

Review article

Artificial intelligence for dental implant classification and peri-implant pathology identification in 2D radiographs: A systematic review

M Bonfanti-Gris^a, E Ruales^b, MP Salido^{a,*}, F Martinez-Rus^a, M Özcan^b, G Pradies^a

^a Department of Conservative Dentistry and Prosthesis, Complutense University of Madrid. Plaza Ramon y Cajal, s/n. 28040, Madrid, Spain

^b Clinic of Masticatory Disorders and Dental Biomaterials, Center for Dental Medicine, University of Zurich. Plattenstrasse, 11, 8032, Zurich, Switzerland



ARTICLE INFO

Keywords:

Artificial Intelligence
Deep learning
Dental implant
Object detection
Panoramic radiograph
Periapical radiograph
Periimplantitis

ABSTRACT

Objective: This systematic review aimed to summarize and evaluate the available information regarding the performance of artificial intelligence on dental implant classification and peri-implant pathology identification in 2D radiographs.

Data Sources: Electronic databases (Medline, Embase, and Cochrane) were searched up to September 2024 for relevant observational studies and both randomized and controlled clinical trials. The search was limited to studies published in English from the last 7 years. Two reviewers independently conducted both study selection and data extraction. Risk of bias assessment was also performed individually by both operators using the Quality Assessment Diagnostic Tool (QUADAS-2).

Study Selection: Of the 1,465 records identified, 29 references were selected to perform qualitative analysis. The study characteristics were tabulated in a self-designed table. QUADAS-2 tool identified 10 and 15 studies to respectively have a high and an unclear risk of bias, while only four were categorized as low risk of bias. Overall, accuracy rates for dental implant classification ranged from 67 % to 99 %. Peri-implant pathology identification showed results with accuracy detection rates over 78,6 %.

Conclusions: While AI-based models, particularly convolutional neural networks, have shown high accuracy in dental implant classification and peri-implant pathology detection, several limitations must be addressed before widespread clinical application. More advanced AI techniques, such as Federated Learning should be explored to improve the generalizability and efficiency of these models in clinical practice.

Clinical Significance: AI-based models offer can and clinicians to accurately classify unknown dental implants and enable early detection of peri-implantitis, improving patient outcomes and streamline treatment planning.

1. Introduction

The discovery of osseointegration marked a significant advancement in the field of dental prosthetics, providing a viable alternative for the replacement of missing teeth [1]. This approach not only offers a stable solution compared to removable prostheses but also preserves the integrity of adjacent teeth (Kim et al., 2022). Numerous clinical studies on osseointegrated implants have consistently reported higher success rates and significant improvements in patient quality of life [2].

Their benefits have caused both an increase in demand and a rushing necessity to deepen clinicians' knowledge in the field. Hence, since their introduction in the 1960s, a growing number of manufacturers and

implant designs have been identified [3]. However, even if dental implants have shown significant survival rates in the long term [4,5], they can suffer both biological and mechanical complications over the years – which, depending on their severity, may or not demand their removal. On such occasions, it is indispensable to have the essential implant information – such as its manufacturer, model, and dimensions – to appropriately solve the patient's condition [6]. Even if these situations are undoubtedly frequent in clinical practice, the importance of dental implant identification has also been reported for forensic purposes, especially during major disasters or incinerated individuals [7,8].

Clinicians' failure to identify dental implants may be due to several motives. From them, the lack or loss of traceability documents provided

* Corresponding author at: Department of Conservative Dentistry and Prosthesis, Complutense University of Madrid. Plaza Ramon y Cajal, s/n. 28040, Madrid, Spain.

E-mail addresses: monibonf@ucm.es (M. Bonfanti-Gris), edwinruales@gmail.com (E. Ruales), mpsalido@ucm.es (M. Salido), framarti@ucm.es (F. Martinez-Rus), mutlu.ozcan@zsm.uzh.ch (M. Özcan), gipradie@ucm.es (G. Pradies).

<https://doi.org/10.1016/j.jdent.2024.105533>

Received 18 September 2024; Received in revised form 12 November 2024; Accepted 13 December 2024

Available online 15 December 2024

0300-5712/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

to patients can be highlighted. Even so, the experience of the professional experience will always be an important consideration to bear in mind – as novel dental practitioners will struggle with this task for natural reasons [3].

Generally, electronic medical records can be used to confirm the type of implant that was inserted but, in cases in which the patient comes from another office and no records can be extracted, both panoramic and periapical radiographs remain the only identification alternatives [9]. As this has been a rising problem during the last two decades, several solutions have emerged from it - including computational methods for both implant morphology [10] and radiographic identification databases [11]. In this field, Artificial Intelligence (AI) appears to be evolving quite rapidly, as its algorithms allow complex data analysis and classification through human-like imitation patterns [12].

Since its first introduction in 1995, AI has revolutionized all healthcare-related fields [13]. Typically, AI is conceived to be divided into Machine (ML) and Deep Learning (DL), depending on the neural network's pre-processing steps for learning. Convolutional Neural Networks (CNNs) can be defined as computer models trained to recognize certain patterns present in a specific dataset, constituting the fundamentals of DL –These systems have already been applied to several image-based medical and dental domains such as photographic melanoma identification [14] or radiographic caries detection [15], showing excellent results.

As for its use in implant dentistry, several studies have reported excellent results for dental implant identification and pathology prediction [3,16]. Systematic and scoping reviews have also been published, but none of them have included both forms of radiographic identification alternatives (panoramic and periapical radiographs) for both purposes [17].

Thus, the primary objective of this systematic review was to answer the following focused PICO question: What is the performance accuracy (Outcome) of Artificial Intelligence-based systems (Intervention) when classifying dental implants and identifying pathological areas (Context) on both panoramic and periapical radiographs (Population) compared to the expert's opinion or reference standards?

2. Material and methods

2.1. Protocol development and focused question

This diagnostic accuracy systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [18]. The review protocol was registered and allocated with the CRD42024575604 in the PROSPERO International Prospective Register of Systematic Reviews hosted by the Centre of Reviews and Dissemination (University of York, National Institute for Health Research, United Kingdom). The objective of this review was to appraise Journal Citation Reports (JCR) studies that used any form of AI-based systems to automatically classify dental implants and detect peri-implant pathologies on 2D radiographic images (panoramic and periapical radiographs).

2.2. Eligibility criteria

2.2.1. Inclusion criteria

Studies reporting the following criteria were included:

- **Population:** 2D radiographic images, including both panoramic and periapical radiographs. Patient age or sex will not be considered.
- **Intervention:** Any form of AI-based models employing dental implant system classification and peri-implant pathology identification tasks (Machine Learning, Deep Learning, Transfer Learning, Collaborative Learning, Federated Learning, etc.).
- **Comparative** Ground Truth or Reference Test established by experts.

- **Outcomes:** Accuracy (A), Sensitivity (S), Specificity (E), Precision (P), Recall (R), Mean Average Precision and Recall (mAP and mAR), and F1 Score will be sought. Nevertheless, performance measurements like Dice Coefficient, Positive Predictive Values (PPV), Negative Predictive Values (NPV), Area Under the Curve (AUC), and confusion matrix results will also be considered if available.
- **Study Design:** Observational Studies, Randomized and Controlled Clinical Trials (RCT/CCT).

2.2.2. Exclusion criteria

Studies with the following criteria were excluded:

- Studies published before 2017.
- Studies not written in English.
- Studies registered as protocols only.
- Literature and Scoping reviews, unpublished manuscripts, and conference abstracts.
- Studies without sufficient details on the kind of AI-based method used.
- Studies without sufficient details concerning the methodology employed.

2.3. Information sources and search strategy

A comprehensive literature search was conducted in August 2024 to identify relevant literature in PubMed (Medline), Embase, and Cochrane databases. The search satisfied the inclusion criteria, including (MeSH) terms specified in Table 1 – where each database source search strategy can also be found. A second electronic search was performed in September 2024 to ensure updated conclusions.

2.4. Screening methods

For managing the citations, Zotero v.6.0.26 (Center for History and New Media, George Mason University, Virginia, USA) was used. Nevertheless, the total number of citations was transferred to Rayyan's online service (<https://www.rayyan.ai>) to ease the first selection process based on the article's title and abstract. Also, cross-references were detected and solved using this software. The screening process was performed by two independent reviewers (M.B-G and E.R). Articles' reasons for exclusion were categorized and registered (Table 2). Full-text studies were considered if they followed the inclusion criteria or if they did not provide clear information regarding their methodology and purpose in the abstract. If disagreement was found at any stage of the selection process, it was solved through consensus with a third reviewer (G.P). The inter-reviewer reliability was registered by calculating the kappa correlation coefficient for both screenings.

2.5. Data extraction and analysis

Two different reviewers individually collected data from each of the studies included in the present systematic review (M.B-G and E.R.). A third reviewer revised data extraction for discrepancies and disagreements (M.Ö). The following data items were extracted: Author, Radiographic Image, objective, CNN Architecture, Initial Dataset, Data Augmentation, Final Dataset, Dataset Distribution, Validation Method and Results.

When data were incomplete or missing, corresponding authors were contacted for clarification. If the information could not be obtained in time, the reference was excluded until notification was available. If the results of a given study were published more than once, only the most recent one was included. If more than one neural network was employed in the same study, all given results were reported to enable objective conclusions. In case a study included performance evaluation of a CNN for several variables, only the ones considering implant classification and/or peri-implant pathology identification were considered.

Table 1

Search execution strategy followed on PubMed, Embase and Cochrane databases.

