Arzani et al. BMC Medical Imaging (2025) 25:174
https://doi.org/10.1186/s12880-025-01719-9

BMC Medical Imaging

SYSTEMATIC REVIEW Open Access

Detection of carotid artery calcifications using artificial intelligence in dental radiographs: a systematic review and meta-analysis

Sarah Arzani[1*], Parisa Soltani[2], Ali Karimi[3*], Maryam Yazdi[1], Ashraf Ayoub[4], Zohaib Khurshid[5,6], Domenico Galderisi[7] and Hugh Devlin[8,9]

Abstract

Background Carotid artery calcifications are important markers of cardiovascular health, often associated with atherosclerosis and a higher risk of stroke. Recent research shows that dental radiographs can help identify these calcifications, allowing for earlier detection of vascular diseases. Advances in artificial intelligence (AI) have improved the ability to detect carotid calcifications in dental images, making it a useful screening tool. This systematic review and meta-analysis aimed to evaluate how accurately AI methods can identify carotid calcifications in dental radiographs.

Materials and methods A systematic search in databases including PubMed, Scopus, Embase, and Web of Science for studies on AI algorithms used to detect carotid calcifications in dental radiographs was conducted. Two independent reviewers collected data on study aims, imaging techniques, and statistical measures such as sensitivity and specificity. A meta-analysis using random effects was performed, and the risk of bias was evaluated with the QUADAS-2 tool.

Results Nine studies were suitable for qualitative analysis, while five provided data for quantitative analysis. These studies assessed AI algorithms using cone beam computed tomography ($n=3$) and panoramic radiographs ($n=6$). The sensitivity of the included studies ranged from 0.67 to 0.98 and specificity varied between 0.85 and 0.99. The overall effect size, by considering only one AI method in each study, resulted in a sensitivity of 0.92 [95% CI 0.81 to 0.97] and a specificity of 0.96 [95% CI 0.92 to 0.97].

Conclusions The high sensitivity and specificity indicate that AI methods could be effective screening tools, enhancing the early detection of stroke and related cardiovascular risks.

Clinical trial number Not applicable.

Keywords Artificial intelligence, Atherosclerotic plaque, Dental radiograph, Carotid artery thrombosis, Machine learning

*Correspondence:
Sarah Arzani
sa.arzan@yahoo.com
Ali Karimi
a2022karimi@gmail.com
Full list of author information is available at the end of the article

Introduction

Stroke is the third most common cause of death globally, responsible for about 10% of total fatalities according to the World Health Organization (WHO) [1]. A considerable contributor to stroke risk is atherosclerotic carotid plaques, which are present in about 15% of cases [2]. Carotid artery calcification appears on various radiographs like lateral cephalometric, panoramic or lateral neck radiographs as an irregular nodular radiopacity situated just below the mandibular angle and the hyoid bone, near the cervical vertebrae at the C3–C4 intervertebral space [3]. These calcifications are notable indicators of cardiovascular disease, often pointing to the presence of atherosclerosis and a higher risk of stroke and other vascular problems [3, 4].

Early detection of atherosclerotic carotid plaques is crucial for preventing strokes, enabling healthcare providers to manage at-risk patients before serious damage occurs [5]. Traditionally, identifying carotid calcifications has relied on imaging techniques like ultrasound and computed tomography (CT) [5, 6]. However, recent studies have indicated that dental radiographs, including panoramic images and cone beam computed tomography (CBCT) scans, can reveal important incidental findings, such as carotid calcifications, making dental imaging a valuable tool for screening cardiovascular risk [7, 8]. Since these radiographic examinations are commonly performed by dentists and oral and maxillofacial surgeons, they offer a practical way to identify potential health concerns.

Advancements in artificial intelligence (AI) are revolutionizing diagnostic practices across various fields of medical and dental sciences [9–11]. Recent advances in machine learning, particularly deep learning architectures such as convolutional neural networks (CNNs), have demonstrated remarkable capabilities in analyzing complex medical imaging data with precision that frequently exceeds human performance. These technologies extract detailed and objective information from images, significantly improving disease detection accuracy [9, 12–14]. Decision-making models have the potential to enhance computerized analysis, enabling the acquisition of accurate and consistent data quickly, which can then be used to develop treatment strategies. However, this approach to computerized diagnosis and treatment planning remains in its early stages, despite various technological advancements in AI [13].

Several research papers have used AI to automatically segment carotid artery calcification in magnetic resonance (MR), CT, and ultrasound images [15–17]. In dental radiology, AI applications now encompass the detection of pathological conditions, identification of anatomical structures, and recognition of incidental findings, offering substantial potential to enhance diagnostic accuracy and patient outcomes [14, 18, 19]. By applying AI to dental radiographs, the identification of carotid calcifications can be improved, utilizing readily available dental imaging resources.

This systematic review aimed to assess the performance of AI-based detection methods for carotid calcifications in dental radiographs, synthesizing findings from existing literature to answer the question "Are AI algorithms effective for the detection of carotid artery calcifications?" by including a meta-analysis and reporting diagnostic accuracy, sensitivity, and specificity. The authors hypothesize that AI-assisted detection would demonstrate clinically acceptable performance for identifying carotid calcifications, potentially serving as a valuable decision-support tool in dental practice.

Materials and methods

The protocol for this review was registered with PROSPERO under the registration number #CRD42024595866. The review methodology and results were reported accurately and transparently by adhering to the Preferred Reporting Items for Overviews of Reviews (PRIOR), the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies (PRISMA-DTA), and PRISMA-AI criteria [20–22].

Eligibility criteria

"Are artificial intelligence (AI) algorithms effective for carotid calcifications detection in dental radiographs?" was the primary research question formulated for this study based on the PIRD criteria (the population, index test, reference test, and diagnosis of interest) which are used to choose studies on diagnostic accuracy [23]. This review used the following PIRD framework: Population: Adults who have carotid atherosclerosis and underwent dental imaging. Index Test: The use of AI algorithms for detecting carotid calcifications in dental radiographs (Fig. 1). Reference Test: Assessment of carotid calcifications based on ultrasound images or dental radiographs by radiologists or experienced dentists. Diagnosis of Interest: The diagnostic accuracy of detecting atherosclerotic plaques using AI models, including sensitivity, specificity, log diagnostic odds ratio, and the area under the receiver operating characteristic (ROC) curve (AUC).

The inclusion criteria for this review consisted of studies utilizing AI algorithms specifically for detecting carotid calcifications in dental radiographs including panoramic, CBCT, and cephalometry. Studies that did not focus on carotid calcifications, non-diagnostic studies, case reports, and reviews were excluded. Additionally, guidelines, comments, editorials, duplicate publications, conference papers, and abstracts lacking full-text availability were also omitted from our analysis.

