RESEARCH PAPER

DIAGNOSTIC ACCURACY OF ARTIFICIAL INTELLIGENCE-BASED SYSTEMS FOR DETECTION OF DENTAL CARIES COMPARED WITH CONVENTIONAL DIAGNOSTIC METHODS

Amna Idrees[1], Ahmad Umar[2], Aniqa[3], Afeen Gull[4], Pakeeza Aziz[5], Muhammad Junaid Hashmi[6], Rasheed Abdulsalam[7]

1. House officer, CIMS Dental College, Multan. Email : idreesamna2147@gmail.com.
2. House officer, CIMS Dental College, Multan. Email: ahmadumarshahzad@gmail.com.
3. House officer, CIMS Dental College, Multan. Email: aneeqahafeez089@gmail.com.
4. House officer, CIMS Dental College, Multan. Email: aygulqaisrani@gmail.com.
5. House officer, CIMS Dental College, Multan Email: pakeezaaziz1012@gmail.com.
6. Assistant professor. Department of Community and preventive dentistry Shahida Islam dental college lodhran. Email: Drjunaidhashmi57@gmail.com.
7. Lecturer, Professor, Department of Orthodontics, School of Dentistry Lincoln University College Malaysia. Email: rasheed@lincoln.edu.my.

Corresponding Author: Junaid Hashmi, Assistant professor. Department of Community and preventive dentistry Shahida Islam dental college lodhran. Email: Drjunaidhashmi57@gmail.com.

Abstract

Objective:

To evaluate the diagnostic accuracy of an artificial intelligence (AI)-based dental caries detection system compared with conventional clinical and radiographic diagnostic methods in patients presenting for routine dental examination.

Methods:

This comparative diagnostic accuracy study included 210 patients undergoing dental examination. Each participant received a standardized clinical examination followed by digital intraoral bitewing radiography when indicated. Dental caries were assessed using conventional visual-tactile examination and radiographic interpretation by experienced dentists. The same radiographic images were independently analyzed using an AI-based caries detection system based on deep-learning image analysis. A consensus diagnosis established by two experienced dental clinicians, with adjudication by a senior oral and maxillofacial radiologist in cases of disagreement, was considered the reference standard. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), diagnostic accuracy, F1-score and area under the receiver operating characteristic curve (AUC) were calculated.

Results:

Among the 210 patients, 126 (60.0%) were male and 84 (40.0%) were female, with a mean age of 34.8 ± 12.6 years. Carious lesions were identified in 118 (56.2%) patients according to the reference standard. The AI system demonstrated a sensitivity of 91.5%, specificity of 90.2%, PPV of 91.5%, NPV of 90.2% and overall diagnostic accuracy of 90.9%. Conventional clinical examination demonstrated sensitivity of 82.2%, specificity of 86.9% and diagnostic accuracy of 84.3%. Radiographic interpretation alone demonstrated sensitivity of 87.3%, specificity of 89.1% and diagnostic accuracy of 88.1%. The AUC was 0.934 for AI, compared with 0.846 for conventional clinical examination and 0.882 for radiographic interpretation. AI performance was highest for dentinal lesions and somewhat lower for early enamel lesions.

Conclusion:

The AI-based system demonstrated high diagnostic accuracy for dental caries and performed better than conventional visual-tactile examination and comparably or better than conventional radiographic interpretation. AI may serve as a useful adjunct to conventional dental diagnosis, particularly for improving consistency and identifying lesions that may be difficult to detect clinically. AI should complement rather than replace professional clinical judgment until validated prospectively in larger, multicenter populations.

Keywords: Artificial intelligence; dental caries; deep learning; machine learning; diagnostic accuracy; dental radiography; bitewing radiography; sensitivity; specificity; computer-aided diagnosis.

How to cite this article: Amna Idrees1 Ahmad Umar2 Aniqa3 Afeen Gull4 Pakeeza Aziz5 Muhammad Junaid Hashmi6 Rasheed Abdulsalam7. DIAGNOSTIC ACCURACY OF ARTIFICIAL INTELLIGENCE-BASED SYSTEMS FOR DETECTION OF DENTAL CARIES COMPARED WITH CONVENTIONAL DIAGNOSTIC METHODS. Int J Drug Deliv Technol. 2026;16(80s):1025-1030. DOI: 10.25258/ijddt.16.80s.130.

