistance of AI diagnostic models, radiologists are hoping to not only be relieved from reading and reporting on a large number of medical images, but also to improve work efficiency and to achieve more precise outcomes regarding the final diagnosis of various diseases.[6,20]

In the field of dental and maxillofacial radiology (DMFR), pre-clinical studies have been reporting on AI diagnostic models to exactly locate root canal orifices[21,22] and detect vertical root fractures[23] and proximal dental caries[24] with generally favorable findings. These initial data were certainly encouraging further studies on AI diagnostic models to transfer pre-clinical findings into clinical applications. However, the available evidence of AI used in DMFR / diagnostic imaging has not been assessed yet. Therefore, the aim of this study was to systematically investigate the literature related to clinical applications of AI in DMFR, and to provide a comprehensive update on the current diagnostic performance of AI in DMFR and diagnostic imaging.

Methods and material

This systematic review was conducted in accordance with the guidance of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for diagnostic test accuracy (PRISMA-DTA).[25] The study protocol was prospectively registered in "PROSPERO: International prospective register of systematic reviews" (registration number: PROSPERO 2018 CDR42018091176).

Focused question and study eligibility

The focused question used for the present literature search was "What are current clinical applications and diagnostic performance of AI in DMFR?", and studies were deemed eligible if they addressed this question. The PICO elements are listed in Table 1.

Inclusion criteria for the studies were as follows

(1) Original articles published in English;
(2) Radiology-/imaging-based clinical studies using AI models for automatic diagnosis of disease, detection of abnormalities/pathologies, measurement of pathological area/volume or identification of anatomical structures in the dental and maxillofacial region;
(3) Studies that enable an assessment of the performance of AI models.

Exclusion criteria for the studies were as follows

(1) Review articles, letter to editors and case reports/case series involving less than 10 cases;
(2) Full-text is not available or accessible.

Study search strategy and process

An electronic search was performed in August 2018 via databases including PubMed, EMBASE and Medline via Ovid, Cochrane Central Register of Controlled Trials, and Scopus. The study search strategy was determined after consulting with the librarian (S.L.) at the Dental Library of the Faculty of Dentistry, The University of Hong Kong, and the keywords used for the search were combinations of Medical Subject Headings (MeSH) terms for each database, respectively (Table 1). Vocabulary and syntax were adjusted across databases. There was no restriction on the publication period, but the search was restricted to clinical studies in English, excluding experimental articles (*ex-vivo* / *in-vitro* / animal models). Records were collated in a reference manager software (Endnote™; Version: X7, Clarivate Analytics, New York, NY, USA), and the titles were screened for duplicates. An additional manual search was performed in selected core journals in the field of DMFR ("DentoMaxilloFacial Radiology", "Oral Surgery, Oral Medicine, Oral Pathology, and Oral Radiology", "Oral Radiology", and "Imaging Science in Dentistry"), and

Table 1 Description of the "P (population) I (intervention) C (comparator) O (outcomes)" elements used in framing the research question and the search strategy

Criteria	Specification
Focus question	"What are current clinical applications and diagnostic performance of AI in DMFR?"
Population	Clinical images obtained from human subjects in the dental and maxillofacial region;
Intervention	Diagnostic model based on AI algorithms;
Comparator	Reference standard, such as expert's judgment, clinical/pathological examination, etc;
Outcome	Diagnostic performance of the proposed AI model, such as accuracy, sensitivity, specificity, PPV/NPV, AUC and mean difference from reference.
Search strategy	Artificial Intelligence[Mesh] OR Diagnosis, Computer-Assisted[Mesh] OR Neural Networks (Computer)[Mesh] OR AI OR CNN OR Machine learning OR Deep learning OR Convolutional OR Automatic OR Automated AND Diagnostic imaging[Mesh] AND Dentistry[Mesh]

AI, artificial intelligence; AUC, area under the receiver operating characteristic curve; CNN, convolutional neural networks; DMFR, dental and maxillofacial radiology; PPV/NPV, positive/negative predictive value.

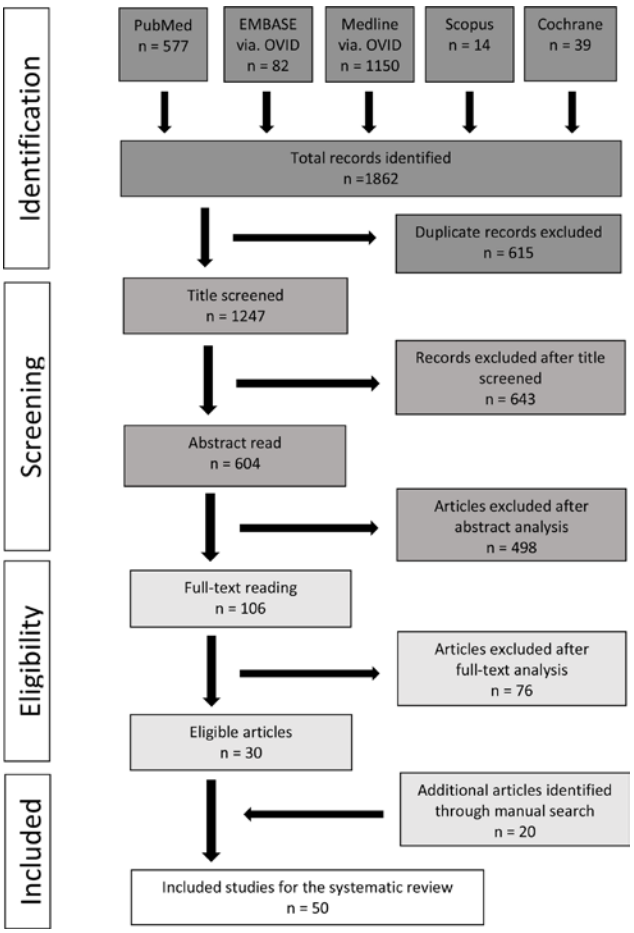

Figure 1 PRISMA flowchart illustrating the study selection process

the reference lists were completed with these studies included. Titles of all records were manually screened independently by two reviewers (K.H. and C.M.), and abstracts of the records were read to identify studies for further full-text reading. Any mismatch at this stage was resolved by discussion. Selected full texts were similarly read, and suitability for inclusion was verified independently by a third reviewer (M.B.). Cohen's κ values were calculated to assess the inter reviewer agreement for the selection of titles, abstracts, and full texts.

Data extraction and outcome of interest

Two reviewers (K.H. and C.M.) extracted data from the studies included and tabulated this information using standardized templates. Data items included author title, year of publication, imaging modality, application of AI technique, the workflow of the AI model, the training/testing data sets, validation technique, type of reference standard, and performance of the AI model. The primary outcome of interest was the current clinical applications of AI in the dental and maxillofacial area, and the performance of these AI models in DMFR and diagnostic imaging.

Quality analysis

The methodological quality of the selected studies was graded independently by two reviewers (K.H. and C.M.), and any mismatch was resolved by discussion.

The risks of bias and applicability of the included studies were assessed by using customized assessment criteria based on the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) (Supplementary Table).[26] Studies were rated on a 3-point scale, reflecting concerns about applicability and risk of bias as low (+), high (−), or unclear (?). For "applicability to the research question", studies proposing AI models based only on rule-/knowledge-based algorithms were not considered for a "high" rating.

Results

Study selection

Initially, a total of 1862 titles were identified via the five databases. After removing 615 duplicates, 1247 titles were considered further for screening. Manual screening of these titles led to the selection of 604 records for abstract reading. After abstracts were read, 106 studies were considered suitable for full-text reading. From these, 30 studies met the inclusion criteria and were included. An additional 20 studies were identified through a manual search based on the reference lists of the studies included and the selected core journals. Thus, a total of 50 studies were included in this systematic review. The PRISMA flowchart exhibiting the study selection process is presented in Figure 1. The two reviewers exhibited excellent inter reviewer agreement for the study selection process with Cohen's κ values of 0.81 for the title screening, 0.85 for abstract reading, and 0.97 for the full-text reading.

Study quality assessment

The study quality assessment of the 50 studies included is presented in Supplementary Figure. Concerns regarding applicability were manifested in the domain of subject selection for nine studies that were rated as having a "high" or an "unclear" risk of concern because of a lack of a detailed description of data sets. With regard to the selection of reference standard, most studies were considered as having a "low" risk of concern. Concerns regarding the risk of bias were relatively high in the domain of index test as nearly half of the studies included did not test their AI models on unused independent images.

Study characteristics

The 50 studies reporting applications of AI models to analyze 2D/3D images in the dental and maxillofacial region covered a period from November 1992 to January 2019 (Figure 2). In these studies, six[27–32] used periapical radiographs,14[33–46] used panoramic radiographs, one[47] used both intraoral and panoramic radiographs, ten[48–57]

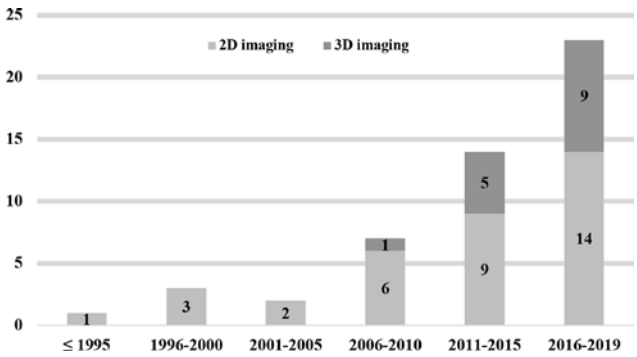

Figure 2 The distribution of the included studies published from November 1992 to January 2019, and the proportion of different image modalities (2D/ 3D imaging) used to develop AI models

used cephalometric radiographs, 14[58–71] used CBCT images, one[72] used intraoral photographs, one[73] used intraoral fluorescent images, and one[74] used undescribed dental X-ray images to develop their AI models. Additionally, there are two studies that used 3D images (CT/MRI)[75] and cephalometric radiographs[76] to test the performance of an available computer-aided diagnostic software (Table 2).