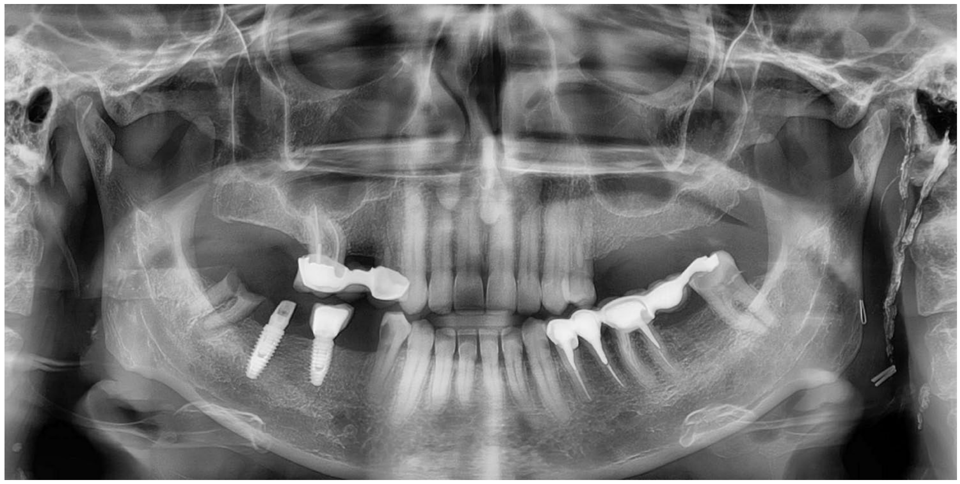

Fig. 1 Panoramic image with severe carotid artery calcification on the left side accompanied by vascular clips

Search strategy

A comprehensive literature search (Table S1) was conducted across five databases up to November 11, 2024, including PubMed (Medline), Scopus, Embase, Cochrane Library and Web of Science. The search strategy combined keywords with syntax adapted to each platform's search rules (e.g., MeSH terms for PubMed, Emtree for Embase) to create a highly sensitive method for identifying potential records. Additionally, the reference lists of the selected research articles and relevant previous studies were reviewed, along with a supplementary search on Google Scholar to locate any other potentially qualifying studies.

Study selection

Two independent reviewers (SA and AK) conducted title and abstract screening after removing duplicate papers using Endnote 21 (Clarivate, Philadelphia, USA). Full-text relevant records were obtained and reviewed based on inclusion and exclusion criteria. During the screening process, text mining was performed using the SWIFT-Review software (Sciome LLC, NC, USA), which automatically groups abstracts on similar topics through machine learning techniques [24]. We utilized AI algorithms within SWIFT-Review to search, categorize, and prioritize a large volume of primary studies during this stage [11]. However, final inclusion decisions were made based on human judgment. Disagreements during the screening process were addressed by reaching a consensus between the two reviewers or by consulting a third author (PS).

Data extraction

Data extraction was performed by the same two independent authors (SA and AK). Various details from each included study were collected, such as the first author and year of publication, study objective, type of dental imaging modality used, sample size, image augmentations applied, software or hardware utilized, the formulation task applied, type of AI algorithm employed, and key findings, including statistical measures like sensitivity, specificity, and accuracy. Disputes in the data extraction phase were settled with the help of a third author (MY).

Statistical analysis

For the diagnostic meta-analysis, true positive (TP), true negative (TN), false positive (FP), and false negative (FN) counts from each eligible study were extracted. For those studies that reported sensitivity, specificity, and standard errors, total negative and positive values were calculated, and other required measures were derived. The summary receiver operating characteristic (SROC) curve was obtained using the "metandi" command in Stata 17. Synthetizing effect sizes was done using random effects meta-analysis considering just one effect size for every study. Heterogeneity across the enrolled studies was evaluated by Cochran Q-statistic and I^2 statistic [18]. Subgroup analyses and sensitivity analyses were also performed according to image modality, leaving out one study per time. Publication bias was evaluated using Egger's and Begg's tests. Significant probability level was considered as $p < 0.05$. Stata Statistical Software version 17 (College Station, TX: StataCorp LLC) was used for meta-analysis.

Risk of bias assessment

The risk of bias in the included studies was assessed by two independent researchers (SA and MY) utilizing the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool. This evaluation focused on key domains including patient selection, index test, reference standard, and flow and timing. A third reviewer (ZK) was brought in to help resolve conflicts between the assessors.

Results

Study selection

The initial search identified 1430 studies. After removing duplicates and screening titles and abstracts, 14 records were deemed suitable for full-text review. Ultimately, after excluding 5 records (Table S2), 9 studies met the eligibility criteria for qualitative analysis, while 5 of these provided data for quantitative analysis (Fig. 2).

Study characteristics

The characteristics of the studies are summarized in Table 1. The included studies covered a range of publication years, with one study conducted before 2022 [5], three conducted in 2022 [4, 25, 26], two conducted in 2023 [6, 9], and three conducted in 2024 [27–29]. These studies assessed the performance of AI algorithms in detecting carotid calcification using CBCT ($n = 3$) [9, 25, 28] and panoramic radiographs ($n = 6$) [4–6, 26, 27, 29].

The reviewed studies employed diverse deep learning approaches for carotid calcification detection, with convolutional neural networks (CNNs) emerging as the predominant methodology. Among these, architectures such as Faster R-CNN (ResNet-based), U-Net, and Inception-based models were most frequently implemented [4–6, 25–29]. The field has seen growing adoption of transfer learning techniques utilizing pre-trained models including ResNet, EfficientNet, and DenseNet architectures [4, 6, 25, 27]. Hybrid architectures demonstrated superior performance, as shown by Inception V3 + U-Net (accuracy: 96.35%; sensitivity: 94.2%) [25] and EfficientNet-B4