	MeSH Terms	Artificial Intelligence; Deep Learning; Dental Implant; Dentistry; Machine Learning; Neural Networks, computer; Panoramic; Radiography, dental, digital; Radiography, dental; Radiography, periapical;
Full-line Search Strategy	PubMed MEDLINE (18/02/2024)	(("Artificial Intelligence"[MeSH Terms] OR "artificial intelligence*" [Text Word] OR "computational intelligence*" [Text Word] OR "machine intelligence*" [Text Word] OR "Machine Learning"[MeSH Terms] OR "machine learning*" [Text Word] OR "computer vision*" [Text Word] OR "Deep Learning"[MeSH Terms] OR "deep learning*" [Text Word] OR "neural networks, computer"[MeSH Terms] OR "neural network*" [Text Word]) AND ("radiography, dental"[MeSH Terms] OR "dental radiograph*" [Text Word] OR "dental radiolog*" [Text Word] OR "radiography, panoramic"[MeSH Terms] OR "panoramic radiograph*" [Text Word] OR "panoramic radiolog*" [Text Word] OR "pantomograph*" [Text Word] OR "orthopantomograph*" [Text Word] OR "radiography, dental, digital"[MeSH Terms] OR "dental digital*" [Text Word] OR "periapical radiograph*" [Text Word] OR "Dentistry"[MeSH Term] OR dentist*[tw]) AND ("dental implant"[MeSH Terms] OR "dental implant*" [Text Word]) NOT ("editorial"[Publication Type] OR "editorial*" [Title] OR "letter"[Publication Type] OR "letter*" [Title] OR "Comment"[Publication Type] OR "comment*" [Title] OR "hascommenton"[All Fields] OR "reply*" [Title] OR "Historical Article"[Publication Type] OR "congress"[Publication Type] OR "Congresses as Topic"[MeSH Terms] OR "proceedings*" [Title] OR "Retracted Publication"[Publication Type] OR "Retraction of Publication"[Publication Type] OR "Retraction of Publication as Topic"[MeSH Terms] OR "retract*" [Title] OR "withdrawn*" [Title] OR "Published Erratum"[Publication Type] OR "erratum*" [Title] OR "errata*" [Title] OR "hasretractionin"[All Fields] OR "hasretractionof"[All Fields] OR "pubmed books"[Filter])
	Embase (18/02/2024)	('artificial intelligence'/exp OR 'artificial intelligence' OR 'artificial intelligenc*' OR 'computational intelligenc*' OR 'machine intelligenc*' OR 'computer vision*' OR 'deep learning'/exp OR 'deep learning' OR 'deep learning*' OR 'artificial neural network'/exp OR 'artificial neural network' OR 'neural network*') AND ('dental radiograph*' OR 'dental radiolog*' OR 'panoramic radiography'/exp OR 'panoramic radiography' OR 'panoramic radiograph*' OR 'panoramic radiolog*' OR pantomograph* OR orthopantomograph* OR 'intraoral x ray system'/exp OR 'intraoral x ray system' OR 'dental digital*' OR 'periapical radiography'/exp OR 'periapical radiography' OR 'periapical radiograph*' OR 'dentistry'/exp OR 'dentistry' OR dentist*) AND ('dental implant*') NOT ([article]/lim OR [article in press]/lim OR [data papers]/lim OR [note]/lim OR [review]/lim OR [short survey]/lim) ([MeSH descriptor: [Artificial Intelligence] OR [Neural Networks, Computer] OR [Deep Learning] OR [Machine Learning] explode all
	Cochrane (18/02/2024)	

Table 1 (continued)

MeSH Terms	Artificial Intelligence; Deep Learning; Dental Implant; Dentistry; Machine Learning; Neural Networks, computer; Panoramic; Radiography, dental, digital; Radiography, dental; Radiography, periapical;
	trees' OR '(artificial intelligenc* OR 'machine intelligenc*' OR 'computer vision*' OR 'deep learning*' OR 'neural network*')): ti, ab, kw]) AND ([MeSH descriptor: [Radiography, Dental] OR [Dentistry] OR [Dental Implant] explode all trees' OR '(dental radiolog* OR 'dental radiograph*' OR 'panoramic radiograph*' OR 'orthopantomography*' OR 'periapical radiograph*' OR 'dentist*')):ti, ab, kw]]

Table 2

Full-text exclusion reasons and comments.

Author	Year	Reason for Exclusion	Comments
Alharbi et al.	2022	Not aligned with objective	The article does not classify implants nor detects pathologies, it predicts the patient's need to receive implant treatment
Jang et al.	2022	Not aligned with objective	The articles does not classify dental implant systems, it only detects implant from periapical radiographs. Also, peri-implant tissue detection results are not shown nor described in the text.
Lee et al.	2021	Not aligned with objective	The articles does not classify implant systems nor detects pathologies, it detects whether or not a dental implant was fractured.
Lee et al.	2022	Not aligned with objective	The article does not classify dental implant systems nor detects pathologies, it enables a critical analysis on whether or not AI-based systems actually help clinicians with these tasks.
Oh et al.	2023	Not aligned with objective	The article does not classify dental implant systems nor detects pathologies, it predicts the level of implant osseointegration based on panoramic or periapical radiographs.
Rekawek et al.	2023	Not aligned with objective	The article does not classify dental implant systems nor detects pathologies, it predicts the patient's probability to suffer implant failure or peri-implantitis bearing in mind different dummy variables

When referring to data analysis, the main information regarding Accuracy (A), Sensitivity (S), Specificity (E), Precision (P), Recall (R), Mean Average Precision and Recall (mAP and mAR), and F1 Score will be sought. Nevertheless, performance measurements like Dice Coefficient, PPV, NPV, AUC, and confusion matrix results will also be considered if available. If a study did not exhibit the results for the aforementioned but did report other metrics with which their calculation could be possible, this was carried out by the authors in the present study (M.B-G and E.R).

2.6. Risk of bias assessment

This process was independently performed by two reviewers (M.B-G and E.R). If any disagreements were found, a third assessor was contacted to resolve discrepancies (M.Ö). The risk of bias and methodological quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies QUADAS-2. This tool is designed to assess the quality of primary diagnostic accuracy studies, consisting of four key domains that discuss Patient Selection, Index Test,

Reference Standard, and Flow and Timing [19].

In this systematic review, the QUADAS-2 checklist was modified, and several questions related to data clarification of the studies' diagnostic accuracy methodologies were included [20] (Table 3). These questions were answered as positively (Yes), negatively (No), unclear, and occasionally, as NA (Not Applicable). Based on the final results, studies were rated for high (HR), low (LR), or unclear risk (UR) of bias. Applicability concerns were also registered based on the three first domains, being categorized as "Low", "High" or "Unclear" [19].

Overall, if the study presented limited information regarding data collection, did not include images belonging to different dental clinics, or was unclear about validation strategies, a high risk of bias was assessed for "Patient Selection". For "Index Test", if poor test reproducibility was detected, or there was a lack of information regarding models' construction and robustness analysis, a high risk of bias was also reported. This was also given for "Reference Standard" when insufficient information on the reference standard definition/ground truth was given or when just one operator established the reference test. Finally, a high risk of bias in "Flow and Timing" was given when studies employed different reference standards, did not include the total dataset in their final results, or inappropriate intervals between index test data acquisition and reference standard were observed.

Accordingly, if a study was found to have a high risk of bias in a single QUADAS-2 domain, it was automatically classified as HR. Similarly, if UR was observed in more than two domains, the authors considered the article highly biased (HR). If a study showed to have both a single domain with unclear risk of bias or a drawing situation between low risk and unclear risk of bias on four of the domains, the study was classified as UR. Only studies with all domains clear of bias were categorized as LR.

Table 3
QUADAS-2 original and modified (*) leading questions for critical appraisal.

QUADAS-2 Domain	Leading Questions
Patient Selection	Was a consecutive or random sample of patients enrolled? Did the dataset include radiographs from different institutions, or different X-Ray machines, allowing generalizability? * Did the study avoid inappropriate exclusions? Were data augmentation or cross-validation techniques applied (if necessary) to overcome overfitting? * Were data imbalances (if any) addressed in the study? * Was the test dataset separate from the training and validation datasets? *
Index Test	Where the deep learning method results interpreted without knowledge of the results of the reference standard? Was the method described in sufficient detail to reproduce the presented results? * Did the study perform any robustness or sensitivity analysis of their model? * If a threshold was used, was it prespecified?
Reference Standard	Is the reference standard likely to correctly classify the target condition? Were the reference standard results interpreted without knowledge of the results of the index test? Did the study use a ground truth reference? * If not, was the annotation procedure described in the study and found to minimize bias? * Did the study sufficiently report their limitations, biases, or issues around generalizability? *
Flow and Timing	Was there an appropriate interval between index tests and reference standard? Did all patients receive a reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?

3. Results

3.1. Search

The selection process and flowchart are depicted in Fig. 1. The search strategy rendered a total of 1.465 titles. After duplicated records removal, 1.346 final articles were finally included for screening. Subsequently, passed title and abstract screening, 40 articles were sought for full-text analysis. From these, a final number of 29 articles were finally included for data extraction. Reasons for exclusion are specified in PRISMA Flow Diagram. Kappa Cohen's Index for both title and abstract and full-text screenings were 0.980 and 1.000, respectively (P < 0.001).

3.2. Description of the selected studies

The characteristics of the included studies are described alphabetically in Table 4. All eligible articles were published in the last four years (2020–2024), being some conducted by the same research group: Kong et al. [9,21,22], Lee et al. [23,24], Park et al. [25–27], and Sukegawa et al. [28–30]. Of the included 29 articles, 24 described the use of AI-based techniques for dental implant classification while only 5 did so for peri-implant disease detection [31–35]. Only one article performed implant detection and segmentation analysis [36]. Articles provided their results using periapical [1,22,24,25,31,33,34,36–40], panoramic [2,9,21,28–30,41–44], or both types of radiographs [3,23,24,26,27,35,45].

Of the included 29 studies, thirteen reported experienced operators who interpreted the reference standard (medical records) and established ground truth [22,23,27,31–35,38,42,43]. Of the sixteen left, ten did not provide information regarding the reference standard [1–3,25,26,37,39,40,44,45], and six specified this information was only obtained from the patient's medical records [21,22,28–30,41].

A total number of 467.252 implants were included along the initial datasets of all 29 studies, and 871.134 of them were finally analyzed after data augmentation techniques. Five studies did not report full initial or final dataset information [3,9,37,41,42]. Dental Implant Systems (DIS) analyzed through all the included studies are depicted in Table 4 but, overall, models from Dentium (*Dentium CO*, California, USA), Dentsply Sirona (*Dentsply Sirona Inc.*, Charlotte, USA), Megagen (*Megagen Implant Co*, Seoul, Korea), Nobel Biocare (*Nobel Biocare Services AG*, Kloten, Switzerland), Osstem (*Osstem Implant Co*, Seoul, Korea), Straumann (*Straumann Group*, Basel, Switzerland) and Zimmer (*Zimmer Biomet Holdings, Inc.*, Indiana, USA) were the most frequently included.

Several convolutional neural networks were analyzed across the studies (Table 4). Overall, although all studies used deep learning methods, five of them utilized this approach combined with ensemble [21] or transfer learning [30], attention branch network [28], fine-tuning [30,35], or other optimization techniques [25]. Also, most of them used a combination of several architectures to approach their research question [22,29,31,38,39,41–45]. Only eleven studies analyzed just one neural network [1,2,22–24,26,33,34,36,37,45].