With regard to the applications of these AI models, 19[48–65,75] reported on localization/measurement of cephalometric landmarks, nine[35–43] on diagnosis of osteoporosis, six[33,34,67–69,76] on classification/segmentation of maxillofacial cysts and/or tumors, three[27–29] on identification of alveolar bone resorption, three[30,31,66] on classification of periapical lesions, two[47,74] on the diagnosis of multiple dental diseases, and two[46,72] on classification of tooth types. Furthermore, single studies were found for the identification of root canals,[71] diagnosis of the maxillary sinusitis,[45] identification of inflamed gingiva,[73] identification of dental plaque,[70] detection of dental caries,[32] and for classification of the stages of the lower third molar (Figure 3).[44] The main study characteristics are summarized in Table 2.

With regard to the methods used for validating the performance of the AI models, 13 studies[29,32,35,39,42,46,51,57,65,70,72–74] used split sample validation, ten[28,30,41,48,50,52,54,62,63,66] used leave-one-out cross-validation (LOOCV), two[34,68] used 3-fold cross-validation (CV), one[56] used 4-fold CV, three[38,40,44] used 5-fold CV, three[33,36,47] used 10-fold CV, one[43] used out-of-bag estimate, 12[31,45,49,55,58–61,64,71,75,76] used independent sample validation, three[27,37,75] used multiple validation techniques, and two[53,67] did not describe the validation method (Table 2).

With regard to the performance of these AI models, 19 studies[48–65,75] regarding cephalometric analysis reported the mean/median localization/measurement deviation on all landmarks/measurements, the range of the mean/median deviation on all landmarks/measurement, and/or successful localization rates based on different precision ranges (Table 3). In these 19 studies, only four[48,49,54,58] compared the performances between automatic and manual localizations. Nine studies[35–43] regarding the diagnosis of osteoporosis reported accuracy, sensitivity, specificity, positive/negative predictive value (PPV/NPV), likelihood ratios, and/or area under the receiver operating characteristic curve (AUC) (Table 4). The remaining 22 studies[27–34,44–47,66–74,76] reported accuracy, sensitivity, specificity, AUC, mean difference from reference, and/or correlation against reference (Table 5).

Discussion

The aim of the present systematic review was to provide a comprehensive overview of the current use of AI in DMFR and diagnostic imaging for various indications in dental medicine. The study was performed to identify areas of specific interest for AI applications using DMFR methodologies and devices. According to the results, the number of studies using clinical images in the dental and maxillofacial region to develop AI models has been significantly increasing since 2006 (Figure 2). Because imaging modalities used in DMFR are mostly based on X-rays that are mainly used to judge the hard tissue conditions, the majority of the AI models proposed in these studies were developed to solve clinical issues regarding teeth and jaws. Initially, 2D images including periapical, panoramic and cephalometric radiographs were predominately used to build computer-aided programs for the assistance of clinical diagnosis. In 2009, Flores et al[66] proposed an AI model using CBCT images obtained from patients to distinguish periapical cysts from granuloma. Afterwards, an increasing number of studies attempted to develop AI models based on CBCT images to solve various clinical issues (Figure 2).

According to this review, reports on AI techniques have been increasing for various aspects in DMFR for more than a decade, and most studies focus on four main applications including automatic localization of cephalometric landmarks, diagnosis of osteoporosis, classification/segmentation of the maxillofacial cysts and/or tumors, and identification of periodontitis/periapical disease (Figure 3; Table 2).

AI applications to localize cephalometric landmarks

One-third of the studies assessed proposed AI models to automatically localize cephalometric landmarks for improving the efficiency in orthodontic treatment planning. In current clinical practice, cephalometric analysis can be performed by a manual tracing approach or a computer-aided digital tracing approach. As manual cephalometric analysis is tedious and time-consuming, most orthodontists prefer to perform analyses with the aid of a digital tracing software.[63] However, although a digital tracing software can automatically complete cephalometric measurement, it still requires orthodontists to manually locate cephalometric landmarks on

Table 2 Characteristics of the AI models proposed in the studies included

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Cephalometric landmarks								
Rudolph[48] (1998)	Localization of 15 cephalometric landmarks	Cephalometric radiographs	Spatial spectroscopy	Noise removal; Pixels labelling according to the edginess; Pixels connection and edges labelling; Localization of landmarks based on position or relationship to a labeled edge	14 images from the department of orthodontics	NA	LOOCV	Expert's localization
Liu[49] (2000)	Localization of 13 cephalometric landmarks	Cephalometric radiographs	Knowledge-based algorithm	Determination of the reference point (manual); Dividing the image into eight regions containing all 13 landmarks; Detection of the edges; Localization of the landmarks	NA	ten images from the department of orthodontics	Independent sample validation	Expert's localization
Hutton[50] (2000)	Localization of 16 cephalometric landmarks	Cephalometric radiographs	Active shape model	Searching the image near each point to find the optimum fit; Application of the required translation, rotation, scaling and deformation to the template; Repeat of the fitting procedure until average movement is less than a pixel; Generation of the final configuration and localization of the landmarks	63 images from the department of orthodontics	NA	LOOCV	Expert's localization
Grau[51] (2001)	Localization of 17 cephalometric landmarks	Cephalometric radiographs	Pattern detection algorithm	Localization of significant lines; Calculation of landmark Identification areas; Calculation of landmark models	20 images	20 images	Split sample validation	Expert's localization
Rueda[52] (2006)	Localization of 43 cephalometric landmarks	Cephalometric radiographs	Active shape model; Principal component analysis	Using the template model to generate new images to fit the test image in the best way; Refinement of the match and localization of the landmarks	96 images	NA	LOOCV	Expert's localization
Sommer[75] (2009)	Measurement of 12 decision-relevant cephalometric angles	Cephalometric radiographs	An available orthodontic software (Orthometic®)	Localization of the landmarks and the soft tissue profile; Correction of the location of the landmarks (manual; optional); Calculation of the cephalometric angles	72 images from 46 female and 26 male subjects aged 5–49 years	NA	Independent sample validation	Expert's localization and measurement

(Continued)

Table 2 (Continued)

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Leonardi[53] (2009)	Localization of 10 cephalometric landmarks	Cephalometric radiographs	Cellular neural network	Assessment of relevant region in input image; Image processing by cellular neural networks templates; Landmark search by knowledge-based algorithms; Check of anatomical constraints; Output of landmark coordinates	41 images from subjects aged 10–17 years	NA	NA	Expert's localization
Vucinic[54] (2010)	Localization of 17 cephalometric landmarks	Cephalometric radiographs	Active shape model	Initial positioning of the active appearance model; Search through different resolution levels; Generation of the final convergence of the model and image	60 images from subjects aged 7.2–25.6 years	NA	LOOCV	Expert's localization
Cheng[65] (2011)	Detection of the odontoid process of the epistropheus	CBCT	Random forest	Extraction of features; Identification and localization of the odontoid process of the epistropheus	50 images	23 images	Split sample validation	Expert's localization
Shahidi[55] (2013)	Localization of 16 cephalometric landmarks	Cephalometric radiographs	Template-matching algorithm	Using the template-matching technique consisted of model-based and knowledge-based approaches to locate the landmarks	NA	40 images from a private oral and craniofacial radiology center	Independent sample validation	Expert's localization
Shahidi[58] (2014)	Localization of 14 3D cephalometric landmarks	CBCT	Feature-based and voxel similarity-based algorithms	Adaptive thresholding and conversion to binary images; Volume construction; Centroid and principal axis calculation; Scaling, rotation and transformation; Volume matching; Landmarks transferring	eight images from subjects aged 10–45 years	20 images from subjects aged 10–45 years	Independent sample validation	Expert's localization
Gupta[59] (2015)	Localization of 20 3D cephalometric landmarks	CBCT	Knowledge-based algorithm	Searching of a seed point; Find an empirical point through distance vector from the seed point; Define VOI around the empirical point; Detect a contour on the anatomical structure of VOI; Landmark detection on the contour	NA	30 images from the orthodontic treatment clinic database irrespective of age, gender and ethnicity	Independent sample validation	Expert's localization

(Continued)

Table 2 (Continued)

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Gupta[60] (2016)	Localization of 21 3D cephalometric landmarks and measurement of 51 cephalometric parameters	CBCT	Knowledge-based algorithm	Searching of a seed point; Find an empirical point through distance vector from the seed point; Define VOI around the empirical point; Detect a contour on the anatomical structure of VOI; Landmark detection on the contour; Cephalometric measurement	NA	30 images from the orthodontic treatment clinic database irrespective of age, gender and ethnicity	Independent sample validation	Expert's localization and measurement
Lindner[56] (2016)	Localization of 19 cephalometric landmarks and classification of skeletal malformations	Cephalometric radiographs	Random forest regression-voting	Object detection; Principal component analysis to the aligned shape; Regularizing the output of the individual landmark predictors; Coarse-to-fine estimate the position of landmarks; Detection of cephalometric landmark positions; Calculation of the measurements between landmark positions; Classification of skeletal malformations	400 images from 235 female and 165 male subjects aged 7–76 years	NA	4-fold CV	Expert's localization and classification
Arik[57] (2017)	Localization of 19 cephalometric landmarks and assessment of craniofacial pathologies	Cephalometric radiographs	Deep convolutional neural network	Localization of the image patch centered at landmark; Recognition of that pixel represents the landmark; Landmark location estimation; Refining the estimations; Landmark detection; Assessment of craniofacial pathologies	150 images from subjects aged 6–60 years	250 images from subjects aged 6–60 years	Split sample validation	Expert's localization and classification
Codari[61] (2017)	Localization of 21 3D cephalometric landmarks	CBCT	Adaptive cluster-based and intensity-based algorithm	Initialization; Threshold optimization; Thresholding; Registration; labeling	NA	18 images from female Caucasian subjects aged 37–74 years	Independent sample validation	Expert's localization

(Continued)

Table 2 (Continued)