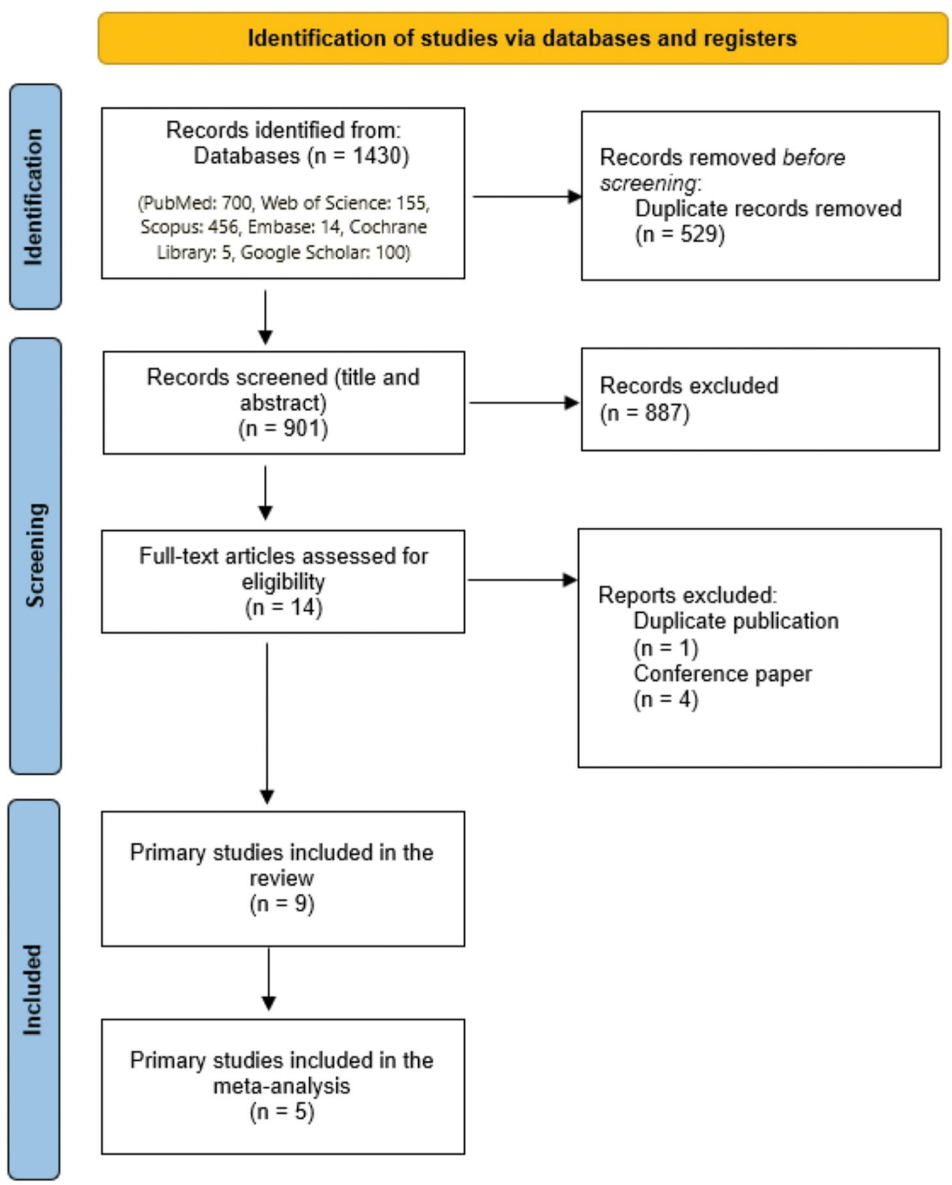

Fig. 2 Systematic review flowchart diagram according to PRISMA 2020

Table 1 Study characteristics of included primary studies

Study	Aim	Image modality	Sample size (Train/Valid/Test)	Reference standard	Augmentations	Formulation task (classification/detection/segmentation)	AI techniques	Findings
Kats L et al. [5], 2018	Detection of atherosclerotic carotid plaques on panoramic radiographs using DL	Panoramic radiographs	65 OPGs (Train: 70%, Valid: 10%, Test: 20%)	Radiologist	Brightness adjustments, cropping, random flipping, and random rotation	Segmentation	Faster R-CNN (Based on ResNet101)	Acc: 83.0%, Se: 75.0%, Sp: 80.0%, AUC: 83.0%
Ajami M et al. [25], 2022	Detection of cervical carotid artery calcifications in CBCT images using DCNN	CBCT	56 CBCTs, axial slices: 15,257 (Train: 40 CBCTs (11,252 axial slices), Test: 16 CBCTs (4005 axial slices)	Radiologist	Random rotation, skewing, flipping, scaling, and brightness adjustments	Classification and segmentation	Inception V3 + U-Net	Acc: 96.35%, Se: 94.2%, Sp: 96.5%, Precision: 56.9%, NPV: 99.7%
Bin Song Y et al. [4], 2022	Comparison of detection performance of soft tissue calcifications using AI in panoramic radiography	Panoramic radiographs	176 OPGs (Test)	Definitive previous diagnosis	NA	Classification and detection	Fast R-CNN (Based on ResNet)	Acc: 72.7%, Se: 77.4%, Sp: 71.7%
Lee S et al. [26], 2022	Detecting 17 fine-grained dental anomalies from panoramic radiographs using AI	Panoramic radiographs	22,999 OPGs (NA)	Radiologist	NA	Classification and detection	Faster R-CNN	Se: 95.8%, Sp: 95.9%, Precision: 42.3%
Amitay M et al. [6], 2023	Screening carotid calcification in dental panoramic radiographs using DCNN	Panoramic radiographs	500 patients, 480 clean and 179 CAC corners (Test: 24 CAC and 120 clean corners, Valid: 24 CAC and 120 clean corners, Train: 131 CAC and 240 clean corners)	Radiologist	Width shift, zoom, rotation, and shear	Classification	DenseNet169	Acc: 93.0%, Se: 80.0%, Sp: 96.0%, Precision: 81.0%, F1-score: 79.0%
							InceptionResNetV2	Acc: 94.0%, Se: 82.0%, Sp: 97.0%, Precision: 84.0%, F1-score: 82.0%

Table 1 (continued)

Study	Aim	Image modality	Sample size (Train/Valid/Test)	Reference standard	Augmentations	Formulation task (classification/ detection/ segmentation)	AI techniques	Findings	
							InceptionResNetV2 + XG-Boost	Acc: 91.0%, Se: 68.0%, Sp: 91.0%, Precision: 73.0%, F1-score: 80.0%	
							EfficientNetV2M	Acc: 93.0%, Se: 78.0%, Sp: 96.0%, Precision: 80.0%, F1-score: 79.0%	
Alajaji SA et al. [9], 2023	Detection of extracranial and intracranial calcified carotid artery atheromas in CBCT using DCNN	CBCT	137 CBCTs (Train: 60%, Valid: 10%, Test: 30%)	Radiologist	Shift and rotation	Classification and segmentation	DL CNN model: Transfer learning via a U-Net-based neural network architecture (Olympus TruAI Deep Learning)	Extracranial CAC: Acc: 55.0%, Se: 92.0%, Sp: 52.0%, AUC: 72.0%, Precision: 13.0%, NPV: 99.0%, JI: 12.0%, F1-score: 22.0% Intracranial CAC: Acc: 17.0%, Se: 96.0%, Sp: 15.0%, AUC: 55.0%, Precision: 02.0%, NPV: 99.0%, JI: 02.0%, F1-score: 04.0%	
Yoo SW et al. [27], 2024	Classification and segmentation of carotid artery calcification on panoramic radiographs using DL	Panoramic radiographs	400 OPGs (Train: 60%, Valid: 20%, Test: 20%)	CT reference images	Cropping, flipping, rotation, width and height shift in horizontal and vertical axes, and zoom	Classification	Cascaded deep learning network (CACSNet)	VGG16	Acc: 92.3%, Se: 88.0%, Sp: 95.0%, AUC: 97.2%

Table 1 (continued)

Study	Aim	Image modality	Sample size (Train/Valid/Test)	Reference standard	Augmentations	Formulation task (classification/detection/segmentation)	AI techniques	Findings	
								MobileNet v2	Acc: 96.2%, Se: 92.0%, Sp: 98.8%, AUC: 99.6%
								ResNet101	Acc: 95.4%, Se: 94.0%, Sp: 96.3%, AUC: 99.2%
								DenseNet121	Acc: 96.9%, Se: 94.0%, Sp: 98.8%, AUC: 99.5%
								EfficientNet-B4	Acc: 98.5%, Se: 98.0%, Sp: 98.8%, AUC: 99.6%
						Segmentation		VGG16	Se: 70.8%, Precision: 67.4%, JI: 51.4%, F1-score: 64.3%, ICC: 56.9%
								ResNet101	Se: 71.9%, Precision: 74.3%, JI: 55.8%, F1-score: 68.7%, ICC: 77.2%

Table 1 (continued)

Study	Aim	Image modality	Sample size (Train/Valid/Test)	Reference standard	Augmentations	Formulation task (classification/ detection/ segmentation)	AI techniques	Findings	
								DenseNet121	Se: 71.9%, Precision: 69.7%, JI: 53.8%, F1-score: 66.2%, ICC: 67.8%
								SENet	Se: 72.6%, Precision: 70.6%, JI: 56.6%, F1-score: 68.8%, ICC: 66.0%
								EfficientNet-B4	Se: 75.6%, Precision: 74.9%, JI: 59.5%, F1-score: 72.2%, ICC: 77.2%
Nelson J et al. [28], 2024	Detection of external carotid artery calcifications using CNN	CBCT	427 CBCTs (Train: 80% (342), Test: 20% (85))	Radiologist	Rotation	Binary classification	CNN	External CAC: Validation Acc: 76.0%, Se: 66.0%, Precision: 79.0%, F1-score: 72.0%, MCC: 53.0%	
Vinayahalingam S et al. [29], 2024	Detection of carotid plaques on panoramic radiographs using DL	Panoramic radiographs	370 (Train: 80%, Valid: 10%, Test: 10%)	Radiologist + clinical and additional radiological examinations	Random horizontal flipping and resizing	Binary classification	Fast R-CNN + ResNet-50	Se: 32.4%, Sp: 62.2%, Precision: 46.2%, F1-score: 38.1%, AUC: 46.5%	

Table 1 (continued)

Study	Aim	Image modality	Sample size (Train/Valid/Test)	Reference standard	Augmentations	Formulation task (classification/detection/segmentation)	AI techniques	Findings
							Faster R-CNN + ResNet-50	Se: 75.7%, Sp: 77.3% Precision: 76.9%, F1-score: 76.3%, AUC: 84.2%
							Faster R-CNN + Swin-T	Se: 88.1%, Sp: 89.7% Precision: 89.5%, F1-score: 88.8%, AUC: 95.0%

Abbreviations: DL, Deep Learning; CNN, Convolutional Neural Network; DCNN, Deep Convolutional Neural Network; OPG, Ortho Pantomo Gram; CBCT, Cone-Beam Computed Tomography; OPG, Orthopantomogram; Faster R-CNN, Faster Region-based Convolutional Neural Network; Fast R-CNN, Fast Region-based Convergence Neural Network; Swin-T, Swin Transformer; NA, Not Available; CAC, Carotid Artery Calcification; Acc, Accuracy; Se, Sensitivity; Sp, Specificity; AUC, Area Under Curve; NPV, Negative Predictive Value; JI, Jaccard Index; ICC, Intraclass Correlation Coefficient; MCC, Matthews Correlation Coefficient

which achieved the highest standalone model performance (accuracy: 98.5%; sensitivity: 98.0%) [27].