Fifteen authors used data augmentation and imbalance correction techniques to prevent overfitting. These included random flipping, rotation, zooming, cropping, and image-changing parameters. Of the fourteen studies left, only three had a less than 1.000 image dataset [24,25,38].

In terms of validation methods, most of the included records obtained their results from normalized split datasets. Only seven studies included 3-fold [9], 4-fold [28–30] or 5-fold [35,39,44] cross-validation methods.

Excellent performance was found for peri-implant pathology detection, with accuracy values over 78,6 % [35]. The overall accuracy of the convolutional neural networks in dental implant classification ranged from 67 % [41] to 99 % [29]. Nevertheless, most included studies

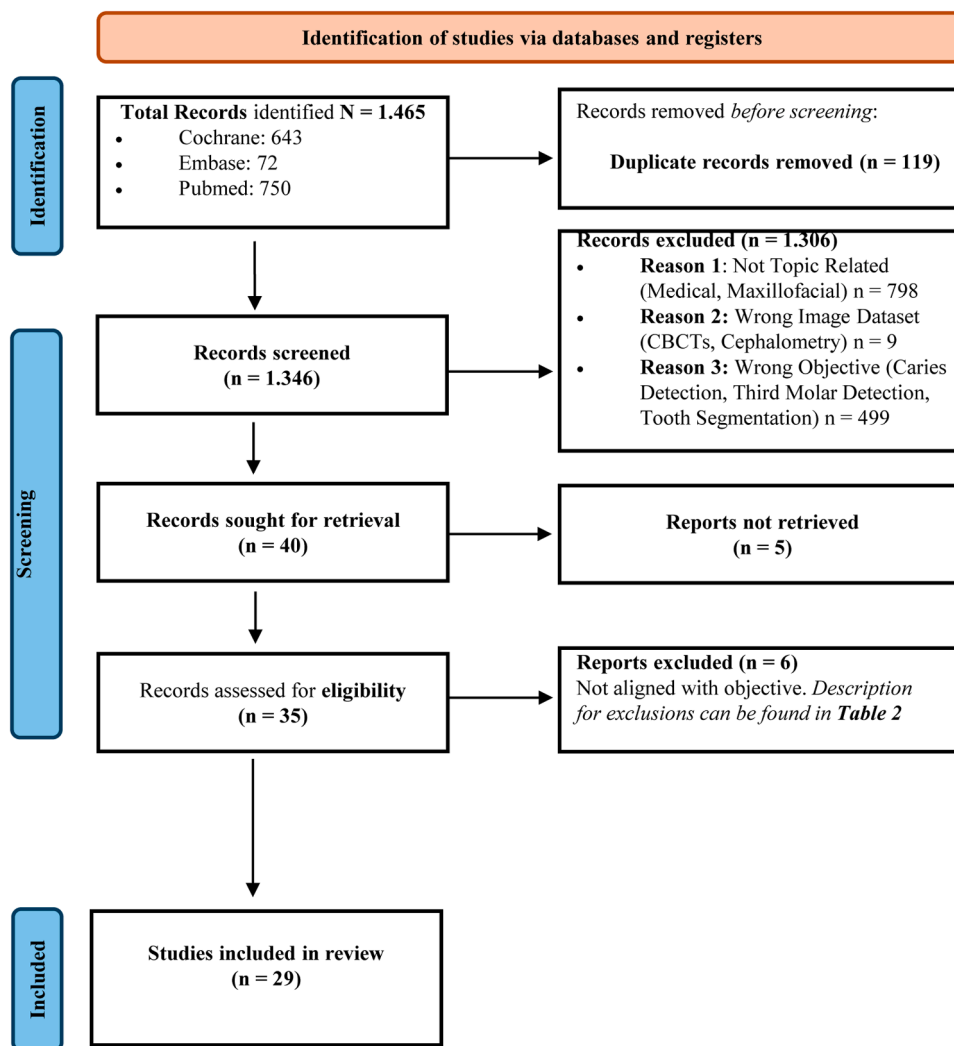


Fig. 1. PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only.

reported mean accuracy values over 85 % [1,3,22–25,27–29,32,36,38–40,42–44].

Authors that analyzed CNN’s performance with and without alternative learning techniques or fine-tuning procedures reported better results when these last ones were applied [21,30,35]. Only one group of researchers reported doubtful improvement of the results when applying an attention branch network [28]. From the studies that analyzed their results using individual and combined periapical and panoramic radiograph datasets, two reported the best results when using the intraoral radiograph dataset alone [23,26], while two described that combined datasets outperformed the others [27,35].

3.3. Risk of bias assessment

QUADAS-2 tool was used to score the reporting of diagnostic accuracy studies evaluating both dental implant classification and peri-implant pathology identification. Detailed information regarding the risk of bias assessment and applicability concerns is reported in Table 5 and Fig. 2. Overall, its critical appraisal checklist for diagnostic accuracy studies showed an unclear risk of bias among the studies.

Among the included studies, 19 (65 %) were found to have a Low or Unclear risk of bias. From these, only 4 (14 %) had a LR in all four domains of the QUADAS-2 tool. From the resulting 15 (52 %) titles, unclear bias was determined due to misinformation regarding leading questions belonging to different domains. On the other hand, 10 (34 %) studies

showed a high risk of bias. Overall, the most debatable domain was “Patient Selection” due to misinformation regarding sampling procedures, only high-quality image inclusion, and the lack of multi-center studies that could allow generalizability. This way, applicability concerns for this domain were unclear, unlike “Index Test” and “Reference Standard”, which resulted in an overall low concern.

4. Discussion

The use of Artificial intelligence in healthcare has increasingly been described over the past years. In the dental field, this technology has been applied from pathology identification [46,47] to implant osseointegration prediction [48]. Nevertheless, due to the growing dental implant-related possible complications, AI-based models have also been described as being able to recognize several dental implant systems and models [42] – which has undoubtedly been a turning point when missing their information.

The current systematic review involved all recently published articles concerning dental implant classification and peri-implant pathology detection using AI-based methods. Overall, this review showed that convolutional neural networks are a reliable and accurate method to help dental professionals in their daily routines for both purposes.

Overall, high accuracy values were reported by all authors for both dental implant classification and peri-implant pathology detection, confirming the AI-based system’s applicability to support dental

Table 4
Characteristics of the included studies.

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
Adnan et al., 2024	Periapical Radiographs	Implant Detection and Segmentation	U-Net	300	Flipping, Rotation and Brightness Modifications	1.294	TrS = 1035; VS = 129; TeS = 130	Split Dataset	Two Operators	Accuracy = 93.8 %; Precision = 90 %; Recall = 83 %; F1 Score = 86 %; IoU = 86.4 %; Loss = 21 %
Benakatti et al., 2021	Panoramic Radiographs	Implant Classification Osstem <i>TS III SA Regular</i> , Osstem <i>TS III SA Medium</i> and Noris <i>Medical Tuff</i>	Support Vector Machine (SVM), Logistic Regression (LR), K Nearest Neighbor (KNN) and Gradient Boosting (X Boost)	Not Specified	No	Not Specified	TrS = 80 % of dataset; TeS= 20 % of dataset	Split Dataset	Patient Records (Not clearly specified how Ground Truth was obtained for comparison with the neural networks)	Accuracy Based on Hu Classifier: SVM = 0.47; KNN = 0.33; LR = 0.50; and X Boost = 0.33. Accuracy Based on Eigenvalues: SVM = 0.67; KNN = 0.17; LR = 0.67; and X Boost = 0.67. YOLOv2 Object Detector: Accuracy = 89.3 %.
Chen et al., 2023	Periapical Radiographs	Periimplant Disease Detection	Algorithm, YOLO v2 for Implant Detection, Linear Regression Algorithm for Implant Cropping and AlexNet for Pathology Detection	406	Horitontal and Vertical Flipping	1.384	TrS = 648 healthy and 656 damaged (total 1304); VS = 40 healthy and 40 damaged (total 80)	Split Dataset	Three Operators with at least 5 years of clinical experience	AlexNet Image Classification: Accuracy = 90.4 %
Guo et al., 2022	Panoramic and Periapical Radiographs	Implant Classification Bego, Bicon and Straumann	VGG16, VGG16-GAP and TVGG-15	2.269	Rotation, 20 % rande object vertical and horizontal shifting, 20 % range cropping and zooming, and horizontal flipping	2.805	TrS = 550, VS = 185 and TeS = 200	Split Dataset	Not Specified	BEGO VGG16 Precision = 0.33; Recall = 1.00 and F1 Score = 0.50; VGG16-GAP Precision = 0.66; Recall = 0.90 and F1 Score = 0.76; TVGG15 Precision = 0.70; Recall = 0.92 and F1 Score = 0.79. BICON VGG16 Precision = 0.00; Recall = 0.00 and F1 Score = 0.00; VGG16-GAP Precision = 0.98; Recall = 0.57 and F1 Score = 0.73; TVGG15 Precision = 0.96; Recall = 0.68 and F1 Score = 0.80. STRAUMANN VGG16 Precision = 0.00; Recall = 0.00 and F1 Score = 0.00; VGG16-GAP Precision = 0.78; Recall =

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
										0.82 and F1 Score = 0.80; TVGG15 Precision = 0.80; Recall = 0.78 and F1 Score = 0.79
Hassan, 2023	Periapical Radiographs	Implant Classification Dentium, Fortis, NucleOSS, ROOTT, Straumann and Zimmer	Yolov8m-seg	Not Specified	Horizontal and Vertical Flipping, 90° rotations in all directions,	2.573	TrS = 2059, VS = 257 and TrS = 257	Split Dataset	Not Specified	DENTIUM Precision = 1.000, Recall = 0.972, F1 Score = 0.986 and mAP = 0.995; FORTIS Precision = 0.915, Recall = 0.727, F1 Score = 0.810 and mAP = 0.965; NUCLEOSS Precision = 0.735, Recall = 0.750, F1 Score = 0.742 and mAP = 0.791; ROOTT Precision = 0.924, Recall = 1.000, F1 Score = 0.960 and mAP = 0.995; STRAUMANN Precision = 0.930, Recall = 0.778, F1 Score = 0.847 and mAP = 0.943; ZIMMER Precision = 0.991, Recall = 1.000, F1 Score = 0.995 and mAP = 0.995
Hsiao et al., 2023	Periapical Radiographs	Implant Classification BioHorizons, Straumann and Nobel Biocare	MnasNET, SuffleNet, MobileNet, AlexNet, VGG, ResNet, DenseNet, SqueezeNet, ResNeXt, Wide ResNet	657	No	657	TrS =506, TeS =151	Split Dataset	One board-certified periodontist	Overall Test Accuracy for each NN: MnasNet = 81.89 %; SuffleNet = 96.85 %; MobileNet = 92.68 %; AlexNet = 94.35 %; VGG = 92.94 %; ResNet = 96.43 %; DenseNet = 96.41 %; SqueezeNet = 91.55 %; ReNeXt =93.90 % and Wide ResNet = 92.01 %
Kim et al., 2022	Periapical Radiographs	Implant Classification Superline, Osstem TS III, Straumann Bone Level	Yolov3	355	No	355	TrS = 284, TeS = 71	Split Dataset	Not Specified	Yolov3 Performance according to number of training epochs: (N= 100) Superline Sensitivity = 80.8 %, Specificity =