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Montufar[62] (2018)	Localization of 18 3D cephalometric landmarks	CBCT	Active shape model	Computing coronal and sagittal digitally reconstructed radiographs projections; Initializing active shape model search by clicking close to sella (manual); Coronal and sagittal planes and landmark correlations; Definition of 3D landmarks in CBCT volume	24 images	NA	LOOCV	Expert's localization
Montufar[63] (2018)	Localization of 18 3D cephalometric landmarks	CBCT	Active shape model	Model-based holistic landmark search; Subvolume cropping; Knowledge-based local landmark search; Automatic 3D landmark annotation on CBCT volume voxels	24 images	NA	LOOCV	Expert's localization
Neelapu[64] (2018)	Localization of 20 3D cephalometric landmarks	CBCT	Template matching algorithm	Bone segmentation; Detection of mid sagittal plane based on symmetry; Reference landmark detection; Partitioning mid sagittal plane; VOI cropping; Extraction of contours; Detection of landmarks based on the definition on the contours	NA	30 images from the postgraduate orthodontic clinical database irrespective of age, gender and ethnicity	Independent sample validation	Expert's localization
Osteoporosis								
Allen[35] (2007)	Detection of low BMD based on MCW	Panoramic radiographs	Active shape model	Identification of four points on the mandible edge (manual); Delineation of the upper and lower bounds of the inferior mandibular cortex; Measurement of MCW	132 images from normal, osteopenic and osteoporotic female subjects aged 45–55 years	100 images from 50 normal and 50 osteoporotic female subjects aged 45–55 years	Split sample validation	DXA examination and expert's measurement
Nakamoto[42] (2008)	Diagnosis of low BMD and osteoporosis based on mandibular cortical erosion	Panoramic radiographs	Discriminant analysis	ROI selection (manual); Adjustment of the image position; Extraction of the morphological skeleton; Classification of normal cortex and eroded cortex	100 images from normal, low BMD and osteoporotic female subjects aged ≥50 years	100 images from normal, low BMD and osteoporotic female subjects aged ≥50 years	Split sample validation	DXA examination

(Continued)

Table 2 (Continued)

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Kavitha[39] (2012)	Diagnosis of osteoporosis based on MCW	Panoramic radiographs	SVM	ROI selection (manual); Image enhancement; Identification of cortical margins; Measurement of MCW; Classification of normal and osteoporotic subjects	60 images from normal and osteoporotic female subjects aged ≥50 years	40 images from normal and osteoporotic female subjects aged ≥50 years	Split sample validation	DXA examination and expert's measurement
Roberts[43] (2013)	Diagnosis of osteoporosis based on cortical texture and MCW	Panoramic radiographs	Random forest	Identification of the mandibular cortical margins; Image normalization; Extraction of texture features and/or measurement of MCW; Classification of normal and osteoporotic subjects	663 images from 523 normal and 140 osteoporotic female subjects	NA	Out-of-bag estimation	DXA examination and expert's measurement
Muramatsu[41] (2013)	Diagnosis of osteoporosis based on MCW	Panoramic radiographs	Active shape model	Detection of mandibular edges; Selection of a reference contour model and fitting of the model; Measurement of MCW;	100 images from 74 normal and 26 osteoporotic male and female subjects	NA	LOOCV	DXA examination and expert's measurement
Kavitha[38] (2013)	Diagnosis of osteoporosis based on MCW	Panoramic radiographs	Back propagation neural network	ROI selection (manual); Segmentation of cortical margins; Measurement of MCW; Classification of normal and osteoporotic subjects	100 images from normal and osteoporotic female subjects aged ≥50 years	NA	5-fold CV	DXA examination and expert's measurement
Kavitha[37] (2015)	Diagnosis of osteoporosis based on textural features and MCW	Panoramic radiographs	Naïve Bayes; k-NN; SVM	ROI selection (manual); Segmentation cortical margins; Evaluation of eroded cortex; Measurement of MCW; Analysis of textual features; Classification of normal and osteoporotic subjects	141 images from normal and osteoporotic female subjects aged 45–92 years	NA	LOOCV/5-fold CV	DXA examination and expert's measurement
Kavitha[40] (2016)	Diagnosis of osteoporosis based on attributes of the mandibular cortical and trabecular bones	Panoramic radiographs	NN	Extraction of attributes based on mandibular cortical and trabecular bones; Analysis of the significance of the extracted attributes; Generation of classifier for screening osteoporosis; Classification of normal and osteoporotic subjects	141 images from normal and osteoporotic female subjects aged 45–92 years	NA	5-fold CV	DXA examination
Hwang[36] (2017)	Diagnosis of osteoporosis based on strut analysis	Panoramic radiographs	Decision tree; SVM	ROIs selection (manual); Imaging processing; Analysis of texture features; Classification of normal and osteoporotic subjects	454 images from 227 normal and 227 osteoporotic male and female subjects	NA	10-fold CV	DXA examination

(Continued)

Table 2 (Continued)

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Maxillofacial cysts and tumors								
Mikulka[33] (2013)	Classification of follicular cysts and radicular cysts	Panoramic radiographs	Decision tree; Naïve Bayes; Neural network; k-NN; SVM; LDA	Cyst identification (manual); Cyst segmentation; Extraction of texture features; Classification of the jawbone cysts	26 images from subjects with 13 follicular cysts and 13 radicular cysts	NA	10-fold CV	Expert's judgement
Nurtanio[34] (2013)	Classification of various maxillofacial cysts and tumors	Panoramic radiographs	SVM	Lesion segmentation (manual); Extraction of texture features; Classification of the lesions	133 images from subjects with various cysts and tumors	NA	3-fold CV	Expert's judgement
Rana[76] (2015)	Segmentation and measurement of keratocysts	3D images (MRI/ CT)	An available navigation software (Brainlab)	Identification of keratocysts (manual); Lesion segmentation; Measurement of the lesion volume	NA	38 images from subjects with keratocysts	Independent sample validation	Expert's segmentation and measurement
Abdolali[67] (2016)	Segmentation of maxillofacial cysts	CBCT	Asymmetry analysis	Image registration; Detection of asymmetry; Segmentation of the cysts	97 images from subjects with 39 radicular cysts, 36 dentigerous cysts and 22 keratocysts	NA	NA	Expert's segmentation
Yilmaz[69] (2017)	Classification of periapical cysts and keratocysts	CBCT	k-NN; Naïve Bayes; Decision tree; Random forest; NN; SVM	Lesion detection and segmentation (manual); Extraction of texture features; Classification of periapical cysts and keratocysts	50 images from subjects with 25 cysts and 25 tumors	NA	10-fold CV/ LOOCV	Expert's judgement, radiological and histopathologic examinations
					25 images from subjects with cysts/tumors	25 images from subjects with cysts/tumors	Split sample validation	
Abdolali[68] (2017)	Classification of radicular cysts, dentigerous cysts and keratocysts	CBCT	SVM; SDA	Lesion segmentation; Extraction of texture features; Classification of lesions	96 images from patients with 38 radicular cysts, 36 dentigerous cysts and 22 keratocysts	NA	3-fold CV	Histopathological examination
Alveolar bone resorption								
Lin[27] (2015)	Identification of alveolar bone loss area	Periapical radiographs	Naïve Bayes; k-NN; SVM	ROI identification (manual); Fusion of texture features; Coarse segmentation of the bone loss area; Fine segmentation of the bone loss area	28 images from subjects with periodontitis	three images from subjects with periodontitis	LOOCV/ Split sample validation	Expert's judgement
Lin[28] (2017)	Measurement of alveolar bone loss degree	Periapical radiographs	Naïve Bayes	Teeth segmentation; Identification of the bone loss area; Identification of the landmarks; Measurement of the bone loss degree	18 images from subjects with periodontitis	NA	LOOCV	Expert's judgement

(Continued)

Table 2 (Continued)

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Lee[29] (2018)	Identification and prediction of periodontally compromised teeth	Periapical radiographs	Deep convolutional neural network	Image augmentation; Extraction of texture features; Classification of the healthy and periodontally compromised premolars/molars	1392 images exhibiting healthy/periodontally compromised premolars and molars	348 images exhibiting healthy/periodontally compromised premolars and molars	Split sample validation	Clinical and radiological examinations
Periapical disease								
Mol[31] (1992)	Classification of periapical condition based on the lesion range	Mandibular periapical radiographs	Rule-based algorithm	Apex identification (manual); Analysis of texture features; Segmentation of the periapical lesion; Classification of the periapical lesion	NA	111 images exhibiting 45 healthy and 66 pathological mandibular teeth	Independent sample validation	Expert's judgement
Carmody[30] (2001)	Classification of periapical condition into no lesion and mild, moderate and severe lesion	Periapical radiographs	A machine learning classifier	Apex identification (manual); Segmentation of the periapical lesion; Classification of the periapical lesion	32 images exhibiting four different periapical conditions	NA	LOOCV	Expert's judgement
Flores[66] (2009)	Classification of periapical cysts and granuloma	CBCT	LDA; AdaBoost	Lesion segmentation (manual); Extraction of texture features; Classification of the periapical cyst and granuloma	17 images exhibiting periapical cysts or granuloma	NA	LOOCV	CBCT examination and histopathologic examinations
Multiple dental diseases								
Ngan[74] (2016)	Diagnosis of cracked dental root, incluse teeth, decay, hypoodontia and periodontal bone resorption	Dental X-ray images	Affinity propagation clustering	Extraction of dental features; Image segmentation; Extraction of features of the segments; Determination of diseases of the segments; Synthesis of the segments; Classification of diseases	NA	66 images exhibiting cracked dental root, impacted teeth, decay, hypoodontia or periodontal bone resorption	Split sample validation	Expert's judgement
Son[47] (2018)	Diagnosis of root fracture, incluse teeth, decay, missing teeth and periodontal bone resorption	Intraoral and panoramic radiographs	Affinity propagation clustering	Extraction of dental features; Image segmentation; Extraction of features of the segments; Determination of diseases of the segments; Synthesis of the segments; Classification of diseases	87 images exhibiting 16 root fracture, 19 incluse teeth, 17 decay, 16 loss of teeth and 19 periodontal bone resorption	NA	10-fold CV	Expert's judgement
Tooth types								
Miki[70] (2017)	Classification of tooth types	CBCT	Deep convolutional neural network	ROI selection (manual); Image resizing; Classification of tooth types	42 images	ten images	Split sample validation	Ground truth

(Continued)

Table 2 (Continued)