Quality assessment

The quality assessment of the included studies, summarized in Fig. 3 using the QUADAS-2 tool, reveals a generally favorable profile concerning risk of bias and applicability concerns. Most studies exhibit low risk across all evaluated domains, indicating robust methodological quality [4, 6, 9, 26, 27]. Two studies [5, 25] show unclear risk in the patient selection domain of bias, while one study [28] showed an unclear risk in the index test (bias) and reference test (applicability) domains, along with a high risk of bias in the reference standard domain of bias. Additionally, Vinayahalingam et al. [29] demonstrated low risk in most areas, though they were marked with high risk in the timing of flow and patient selection. Overall, the majority of studies demonstrate low risk of bias, supporting their reliability for inclusion in the meta-analysis.

Meta-analysis

Four studies reported the numbers of TP, TN, FP, and FN [25–28], while one study provided sensitivity, specificity, and standard errors [6]. These values were used for the meta-analysis. Figure 4 displays a forest plot summarizing the sensitivity and specificity (with 95% confidence intervals (CI)) across studies employing diverse AI methods. The sensitivity ranged from 0.67 (indicating moderate performance in one study) to 0.98 (excellent detection in others), while specificity remained consistently high (0.85–0.99). Figures 5 and 6 present subgroup analyses stratified by dental radiography type (CBCT and panoramic radiograph), with each study restricted to one AI method to avoid redundancy. Figure 5 specifically displays the diagnostic odds ratio (DOR) for each subgroup, with the overall pooled estimate reaching 241.46 [95% CI 39.53 to 1475.03]. This extreme width suggests substantial heterogeneity between studies, possibly due to varying image resolutions or clinical settings. Figure 6 consists of two complementary forest plots analyzing sensitivity and specificity separately for each radiography subgroup. The overall effect size, by considering only one AI method in each study, resulted in a sensitivity of 0.92 [95% CI 0.81 to 0.97] and a specificity of 0.96 [95% CI 0.92 to 0.97]. Figure 7 presents the summary receiver operating characteristic (SROC) curve, which plots the true positive rate (sensitivity) against the false positive rate (1-specificity) across all studies. Sensitivity analysis by leaving out each study one at a time did not change the pooled DOR substantially (Figure S1). There was no evidence of publication bias in reported DORs (Egger's $p = 0.332$, Begg's $p = 1.0$).

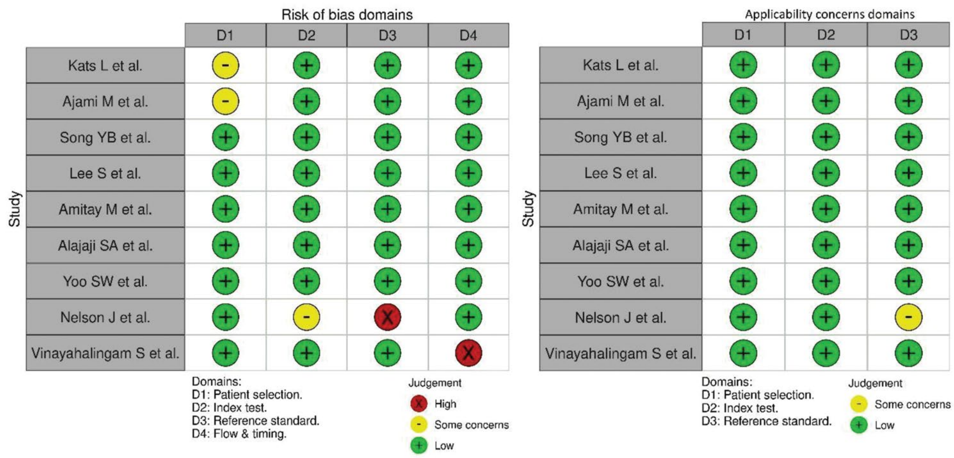

Fig. 3 Quality assessment of included studies according to QUADAS-2

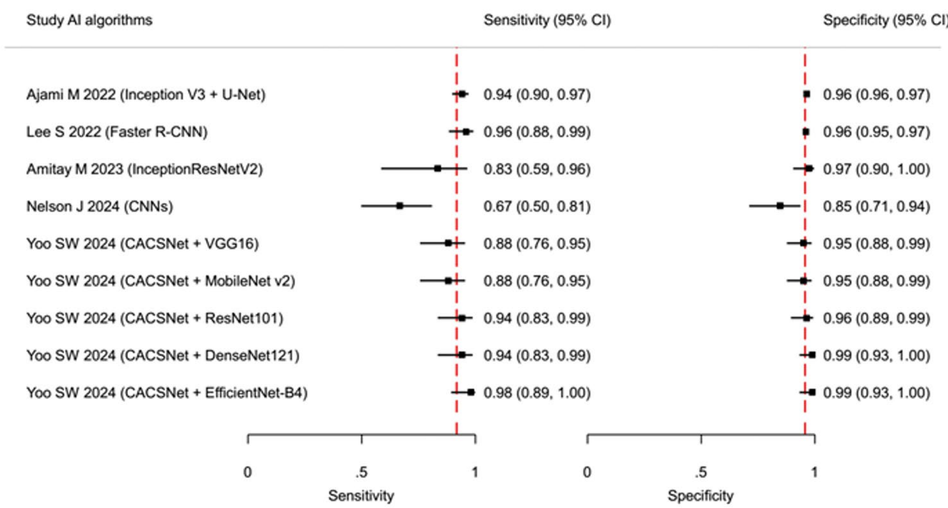

Fig. 4 Forest plot of sensitivity and specificity (95% confidence interval) of AI methods across included studies. Each row represents an individual study, the square markers indicate the point estimates of sensitivity and specificity, while the horizontal lines represent the 95% confidence intervals

Discussion

This systematic review and meta-analysis provide a comprehensive evaluation of the diagnostic performance of AI models in detecting carotid calcification on dental radiographs. The included studies covered various AI techniques, particularly focusing on convolutional neural networks (CNNs). The results indicate that AI models, when applied to CBCT and panoramic radiographs, demonstrate high sensitivity and specificity, reaching 0.92 and 0.96 respectively, in identifying this condition. This substantial diagnostic performance indicates the potential of AI models to accurately identify both true positive and true negative cases of carotid calcification. Additionally, the sensitivity analysis suggests that the pooled DOR of 241.45 is robust and not significantly influenced by any individual study. Moreover, the absence of publication bias in the reported DORs further strengthens the reliability of the findings.