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
										95.6 %, Accuracy = 90.1 %; TS III Sensitivity = 65.2 %, Specificity = 93.8 %, Accuracy = 84.5 %; Bone Level Sensitivity = 81.8 %, Specificity = 98.0 %, Accuracy = 93.0 %; (N= 200) Superline Sensitivity = 88.5 %, Specificity = 97.8 %, Accuracy = 94.4 %; TS III Sensitivity = 95.7 %, Specificity = 95.8 %, Accuracy = 95.8 %; Bone Level Sensitivity = 100.0 %, Specificity = 100.0 %, Accuracy = 100.0 %; (N= 300) Superline Sensitivity = 84.6 %, Specificity = 91.1 %, Accuracy = 85.9 %; TS III Sensitivity = 78.3 %, Specificity = 89.6 %, Accuracy = 85.9 %; Bone Level Sensitivity = 63.6 %, Specificity = 100.0 %, Accuracy = 88.7 %
Kohlakala et al., 2022	Training: Simulated X-Ray images (N= 322.560); Testing: Periapical Radiographs (N= 483)	Implant Classification Anthogyr, Astra, MIS and Nobel Biocare	FCN-1 and FCN-2	483	Applied only in the simulated X-Rays - max. 60° rotation	483	Set 1: TrS= 338; VS = 63 and TeS = 82; Set 2: TrS = 338, VS = 58 TeS = 87.	5-fold cross validation	Not Specified	Overall Detection and Segmentation Accuracy = 94.0 % Set 1: Precision = 74.38 %, Recall 90.98 %, Accuracy 90.43 % and F1 Score = 80.73 %; Set 2: Precision = 68.31 %, Recall 78.64 %, Accuracy 94.06 % and F1 Score = 84.48 %

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
Kong et al., 2023	Panoramic Radiographs	Implant Classification Bicon, BioHorizons, Biomet 3i, Biotem, Dental Ratio, Dentis, Dentium, Dentsply Sirona, Dio Implant, Hi Ossen Implant, Ibs Implant, Keystone Dental, MegaGen, Neobiotech, Nobel Biocare, Osstem Implant, Point Implant, Straumann, Thommen Medical, and Zimmer	EfficientNet, Res2Next and Ensemble Learning	45.909	Resizing and cropping of the image, randomly changing image parameters and horizontally flipping the image	156.252	TrS = 147.124, TeS = 9.128	Split Dataset	Medical Record System	EfficientNet Top1 Accuracy = 73.83 %, Top5 Accuracy = 93.84 %, Precision = 74.61 %, Recall = 73.83 %, F1 Score = 72.02; ResNext2 Top1 Accuracy = 73.09 %, Top5 Accuracy = 93.60 %, Precision = 77.79 %, Recall = 73.08 %, F1 Score = 73.55 %; Ensemble Learning Top1 Accuracy = 75.27 %, Top5 Accuracy = 95.02 % Precision = 78.84 %, Recall = 75.27 % F1 Score = 74.89
Kong et al., 2023	Periapical Radiographs	Implant Classification Osstem TS III, Osstem US II, Biomet 3i Osseotite External and Dentsply Sirona Xive S Plus	AutoML Vision	4.800	No	4.800	TrS = 3.836, VS = 478, TeS = 486	Split Dataset	Medical Record System	Osstem TS III Accuracy = 1.000, Precision = 1.000, Recall = 1.000, Specificity = 1.000 and F1 Score = 1.000; OsstemUS II Accuracy = 0.969, Precision = 0.957, Recall = 0.925, Specificity = 0.976 and F1 Score = 0.941; Biomet 3i Osseotite Accuracy = 0.965, Precision = 0.927, Recall = 0.935, Specificity = 0.975 and F1 Score = 0.931; Dentsply Sirona Xive S Plus Accuracy = 0.988, Precision = 0.967, Recall = 0.983, Specificity = 0.989 and F1 Score = 0.975.
Kong et al., 2023	Panoramic Radiographs	Implant Classification According to implant designs bearing in mind the coronal, middle and	YOLOv5 and YOLOv7	14.037	Image Processing, random scaling and horizontal flipping, random copy-paste using a real-enhanced	Not Specified	TrS = 80 % of final dataset; TeS = 20 % of final dataset	3-fold cross validation	Two experienced prosthodontists	Yolov5 mAP Values Dataset 1 = 0.929, Dataset 2 = 0.940 and Dataset 3 = 0.873; Yolov7 mAP Values Dataset 1 =

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
		apical thirds. No implant model specified.			super-resolution generative adversarial network					0.931, Dataset 2 = 0.984 and Dataset 3 = 0.884
Kurtulus et al., 2024	Panoramic Radiographs	Implant Classification Bilimplant, Dentium, Megagen, Dyna, Implantance and Straumann.	ConvNeXt, VGG16, ResNet-50, EfficientNet (and others)	6.586	Image Scaling, flipping, rotation and hue modification	Not Specified	TrS = 80 % of final dataset; TeS = 20 % of final dataset	Split Dataset	Medical Record System and Experienced Physicians	ConvNeXt Accuracy = 95.74 %, Precision = 96.01 %, Recall = 94.72 %, F1 Score = 95.22 %; VGG16 Accuracy = 92.61 %, Precision = 91.53 %, Recall = 92.25 %, F1 Score = 91.81 %; ResNet-50 Accuracy = 92.61 %, Precision = 91.53 %, Recall = 92.25 %, F1 Score = 91.81 %; EfficientNet Accuracy = 77.63 %, Precision = 78.43 %, Recall = 77.25 %, F1 Score = 77.06 %
Lee et al., 2020	Panoramic and Periapical Radiographs	Implant Classification Dentsply Sirona <i>Astra OsseoSpeed TX</i> ; Dentium <i>Implantium</i> ; Dentium <i>SuperLine</i> ; Osstem <i>TSIII</i> , Straumann <i>SLActive BL</i> , Straumann <i>SLActive BLT</i>	Not Specified	11.980	No	11.980	TrS = 9.584; TeS = 2.396	Split Dataset	Medical Record System and Five Periodontal Residents + Three board-certified periodontists	Dentsply Sirona Astra OsseoSpeed TX Accuracy = 0.945, Youden Index = 0.766, Sensitivity = 0.933 and Specificity 0.833; Dentium Implantium Accuracy = 0.908, Youden Index = 0.780, Sensitivity = 0.933 and Specificity 0.847; Dentium SuperLine Accuracy = 0.903, Youden Index = 0.786, Sensitivity = 0.833 and Specificity 0.954; Osstem TSIII Accuracy = 0.937, Youden Index = 0.813, Sensitivity = 0.900 and Specificity 0.913; Straumann SLActive BL Accuracy = 0.974, Youden Index = 0.833,

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
Lee et al., 2020	Periapical Radiographs	Periimplant Disease Detection	Yolov7	800	No	800	TrS = 600, VS = 100, TeS = 100	Split Dataset	One experienced Operator	Sensitivity = 0.967 and Specificity 0.867, Straumann SLActive BLT Accuracy = 0.981, Youden Index = 0.880, Sensitivity = 0.900 and Specificity 0.980 Bone Loss Detection: Accuracy = 96.18 %, Precision = 100 %, Sensitivity = 95.83 %, Specificity = 100 % and F1 Score = 97.86 %; Non-Bone Loss Detection: Accuracy = 93.42 %, Precision = 100 %, Sensitivity = 93.06 %, Specificity = 100 % and F1 Score = 96.43 %
Lee et al., 2024	Panoramic and Periapical Radiographs	Implant Classification Osstem TS III, Dentium Superline, Straumann SLActive BLT	GoogleNet Inception v3	10.770	Applied only to the Training Set: Random Scaling (up to x10), Horizontal and Vertical Flipping, 10° Rotation, Zooming and Width/Height Shifting	68.928	TrS = 64.620, VS = 2.154, TeS = 2.154	Split Dataset	Medical Record System and Three Periodontal Residents	Overall Classification Performance A) combining Panoramic and Periapical Radiographs Accuracy = 0.971, Sensitivity = 95.3 %, Specificity = 97.6 %; B) Only panoramic radiographs Accuracy = 0.956, Sensitivity = 93.6 %, Specificity = 95.7 %; C) Only Periapical Radiographs: Accuracy = 0.979, Sensitivity = 97.1 %, Specificity = 95.8 %;
Liu et al., 2022	Periapical Radiographs	Periimplant Disease Detection	Inception ResNet v2	1.670	No	1.670	TrS = 1.370, VS = 150, TeS = 150	Split Dataset	Experienced Operator (>5 years) and a maxillofacial radiologist	Implant Detection Average Precision = 0.99; Bone Loss Implant Detection Sensitivity = 67 %, Specificity = 87 %, Mistake

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
										Diagnostic Rate = 13 %, Ommision Diagnostic Rate = 33 %, Positive Predictive Value = 81 %; Bone Loss Sites Detection Sensitivity = 75 %, Specificity = 83 %, Mistake Diagnostic Rate = 17 %, Ommision Diagnostic Rate = 25 %, Positive Predictive Value = 87 %
Lubbad et al., 2024	Panoramic Radiographs	Implant Classification Straumann, Bilimplant, Dyna, MegaGen, and Dentium	RegNet, DenseNet, VGG, EfficientNet, ResNet, ConvNext, MobileNet, and Transformers Architectures	1.258	Rotation, Scaling, Image Flipping and Hue modifications	14.811	TrS = Not Specified, VS = Not Specified, TeS = Not Specified.	Split Dataset	Medical Record System + Expert Verification	Mean Values all CNN versions RegNet Accuracy = 0.935, Precision = 0.926, Recall = 0.929 and F1 Score = 0.928; DenseNet Accuracy = 0.932, Precision = 0.934, Recall = 0.916 and F1 Score = 0.922; VGG Accuracy = 0.925, Precision = 0.926, Recall = 0.908 and F1 Score = 0.917; EfficientNet Accuracy = 0.902, Precision = 0.905, Recall = 0.808 and F1 Score = 0.895; ResNet Accuracy = 0.864, Precision = 0.862, Recall = 0.849 and F1 Score = 0.853; ConvNext Accuracy = 0.905, Precision = 0.910, Recall = 0.880 and F1 Score = 0.876; MobileNet Accuracy = 0.844, Precision = 0.834, Recall = 0.816 and F1 Score = 0.824; and Transformers Architectures