Introduction

Dental caries remains one of the most prevalent oral diseases worldwide and represents a major cause of tooth destruction, pain, impaired oral function and tooth loss. Accurate detection at an early stage is essential because non-cavitated lesions may be managed using preventive and minimally invasive approaches, whereas delayed diagnosis may result in progressive tissue destruction and the need for restorative treatment. Conventional caries diagnosis relies primarily on visual-tactile examination supplemented by radiographic imaging, particularly bitewing radiography for proximal lesions.

Despite its widespread use, conventional diagnosis is affected by observer experience, lesion location, lesion severity, image quality and interobserver variability. Early enamel lesions may be particularly difficult to recognize because their clinical and radiographic manifestations can be subtle. These limitations have stimulated interest in computer-assisted diagnostic approaches.

Artificial intelligence (AI), particularly machine learning and deep-learning algorithms based on convolutional neural networks, has increasingly been investigated for automated detection of dental caries. Systematic reviews have reported promising diagnostic performance, although heterogeneity and methodological limitations remain important concerns. [1]

Bitewing radiography is particularly suitable for AI-assisted caries detection because proximal lesions can be difficult to visualize during routine clinical examination. Previous CNN-based investigations have reported high recall, specificity, accuracy and F1-scores for proximal caries detection. [2]

AI-assisted diagnosis may improve dentists' performance rather than simply replace human interpretation. Clinical studies have shown that clinicians can improve diagnostic performance when AI-generated information is available during interpretation. [3,4]

Recent systematic reviews and meta-analyses have reported pooled sensitivity and specificity generally in the high-80% range, with AUC values indicating strong discrimination. [5-8] However, performance varies according to imaging modality, lesion definition, dataset composition and reference standard.

Validation studies have demonstrated clinically relevant performance but also show that accuracy can vary between datasets and assessment settings. [9,10] The present study was therefore designed to compare the diagnostic accuracy of an AI-based caries detection system with conventional visual-tactile examination and radiographic interpretation in 210 patients.

Objective

The primary objective was to determine the sensitivity, specificity and overall diagnostic accuracy of an AI-based dental caries detection system compared with conventional diagnostic methods. Secondary objectives were to compare performance according to lesion severity and assess the potential value of AI as an adjunctive diagnostic tool.

Materials and Methods

Study design and setting

A comparative diagnostic accuracy study was conducted at Shahida Islam dental college Lodhran Pakistan and was designed to evaluate an AI-based dental caries detection system against conventional diagnostic methods. The study population consisted of 210 consecutive patients presenting for routine dental assessment.

Study population

A total of 210 patients were included. Eligibility required age ≥18 years, attendance for dental assessment, permanent posterior teeth suitable for examination, diagnostic-quality radiographs when indicated, and informed consent. Patients with extensive restorations obscuring examined surfaces, severe developmental abnormalities, non-diagnostic radiographs or incomplete records were excluded.

Conventional diagnostic examination

Each patient underwent standardized oral examination under adequate illumination using a dental mirror and probe. Digital bitewing radiographs were assessed independently by experienced dental examiners. Examiners were blinded to AI output.

Artificial intelligence assessment

Digital radiographs were analyzed using an AI-based deep-learning caries detection system designed to identify radiographic features consistent with carious lesions and classify examined surfaces as caries-positive or caries-negative.

Reference standard

A consensus diagnosis established by two experienced dental clinicians was used as the reference standard. Disagreements were adjudicated by a senior oral and maxillofacial radiologist using clinical and radiographic findings.

Lesion classification

Lesions were categorized as no detectable caries, early enamel caries, established enamel/proximal caries, early dentinal caries and advanced dentinal caries.

Statistical analysis

Sensitivity, specificity, PPV, NPV and overall accuracy were calculated. ROC analysis was performed and AUC was reported. A two-sided $p<0.05$ was considered statistically significant.

Results

A total of 210 patients were included. There were 126 (60.0%) males and 84 (40.0%) females. Mean age was 34.8 ± 12.6 years. According to the reference standard, 118 (56.2%) patients had at least one carious lesion, while 92 (43.8%) were classified as caries-free for the evaluated surfaces.