Atherosclerosis in the carotid arteries is a complex process that involves several stages and mechanisms. This process begins with damage to the endothelial cells of the arterial walls, caused by factors such as high blood pressure, smoking, high cholesterol, and diabetes [30, 31]. Once the endothelium is damaged, low-density lipoproteins (LDL) penetrate the arterial wall and become oxidized, forming oxidized LDL (oxLDL), which in turn initiates an inflammatory response. Immune cells, particularly macrophages, engulf the oxLDL, becoming foam cells [32]. These foam cells accumulate and form fatty streaks, which are the earliest visible signs of atherosclerosis. Over time, the fatty streaks evolve into more complex atherosclerotic plaques. These plaques consist of a lipid core, a fibrous cap, and calcified areas [33]. The plaques can grow large enough to obstruct blood flow or become unstable. Unstable plaques can rupture, exposing the thrombogenic material within the plaque to the

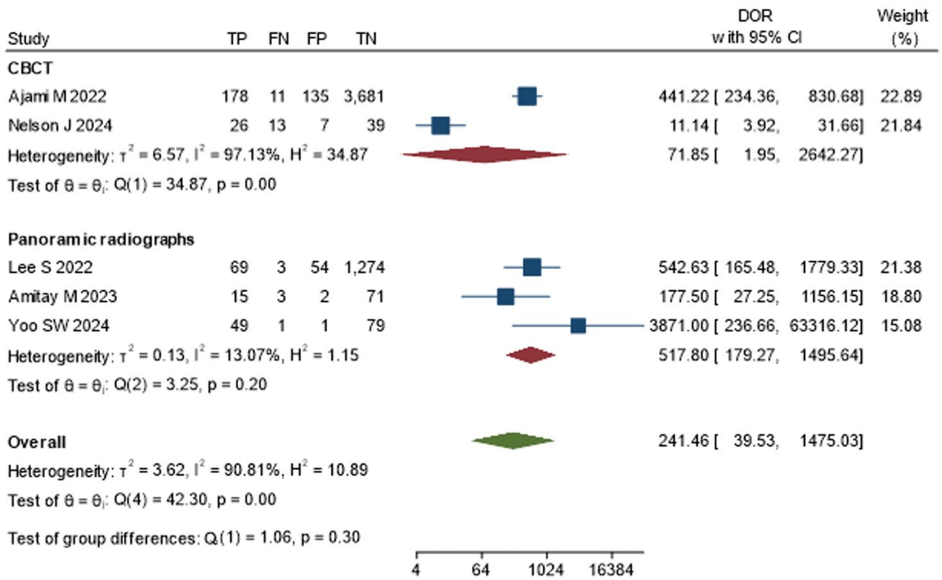

Fig. 5 Forest plot of diagnostic odds ratios (DOR) with 95% confidence intervals for AI-based detection of carotid artery calcifications across different dental radiograph types (CBCT and panoramic). Each row lists an included study with a corresponding blue square representing the point estimate of DOR and a horizontal line indicating its 95% confidence interval. The size of the square reflects the study's weight in the meta-analysis. Subgroup results (CBCT, panoramic radiographs) are shown with red diamonds, and the overall pooled estimate is shown with a green diamond. Larger DOR values indicate better diagnostic performance. Heterogeneity statistics (τ^2, I^2, H^2) are provided for each group and overall

bloodstream. This can lead to the formation of a thrombus, which can further obstruct blood flow or break off and travel to the brain, causing a stroke [34]. Carotid artery calcification is a significant risk factor for cardiovascular disease. Traditional methods for detecting these calcifications, such as ultrasound, CT angiography, and magnetic resonance imaging (MRI), can be invasive and costly [35, 36]. Dental panoramic radiographs, on the other hand, are widely available and non-invasive.

Before the advent of deep learning, traditional methods were employed to detect coronary artery calcifications on panoramic radiographs. These methods included fuzzy image contrast enhancement and algebraic image operators, which aimed to highlight calcified regions. These approaches achieved a detection rate of 50% [37]. Harada et al. [38] improved detection accuracy by incorporating a support vector machine, reducing false positives by 75%. Another method by Sawagashira et al. [39], using a top-hat filter, achieved a sensitivity of 93.6% with 4.4 false positives per image by using a rule-based approach and support vector machine to minimize false positives.

Recent advancements in AI, specifically the introduction of deep learning paved the way for automating the detection of carotid artery calcifications on dental radiographs. Among the included studies in this review, the study by Kats et al. [5] in 2018 was one of the pioneering efforts to use deep learning for detecting atherosclerotic carotid plaques in panoramic radiographs. Despite working with a relatively small dataset of 65 panoramic images, the Faster R-CNN model achieved an accuracy of 83.0%, a sensitivity of 75.0%, a specificity of 80.0%, and an AUC of 83.0%. Future studies utilized larger datasets and explored the use of CBCT images to achieve more reliable diagnoses in three-dimensional images. In general, CNNs, particularly with architecture like Faster R-CNN, U-Net, and Inception-based models were the preferred deep-learning approach [6, 9, 25, 26]. Additionally, the use of transfer learning by employment of pre-trained models such as ResNet and EfficientNet is gaining prominence [4, 6, 27].

Considering imaging modalities, six studies used datasets containing panoramic radiographs for the detection of carotid artery calcifications [4–6, 26, 27, 29]. The datasets used in these studies for developing models used numbers of images that varied from 65 panoramic images used by Kats et al. [5] to 22,999 radiographs in the study of Lee et al. [26]. Interestingly, Yoo et al. [27] developed a cascaded deep learning network (CACSNet) consisting of classification and segmentation networks for carotid calcifications on panoramic images. CACSNet with EfficientNet-B4 outperformed other tested models, with a classification accuracy of 98.5% and segmentation sensitivity and precision of 75.6% and 74.9%, respectively. Moreover, Vinayahalingam et al. [29] compared the performance of different backbone networks (ResNet-50 and Swin-T) with Fast R-CNN and Faster R-CNN on a dataset of 370 panoramic radiographs. Faster R-CNN + Swin-T demonstrated superior performance with an F1 score of 88.8%, sensitivity of 88.1%, and specificity of 89.7%. Three studies used CBCT images to detect carotid