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
Park et al., 2023	Periapical Radiographs	Implant Classification Staumann <i>Bone Level</i>	Deep Learning: VGG16; Clustering Analysis: K-Means ++ algorithm	1.320	No	1.320	Deep Learning: TrS = 960, VS = 180, TeS = 180; Clustering Analysis: TrS = 1.140, TeS = 180	Split Dataset	Not Specified	Accuracy = 0.900, Precision = 0.906, Recall = 0.806 and F1 Score = 0.894. Deep Learning Accuracy = 0.994, Sensitivity = 0.950, Specificity = 0.994, F1 Score = 0.974, PPV = 0.952, NPV = 0.994, AUC-ROC = 0.975; Clustering Analysis Accuracy = 0.983, Sensitivity = 0.900, Specificity = 0.988, F1 Score = 0.923, PPV = 0.909, NPV = 0.988, AUC-ROC = 0.947
Park et al., 2023	Panoramic and Periapical Radiographs	Implant Classification Neobiotech, Nobel Biocare, Dentsply Sirona, Dentium, Dioimplant, MegaGen, Straumann, Shinhung, Osstem and Warantec	Not Specified	165.700	Random Rotation (90°), hue, brightness, saturation and contrast modification, noise reduction and horizontal/vertical flipping	156.956	TrS = 80 % of final dataset, VS = 10 % of final dataset, TeS = 10 % of final dataset	Split Dataset	Medical Records and Maxillofacial Radiologist + Physician specialized in Implantology Validation	Overall Classification Performance A) combining Panoramic and Periapical Radiographs Accuracy = 88.53 %, Precision = 85.70 %, Recall = 82.30 %, F1 Score = 84.00 % B) Only panoramic radiographs Accuracy = 87.89 %, Precision = 85.20 %, Recall = 81.10 %, F1 Score = 83.10 % C) Only Periapical Radiographs: Accuracy = 86.87 %, Precision = 84.40 %, Recall = 81.70 %, F1 Score = 83.00 %
Park et al., 2023	Panoramic and Periapical Radiographs	Implant Classification Neobiotech, Nobel Biocare, Dentsply Sirona, Dentium, Dioimplant, MegaGen, Straumann, Shinhung, Osstem and Warantec	ResNet-50	150.733	No	150.733	TrS = 120.585, VS = 15.074, TeS = 15.074	Split Dataset	Not Specified	Overall Classification Performance A) combining Panoramic and Periapical Radiographs Accuracy = 82.30 %, AUC = 0.823, Sensitivity = 80 %, Specificity = 84.50 %, PPV = 83.80 %, NPV

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
										= 80.90 % B) Only panoramic radiographs Accuracy = 73.30 %, AUC = 0.733, Sensitivity = 71.5 %, Specificity = 75 %, PPV = 74.10 %, NPV = 72.50 %; C) Only Periapical Radiographs: Accuracy = 83.80 %, AUC = 0.838, Sensitivity = 81.5 %, Specificity = 86 %, PPV = 83.30 %, NPV = 82.30 %
Said et al., 2020	Panoramic and Periapical Radiographs	Implant Classification Nobel Biocare and Branemark System, Straumann Bone Level and Tissue Level, and Zimmer Tapered Screw Vent and Swiss Plus	GoogleNet Inception v3	Not Specified	Rotation, angulation, tone, brightness, contrast, blur, sharpness and gamma correction were applied	1.206	TrS and VS = 965, TeS = 241.	Split Dataset	Not Specified	Classification Results for each Dental Implant Model: Nobel Biocare A) Accuracy = 93 %, Sensitivity = 95.5 %, Specificity = 97.1 %, Positive Predictive Value = 94.2 %, Negative Predictive Value = 90.3 %, AUC = 0.918; B) Branemark System Accuracy = 95.8 %, Sensitivity = 91.5 %, Specificity = 94.8 %, Positive Predictive Value = 92 %, Negative Predictive Value = 93.7 %, AUC = 0.922; Straumann A) Bone Level Accuracy = 92.9 %, Sensitivity = 93 %, Specificity = 95.2 %, Positive Predictive Value = 90.4 %, Negative Predictive Value = 89.9 %, AUC =

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
										0.909; B) Tissue Level Accuracy = 93.6 %, Sensitivity = 90.7 %, Specificity = 94.1 %, Positive Predictive Value = 93.2 %, Negative Predictive Value = 92.6 %, AUC = 0.890; Zimmer A) Tapered Scrw Vent Accuracy = 96.2 %, Sensitivity = 94.6 %, Specificity = 93.6 %, Positive Predictive Value = 92.3 %, Negative Predictive Value = 92.6 %, AUC = 0.931, B) Swiss Plus Accuracy = 91.8 %, Sensitivity = 95.6 %, Specificity = 90.3 %, Positive Predictive Value = 90.1 %, Negative Predictive Value = 91.1 %, AUC = 0.911.
Santos et al., 2021	Periapical Radiographs	Implant Classification Straumann, Neodent and SIN Implants	Keras and Tensorflow Google Frameworks	1.800	Only applied to training dataset: Rotation, angulation, blur, contrast, and brightness modifications	402.360	TrS = 402.000, TeS= 360	Split Dataset	Not Specified	Accuracy = 85.3 %, Sensitivity = 89.9 %, Specificity = 82.4 %, Positive Predictive Value = 82.6 %, Negative Predictive Value = 88.5 %
Sukegawa et al., 2020	Panoramic Radiographs	Implant Classification Zimmer <i>Full Osseotite</i> , Dentsply <i>Astra EV, Astra TX, and Astra MicroThread</i> , Nobel Biocare <i>Branemark MKIII and Nobel Replace, Replace Select Tapered</i> , Straumann <i>Tissue Level</i>	Basic CNN with three convolution layers, VGG16 and VGG19. Transfer Learning and fine tuning were performed to each VGG16 and VGG19	8.859	No	8.859	TrS = 6.419; TeS = 2.140	4-fold cross validation	Medical Record System	Basic CNN Recall = 0.802, Precision = 0.842, Accuracy = 0.860 and F1-Score = 0.819; VGG16 Transfer Recall = 0.864, Precision = 0.888, Accuracy = 0.899 and F1-Score = 0.874; VGG16 Fine Tuning Recall

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
		and Kyocera Finesia.								= 0.907, Precision = 0.928, Accuracy = 0.935 and F1-Score = 0.916; VGG19 Transfer Recall = 0.840, Precision = 0.873, Accuracy = 0.880 and F1-Score = 0.853; VGG19 Fine Tuning Recall = 0.894, Precision = 0.913, Accuracy = 0.927 and F1-Score = 0.902.
Sukegawa et al., 2021	Panoramic Radiographs	Implant Classification Zimmer Full Osseotite, Dentsply Astra EV, Astra TX, and Astra MicroThread, Nobel Biocare Branemark MKIII and Nobel Replace, Replace Select Tapered, Straumann Tissue Level and Bone Level, and Kyocera Finesia	ResNet-18, 34, 50, 101, and 152	9.767	No	9.767	TrS = 7.813; TeS = 1.954	4-fold cross validation	Medical Record System	Overall Multi-task CNN's Results: ResNet18 Accuracy = 0.980, Precision = 0.978, Recall = 0.972, F1 Score = 0.974 and AUC = 0.999; ResNet34 Accuracy = 0.985, Precision = 0.984, Recall = 0.980, F1 Score = 0.982 and AUC = 0.999; ResNet50 Accuracy = 0.986, Precision = 0.987, Recall = 0.981, F1 Score = 0.983 and AUC = 0.999; ResNet101 Accuracy = 0.989, Precision = 0.989, Recall = 0.986, F1 Score = 0.987 and AUC = 0.999; ResNet152 Accuracy = 0.990, Precision = 0.991, Recall = 0.988, F1 Score = 0.989 and AUC = 0.999. Overall Single-task CNN's Results: ResNet18 Accuracy = 0.978, Precision = 0.973, Recall = 0.972, F1 Score = 0.972 and

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
										AUC = 0.999; ResNet34 Accuracy = 0.980, Precision = 0.979, Recall = 0.974, F1 Score = 0.976 and AUC = 0.999; ResNet50 Accuracy = 0.980, Precision = 0.981, Recall = 0.974, F1 Score = 0.977 and AUC = 0.999; ResNet101 Accuracy = 0.984, Precision = 0.982, Recall = 0.979, F1 Score = 0.980 and AUC = 0.999; ResNet152 Accuracy = 0.985, Precision = 0.983, Recall = 0.980, F1 Score = 0.982 and AUC = 0.999.
Sukegawa et al., 2022	Panoramic Radiographs	Implant Classification Zimmer Full Osseotite, Dentsply Astra EV, Astra TX, and Astra MicroThread, Nobel Biocare Branemark MKIII and Nobel Replace, Replace Select Tapered, Straumann Tissue Level and Bone Level, and Kyocera Finesia and Poi Ex	ResNet-18, 50 and 152 + Attention Branch Network (ABN)	10.191	No	10.191	TrS = 8.157, TeS = 2.034	4-fold cross validation	Medical Record System	ResNet-18 Accuracy = 0.984, Precision = 0.944, Recall = 0.933, F1 Score = 0.938, and AUC = 0.997; ResNet-18 ± ABN Accuracy = 0.971, Precision = 0.968, Recall = 0.962, F1 Score = 0.965, and AUC = 0.999; ResNet-50 ± ABN Accuracy = 0.951, Precision = 0.947, Recall = 0.938, F1 Score = 0.941, and AUC = 0.997; ResNet-152 ± ABN Accuracy = 0.957, Precision = 0.954, Recall = 0.947, F1 Score = 0.949, and AUC = 0.998; ResNet-50 ± ABN Accuracy = 0.951, Precision = 0.947, Recall = 0.938, F1 Score = 0.941, and AUC = 0.997; ResNet-152 ± ABN Accuracy = 0.962, Precision = 0.957, Recall = 0.950, F1 Score = 0.953, and AUC = 0.998; ResNet-152 ± ABN Accuracy =