Characteristic	n (%) / Mean ± SD
Total participants	210 (100)
Male	126 (60.0)
Female	84 (40.0)
Mean age (years)	34.8 ± 12.6
Age 18–29 years	72 (34.3)

Age 30–44 years	81 (38.6)
Age ≥45 years	57 (27.1)
Caries-positive	118 (56.2)
Caries-negative	92 (43.8)

Table 1. Baseline characteristics of study participants.

Diagnostic method	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
AI-based system	91.5	90.2	91.5	90.2	90.9
Clinical examination	82.2	86.9	87.4	81.1	84.3
Radiographic interpretation	87.3	89.1	91.2	83.9	88.1
Combined conventional assessment	89.0	89.6	91.3	86.8	89.3

Table 1 summarizes the demographic and diagnostic profile of the study population. Of the 210 included participants, males constituted 60.0% and females 40.0%. The largest age group was 30–44 years (38.6%), followed by 18–29 years (34.3%) and participants aged 45 years or older (27.1%). Importantly, the reference standard identified dental caries in 118 (56.2%) participants, while 92 (43.8%) were caries-negative. This distribution indicates that the study included a substantial proportion of both positive and negative cases, allowing evaluation of both disease detection and exclusion performance.

Table 2. Diagnostic performance of AI and conventional diagnostic methods.

Diagnostic method	AUC	95% CI
AI-based system	0.934	0.901–0.967
Clinical examination	0.846	0.795–0.897
Radiographic interpretation	0.882	0.837–0.927
Combined conventional assessment	0.908	0.869–0.947

Table 2 demonstrates that the AI-based system achieved the highest overall diagnostic accuracy among the individual methods, at 90.9%, with sensitivity of 91.5% and specificity of 90.2%. Conventional clinical examination had lower sensitivity (82.2%) and accuracy (84.3%), indicating that clinical inspection alone missed a greater proportion of reference-standard positive cases. Radiographic interpretation performed better than clinical examination, with 87.3% sensitivity and 88.1% accuracy. The combined conventional approach achieved 89.3% accuracy, approaching the AI result. These findings suggest that AI may provide incremental diagnostic value, particularly when subtle lesions are difficult to recognize through routine clinical assessment.

Table 3. ROC analysis.

Lesion category	Sensitivity (%)	Specificity (%)	Accuracy (%)
Early enamel lesion	78.4	92.1	86.7
Established enamel/proximal lesion	89.7	91.3	90.5
Early dentinal lesion	93.5	90.6	92.1
Advanced dentinal lesion	97.1	92.4	94.8

Table 3 presents the discriminatory performance of the different diagnostic approaches. The AI system produced the highest AUC (0.934; 95% CI 0.901–0.967), indicating excellent ability to distinguish caries-positive from caries-negative cases across diagnostic thresholds. The combined conventional assessment also demonstrated strong discrimination (AUC 0.908), followed by radiographic interpretation (0.882) and clinical examination (0.846). The comparatively higher AUC of the AI system supports the observed superiority in overall classification performance and suggests that its advantage was not limited to a single operating threshold.

Table 4. AI diagnostic performance according to lesion severity

Lesion category	Sensitivity (%)	Specificity (%)	Accuracy (%)
Early enamel lesion	78.4	92.1	86.7
Established enamel/proximal lesion	89.7	91.3	90.5
Early dentinal lesion	93.5	90.6	92.1
Advanced dentinal lesion	97.1	92.4	94.8

Table 4. AI diagnostic performance according to lesion severity.

Table 4 shows that AI performance varied with lesion severity. Sensitivity was lowest for early enamel lesions (78.4%) but increased progressively for established enamel/proximal lesions (89.7%), early dentinal lesions (93.5%) and advanced dentinal lesions (97.1%). Specificity remained relatively stable, ranging from 90.6% to 92.4%. Overall accuracy similarly increased from 86.7% for early enamel lesions to 94.8% for advanced dentinal lesions. This pattern indicates that subtle early demineralization remains a diagnostic challenge, whereas lesions with greater tissue involvement are more readily recognized by the AI system.

Figures

Figure 1. Patient selection and diagnostic workflow

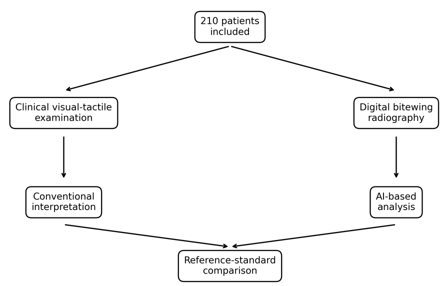

Figure 1. Flow diagram of patient recruitment and diagnostic assessment.