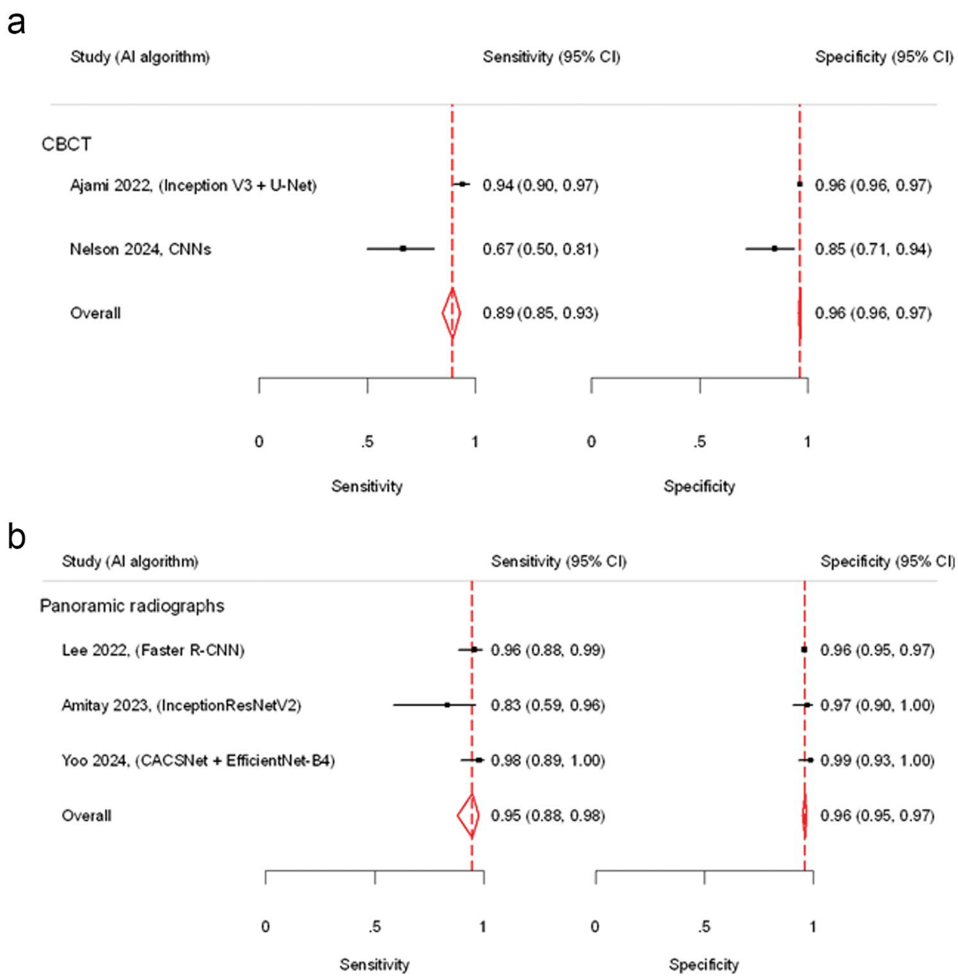

Fig. 6 **(a)** Forest plot of sensitivity and specificity (95% CI) for studies using AI-based methods on CBCT images to detect carotid artery calcifications. The black square represents the point estimate of sensitivity or specificity, and the horizontal line shows its 95% confidence interval. The size of the square reflects the study's weight in the analysis. The red diamond represents the pooled estimate across all studies, and the red dashed line indicates the summary value. **(b)** Forest plot of sensitivity and specificity (95% CI) for studies using AI-based methods on Panoramic radiographies to detect carotid artery calcifications. The black square represents the point estimate of sensitivity or specificity, and the horizontal line shows its 95% confidence interval. The size of the square reflects the study's weight in the analysis. The red diamond represents the pooled estimate across all studies, and the red dashed line indicates the summary value

calcifications. Firstly, Ajami et al. [25] combined Inception V3 and U-Net to classify and segment calcifications in the cervical carotid artery in CBCT images, reporting an accuracy of 96.35%. In addition, Alajaji et al. [9] in 2024 used a U-Net-based neural network to segment and classify extracranial and intracranial carotid artery atheromas in CBCT images. While the model achieved high sensitivity values for both regions (92% and 96%, respectively), other diagnostic metrics including precision, accuracy, and specificity were relatively low, particularly for intracranial calcifications. Finally, Nelson et al. [28] employed a CNN to classify external carotid artery calcifications in CBCT images, achieving an accuracy of 76%. Panoramic imaging is a readily available and widely used radiation technique, offering a relatively low cost and radiation dose of nearly 0.02 mSv. Given its prevalence, it is important to use its potential in older adults to monitor for stroke risk factors, such as carotid artery calcifications. In contrast, CBCT images, while exposing the patient to a relatively higher radiation dose, provide superior resolution for detecting subtle calcified lesions due to their three-dimensional nature. Interestingly, the included studies indicated that panoramic radiographs had a higher diagnostic odds ratio (DOR) than CBCTs; however, the difference was not statistically significant ($P = 0.30$). Both panoramic images and CBCT scans are commonly obtained in older individuals experiencing tooth loss, primarily for treatment planning related to oral rehabilitation. By utilizing these imaging techniques, clinicians can plan dental treatments effectively and also concurrently assess and monitor for other significant

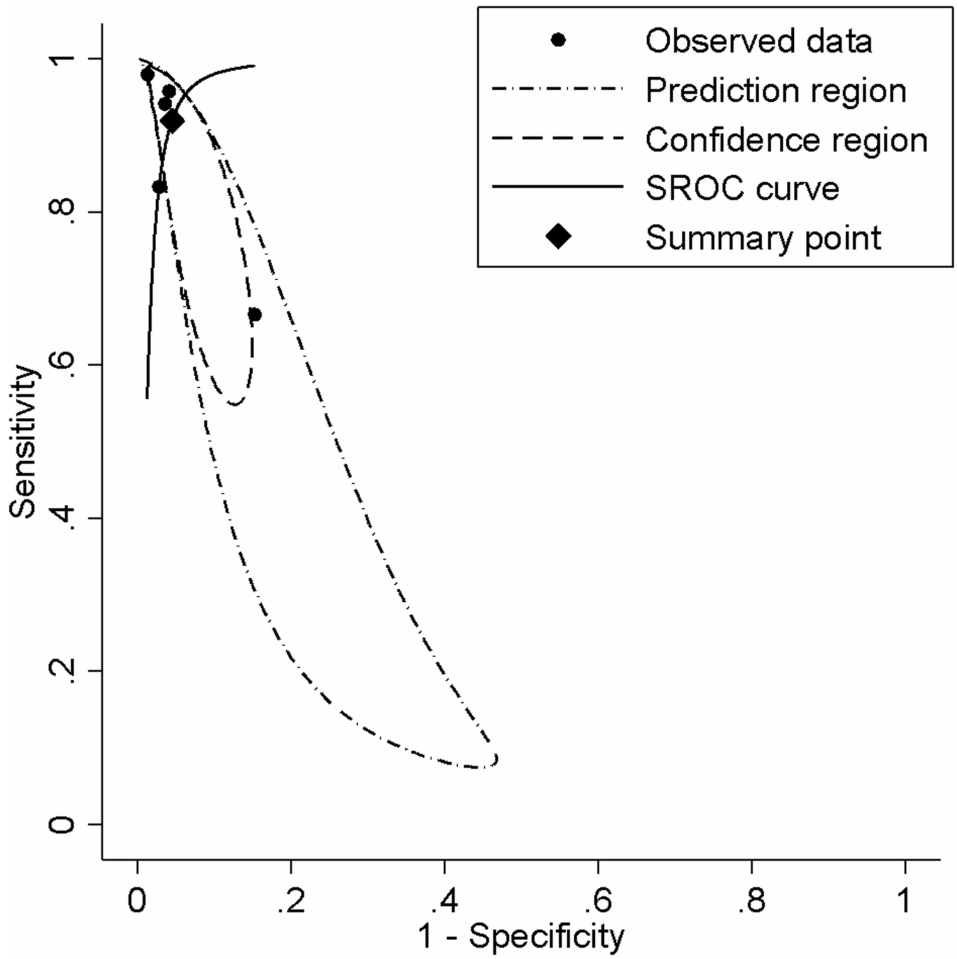

Fig. 7 Summary Receiver Operating Characteristic (SROC) curve with 95% confidence and prediction regions for AI-based detection of carotid artery calcifications. Each black dot represents the observed sensitivity and 1-specificity from an individual study. The solid line shows the SROC curve, which summarizes the overall diagnostic accuracy. The diamond marks the pooled (summary) sensitivity and specificity point. The dashed line represents the 95% confidence region around this summary point, indicating the precision of the pooled estimate. The dot-dash line shows the 95% prediction region, reflecting expected variability in future studies. The closer the curve is to the top-left corner, the better the overall diagnostic performance

health conditions, such as cardiovascular disease and stroke risk.