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
Takahashi et al., 2020	Panoramic Radiographs	Implant Classification Nobel Biocare MKIII, MK III Groovy, MKIV and Speedy Groovy; Straumann Bone Level and GC Genesis Plus ST	Yolov3	1.282	No	1.282	TrS = 1.026, TeS = 256	Split Dataset	Not Specified	= 0.956, Precision = 0.951, Recall = 0.945, F1 Score = 0.947, and AUC = 0.995 Average Precision MKIII/III Groovy = 0.85, MKIV/Speedy Groovy = 0.78, Bone Level = 0.69 ad Genesis = 0.51. Sensitivity = MKIII/III Groovy = 0.82 MKIV/Speedy Groovy = 0.74, Bone Level = 0.74 ad Genesis = 0.50.
Tiryaki et al., 2023	Panoramic Radiographs	Implant Classification Bego Implant Systems, Impliance Dental Implant System, Implant Direct, Nucleoss Implant and Osstem Implant	VGG-16, VGG-19, ResNet-50, ResNet-101 and GoogleNet	11.904	No	11.904	Not Specified	5-fold cross validation	Not Specified	Overall CNN Fusion Implant Detection Accuracy Results. Bego = 95.7 %, Direct = 98.9 %, Impliance = 99.2 %, Nucleoss = 98.7 %, Osstem = 99.2 %.
Vera et al., 2023	Periapical and Bitewing Radiographs	Implant Detection/ Segmentation and Periimplant Disease Detection	Yolov3	2.336	Image enhancement and 10° rotation angles	2.920	TrS= 1.314, VS = 146, TeS = 1.460	Split Dataset	Several Specialists (no number specified)	Mean Object Detection Results: A) Periapical: Mean Average Precision = 0.844; B) Bitewing: Mean Average Precision = 0.743
Zhang et al., 2023	Panoramic and Periapical Radiographs	Periimplant Disease Detection	ResNet-50 + Fine Tuning	1.080	Random flipping and rotating (up to 30°)	1.800	TrS = 1.440, TeS = 360	5-fold cross validation	Two certified specialist in implant dentistry with more than 7 uears of experience	Overall Periimplant Pathology Detection Performance A) combining Panoramic and Periapical Radiographs Accuracy = 87 %, Precision = 0.85, Recall = 0.88 and F1 Score = 0.85 B) Only panoramic radiographs Accuracy = 78.7 %, Precision = 0.87, Recall = 0.63 and F1 Score = 0.66; C) Only Periapical Radiographs: Accuracy =

(continued on next page)

Table 4 (continued)

Author	Radiographic Image	Objective	CNN Architecture	Initial Dataset	Data Augmentation	Final Dataset	Dataset Distribution	Validation Method	Reference Standard	Results
										78.6 %, Precision = 0.84, Recall = 0.73 and F1 Score = 0.75.

Table 5

QUADAS-2 Risk of Bias Assessment Results and Comments. QUADAS-2 Risk of Bias Assessment Results and Comments – A, Low generalizability due to the inclusion of just one institution; B, Patient Selection Bias (only good quality images or with just one dental implant were included); C, No information regarding data sampling (random or consecutive); D, No information regarding total dataset; E, No description of how overfitting or data imbalance was tackled when using reduced or disproportionated datasets; F, No information was provided concerning whether or not the operators that analyzed the CNN’s results are the same that the ones that labeled the images based on the patient’s medical records; G, Methodology was not clearly described (Labelling, Reference Standard/Ground Truth establishment); and H, Reference Standard not clearly specified.

STUDY	RISK OF BIAS							APPLICABILITY CONCERNS			
	OVERALL	PATIENT SELECTION	Comments	INDEX TEST	Comments	REFERENCE STANDARD	Comments	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Adnan et al.	Some Concerns	Some Concerns	A, B	Low	None	Low	None	Low	Unclear	Low	Low
Benakatti et al.	High	High	A, B, C, D, E	High	F, G	Low	H	Low	High	Low	Low
Chen et al.	Some Concerns	Some Concerns	C	Some Concerns	F, G	Low	H	Low	Unclear	Low	Low
Guo et al.	High	Some Concerns	A, B, C	High	G	High	H	Unclear	Unclear	Low	Low
Hassan	Some Concerns	Some Concerns	B, C	Some Concerns	G	Low	None	Low	Unclear	Low	Low
Hsiao et al.	Low	Low	A, E	Low	F	Low	None	Low	Low	Low	Low
Kim et al.	High	Low	A,E	High	F, G	Some Concerns	H	Low	Low	Low	Unclear
Kohlakala et al.	Some Concerns	Some Concerns	C	Low	None	Low	None	Low	Unclear	Low	Low
Kong et al.	Some Concerns	Low	B, E	Some Concerns	F, G	Some Concerns	H	Some Concerns	Low	Unclear	Unclear
Kong et al.	High	Some Concerns	A, C	Some Concerns	G	Some Concerns	H	Low	Unclear	Unclear	Unclear
Kong et al.	Some Concerns	Some Concerns	B, E	Low	F	Low	None	Low	Low	Low	Low
Kurtulus et al.	Some Concerns	Some Concerns	B	Low	F	Low	None	Low	Unclear	Low	Low
Lee et al.	Some Concerns	Some Concerns	A, C, E	Low	F	Low	None	Low	Unclear	Low	Low
Lee et al.	Some Concerns	Some Concerns	A	Low	F	Low	None	Low	Unclear	Low	Low
Lee et al.	Some Concerns	Some Concerns	A, B, C	Low	None	Low	None	Low	Unclear	Low	Low
Liu et al.	Some Concerns	Some Concerns	B, C	Low	None	Low	None	Low	Unclear	Low	Low
Lubbad et al.	High	High	A, B, C, E	Low	None	Low	None	Low	High	Low	Low
Park et al.	Some Concerns	Some Concerns	B, C	Low	None	Low	None	low	Unclear	Low	Low
Park et al.	Low	Low	C, E	Low	None	Low	None	Low	Low	Low	Low
Park et al.	High	Some Concerns	C, E	Some Concerns	G	Some Concerns	H	Low	Unclear	Unclear	Unclear
Said et al.	High	Some Concerns	B, C, E	High	F, G	Some Concerns	H	Low	Unclear	Low	Unclear
Santos et al.	High	High	A, B, C	High	F, G	Some Concerns	H	Low	High	Unclear	Unclear
Sukegawa et al.	Some Concerns	Some Concerns	A, C, E	Some Concerns	F	Low	None	Low	Unclear	Unclear	Low
Sukegawa et al.	Some Concerns	Some Concerns	A, C, E	Low	None	Low	None	Low	Unclear	Low	Low
Sukegawa et al.	Low	Low	C	Low	None	Low	None	Low	Low	Low	Low
Takahashi et al.	High	Some Concerns	A, C, E	Some Concerns	G	Some Concerns	H	Low	Unclear	Unclear	Unclear
Tiryaki et al.	High	Some Concerns	C, B, E	Some Concerns	G	Some Concerns	H	Low	Unclear	Unclear	Unclear
Vera et al.	Low	Low	A	Low	None	Low	None	Low	Low	Low	Low
Zhang et al.	Some Concerns	Some Concerns	A, B, C	Low	None	Low	None	Low	Unclear	Low	Low

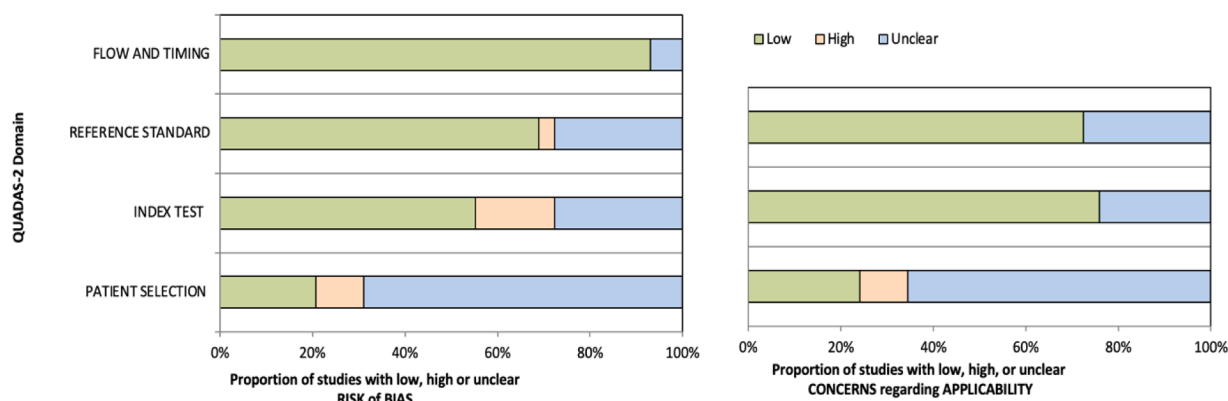


Fig. 2. QUADAS-2 Tool overall bias assessment and concerns regarding applicability graphs.

clinicians’ decision-making processes when biological, mechanical or both complications occur. Despite their promising results, this review identified several underpinnings that could have influenced the obtained outcomes.

First, datasets were found to include only high-quality images, excluding potentially misleading radiographs that could hinder CNN’s results. Also, data pre-processing was often carried out to enhance deep learning models’ accuracy rates [3,23]. This way, blurred images as well as not parallelized-taken radiographs were eliminated from the final samples. Bearing in mind that these situations are commonly found in the daily routine, AI-based models could not be sufficiently prepared to be applied in the clinical scenario. Besides this, only ten studies included multi-center datasets or images acquired by different X-ray machines, allowing generalizability [3,9,21,24,26,27,30,31,33,37].

Related to the aforementioned, the authors detected a wide variation in sample sizes across the included studies going from 355 [1] to 402.360 images [40]. Studies also differed in the number of dental implant systems that were analyzed – observing that, overall, the best accuracy results were related to a lower number of included DIS [1,37,38,45]. Similarly, when using different implant models, only a few studies ensured equally distributed data for each one of them. Thus, data imbalance across the included DIS was frequently observed, negatively influencing the results. Only five studies specifically reported to have overcome this problem [23,25,35,39,45].

To overcome both overfitting and data imbalance, data augmentation or cross-validation techniques were often used [3,9,21,27,31,32,34–37,39,40,42,43,45]. Overfitting occurs when the neural network is not able to generalize and fits too closely to the training data. These techniques applied while training allow the model to recognize dental implant systems in unseen data, thus improving validation results.