Author (year)	Application	Imaging modality	AI technique	Workflow of AI model	Data set used to develop the AI model	Independent testing data set	Validation technique	Reference standard
Tuzoff[46] (2019)	Tooth detection and numbering	Panoramic radiographs	Deep convolutional neural network	Define the boundaries of each tooth; Crop the image based on the predicted bounding boxes; Classify each cropped region; Output the bounding boxes coordinates and corresponding teeth numbers	1352 images	222 images	Split sample validation	Expert's judgement
Others								
Benyó[71] (2012)	Identification of root canal	CBCT	Decision tree	Classification of tooth type (manual); Image segmentation; Reconstruction of the shape of the tooth and root canal; Extraction of the medial line of the root canal	NA	36 images	Independent sample validation	Expert's judgement
Ohashi[45] (2016)	Detection of the maxillary sinusitis	Panoramic radiographs	Asymmetry analysis	Edge extraction; Image registration; Detection of maxillary sinusitis; Decision making (manual)	NA	98 images from 49 subjects with maxillary sinusitis and 49 subjects with healthy sinuses	Independent sample validation	Clinical symptom and CT images
Rana[73] (2017)	Identification of inflamed gingiva	Intraoral fluorescent images	A machine learning classifier	Segmentation of lesions; Classification of inflamed and non-inflamed gingiva	258 images from subjects with gingivitis	147 images from subjects with gingivitis	Split sample validation	Expert's judgement
Yauney[72] (2017)	Identification of dental plaque	Intraoral photographs	Convolutional neural network	Segmentation of plaque; Classification of plaque presence/absence	33 images taken from CD and 35 images taken from RD	14 images taken from CD and 14 images taken from RD	Split sample validation	Expert's judgement and biomarker label
De Tobel[44] (2017)	Classification of the stages of the lower third molar development	Panoramic radiographs	Deep convolutional neural network	ROI selection (manual); Feature extraction; Classification of the stage of the third molar	400 images consisted of 20 images per stage per sex	NA	5-fold CV	Expert's judgement based on a modified Demirjian's staging technique
Lee[32] (2018)	Detection of dental caries	Periapical radiographs	Deep convolutional neural networks	ROI selection (manual); Feature segmentation; lesion detection; Diagnosis of dental caries	2400 images exhibiting 1200 caries and 1200 non-caries in the maxillary premolars/molars	600 images exhibiting 300 caries and 300 non-caries in the maxillary premolars/molars	Split sample validation	Expert's judgement

AUC, area under the receiver operating characteristic curve; BMD, bone mineral density; CBCT, cone-beam CT; CD, commercial device; CV, cross-validation; DXA, dual-energy X-ray absorptiometry; LDA, linear discriminant analysis; LOOCV, leave-one-out cross-validation; MCW, mandibular cortical width; NA, not available; NN, neural network; PCT, periodontally compromised teeth; RD, research device; ROI, region of interest; SDA, sparse discriminant analysis; SVM, support vector machine; VOI, volume of interest; k-NN, k-nearest neighbors.

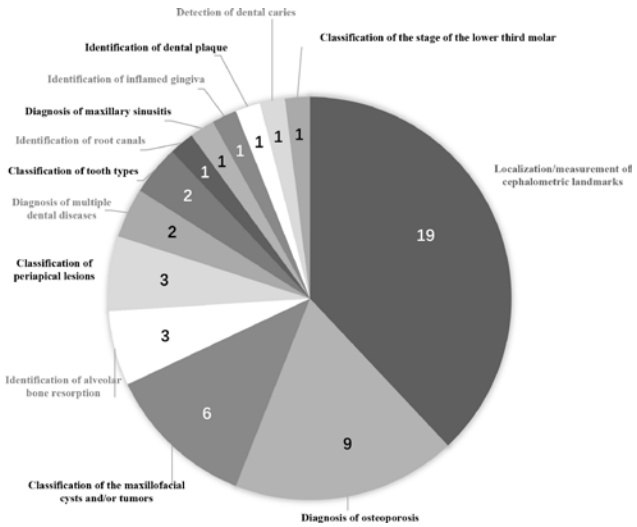

Figure 3 Pie chart for the clinical applications of AI proposed in the included studies

the monitor. Additionally, cephalometric analyses on digital tracing software are still prone to manual errors due to the deviation of landmark localization between different observers.[77] To overcome these shortcomings, an automated approach based on AI techniques was proposed and has been improved gradually by using different algorithms.[77] 19 studies[48–65,75] in this review reported the accuracy of AI models developed for cephalometric analysis (Table 3). In these proposed models, the number of localizable landmarks varies from 10 to 43 (Table 2). Most of these studies reported the localization deviation for individual landmarks and the mean deviation for all landmarks. According to these studies, the mean localization deviations in these AI models ranged from 1.1 to 4.09 mm (Table 3). Automatic landmark localization is considered to be successful if the difference between the location of a landmark localized by a model and by an expert's annotation is ≤2 mm.[77] Only half of these studies reported on successful rates ranging from 35 to 84.70% in accordance with the aforementioned standard (Table 3).

Since Cheng et al[65] proposed an AI model to automatically localize one key landmark on CBCT images in 2011, cephalometric radiographs were gradually replaced with CBCT images to develop models for cephalometric analysis. Cephalometric analysis on CBCT images is considered as a more versatile approach capable of performing 3D measurements, but the automatic localization performance on the existing models is still not accurate enough to meet clinical requirements. Therefore, the existing models/software can be recommended for the use of a preliminary localization of the cephalometric landmarks, but a manual correction is still necessary prior to further cephalometric analyses.

AI applications for the diagnosis of osteoporosis

The diagnosis of low bone mineral density (BMD) and osteoporosis is also considered as a potential area for AI applications. Both conditions mentioned are discussed as relevant diagnostic findings in dental medicine including implant dentistry.[78,79]Patients with osteoporosis are more susceptible to have marginal bone loss around dental implants,[78] and those treated with antiresorptive medications are more at risk for osteonecrosis of the jaws following oral surgery.[79]Nine studies in this review proposed AI models to classify normal and osteoporotic subjects on panoramic radiographs based on reduction of mandibular cortical width and erosion of mandibular cortex that are both relevant to low skeletal BMD, high bone turnover rate and a higher risk of osteoporotic fracture.[80,81] The diagnostic performance of these models has been gradually increasing due to multiple improvements. In recent studies, the accuracy, sensitivity, and specificity all reached 95% and above demonstrating that these models might be used in clinical practice in the near future (Table 4).

AI application to classify/segment maxillofacial cysts and/or tumors

Accurate segmentation and diagnosis of various maxillofacial cysts and/or tumors are challenges to general dental practitioners. In some complicated cases, even radiologists can only provide tentative diagnoses, and have to refer patients for a biopsy examination to reach a final diagnosis. Therefore, the application of AI for automated diagnosis of various jaw cysts and/or tumors will be of great value in clinical practice. In this review, six studies reporting on automatic segmentation/classification of various maxillofacial cysts and/or tumors have been included. Abdolali et al[67] proposed a model based on asymmetry analysis to automatically segment radicular cysts, dentigerous cysts and keratocysts. Rana et al[76] used an available surgical navigation software (iPlan, Brainlab AG, Feldkirchen Germany) to automatically segment keratocysts and measure their volume. The remaining four studies[33,34,68,69] proposed AI models trained with 2D/3D images to classify various maxillofacial cysts and/or tumors (Table 2). Technically, the procedure of an AI model to classify cysts and/or tumors follows four main steps that are lesion detection, segmentation, extraction of texture features and subsequent classification. Currently, the first step of lesion detection in these models is still required to be performed manually so that these models can automatically perform the following steps. It still remains a challenge to develop a fully automated model that can identify cysts and/or tumors.

AI applications to identify periodontitis/periapical disease

The features of alveolar bone resorption and periapical radiolucency can both contribute to the development of AI models for diagnosis of periodontitis and periapical

Table 3 List of the performance reported in the studies included regarding automatic localization of cephalometric landmarks

	Automatic localization			*Manual localization*	
Author (year)	Mean deviation (SD) or median deviation (IQR) on all landmarks/ measurements	Range of the deviation on all landmarks/measurements	Successful localization rates based on different precision ranges	Mean deviation (SD) on all landmarks/measurements	Range of the deviation on all landmarks/measurements
Rudolph[48] (1998)	3.07 mm (3.09 mm)	1.85 mm-5.67 mm	NA	3.09 mm (4.76 mm)	1.99 mm-4.93 mm
Liu[49] (2000)	2.86 mm (1.24 mm)	0.94 mm-5.28 mm	NA	1.48 mm (0.59 mm)	0.63 mm-2.57 mm
Hutton[50] (2000)	4.09 mm	2.7 mm-7.3 mm	13% (deviation <1 mm); 35% (deviation <2 mm); 74% (deviation <5 mm)	NA	NA
Grau[51] (2001)	1.1 mm	0.39 mm-1.92 mm	NA	NA	NA
Rueda[52] (2006)	2.48 mm (1.66 mm)	1.52 mm-3.88 mm	50.04% (deviation <2 mm); 72.62% (deviation <3 mm); 91.44% (deviation <5 mm)	NA	NA
Sommer[75] (2009)	1.26°/semi-automatic 8.00°/fully automatic	0.53°−3.35°/semi-automatic 2.83°−15.00°/fully automatic	NA	NA	NA
Leonardi[53] (2009)	NA	0.003 mm-0.596 mm	NA	NA	NA
Vucinic[54] (2010)	1.68 mm	1.07 mm-3.27 mm	28% (deviation <1 mm); 61% (deviation <2 mm); 95% (deviation <5 mm)	0.49 mm	0.14 mm-1.07 mm
Cheng[65] (2011)	3.15 mm	0.84 mm-6.27 mm	NA	NA	NA
Shahidi[55] (2013)	2.59 mm (3.45 mm)	1.71 mm-4.44 mm	12.5% (deviation <1 mm); 43.75% (deviation <2 mm); 93.75% (deviation <4 mm); 100% (deviation <5 mm)	NA	NA
Shahidi[58] (2014)	3.40 mm	3.00 mm-3.86 mm	63.57% (deviation <3 mm)	1.41 mm	0.59 mm-2.15 mm
Gupta[59] (2015)	2.01 mm (1.23 mm)	1.17 mm-3.20 mm	64.67% (deviation <2 mm); 82.67% (deviation <3 mm); 90.33% (deviation <4 mm)	NA	NA
Gupta[60] (2016)	1.60 mm /3 landmarks 1.28 mm /28 linear measurements 0.94°/16 angular measurements	1.52 mm-1.70 mm /3 landmarks 0.46 mm-2.63 mm /28 linear measurements 0.37°−2.12°/16 angular measurements	NA	NA	NA
Lindner[56] (2016)	1.20 mm (0.06 mm)	0.65 mm-2.69 mm	84.70% (deviation <2 mm); 89.38% (deviation <2.5 mm); 92.62% (deviation <3 mm) 96.30% (deviation <4 mm)	NA	NA