It is important to consider that in six of the reviewed papers, the ground truth was provided by radiologists annotating the same CBCT or panoramic images. While only three papers employed a more definitive approach for diagnosis, the exact method for confirming the presence of carotid artery calcifications was not clarified in any of the studies. This lack of methodological transparency can lead to variability in the accuracy and reliability of the results. The gold standard for diagnosing carotid artery calcifications typically includes digital subtraction angiography, which offers high-resolution images and precise measurements of arterial stenosis. Additionally, duplex ultrasonography and contrast-enhanced CT and MRI are widely used for detecting carotid artery calcifications due to their high sensitivity and specificity. However, the use of less definitive methods for establishing ground truth can introduce biases and reduce the validity of the findings. The reliance on weaker ground truths, such as annotations by radiologists without standardized protocols, can lead to inconsistencies and limit the generalizability of the results. To improve the robustness of future studies, it is crucial to adopt standardized and validated methods for establishing ground truth. This includes clear documentation of the diagnostic criteria and techniques used, as well as the involvement of multiple experts to minimize inter-observer variability. By ensuring methodological rigor, researchers can enhance the reliability and reproducibility of their findings, ultimately contributing to more accurate and clinically relevant conclusions.

The integration of AI-based detection systems into clinical practice presents several challenges that must be addressed to ensure effective utilization. One key challenge is the need for user-friendly interfaces that allow dental practitioners to easily interact with AI models without requiring extensive technical expertise. Additionally, interoperability between AI systems and existing dental imaging software is critical for seamless integration into routine workflows.

Another important consideration is the variability in AI model performance across different imaging modalities and datasets. Panoramic radiographs exhibited lower heterogeneity and a consistently high diagnostic performance across studies, with models such as CACSNet paired with EfficientNet-B4 achieving standout diagnostic potential with a pooled sensitivity and specificity of 98% and 99%, respectively [27]. Faster R-CNN, utilized by Lee et al. [26], also performed robustly with a pooled sensitivity of 92% and specificity of 96%. These models benefit from transfer learning and deep feature extraction, enhancing their ability to accurately detect carotid calcifications in large datasets. Models applied to CBCT images, despite their superior resolution, displayed higher heterogeneity, suggesting performance inconsistencies. The approach by Ajami et al. [25] using Inception V3 and U-Net yielded impressive sensitivity and specificity (94% and 97%, respectively), but Nelson et al.'s CNN showed lower sensitivity (67%) [28]. This disparity indicates that while CBCT offers detailed imaging, the effectiveness of AI models depends heavily on dataset characteristics and model architecture.

While the findings of the present review are encouraging, it is important to acknowledge the limitations of this study. First, the included studies used diverse AI techniques and imaging modalities, which could introduce heterogeneity into the analysis. Second, the relatively small number of studies included in the meta-analysis may limit the generalizability of the findings. Finally, further validation of these AI models in larger, independent datasets is mostly missing in the included studies. Future research should focus on the integration of these AI-based models into clinical practice by developing user-friendly interfaces and guidelines for incorporating AI into routine dental practice. Additionally, validation of these AI models is crucial to confirm their performance and clinical utility.

The differential diagnosis of carotid artery calcifications includes calcification of triticeal cartilage, stylohyoid ligament, thyroid cartilage, and epiglottis. It is crucial to differentiate these structures from true carotid artery calcifications to avoid misdiagnosis and ensure appropriate patient management. Further imaging studies such as ultrasound or CT scans can be helpful for confirmation of diagnosis. Results of a study by Almog et al. [40] indicated that human examiners, after completing a specialized training program, achieved a positive predictive value (PPV) of 34.6% for identifying carotid artery calcifications, with good inter-examiner agreement ($\kappa = 0.87$). In contrast, our meta-analysis of AI performance showed significantly higher accuracy, with a pooled sensitivity of 0.92 and a pooled specificity of 0.96. These findings suggest that AI systems not only offer superior diagnostic precision but also minimize variability in performance between different examiners, thus providing more consistent and reliable results in detecting carotid artery calcifications on dental images. Additionally, AI can serve as an educational or training tool for young dentists and dental students, enhancing their diagnostic skills and providing valuable learning experiences through exposure to a wide range of cases and expert-level analyses.

Conclusion

This systematic review and meta-analysis revealed promising results. The included studies demonstrated high accuracy in identifying carotid calcifications using AI algorithms, particularly in CBCT and panoramic radiographs. The pooled sensitivity and specificity values were notably high, indicating the potential of AI-based methods as effective novel screening tools. These results highlight the importance of using AI in dental imaging to increase early detection of stroke and other atherosclerosis-related risks, leading to better patient outcomes and preventative treatments. There is great potential for improving the field of cardiovascular risk assessment and management through additional research and application of AI-driven techniques in dental radiology.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12880-025-01719-9.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3: Figure S1. Sensitivity analysis showing the pooled Diagnostic Odds Ratio (DOR) after omitting each included study one at a time. This "leave-one-out" analysis assesses the influence of individual studies on the overall diagnostic performance of AI methods. Each point on the plot represents the recalculated pooled DOR after excluding a specific study. Minimal variation in the DOR across points indicates that the overall results are robust and not overly dependent on any single study.

Acknowledgements

The authors wish to express their sincere appreciation to Prof. Roya Kelishadi for her essential guidance and support during the preparation of this manuscript.

Author contributions

S.A.: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. P.S.: Methodology, Investigation, Writing - review and editing, Supervision, Final approval. A.K.: Conceptualization, Methodology, Investigation, Data Curation, Writing - Original Draft, Visualization. M.Y.: Formal analysis, Data Curation, Writing - review and editing, Final approval. A.A.: Investigation, Writing - review and editing, Supervision, Final approval. Z.K: Investigation, Writing - review and editing, Supervision, Final approval. D.G.: Methodology, Writing - review and editing, Supervision, Final approval. H.D: Methodology, Writing - review and editing, Supervision, Final approval. All authors reviewed the manuscript.

Funding
This review was conducted without external funding.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Human ethics and consent to participate
Not applicable.

Declaration of generative AI and AI-assisted technologies
During the revision of this work the authors used ChatGPT-4 to improve the English language in a few paragraphs, not the whole manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Child Growth and Development Research Center, Research Institute for Primordial Prevention of Non-Communicable Disease, Isfahan University of Medical Sciences, Hezar-Jarib Ave, Isfahan 81551-39998, Iran
[2]Department of Neurosciences, Reproductive and Odontostomatological Sciences, University of Naples Federico II, Naples, Italy
[3]Maxillogram Maxillofacial Surgery, Implantology and Biomaterial Research Foundation, Istanbul 8418829912, Turkey
[4]Scottish Craniofacial Research Group, School of Medicine, Dentistry and Nursing, Glasgow University MVLS College, Glasgow University Dental School, Glasgow, UK
[5]Department of Prosthodontics and Dental Implantology, College of Dentistry, King Faisal University, Al Ahsa, Saudi Arabia
[6]Center for Artificial Intelligence and Innovation (CAII), Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand
[7]Department of Medicine, Surgery and Dentistry, University of Salerno, Baronissi, Salerno, Italy
[8]The Dental School, University of Bristol, Bristol, UK
[9]The University of Jordan, Amman, Jordan