Secondly, variations were found in the image annotation process. While some authors manually performed this task using specifically designed tools, others created an independent CNN to automatize this process and complete it efficiently [31]. While it might be helpful, letting implant detection ground truth establishment to a neural network can trigger errors – especially when considering the high rate of False Positive (FP) detections due to the presence of other metallic structures near the region of interest (ROI) [49,50]. On the other hand, if this process was performed manually, detailed instructions should be provided to ensure methodological reproducibility. In this case, only a few articles did not provide this description [2,3,9,22,27,40,41,44,45], making their label acquisition to some degree uncertain.

Similarly, the reference standard was observed to be based on medical records in the majority of included studies [36]. Thus, the quality assessment results were influenced by the lack of professional expert validation. Overall, if dental professionals were involved in image-acquisition analysis and label validation, these were found to be general practitioners with at least 5 years of experience [24,31,33,36,

42,43], prosthodontists [9], board-certified periodontists [32,34,38] or implantologists [35], and maxillofacial radiologists [27,33].

An important issue to bear in mind is the presence of an unbiased index test interpretation. The majority of the included articles did not provide information regarding whether or not the AI-based results were interpreted without the reference standard results, which certainly increased the study’s risk of bias assessment for the Index Test domain.

Furthermore, it should be considered that the accuracy results could have been influenced by the characteristics of the used AI-based models. Specifically, training epochs ranged from 25 [40] to 300 [2,3,23,32,35], and architecture layers, as well as learning rates, varied across studies. Besides this, some authors proposed a modified pre-existing neural network to improve classification performance [45]. Thus, an objective comparison between CNN’s DIS classification performance might be difficult to carry out.

For DIS classification, Sukegawa et al., tested a neural network to single and multi-task performance. Their main purpose was to develop an AI-based model that could simultaneously classify a DIS and detect its treatment stage. The authors analyzed CNN’s performance while carrying out both tasks both individually and simultaneously. Overall, their analysis revealed that multi-tasking improved brand classification and significantly increased all metrics related to treatment phase detection [29]. Alternative methods were also developed and tested by other authors. The addition of the Attention Branch Network was found to decrease the overall performance metrics when using ResNet50 and ResNet152 [28]. On the contrary, ensemble learning outperformed other AI-based methods, achieving significantly better results than other existing algorithms [21]. Noteworthy, the use of different neural networks combined with transfer learning and fine-tuning resulted to achieve outstanding results compared to the use of basic CNNs [1,30].

When referring to peri-implant pathology detection, Chen and collaborators implemented a two-stage NN for both tasks. This way, they proposed an automatized system that allowed image enhancement to improve the visibility of both implants and gums. The overall accuracy rates related to each neural network were 89.3 % (YOLOv2) and 90.4 % (AlexNet), respectively [31]. Contrarily, Lee et al., reported promising results by only using one CNN (YOLOv7). These authors reported overall accuracy and sensitivity rates of over 94.74 % and 94.44 %, respectively [32]. Similarly, Vera et al., described a mean average precision rate of 84.4 % for this task [34]. Nevertheless, other authors reported worse results while using one neural network. Liu et al., described an overall bone loss implant detection rate of 67 % using Inception ResNet v2 [33].

Lastly, although all the studies evaluated the accuracy of different CNNs, only three compared their results with previously trained dental professionals [23,24,26]. Overall, all the references reported that the AI-based models outperformed humans. Park et al., reported an accuracy rate of 82 % for the neural network and 16.8 % and 43.3 % for both general practitioners and specialized clinicians, respectively [26].

Similarly, Lee et al., stated accuracy rates ranging from 90.8 % to 98.1 % for the NN compared to a mean of 72.2 % of all human participants (including board-certified periodontists, periodontology residents, and general practitioners) [24]. In another study, the authors described that, using both panoramic and periapical radiographs, the deep learning model achieved AUC values of 0.956 and 0.979, while the periodontist showed AUC rates of 0.891 and 0.959, respectively [23].

This systematic review only reported results based on 2D image analysis (mainly periapical and panoramic radiographs). Bearing in mind that three-dimensional X-rays are widely used in implantology, these radiographic images should also be considered to be included in further CNN testing and performance analysis. Also, although studies have significantly opened the range of implants included, further studies including vast datasets are needed to ensure generalizability. A meta-analysis was not feasible due to the lack of homogeneity among the selected studies, stemming from significant methodological variability, differences in dataset sizes and augmentation, inconsistent validation methods, diverse reference standards, and imbalanced representation of dental implant systems, all of which compromised the comparability of outcome metrics and overall validity of pooled data.

Lastly, all the technology used in the included studies is based on the use of individual and basic neural networks that impede fast-growing and efficient training and learning processes. Future studies should include the use of new forms of artificial intelligence, such as Federated Learning to ensure both generalizability and time-efficiency.

5. Conclusions

From this study, the following could be concluded:

1. Based on the reviewed literature, AI-based models, particularly convolutional neural networks, showed accuracy values over 78,6 % for peri-implant pathology detection.
2. Overall accuracy of the convolutional neural networks in dental implant classification ranged from 67 % to 99 %.
3. The exclusion of low-quality images and lack of diverse datasets could limit the real-world applicability of the model where variations in image annotation processes and the potential for overfitting highlight the need for standardized methodologies.

Statements relating to ethics and integrity policies

Start Date of the Systematic Review: August 2024.
PROSPERO Registration Number: CRD42024575604.
Authors have no conflict of interest to declare.
This article was carried out without any financial support.

CRedit authorship contribution statement

M Bonfanti-Gris: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **E Ruales:** Writing – review & editing, Validation, Methodology, Formal analysis, Data curation. **MP Salido:** Validation, Supervision, Project administration, Methodology, Conceptualization. **F Martinez-Rus:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **M Özcan:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **G Pradies:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This systematic review did not receive any financial support. Authors have no conflict of interest to declare.

References

- [1] H.-S. Kim, E.-G. Ha, Y.H. Kim, K.J. Jeon, C. Lee, S.-S. Han, Transfer learning in a deep convolutional neural network for implant fixture classification: a pilot study, *Imag. Sci. Dent.* 52 (2022) 219–224, <https://doi.org/10.5624/isd.20210287>.
- [2] T. Takahashi, K. Nozaki, T. Gonda, T. Mamenno, M. Wada, K. Ikebe, Identification of dental implants using deep learning-pilot study, *Int. J. Implant Dent.* 6 (2020) 53, <https://doi.org/10.1186/s40729-020-00250-6>.
- [3] M. Hadj Saïd, M.-K. Le Roux, J.-H. Catherine, R. Lan, Development of an artificial intelligence model to identify a dental implant from a radiograph, *Int. J. Oral Maxillofac. Implants* 36 (2020) 1077–1082, <https://doi.org/10.11607/jomi.8060>.
- [4] B.E. Pjetursson, D. Thoma, R. Jung, M. Zwahlen, A. Zembic, A systematic review of the survival and complication rates of implant-supported fixed dental prostheses (FDPs) after a mean observation period of at least 5 years, *Clin. Oral. Implants Res.* 23 (Suppl 6) (2012) 22–38, <https://doi.org/10.1111/j.1600-0501.2012.02546.x>.
- [5] K. Swierkot, P. Lottholz, L. Flores-de-Jacoby, R. Mengel, Mucoisitis, peri-implantitis, implant success, and survival of implants in patients with treated generalized aggressive periodontitis: 3- to 16-year results of a prospective long-term cohort study, *J. Periodontol.* 83 (2012) 1213–1225, <https://doi.org/10.1902/jop.2012.110603>.
- [6] S. Sukegawa, K. Yoshii, T. Hara, T. Matsuyama, K. Yamashita, K. Nakano, K. Takabatake, H. Kawai, H. Nagatsuka, Y. Furuiki, Multi-task deep learning model for classification of dental implant brand and treatment stage using dental panoramic radiograph images, *Biomolecules* 11 (2021), <https://doi.org/10.3390/biom11060815>.
- [7] A. Byraki, A. Costea, G. Curca, S. Hostiuc, Morphological analysis of dental implants - forensic significance, *Romanian J. Leg. Med.* 18 (2010) 207–212, <https://doi.org/10.4323/rjlm.2010.207>.
- [8] J. Berketa, R. Hirsch, D. Higgins, H. James, Radiographic recognition of dental implants as an aid to identifying the deceased, *J. Forensic Sci.* 55 (2009) 66–70, <https://doi.org/10.1111/j.1556-4029.2009.01226.x>.
- [9] H.-J. Kong, J.-Y. Yoo, J.-H. Lee, S.-H. Eom, J.-H. Kim, Performance evaluation of deep learning models for the classification and identification of dental implants, *J. Prosthet. Dent.* (2023), <https://doi.org/10.1016/j.prosdent.2023.07.009>. S0022-3913(23)00467-5.
- [10] I.G. Sahiwal, R.D. Woody, B.W. Benson, G.E. Guillen, Radiographic identification of threaded endosseous dental implants, *J. Prosthet. Dent.* 87 (2002) 563–577, <https://doi.org/10.1067/mp.2002.124430>.
- [11] G. Michelinakis, A. Sharrock, C.W. Barclay, Identification of dental implants through the use of Implant Recognition Software (IRS), *Int. Dent. J.* 56 (2006) 203–208, <https://doi.org/10.1111/j.1875-595x.2006.tb00095.x>.
- [12] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (2015) 436–444, <https://doi.org/10.1038/nature14539>.
- [13] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, J. Dean, A guide to deep learning in healthcare, *Nat. Med.* 25 (2019) 24–29, <https://doi.org/10.1038/s41591-018-0316-z>.
- [14] J. Gomez Rossi, N. Rojas-Perilla, J. Krois, F. Schwendicke, Cost-effectiveness of artificial intelligence as a decision-support system applied to the detection and grading of melanoma, dental caries, and diabetic retinopathy, *JAMA Netw. Open* 5 (2022) e220269, <https://doi.org/10.1001/jamanetworkopen.2022.0269>.
- [15] I.S. Bayrakdar, K. Orhan, S. Akarsu, Ö. Çelik, S. Atasoy, A. Pekince, Y. Yasa, E. Bilgir, H. Sağlam, A.F. Aslan, A. Odabaş, Deep-learning approach for caries detection and segmentation on dental bitewing radiographs, *Oral Radiol* 38 (2022) 468–479, <https://doi.org/10.1007/s11282-021-00577-9>.
- [16] M.T. Alharbi, M.M. Almutiq, Prediction of dental implants using machine learning algorithms, *J. Healthc. Eng.* 2022 (2022) 7307675, <https://doi.org/10.1155/2022/7307675>.
- [17] A.Y. Alqutaibi, R.S. Algabri, D. Elawady, W.I. Ibrahim, Advancements in artificial intelligence algorithms for dental implant identification: a systematic review with meta-analysis, *J. Prosthet. Dent.* (2023), <https://doi.org/10.1016/j.prosdent.2023.11.027>. S0022391323007837.
- [18] D. Moher, L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, L.A. Stewart, Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement, *Syst. Rev.* 4 (2015) 1, <https://doi.org/10.1186/2046-4053-4-1>.
- [19] F. Whiting Penny, A.W.S. Rutjes, M.E. Westwood, S. Mallett, J.J. Deeks, J. B. Reitsma, Ma.M.G. Leeflang, J.A.C. Sterne, P. Bossuyt, QUADAS-2 Group, QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies, *Ann. Intern. Med.* 155 (2011) 529–536.
- [20] H. Mohammad-Rahimi, S.R. Motamedian, M.H. Rohban, J. Krois, S.E. Uribe, E. Mahmoudinia, R. Rokhsad, M. Nadimi, F. Schwendicke, Deep learning for caries detection: a systematic review, *J. Dent.* 122 (2022) 104115, <https://doi.org/10.1016/j.jdent.2022.104115>.
- [21] H.-J. Kong, S.-H. Eom, J.-Y. Yoo, J.-H. Lee, Identification of 130 dental implant types using ensemble deep learning, *Int. J. Oral Maxillofac. Implants* 38 (2023) 150–156, <https://doi.org/10.11607/jomi.9818>.
- [22] H.J. Kong, Classification of dental implant systems using cloud-based deep learning algorithm: an experimental study, *J. Yeungnam Med. Sci.* 40 (2023) S29–S36, <https://doi.org/10.12701/jyms.2023.00465>.