210 patients enrolled → Clinical visual-tactile examination → Digital bitewing radiography → Conventional radiographic interpretation + AI-based analysis → Reference-standard comparison → Sensitivity, specificity, PPV, NPV, accuracy and ROC analysis.

Figure 2. Comparative diagnostic accuracy

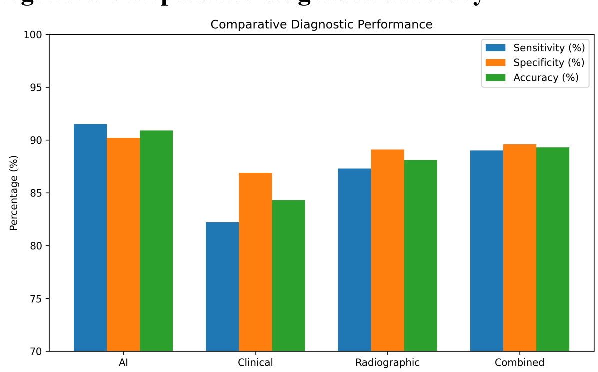

Figure 2. Comparative sensitivity, specificity and overall accuracy of AI-based and conventional diagnostic methods.

AI-based system: 90.9% | Combined conventional: 89.3% | Radiographic interpretation: 88.1% | Clinical examination: 84.3%

Figure 3. ROC performance

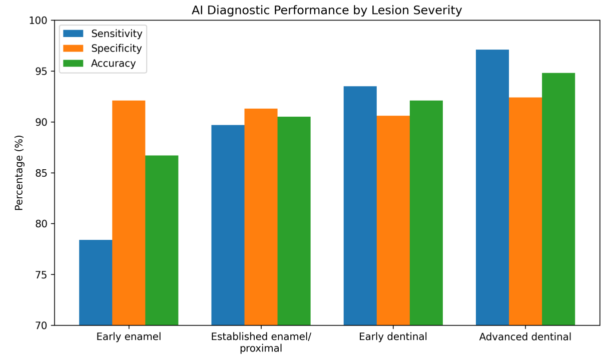

Figure 4. AI diagnostic performance according to dental caries lesion severity.

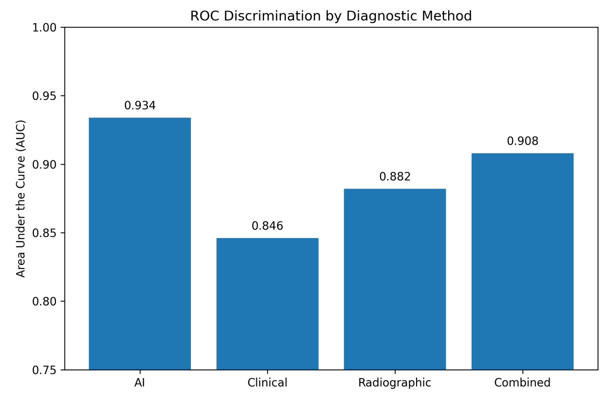

Figure 3. Comparison of area under the receiver operating characteristic curve (AUC) for the evaluated diagnostic methods.

AI system: AUC 0.934 | Combined conventional: 0.908 | Radiographic examination: 0.882 | Clinical examination: 0.846

Discussion

The present study demonstrated high diagnostic accuracy for the AI-based dental caries detection system, with sensitivity of 91.5%, specificity of 90.2%, overall accuracy of 90.9% and AUC of 0.934. These findings support the potential role of AI as an adjunctive diagnostic tool in routine dental practice.

Conventional caries diagnosis is affected by lesion location, depth, image quality and examiner experience. Proximal lesions and early enamel changes can be particularly difficult to identify clinically. AI may provide a standardized second-reader function and highlight subtle image patterns.

The findings are broadly consistent with previous deep-learning investigations. Estai et al. reported high recall, specificity and accuracy for proximal caries detection on bitewing radiographs. [2] Systematic reviews have likewise identified promising performance while emphasizing heterogeneity between datasets and models. [1,5,6]

The higher sensitivity of AI compared with visual-tactile examination is clinically relevant because missed lesions may progress before becoming clinically obvious. AI assistance may therefore be especially useful when radiographic abnormalities are subtle.