Received: 3 March 2025 / Accepted: 9 May 2025
Published online: 19 May 2025

References

1. WHO. The top 10 causes of death. 2021. Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death
2. Fairhead JF, Mehta Z, Rothwell PM. Population-based study of delays in carotid imaging and surgery and the risk of recurrent stroke. Neurology. 2005;65(3):371–5.
3. Nasseh I, Aoun G. Carotid artery calcification: A digital Panoramic-Based study. Diseases. 2018;6(1).
4. Bin Song Y, Jeong HG, Kim C, Kim D, Kim J, Kim HJ, Park W. Comparison of detection performance of soft tissue calcifications using artificial intelligence in panoramic radiography. Sci Rep. 2022;12(1).
5. Kats L, Vered M, Zlotogorski-Hurvitz A, Harpaz I. Atherosclerotic carotid plaques on panoramic imaging: an automatic detection using deep learning with small dataset. ArXiv. 2018. abs/1808.08093.
6. Amitay M, Barnett-Itzhaki Z, Sudri S, Drori C, Wase T, Abu-El-Naaj I, et al. Deep Convolution neural network for screening carotid calcification in dental panoramic radiographs. PLOS Digit Health. 2023;2(4):e0000081.
7. Prados-Privado M, García Villalón J, Blázquez Torres A, Martínez-Martínez CH, Prados-Frutos JC, Ivorra C. Are panoramic images a good tool to detect calcified carotid atheroma? A systematic review. Biology (Basel). 2022;11(11).
8. Dief S, Veitz-Keenan A, Amintavakoli N, McGowan R. A systematic review on incidental findings in cone beam computed tomography (CBCT) scans. Dentomaxillofac Radiol. 2019;48(7):20180396.
9. Alajaji SA, Amarin R, Masri R, Tavares T, Kumar V, Price JB, Sultan AS. Detection of extracranial and intracranial calcified carotid artery atheromas in cone beam computed tomography using a deep learning convolutional neural network image segmentation approach. Oral Surg Oral Med Oral Pathol Oral Radiol. 2024;138(1):162–72.
10. Thrall JH, Li X, Li Q, Cruz C, Do S, Dreyer K, Brink J. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. J Am Coll Radiol. 2018;15(3):504–8.
11. Fatalla AA, Arzani S, Veseli E, Khademi A, Khandan A, Fahmy MD, et al. Revolutionizing systematic reviews and Meta-analyses: the role of artificial intelligence in evidence synthesis. Dent Hypotheses. 2023;14(4):93–4.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563–77.
13. Londono J, Ghasemi S, Hussain Shah A, Fahimipour A, Ghadimi N, Hashemi S, et al. Evaluation of deep learning and convolutional neural network algorithms accuracy for detecting and predicting anatomical landmarks on 2D lateral cephalometric images: A systematic review and meta-analysis. Saudi Dent J. 2023;35(5):487–97.
14. Dashti M, Ghaedsharaf S, Ghasemi S, Zare N, Constantin EF, Fahimipour A, et al. Evaluation of deep learning and convolutional neural network algorithms for mandibular fracture detection using radiographic images: A systematic review and meta-analysis. Imaging Sci Dent. 2024;54(3):232–9.
15. Li R, Zheng J, Zayed MA, Saffitz JE, Woodard PK, Jha AK. Carotid atherosclerotic plaque segmentation in multi-weighted MRI using a two-stage neural network: advantages of training with high-resolution imaging and histology. Front Cardiovasc Med. 2023;10:1127653.
16. Zhu Y, Chen L, Lu W, Gong Y, Wang X. The application of the nnU-Net-based automatic segmentation model in assisting carotid artery stenosis and carotid atherosclerotic plaque evaluation. Front Physiol. 2022;13:1057800.
17. Deng C, Adu J, Xie S, Li Z, Meng Q, Zhang Q, et al. Automatic segmentation of ultrasound images of carotid atherosclerotic plaque based on Dense-UNet. Technol Health Care. 2023;31(1):165–79.
18. Ahmed N, Abbasi MS, Zuberi F, Qamar W, Halim MSB, Maqsood A, Alam MK. Artificial intelligence techniques: analysis, application, and outcome in Dentistry-A systematic review. Biomed Res Int. 2021;2021:9751564.
19. Abdul NS, Shivakumar GC, Sangappa SB, Di Blasio M, Crimi S, Cicciù M, Minervini G. Applications of artificial intelligence in the field of oral and maxillofacial pathology: a systematic review and meta-analysis. BMC Oral Health. 2024;24(1):122.
20. Pollock M, et al. Preferred reporting items for overviews of reviews (PRIOR): a protocol for development of a reporting guideline for overviews of reviews of healthcare interventions. Syst Rev. 2019;8:335.
21. McInnes MDF, et al. Preferred reporting items for a systematic review and Meta-analysis of diagnostic test accuracy studies. JAMA. 2018;319:388.
22. Cacciamani GE, et al. PRISMA AI reporting guidelines for systematic reviews and meta-analyses on AI in healthcare. Nat Med. 2023;29:14–5.
23. Campbell JM, Klugar M, Ding S, Carmody DP, Hakonsen SJ, Jadotte YT, et al. Diagnostic test accuracy: methods for systematic review and meta-analysis. Int J Evid Based Healthc. 2015;13(3):154–62.
24. Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. Syst Rev. 2019;8:163.
25. Ajami M, Tripathi P, Ling H, Mahdian M. Automated detection of cervical carotid artery calcifications in cone beam computed tomographic images using deep convolutional neural networks. Diagnostics. 2022;12(10).
26. Lee S, Kim D, Jeong H-G. Detecting 17 fine-grained dental anomalies from panoramic dental radiography using artificial intelligence. Sci Rep. 2022;12(1):5172.
27. Yoo SW, Yang S, Kim JE, Huh KH, Lee SS, Heo MS, Yi WJ. CACSNet for automatic robust classification and segmentation of carotid artery calcification on panoramic radiographs using a cascaded deep learning network. Sci Rep. 2024;14(1):13894.
28. Nelson J, Vaddi A, Tadinada A. Can convolutional neural networks identify external carotid artery calcifications? Oral Surg Oral Med Oral Pathol Oral Radiol. 2024;138(1):142–8.
29. Vinayahalingam S, van Nistelrooij N, Xi T, Heiland M, Bressem K, Rendenbach C, et al. Detection of carotid plaques on panoramic radiographs using deep learning. J Dent. 2024;151:105432.

30. Chen X, Shi J, Hu Y, Ma H, Jiang Z, Lou B. Study on risk factors of carotid atherosclerosis in type 2 diabetes mellitus and development of prediction model. Int J Diabetes Developing Ctries. 2024.
31. Fan J, Watanabe T, Atherosclerosis. Known and unknown. Pathol Int. 2022;72(3):151–60.
32. Xiang P, Blanchard V, Francis GA. Smooth muscle cell-Macrophage interactions leading to foam cell formation in atherosclerosis: location, location, location. Front Physiol. 2022;13:921597.
33. Alonso-Herranz L, Albarrán-Juárez J, Bentzon JF. Mechanisms of fibrous cap formation in atherosclerosis. Front Cardiovasc Med. 2023;10:1254114.
34. Jansen MF, Lutgens E, Daemen MJAP. Atherosclerosis of the carotid artery. In: Yuan C, Hatsukami TS, Mossa-Basha M, editors. Vessel based imaging techniques: diagnosis, treatment, and prevention. Cham: Springer International Publishing. 2020;69–91.
35. Meng H, Ruan J, Yan Z, Chen Y, Liu J, Li X, Meng F. New progress in early diagnosis of atherosclerosis. Int J Mol Sci. 2022;23(16).
36. Mantella LE, Liblik K, Johri AM. Vascular imaging of atherosclerosis: strengths and weaknesses. Atherosclerosis. 2021;319:42–50.
37. Shinjo K, Muneyasu M, Hasegawa K, Asano A, Taguchi A, editors. A detection method for carotid artery calcification in dental panoramic radiographs. International Workshop on Smart Info-Media Systems in Asia; 2009.
38. Harada H, Muneyasu M, Asano A, Uchida K, Taguchi A, editors. Improved detection method for carotid artery calcification in dental panoramic radiographs considering local features. ISPACS 2013–2013 International Symposium on Intelligent Signal Processing and Communication Systems; 2013.
39. 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. IEICE Trans Inf Syst. 2013;96(8):1878–81.
40. Almog DM, Tsimidis K, Moss ME, Gottlieb RH, Carter LC. Evaluation of a training program for detection of carotid artery calcifications on panoramic radiographs. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2000;90(1):111–7.

Publisher's note