(Continued)

Table 3 (Continued)

Author (year)	*Automatic localization*			*Manual localization*	
	Mean deviation (SD) or median deviation (IQR) on all landmarks/measurements	Range of the deviation on all landmarks/measurements	Successful localization rates based on different precision ranges	Mean deviation (SD) on all landmarks/measurements	Range of the deviation on all landmarks/measurements
Arik[57] (2017)	NA	NA	67.68–75.58% (deviation <2 mm); 74.16–81.26% (deviation <2.5 mm); 79.11–84.68% (deviation <3 mm) 84.63–88.25% (deviation <4 mm)	NA	NA
Codari[61] (2017)	1.99 mm (1.22 mm, 2.89 mm)	1.40 mm-4.00 mm	63% (deviation <2.5 mm); 90% (deviation <5 mm)	NA	NA
Montufar[62] (2018)	3.64 mm (1.43 mm)	1.62 mm-9.72 mm	NA	NA	NA
Montufar[63] (2018)	2.51 mm (1.60 mm)	1.46 mm-3.72 mm	NA	NA	NA
Neelapu[64] (2018)	1.88 mm (1.10 mm)	0.95 mm-3.78 mm	63.53% (deviation <2 mm); 85.29% (deviation <3 mm); 93.92% (deviation <4 mm)	NA	NA

IQR, interquartile range; NA, not available; SD, standard deviation.

disease. Lin et al proposed two models to identify alveolar bone loss[27] and to measure the degree of bone loss,[28] respectively. Lee et al[29] proposed a model based on deep learning convolutional neural network to identify periodontally compromised premolars and molars, and predict hopeless premolars and molars according to the degree of alveolar bone loss. Regarding the diagnosis of periapical disease, Mol et al[31] and Carmody et al[30] proposed models to classify the extent of periapical lesions. Moreover, Flores et al[66] proposed a model using CBCT images to distinguish periapical cysts from granuloma, which is considered as having high value in clinical practice as periapical granuloma can heal after root canal treatment without surgical intervention.

AI application for the detection of dental caries

Dental caries is a common oral disease, which can be prevented through early detection and treatment.[82] Although only one study in this review proposed a caries detection model developed by using clinical X-ray images,[32] a large number of studies have attempted to develop caries detection models by using non-clinical 2D images obtained from extracted teeth.[24,83–86] The diagnostic performance of the models in these pre-clinical studies exhibited satisfactory results, but these results might be overly optimistic due to the training-testing images presenting obvious lesions from extracted teeth and lacking the image of other oral tissues.[24,84–86]Lee et al[32] proposed a caries detection model based on deep learning algorithms to detect caries in maxillary premolars and molars. The model exhibited high diagnostic performance for both maxillary premolars and molars. However, this caries detection model still has limitations. Because this model was developed using 2D images, it might only detect proximal and occlusal caries, and might not be able to detect buccal and lingual caries. Moreover, this model was trained on images of aligned teeth without restorative materials. Hence, it is still unknown if this model could detect secondary caries or caries on the overlapping teeth. Finally, training image data used in this model only included images of permanent premolars and molars so that its applicability on deciduous teeth remains unknown.

AI applications for other purposes

The remaining studies in this review developed AI models addressing the diagnosis of maxillary sinusitis,[45] the classification of the stages of the lower third molar development[44] and tooth types,[46,70] the identification of root canals orifices,[71] dental plaque[72] and inflamed gingiva,[73] and the diagnosis of multiple dental diseases.[47,74] This further demonstrates that AI techniques are now being explored more and more broadly for various fields in DMFR.

Data collection strategy in the assessed studies

In order to create a high-quality diagnostic AI model, it is imperative to have a strategy to collect image data

Table 4 List of the diagnostic performance reported in the studies included regarding the diagnosis of osteoporosis

Author (year)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	Likelihood ratios (95% CI)	AUC (95% CI)
Allen[35] (2007)	NA	NA	NA	NA	NA	NA	0.81/automatic measurement; 0.61–0.68/manual measurement
Nakamoto[42] (2008)				**Low BMD**			
	74.0%	76.8%	61.1%	90.0% (84.1%, 95.9%)	36.7% (27.3%, 46.1%)	1.98 (1.09, 3.57)	NA
				Osteoporosis			
	62.0%	94.4%	43.8%	48.6% (38.8%, 58.4%)	93.9% (88.4%, 100%)	1.68 (1.33, 2.11)	NA
Kavitha[39] (2012)				**Identifying site: lumbar spine**			
	88% (81.6%, 94.4%)	90.9% (85.3%, 96.5%)	83.8% (76.6%, 91.0%)	71.4% (62.5%, 80.3%)	96.7% (93.8%, 99.6%)	5.5 (4.5, 6.5)	NA
				Identifying site: femoral neck			
	75.0% (66.5%, 83.5%)	90.0% (84.1%, 95.9%)	69.6% (60.1%, 78.6%)	46.6% (38.8%, 56.4%)	96.0% (92.1%, 99.2%)	3.1 (2.2, 4.0)	NA
				Identifying site: combination of lumbar spine and lumbar spine			
	79.0% (71.0%, 86.9%)	90.6% (84.9%, 96.3%)	80.9% (73.2%, 88.6%)	61.2% (51.7%, 70.8%)	96.6% (92.0%, 100%)	4.1 (3.3, 5.1)	NA
Roberts[43] (2013)				**Identifying site: lumbar spine**			
	NA	NA	NA	NA	NA	NA	0.802/based on MCW; 0.730/based on cortical texture; 0.800/based on MCW and texture
				Identifying site: femoral neck			
	NA	NA	NA	NA	NA	NA	0.830/based on MCW; 0.824/based on cortical texture; 0.872/based on MCW and texture
Muramatsu[41] (2013)	NA	88.5%	97.30%	NA	NA	NA	0.96/automatic measurement 0.98/manual measurement
Kavitha[38] (2013)				**Identifying site: lumbar spine**			
	93.0% (88.0%, 98.0%) /SVM; 91.0% (85.4%, 96.6%) /NN	95.8% (91.9%, 99.7%) /SVM; 93.3% (88.0%, 98.0%) /NN	86.6% (79.9%, 93.3%) /SVM; 83.2% (75.6%, 90.4%) /NN	76.8% (68.5%, 5.1%) /SVM; 75.5% (66.5%, 3.4%) /NN	98.5% (96.1%, 100%) /SVM; 98.2% (95.0%, 100%) /NN	7.2 (6.3, 8.1) /SVM; 6.9 (5.9, 7.8) /NN	0.871 (0.804, 0.936)
				Identifying site: femoral neck			
	89.0% (82.9%, 95.1%) /SVM; 86.0% (79.2%, 92.8%) /NN	96.0% (92.2%, 99.8%) /SVM; 93.8% (89.0%, 98.0%) /NN	84.0% (76.8%, 91.2%) /SVM; 82.0% (74.5%, 89.5%) /NN	66.7% (57.5%, 5.9%) /SVM; 66.3% (56.7%, 5.3%) /NN	98.4% (95.9%, 100%) /SVM; 98.0% (95.3%, 100%) /NN	6.0 (5.4, 6.9) /SVM; 5.9 (4.9, 6.8) /NN	0.886 (0.816, 0.944)
Kavitha[37] (2015)				**Identifying site: lumbar spine**			
	91.5%	94.0%	82.8%	NA	NA	NA	0.922 (0.878, 0.966)
				Identifying site: femoral neck			
	93.5%	96.1%	84.7%	NA	NA	NA	0.947 (0.910, 0.984)

(Continued)

Table 4 (Continued)

Author (year)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	Likelihood ratios (95% CI)	AUC (95% CI)
Kavitha[40] (2016)				**Identifying site: lumbar spine**			
	96.0% (90.3%, 95.9%)	95.3% (76.2%, 99.8%)	94.7% (90.5%, 98.6%)	80.0% (59.3%, 93.2%)	98.1% (95.3%, 99.9%)	22.8 (9.6, 54.2)	0.962 (0.912, 0.983)
				Identifying site: femoral neck			
	98.9% (92.0%, 100%)	99.1% (83.2%, 100%)	98.4% (94.2%, 99.8%)	91.2% (70.8%, 98.9%)	99.4% (96.7%, 100%)	60.5 (15, 239)	0.986 (0.942, 0.998)
Hwang[36] (2017)	96.2% / Decision tree; 96.6% /SVM	97.1% / Decision tree; 97.2% /SVM	95.7% / Decision tree; 97.1% /SVM	NA	NA	NA	NA

AUC, area under the receiver operating characteristic curve; BMD, bone mineral density; CI, confidence interval; MCW, mandibular cortical width; NA, not available; NN, neural network; SVM, support vector machine.

prior to model development. Ideally, it is suggested that images used for developing a model should consist of heterogeneous images collected from multiple institutions in different time periods.[87] In addition, both images from subjects with and subjects without the condition of interest should be used. However, all studies in this review only used image data collected from the same institution in one time period, and some classification models were only trained and tested with images from subjects with confirmed diseases.[27,28,47,72–74] These might result in a lack of generalizability and a risk of overfitting.[87]

Model validation techniques used in the assessed studies

The main steps to develop an AI model consist of training a model, tuning the hyperparameters and evaluating the performance of the model.[7,87] Split sample validation and k-fold cross-validation (CV) are two common techniques used to develop AI models.[88] Split validation is recommended when a large number of data are available, and the whole data will be randomly split into three data sets. K-fold CV is more applicable to small to medium-sized data sets. This technique has a parameter named k, and when a specific value for k is chosen, the whole data are to be split into k data sets. One of these data sets is used to validate the model while data in the remaining (k-1) sets serve as training data. This process will be repeated k times until each data set has been served once as a validation set. Afterwards, the performance of the model will be evaluated on the average of each of the separate results obtained. There is a particular situation when the k value equals the number of the whole data (n), which is called leave-one-out cross-validation (LOOCV) and is mainly performed on small-sized data sets.