- [23] J.J. Lee Seong-Nyum, Efficacy of deep convolutional neural network algorithm for the identification and classification of dental implant systems, using panoramic and periapical radiographs, *Open Med* 99 (2020) 1–7.
- [24] J.-H. Lee, Y.-T. Kim, J.-B. Lee, S.-N. Jeong, A performance comparison between automated deep learning and dental professionals in classification of dental implant systems from dental imaging: a multi-center study, *Diagn. Basel Switz.* 10 (2020), <https://doi.org/10.3390/diagnostics10110910>.
- [25] J.-H. Park, H.S. Moon, H.-I. Jung, J.J. Hwang, Y.-H. Choi, J.-E. Kim, Deep learning and clustering approaches for dental implant size classification based on periapical radiographs, *Sci. Rep.* 13 (2023), <https://doi.org/10.1038/s41598-023-42385-7>.
- [26] L.J.H. Park W Schwendicke F, Krois J, Huh JK, Identification of Dental Implant Systems Using a Large-Scale Multicenter Data Set, (n.d.), <https://doi.org/10.1177/00220345231160750>.
- [27] W.-S. Park, J.-K. Huh, J.-H. Lee, Automated deep learning for classification of dental implant radiographs using a large multi-center dataset, *Sci. Rep.* 13 (2023), <https://doi.org/10.1038/s41598-023-32118-1>.
- [28] S. Sukegawa, K. Yoshii, T. Hara, F. Tanaka, K. Yamashita, T. Kagaya, K. Nakano, K. Takabatake, H. Kawai, H. Nagatsuka, Y. Furuki, Is attention branch network effective in classifying dental implants from panoramic radiograph images by deep learning? *PLoS ONE* 17 (2022) <https://doi.org/10.1371/journal.pone.0269016>.
- [29] S. Sukegawa, K. Yoshii, T. Hara, T. Matsuyama, K. Yamashita, K. Nakano, K. Takabatake, H. Kawai, H. Nagatsuka, Y. Furuki, Multi-task deep learning model for classification of dental implant brand and treatment stage using dental panoramic radiograph images, *Biomolecules* 11 (2021), <https://doi.org/10.3390/biom11060815>.
- [30] S. Sukegawa, K. Yoshii, T. Hara, K. Yamashita, K. Nakano, N. Yamamoto, H. Nagatsuka, Y. Furuki, Deep neural networks for dental implant system classification, *Biomolecules* 10 (2020), <https://doi.org/10.3390/biom10070984>.
- [31] YC Chen, PAR. Abu, MY Chen, TY Chen, ML Chan, YY Huang, YL Liu, PT Lee, GJ Lin, TF Li, CA Chen, SL Chen, KC Li, Improving dental implant outcomes: CNN-Based system accurately measures degree of peri-implantitis damage on periapical film, *Bioengineering* 10 (2023), <https://doi.org/10.3390/bioengineering10060640>.
- [32] W.-F. Lee, M.-Y. Day, C.-Y. Fang, V. Nataraj, S.-C. Wen, W.-J. Chang, N.-C. Teng, Establishing a novel deep learning model for detecting peri-implantitis, *J. Dent. Sci.* 19 (2024) 1165–1173, <https://doi.org/10.1016/j.jds.2023.11.017>.
- [33] M. Liu, S. Wang, H. Chen, Y. Liu, A pilot study of a deep learning approach to detect marginal bone loss around implants, *BMC Oral Health* 22 (2022) 11, <https://doi.org/10.1186/s12903-021-02035-8>.
- [34] M. Vera, M.J. Gómez-Silva, V. Vera, C.I. López-González, I. Aliaga, E. Gascó, M. Pedrera-Canal, E. Besada-Portas, G. Pajares, Artificial intelligence techniques for automatic detection of peri-implant marginal bone remodeling in intraoral radiographs, *J. Digit. Imaging* 36 (2023) 2259–2277, <https://doi.org/10.1007/s10278-023-00880-3>.
- [35] G.Y. Zhang C Fan L, Zhang S, Zhao J., Deep learning based dental implant failure prediction from periapical and panoramic films, (n.d.), <https://doi.org/10.21037/qims-22-457>.
- [36] N. Adnan, M. Hanif, K. Khan, F. Faridooon, F. Umer, An Artificial Intelligence model for implant segmentation on periapical radiographs, *JPMA J. Pak. Med. Assoc.* 74 (2024) S5–S9, <https://doi.org/10.47391/JPMA.AKU-9S-02>.
- [37] A.-F.M.A. Hassan NA Kamel AE, Omran AE, Gad MW, Ashraf NM, Ahmed OM, Automated identification of dental implants: a new, fast and accurate artificial intelligence system, (n.d.), https://doi.org/10.1922/EJPRD_2620Hassan06.
- [38] C.Y. Hsiao, H. Bai, H. Ling, J. Yang, Artificial intelligence in identifying dental implant systems on radiographs, *Int. J. Periodontics Restorative Dent.* 43 (2023) 363–368, <https://doi.org/10.11607/prd.5781>.
- [39] A. Kohlakala, J. Coetzer, J. Bertels, D. Vandermeulen, Deep learning-based dental implant recognition using synthetic X-ray images, *Med. Biol. Eng. Comput.* 60 (2022) 2951–2968, <https://doi.org/10.1007/s11517-022-02642-9>.
- [40] R.P.D.M. Santos, H.E.V.O. Prado, I.S.A. Neto, G.A.A. de Oliveira, A.I.V. Silva, E. G. Zenóbio, F.R. Manzi, Automated identification of dental implants using artificial intelligence, *Int. J. Oral Maxillofac. Implants* 36 (2021) 918–923, <https://doi.org/10.11607/jomi.8684>.
- [41] V.B. Benakatti, R.P. Nayakar, M. Anandhalli, Machine learning for identification of dental implant systems based on shape - A descriptive study, *J. Indian Prosthodont. Soc.* 21 (2021) 405–411, <https://doi.org/10.4103/jips.jips.324.21>.
- [42] I. Kurtulus, M. Lubbad, O. Yilmaz, K. Kilic, D. Karaboga, A. Basturk, B. Akay, U. Nalbantoglu, S. Yilmaz, M. Ayata, I. Pacal, A robust deep learning model for the classification of dental implant brands, *J. Stomatol. Oral Maxillofac. Surg.* (2024) 101818, <https://doi.org/10.1016/j.jormas.2024.101818>.
- [43] M. Lubbad, I. Kurtulus, D. Karaboga, K. Kilic, A. Basturk, B. Akay, O. Nalbantoglu, O. Yilmaz, M. Ayata, S. Yilmaz, I. Pacal, A comparative analysis of deep learning-based approaches for classifying dental implants decision support system, *J. Imaging Inform. Med.* (2024), <https://doi.org/10.1007/s10278-024-01086-x>.
- [44] O.I.Y. Tiryaki B Ozdogan A, Guller MT, Miloglu O, Oral EA, Dental implant brand and angle identification using deep neural networks, (n.d.), <https://doi.org/10.1016/j.prosdent.2023.07.022>.
- [45] J. Guo, P.-W. Tsai, X. Xue, D. Wu, Q.T. Van, C.N. Kaluarachchi, H.T. Dang, N. Chintha, TVGG dental implant identification system, *front. Pharmacol.* 13 (2022), <https://doi.org/10.3389/fphar.2022.948283>.
- [46] Á. García-Canas, M. Bonfanti-Gris, S. Paraiso-Medina, F. Martínez-Rus, G. Pradies, Diagnosis of interproximal caries lesions in bitewing radiographs using a deep convolutional neural network-based software, *Caries Res* (2022), <https://doi.org/10.1159/000527491>.
- [47] S. Hajera Begum, P. Vidyullatha, A novel lightweight deep learning hybrid CNN model for automatic detection of oral squamous cell carcinoma (OSCC) using histopathological images, *Neuro Quantol.* 20 (2022) 7028–7046, <https://doi.org/10.14704/nq.2022.20.8.NQ44728>.
- [48] S. Oh, Y.J. Kim, J. Kim, J.H. Jung, H.J. Lim, B.C. Kim, K.G. Kim, Deep learning-based prediction of osseointegration for dental implant using plain radiography, *BMC Oral Health* 23 (2023), <https://doi.org/10.1186/s12903-023-02921-3>.
- [49] H.-R. Choi, T.S. Siadari, J.-E. Kim, K.-H. Huh, W.-J. Yi, S.-S. Lee, M.-S. Heo, Automatic detection of teeth and dental treatment patterns on dental panoramic radiographs using deep neural networks, *Forensic Sci. Res.* 7 (2022) 456–466, <https://doi.org/10.1080/20961790.2022.2034714>.
- [50] A.E. Yüksel, S. Gültekin, E. Simsar, Ş.D. Özdemir, M. Gündoğar, S.B. Tokgöz, İ. E. Hamamcı, Dental enumeration and multiple treatment detection on panoramic X-rays using deep learning, *Sci. Rep.* 11 (2021) 12342, <https://doi.org/10.1038/s41598-021-90386-1>.