The number of data needed to train an AI model depends on many factors, such as input features, algorithm type, number of algorithm parameters, etc.[89] Theoretically, it is suggested that 10 times the number of an algorithm's parameters is a reasonable estimate of the data needed to train a model.[89] In this review, 40% of the studies included used less than 50 images to develop their models, 68% used less than 100 images, and 90% used less than 500. In general, most studies used a small number of images to develop their models. This limited number of images is not suitable to be further split into subsets so that k-fold CV techniques are frequently performed in these studies.

Limitations and future outlook

As the present systematic review focused specifically on the use of AI in DMFR and diagnostic imaging for various purposes in dental medicine, some studies that were using common oral radiology methodologies for AI applications but with no direct relevance for dental medicine were excluded following discussion between the observers. For example, out of such articles excluded in the present study, one evaluated carotid artery calcifications on panoramic radiographs.[89] This focused selection could have introduced a certain bias to the present systematic review. Future reviews that address specific topics in more detail and critically assess their impact and clinical potential are needed.

Despite the promising performance of the AI models described, it is still necessary to verify the generalizability and reliability of these models by using adequate external data that are obtained in newly-recruited patients or collected from other dental institutions.[87] On the other hand, deep learning that is regarded as a more advanced AI technique and is widely used to develop diagnostic AI models in the field of clinical medicine should also be used to expand clinical applications of AI in DMFR.[6] The future goals of AI development in DMFR can not only be expected to improve the performance of AI models on par with experts, but also to detect early lesions that cannot be seen by human eyes.

Conclusion

In summary, AI models described in the studies included exhibited various potential applications for DMFR,

Table 5 List of the diagnostic performance reported in the remaining studies

Author (year)	Accuracy	Sensitivity	Specificity	Mean deviation from reference	AUC	Correlation against reference
Alveolar bone resorption						
Lin[27] (2015)	NA	92.5% /LOOCV; 92.8% /independent sample	86.2% /LOOCV; 85.9% /independent sample	NA	NA	NA
Lin[28] (2017)	NA	NA	NA	9.5%	NA	NA
Lee[29] (2018)	**Identification of PCT** 81.0% /premolar; 76.7% /molar **Prediction of hopeless teeth** 82.8% /premolar; 73.4% /molar	NA	NA	NA	**Prediction of hopeless teeth** 0.826/premolar; 0.734/molar	NA
Periapical lesions						
Mol[31] (1992)	80.2%	83.3%	75.6%	NA	NA	0.67
Carmody[30] (2001)	83.40% /automatic classification; 57.8% /manual classification	NA	NA	NA	NA	0.78/automatic classification; 0.44/manual classification
Flores[66] (2009)	94.1% /reference: CBCT; 88.2% /reference: biopsy	NA	NA	NA	NA	NA
Maxillofacial cysts and tumors						
Mikulka[33] (2013)	85% /follicular and radicular cysts; 81.8% /follicular cysts; 88.9% /radicular cysts	NA	NA	NA	NA	NA
Nurtanio[34] (2013)	87.18%	NA	NA	NA	0.944	NA
Rana[76] (2015)	NA	NA	NA	No significant difference	NA	NA
Abdolali[67] (2016)	NA	NA	NA	NA	0.936	0.83/radicular cysts; 0.87/dentigerous cysts; 0.80/keratocysts
Yilmaz[69] (2017)	100% /10-fold CV; 94% /LOOCV; 96% /split sample	NA	NA	NA	NA	NA
Abdolali[68] (2017)	94.29% /SVM; 96.48% /SDA	NA	NA	NA	NA	NA
Decision making model						
Ngan[74] (2016)	93.02%	NA	NA	NA	NA	NA
Son[47] (2018)	92.74%	NA	NA	NA	NA	NA
Tooth types						
Miki[70] (2017)	88.8%	NA	NA	NA	NA	NA

(Continued)

Table 5 (Continued)

Author (year)	Accuracy	Sensitivity	Specificity	Mean deviation from reference	AUC	Correlation against reference
Tuzoff[46] (2019)	99.87% /teeth numbering	99.94% /teeth detection 98.00% /teeth numbering	99.94% /teeth numbering	NA	NA	NA
Identification of root canals						
Benyó[71] (2012)	91.70%	NA	NA	NA	NA	NA
Detection of maxillary sinusitis						
Ohashi[45] (2016)			**Diagnostic performance of the AI model**			
	73.50%	77.60%	69.40%	NA	NA	NA
			Change of diagnostic performance after the use of the AI model			
	Before *After*	*Before* *After*	*Before* *After*	NA	*Before* *After*	NA
	66.0% /IEDs; 73.4% /IEDs; 79.9% /experts 81.1% /experts	63.4% /IEDs; 71.6% /IEDs; 74.5% /experts 76.0% /experts	68.6% /IEDs; 75.30% /IEDs; 85.2% /experts 86.2% /experts		0.728/IEDs; 0.780/IEDs; 0.871/experts 0.897/experts	
Identification of inflamed gingiva						
Rana[73] (2017)	NA	NA	NA	NA	0.746	NA
Identification of dental plaque						
Yauney[72] (2017)	84.67% /RD; 87.18/CD	NA	NA	NA	0.769/RD; 0.872/CD	NA
Classification of the stages of the lower third molar development						
De Tobel[44] (2017)	51%	NA	NA	0.6 stages	NA	NA
Detection of dental caries						
Lee[32] (2018)	89.0% /premolar; 88.0% /molar; 82.0% /both	84.0% /premolar; 92.3% /molar; 81.0% /both	94.0% /premolar; 84.0% /molar; 83.0% /both	NA	0.917/premolar; 0.890/molar; 0.845/both	NA

AUC, area under the receiver operating characteristic curve; CBCT, cone-beam CT; CD, commercial device; CV, cross-validation; IED, inexperienced dentist; LOOCV, leave-one-out cross-validation; NA, not available; NN, neural network; PCT, periodontally compromised teeth; RD, research device; SDA, sparse discriminant analysis; SVM, support vector machine.

which were mainly focusing on automated localization of cephalometric landmarks, diagnosis of osteoporosis, classification/segmentation of maxillofacial cysts and/or tumors, and identification of periodontitis/periapical disease. Future systematic reviews should focus in more detail on these specific areas in dental medicine to describe and assess the value and impact of AI in daily practice. The diagnostic performance of the AI models varies among different algorithms used, and it is still necessary to verify the generalizability and reliability of these models by using adequate, representative images from multiple institutions prior to transferring and implementing these models into clinical practice.

Acknowledgment

The authors are grateful to Mr Sam Lee, Dental Library, Faculty of Dentistry, The University of Hong Kong, for his advice regarding the study search strategy. This study was funded by departmental funds alone. All authors report no conflict of interest in connection with the submitted manuscript.

References

1. Wong SH, Al-Hasani H, Alam Z, Alam A. Artificial intelligence in radiology: how will we be affected? *Eur Radiol* 2019; **29**: 141–3. doi: https://doi.org/10.1007/s00330-018-5644-3
2. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg* 2018; **268**: 70–6. doi: https://doi.org/10.1097/SLA.0000000000002693
3. Stone P, Brooks R, Brynjolfsson E, Calo R, Etzioni O, Hager Get al. *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*. Stanford, CA: Stanford University; 2016. http://ai100stanfordedu/2016-report.
4. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017; **2**: 230–43. doi: https://doi.org/10.1136/svn-2017-000101
5. Fazal MI, Patel ME, Tye J, Gupta Y. The past, present and future role of artificial intelligence in imaging. *Eur J Radiol* 2018; **105**: 246–50. doi: https://doi.org/10.1016/j.ejrad.2018.06.020
6. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017; **42**: 60–88. doi: https://doi.org/10.1016/j.media.2017.07.005
7. Thrall JH, Li X, Li Q, Cruz C, Do S, Dreyer K, et al. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. *J Am Coll Radiol* 2018; **15**(3 Pt B): 504–8. doi: https://doi.org/10.1016/j.jacr.2017.12.026
8. Kohli M, Prevedello LM, Filice RW, Geis JR. Implementing machine learning in radiology practice and research. *AJR Am J Roentgenol* 2017; **208**: 754–60. doi: https://doi.org/10.2214/AJR.16.17224
9. Liu Y, Balagurunathan Y, Atwater T, Antic S, Li Q, Walker RC, et al. Radiological image traits predictive of cancer status in pulmonary nodules. *Clin Cancer Res* 2017; **23**: 1442–9. doi: https://doi.org/10.1158/1078-0432.CCR-15-3102
10. Ye X, Lin X, Dehmeshki J, Slabaugh G, Beddoe G. Shape-Based computer-aided detection of lung nodules in thoracic CT images. *IEEE Trans Biomed Eng* 2009; **56**: 1810–20.
11. Wang S, Yao J, Summers RM. Improved classifier for computer-aided polyp detection in CT colonography by nonlinear dimensionality reduction. *Med Phys* 2008; **35**: 1377–86. doi: https://doi.org/10.1118/1.2870218
12. Arimura H, Li Q, Korogi Y, Hirai T, Katsuragawa S, Yamashita Y, et al. Computerized detection of intracranial aneurysms for three-dimensional Mr angiography: feature extraction of small protrusions based on a shape-based difference image technique. *Med Phys* 2006; **33**: 394–401. doi: https://doi.org/10.1118/1.2163389
13. Kwak JT, Xu S, Wood BJ, Turkbey B, Choyke PL, Pinto PA, et al. Automated prostate cancer detection using T2-weighted and high-b-value diffusion-weighted magnetic resonance imaging. *Med Phys* 2015; **42**: 2368–78. doi: https://doi.org/10.1118/1.4918318
14. Le MH, Chen J, Wang L, Wang Z, Liu W, Cheng K-TT, et al. Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks. *Phys Med Biol* 2017; **62**: 6497–514. doi: https://doi.org/10.1088/1361-6560/aa7731
15. Schuhbaeck A, Otaki Y, Achenbach S, Schneider C, Slomka P, Berman DS, et al. Coronary calcium scoring from contrast coronary CT angiography using a semiautomated standardized method. *J Cardiovasc Comput Tomogr* 2015; **9**: 446–53. doi: https://doi.org/10.1016/j.jcct.2015.06.001
16. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115–8. doi: https://doi.org/10.1038/nature21056
17. Shah SK, McNitt-Gray MF, Rogers SR, Goldin JG, Suh RD, Sayre JW, et al. Computer aided characterization of the solitary pulmonary nodule using volumetric and contrast enhancement features. *Acad Radiol* 2005; **12**: 1310–9. doi: https://doi.org/10.1016/j.acra.2005.06.005
18. Aoyama M, Li Q, Katsuragawa S, Li F, Sone S, Doi K. Computerized scheme for determination of the likelihood measure of malignancy for pulmonary nodules on low-dose CT images. *Med Phys* 2003; **30**: 387–94. doi: https://doi.org/10.1118/1.1543575
19. Lee H, Tajmir S, Lee J, Zissen M, Yeshiwas BA, Alkasab TK, et al. Fully automated deep learning system for bone age assessment. *J Digit Imaging* 2017; **30**: 427–41. doi: https://doi.org/10.1007/s10278-017-9955-8
20. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer* 2018; **18**: 500–10. doi: https://doi.org/10.1038/s41568-018-0016-5
21. Saghiri MA, Asgar K, Boukani KK, Lotfi M, Aghili H, Delvarani A, et al. A new approach for locating the minor apical foramen using an artificial neural network. *Int Endod J* 2012; **45**: 257–65. doi: https://doi.org/10.1111/j.1365-2591.2011.01970.x
22. Saghiri MA, Garcia-Godoy F, Gutmann JL, Lotfi M, Asgar K. The reliability of artificial neural network in locating minor apical foramen: a cadaver study. *J Endod* 2012; **38**: 1130–4. doi: https://doi.org/10.1016/j.joen.2012.05.004
23. Johari M, Esmaeili F, Andalib A, Garjani S, Saberkari H. Detection of vertical root fractures in intact and endodontically treated premolar teeth by designing a probabilistic neural network: an ex vivo study. *Dentomaxillofac Radiol* 2017; **46**: 20160107. doi: https://doi.org/10.1259/dmfr.20160107
24. Devito KL, de Souza Barbosa F, Felippe Filho WN. An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2008; **106**: 879–84. doi: https://doi.org/10.1016/j.tripleo.2008.03.002
25. McGrath TA, Alabousi M, Skidmore B, Korevaar DA, Bossuyt PMM, Moher D, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. *Syst Rev* 2017; **6**: 194. doi: https://doi.org/10.1186/s13643-017-0590-8

26. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**: 529–36. doi: https://doi.org/10.7326/0003-4819-155-8-201110180-00009
27. Lin PL, Huang PW, Huang PY, Hsu HC. Alveolar bone-loss area localization in periodontitis radiographs based on threshold segmentation with a hybrid feature fused of intensity and the H-value of fractional Brownian motion model. *Comput Methods Programs Biomed* 2015; **121**: 117–26. doi: https://doi.org/10.1016/j.cmpb.2015.05.004
28. Lin PL, Huang PY, Huang PW. Automatic methods for alveolar bone loss degree measurement in periodontitis periapical radiographs. *Comput Methods Programs Biomed* 2017; **148**: 1–11. doi: https://doi.org/10.1016/j.cmpb.2017.06.012
29. Lee J-H, Kim D-H, Jeong S-N, Choi S-H. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J Periodontal Implant Sci* 2018; **48**: 114–23. doi: https://doi.org/10.5051/jpis.2018.48.2.114
30. Carmody DP, McGrath SP, Dunn SM, van der Stelt PF, Schouten E. Machine classification of dental images with visual search. *Acad Radiol* 2001; **8**: 1239–46. doi: https://doi.org/10.1016/S1076-6332(03)80706-7
31. Mol A, van der Stelt PF. Application of computer-aided image interpretation to the diagnosis of periapical bone lesions. *Dentomaxillofac Radiol* 1992; **21**: 190–4. doi: https://doi.org/10.1259/dmfr.21.4.1299632
32. 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* 2018; **77**: 106–11. doi: https://doi.org/10.1016/j.jdent.2018.07.015
33. Mikulka J, Gescheidtova E, Kabrda M, Perina V. Classification of jaw bone cysts and necrosis via the processing of Orthopantomograms. *Radioengineering* 2013; **22**: 114–22.
34. Nurtanio I, Astuti ER, Ketut Eddy Pumama I, Hariadi M, Purnomo MH. Classifying cyst and tumor lesion using support vector machine based on dental panoramic images texture features. *IAENG Int J Comput Sci* 2013; **40**: 29–37.
35. Allen PD, Graham J, Farnell DJJ, Harrison EJ, Jacobs R, Nicopolou-Karayianni K, et al. Detecting reduced bone mineral density from dental radiographs using statistical shape models. *IEEE Trans. Inform. Technol. Biomed.* 2007; **11**: 601–10. doi: https://doi.org/10.1109/TITB.2006.888704
36. Hwang JJ, Lee J-H, Han S-S, Kim YH, Jeong H-G, Choi YJ, et al. Strut analysis for osteoporosis detection model using dental panoramic radiography. *Dentomaxillofac Radiol* 2017; **46**: 20170006. doi: https://doi.org/10.1259/dmfr.20170006
37. Kavitha MS, An S-Y, An C-H, Huh K-H, Yi W-J, Heo M-S, et al. Texture analysis of mandibular cortical bone on digital dental panoramic radiographs for the diagnosis of osteoporosis in Korean women. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2015; **119**: 346–56. doi: https://doi.org/10.1016/j.oooo.2014.11.009
38. Kavitha MS, Asano A, Taguchi A, Heo M-S. The combination of a histogram-based clustering algorithm and support vector machine for the diagnosis of osteoporosis. *Imaging Sci Dent* 2013; **43**: 153–61. doi: https://doi.org/10.5624/isd.2013.43.3.153
39. Kavitha MS, Asano A, Taguchi A, Kurita T, Sanada M. Diagnosis of osteoporosis from dental panoramic radiographs using the support vector machine method in a computer-aided system. *BMC Med Imaging* 2012; **12**: 1. doi: https://doi.org/10.1186/1471-2342-12-1
40. Kavitha MS, Ganesh Kumar P, Park S-Y, Huh K-H, Heo M-S, Kurita T, et al. Automatic detection of osteoporosis based on hybrid genetic Swarm fuzzy classifier approaches. *Dentomaxillofac Radiol* 2016; **45**: 20160076. doi: https://doi.org/10.1259/dmfr.20160076
41. Muramatsu C, Matsumoto T, Hayashi T, Hara T, Katsumata A, Zhou X, et al. Automated measurement of mandibular cortical width on dental panoramic radiographs. *Int J Comput Assist Radiol Surg* 2013; **8**: 877–85. doi: https://doi.org/10.1007/s11548-012-0800-8
42. Nakamoto T, Taguchi A, Ohtsuka M, Suei Y, Fujita M, Tsuda M, et al. A computer-aided diagnosis system to screen for osteoporosis using dental panoramic radiographs. *Dentomaxillofac Radiol* 2008; **37**: 274–81. doi: https://doi.org/10.1259/dmfr/68621207
43. Roberts MG, Graham J, Devlin H. Image texture in dental panoramic radiographs as a potential biomarker of osteoporosis. *IEEE Trans Biomed Eng* 2013; **60**: 2384–92. doi: https://doi.org/10.1109/TBME.2013.2256908
44. De Tobel J, Radesh P, Vandermeulen D, Thevissen PW. An automated technique to stage lower third molar development on panoramic radiographs for age estimation: a pilot study. *J Forensic Odontostomatol* 2017; **2**: 42–54.
45. Ohashi Y, Ariji Y, Katsumata A, Fujita H, Nakayama M, Fukuda M, et al. Utilization of computer-aided detection system in diagnosing unilateral maxillary sinusitis on panoramic radiographs. *Dentomaxillofac Radiol* 2016; **45**: 20150419. doi: https://doi.org/10.1259/dmfr.20150419
46. Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al.;in press Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac Radiol* 2019; **48**: 20180051. doi: https://doi.org/10.1259/dmfr.20180051
47. Son LH, Tuan TM, Fujita H, Dey N, Ashour AS, Ngoc VTN, et al. Dental diagnosis from X-ray images: an expert system based on fuzzy computing. *Biomed Signal Process Control* 2018; **39**: 64–73. doi: https://doi.org/10.1016/j.bspc.2017.07.005
48. Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop* 1998; **113**: 173–9. doi: https://doi.org/10.1016/S0889-5406(98)70289-6
49. Liu JK, Chen YT, Cheng KS. Accuracy of computerized automatic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop* 2000; **118**: 535–40. doi: https://doi.org/10.1067/mod.2000.110168
50. Hutton TJ, Cunningham S, Hammond P. An evaluation of active shape models for the automatic identification of cephalometric landmarks. *Eur J Orthod* 2000; **22**: 499–508. doi: https://doi.org/10.1093/ejo/22.5.499
51. Grau V, Alcañiz M, Juan MC, Monserrat C, Knoll C. Automatic localization of cephalometric landmarks. *J Biomed Inform* 2001; **34**: 146–56. doi: https://doi.org/10.1006/jbin.2001.1014
52. Rueda S, Alcañiz M. An approach for the automatic cephalometric landmark detection using mathematical morphology and active appearance models. *Med Image Comput Comput Assist Interv* 2006; **9**(Pt 1): 159–66.
53. Leonardi R, Giordano D, Maiorana F. An evaluation of cellular neural networks for the automatic identification of cephalometric landmarks on digital images. *J Biomed Biotechnol* 2009; **2009**: 1–12. doi: https://doi.org/10.1155/2009/717102
54. Vucinić P, Trpovski Z, Sćepan I. Automatic landmarking of cephalograms using active appearance models. *Eur J Orthod* 2010; **32**: 233–41. doi: https://doi.org/10.1093/ejo/cjp099
55. Shahidi S, Oshagh M, Gozin F, Salehi P, Danaei SM. Accuracy of computerized automatic identification of cephalometric landmarks by a designed software. *Dentomaxillofac Radiol* 2013; **42**: 20110187. doi: https://doi.org/10.1259/dmfr.20110187
56. Lindner C, Wang CW, Huang CT, Li CH, Chang SW, Cootes TF. Fully automatic system for accurate localisation and analysis of cephalometric landmarks in lateral Cephalograms. *Sci Rep* 2016; **20**: 33581.
57. Arık Sercan Ö, Ibragimov B, Xing L. Fully automated quantitative cephalometry using convolutional neural networks. *J Med Imaging* 2017; **4**: 014501: 014501: . doi: https://doi.org/10.1117/1.JMI.4.1.014501
58. Shahidi S, Bahrampour E, Soltanimehr E, Zamani A, Oshagh M, Moattari M, et al. The accuracy of a designed software for automated localization of craniofacial landmarks on CBCT images. *BMC Med Imaging* 2014; **14**: 32. doi: https://doi.org/10.1186/1471-2342-14-32

59. Gupta A, Kharbanda OP, Sardana V, Balachandran R, Sardana HK. A knowledge-based algorithm for automatic detection of cephalometric landmarks on CBCT images. *Int J Comput Assist Radiol Surg* 2015; **10**: 1737–52. doi: https://doi.org/10.1007/s11548-015-1173-6
60. Gupta A, Kharbanda OP, Sardana V, Balachandran R, Sardana HK. Accuracy of 3D cephalometric measurements based on an automatic knowledge-based landmark detection algorithm. *Int J Comput Assist Radiol Surg* 2016; **11**: 1297–309. doi: https://doi.org/10.1007/s11548-015-1334-7
61. Codari M, Caffini M, Tartaglia GM, Sforza C, Baselli G. Computer-Aided cephalometric landmark annotation for CBCT data. *Int J Comput Assist Radiol Surg* 2017; **12**: 113–21. doi: https://doi.org/10.1007/s11548-016-1453-9
62. Montúfar J, Romero M, Scougall-Vilchis RJ, Scougall VRJ. Automatic 3-dimensional cephalometric landmarking based on active shape models in related projections. *Am J Orthod Dentofacial Orthop* 2018; **153**: 449–58. doi: https://doi.org/10.1016/j.ajodo.2017.06.028
63. Montúfar J, Romero M, Scougall-Vilchis RJ, Scougall V RJ. Hybrid approach for automatic cephalometric landmark annotation on cone-beam computed tomography volumes. *Am J Orthod Dentofacial Orthop* 2018; **154**: 140–50. doi: https://doi.org/10.1016/j.ajodo.2017.08.028
64. Neelapu BC, Kharbanda OP, Sardana V, Gupta A, Vasamsetti S, Balachandran R, et al. Automatic localization of three-dimensional cephalometric landmarks on CBCT images by extracting symmetry features of the skull. *Dentomaxillofac Radiol* 2018; **47**: 20170054. doi: https://doi.org/10.1259/dmfr.20170054
65. Cheng E, Chen J, Yang J, Deng H, Wu Y, Megalooikonomou V, et al. Automatic Dent-landmark detection in 3-D CBCT dental volumes. *Conf Proc IEEE Eng Med Biol Soc* 2011; **2011**: 6204–7. doi: https://doi.org/10.1109/IEMBS.2011.6091532
66. Flores A, Rysavy S, Enciso R, Okada K. Non-Invasive differential diagnosis of dental periapical lesions in cone-beam CT. *IEEE International Symposium on Biomedical Imaging* 2009; 566–99.
67. Abdolali F, Zoroofi RA, Otake Y, Sato Y. Automatic segmentation of maxillofacial cysts in cone beam CT images. *Comput Biol Med* 2016; **72**: 108–19. doi: https://doi.org/10.1016/j.compbiomed.2016.03.014
68. Abdolali F, Zoroofi RA, Otake Y, Sato Y. Automated classification of maxillofacial cysts in cone beam CT images using contourlet transformation and spherical harmonics. *Comput Methods Programs Biomed* 2017; **139**: 197–207. doi: https://doi.org/10.1016/j.cmpb.2016.10.024
69. Yilmaz E, Kayikcioglu T, Kayipmaz S. Computer-Aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography. *Comput Methods Programs Biomed* 2017; **146**: 91–100. doi: https://doi.org/10.1016/j.cmpb.2017.05.012
70. Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput Biol Med* 2017; **80**: 24–9. doi: https://doi.org/10.1016/j.compbiomed.2016.11.003
71. Benyó B. Identification of dental root canals and their medial line from micro-CT and cone-beam CT records. *Biomed Eng Online* 2012; **11**: 81. doi: https://doi.org/10.1186/1475-925X-11-81
72. Yauney G, Angelino K, Edlund D, Shah P. *Convolutional Neural Network for Combined Classification of Fluorescent Biomarkers and Expert Annotations using White Light Images. 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*; 2017. pp. 303–9.
73. Rana A, Yauney G, Wong LC, Gupta O, Muftu A, Shah P. *Automated segmentation of gingival diseases from oral images. 2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)*; 2017. pp. 144–7.
74. Ngan TT, Tuan TM, Son LH, Minh NH, Dey N. Decision making based on fuzzy aggregation operators for medical diagnosis from dental X-ray images. *J Med Syst* 2016; **40**: 280. doi: https://doi.org/10.1007/s10916-016-0634-y
75. Sommer T, Ciesielski R, Erbersdobler J, Orthuber W, Fischer-Brandies H. Precision of cephalometric analysis via fully and semiautomatic evaluation of digital lateral cephalographs. *Dentomaxillofac Radiol* 2009; **38**: 401–6. doi: https://doi.org/10.1259/dmfr/85543699
76. Rana M, Modrow D, Keuchel J, Chui C, Rana M, Wagner M, et al. Development and evaluation of an automatic tumor segmentation tool: a comparison between automatic, semi-automatic and manual segmentation of mandibular odontogenic cysts and tumors. *J Craniomaxillofac Surg* 2015; **43**: 355–9. doi: https://doi.org/10.1016/j.jcms.2014.12.005
77. Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis. *Angle Orthod* 2008; **78**: 145–51. doi: https://doi.org/10.2319/120506-491.1
78. de Medeiros FCFL, Kudo GAH, Leme BG, Saraiva PP, Verri FR, Honório HM, et al. Dental implants in patients with osteoporosis: a systematic review with meta-analysis. *Int J Oral Maxillofac Surg* 2018; **47**: 480–91. doi: https://doi.org/10.1016/j.ijom.2017.05.021
79. Aghaloo T, Pi-Anfruns J, Moshaverinia A, Sim D, Grogan T, Hadaya D. The effects of systemic diseases and medications on implant osseointegration: a systematic review. *Int J Oral Maxillofac Implants* 2019; **34**(Suppl): s35–49. doi: https://doi.org/10.11607/jomi.19suppl.g3
80. Taguchi A, Tsuda M, Ohtsuka M, Kodama I, Sanada M, Nakamoto T, et al. Use of dental panoramic radiographs in identifying younger postmenopausal women with osteoporosis. *Osteoporos Int* 2006; **17**: 387–94. doi: https://doi.org/10.1007/s00198-005-2029-7
81. Vlasiadis KZ, Damilakis J, Velegrakis GA, Skouteris CA, Fragouli I, Goumenou A, et al. Relationship between BMD, dental panoramic radiographic findings and biochemical markers of bone turnover in diagnosis of osteoporosis. *Maturitas* 2008; **59**: 226–33. doi: https://doi.org/10.1016/j.maturitas.2008.01.006
82. Monte-Santo AS, Viana SVC, Moreira KMS, Imparato JCP, Mendes FM, Bonini GAVC. Prevalence of early loss of primary molar and its impact in schoolchildren's quality of life. *Int J Paediatr Dent* 2018; **28**: 595–601. doi: https://doi.org/10.1111/ipd.12416
83. Firestone AR, Sema D, Heaven TJ, Weems RA. The effect of a knowledge-based, image analysis and clinical decision support system on observer performance in the diagnosis of approximal caries from radiographic images. *Caries Res* 1998; **32**: 127–34. doi: https://doi.org/10.1159/000016442
84. Gakenheimer DC. The efficacy of a computerized caries detector in intraoral digital radiography. *J Am Dent Assoc* 2002; **133**: 883–90. doi: https://doi.org/10.14219/jada.archive.2002.0303
85. Wenzel A, Hintze H, Kold LM, Kold S. Accuracy of computer-automated caries detection in digital radiographs compared with human observers. *Eur J Oral Sci* 2002; **110**: 199–203. doi: https://doi.org/10.1034/j.1600-0447.2002.21245.x
86. Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. *Radiology* 2018; **286**: 800–9. doi: https://doi.org/10.1148/radiol.2017171920
87. Steyerberg EW. Validation of prediction models. In: Steyerberg E. W, ed. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2010. pp. 299–310.
88. Kohli MD, Summers RM, Geis JR. Medical image data and datasets in the era of machine Learning-Whitepaper from the 2016 C-MIMI meeting dataset session. *J Digit Imaging* 2017; **30**: 392–9. doi: https://doi.org/10.1007/s10278-017-9976-3
89. Sawagashira T, Hayashi T, Hara T, Katsumata A, Muramatsu C, Zhou X, et al. An automatic detection method for carotid artery calcifications using top-hat filter on dental panoramic radiographs. *Conf Proc IEEE Eng Med Biol Soc* 2011; **2011**: 6208–11. doi: https://doi.org/10.1109/IEMBS.2011.6091533