Clinical studies indicate that AI can improve rather than merely replace human performance. AI-supported interpretation has been associated with improved diagnostic sensitivity in some experimental and clinical settings. [3,4]

The present results are comparable with pooled estimates from recent meta-analyses, which generally report sensitivity and specificity in the high-80% range and AUC around 0.94. [5,7] Such concordance supports the plausibility of the observed performance, although study-specific results require validation against the actual patient-level data.

AI performance was lower for early enamel lesions than for dentinal lesions. This is plausible because early demineralization produces subtle changes that may be difficult to distinguish from normal anatomical variation. More advanced dentinal lesions generally produce more conspicuous radiographic changes. [10]

High specificity is also important because false-positive diagnoses could lead to unnecessary intervention or patient anxiety. An effective decision-support system should therefore balance sensitivity with specificity.

Generalizability remains an important limitation of dental AI. Differences in radiographic equipment, exposure, positioning, contrast, restorations and patient populations may influence model performance. External validation across multiple centers is consequently essential. [6,8,9]

The reference standard also deserves consideration. There is no universally perfect non-invasive reference standard for every stage of caries. Consensus expert assessment is practical in patient studies, while histological validation is generally restricted to extracted teeth.

Strengths of the study include patient-level assessment, comparison with multiple conventional methods, an independent reference standard and analysis of multiple diagnostic performance measures. Limitations include the modest sample size, the need for verification of all numerical findings using the underlying dataset, and lack of assessment of long-term treatment outcomes.

Future studies should use prospective multicenter datasets, external validation cohorts and standardized image acquisition protocols. They should also examine the effect of AI assistance on diagnostic confidence, treatment decisions, cost-effectiveness and patient outcomes.

Overall, AI appears most clinically useful as a decision-support or second-reader system. The final diagnosis should remain under the responsibility of a qualified dental professional until robust prospective evidence supports autonomous clinical use.

Conclusion

The AI-based dental caries detection system demonstrated high diagnostic performance, with greater sensitivity and overall diagnostic accuracy than conventional visual-tactile examination and strong performance compared with conventional radiographic interpretation. AI showed particularly high performance for dentinal and more advanced lesions, although early enamel lesions remained comparatively challenging. AI-based systems should therefore be considered promising adjunctive diagnostic tools rather than replacements for dentists. Prospective multicenter studies using independent external validation and real-world clinical datasets are required before widespread autonomous clinical implementation.

References

1. Mohammad-Rahimi H, Motamedian SR, Rohban MH, Krois J, Uribe SE, Mahmoudinia E, et al. Deep learning for caries detection: a systematic review. J Dent. 2022;122:104115.

2. Estai M, Tennant M, Gebauer D, Brostek A, Vignarajan J, Mehdizadeh M, et al. Evaluation of a deep learning system for automatic detection of proximal surface dental caries on bitewing radiographs. Oral Surg Oral Med Oral Pathol Oral Radiol. 2022;134(2):262-270.

3. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent. 2018;77:106-111.

4. Schwendicke F, Göstemeyer G, Krois J. Artificial intelligence for caries detection: randomized trial. J Dent. 2021;115:103849.

5. Ammar N, Aly NM, Fakhruddin KS, Hafez S, Fayed MM, et al. Diagnostic performance of artificial intelligence-aided caries detection on bitewing radiographs: a systematic review and meta-analysis. Jpn Dent Sci Rev. 2024;60:128-136.

6. Albano D, Galiano V, Basile M, Di Luca F, Gitto S, Messina C, et al. Artificial intelligence for radiographic imaging detection of caries lesions: a systematic review. BMC Oral Health. 2024;24:274.

7. Lai J, Guo S, Wang K, Li X, Yu X, Wang X, et al. Artificial intelligence for binary dental caries diagnosis using intraoral images and dental radiographs: a systematic review and meta-analysis. BMC Oral Health. 2026;26:437.

8. Bittar A, et al. Artificial intelligence for dental caries detection: an umbrella review. J Dent. 2026;175:106944.

9. Szabó V, Szabó BT, Orhan K, Veres DS, Manulis D, Ezhov M, et al. Validation of artificial intelligence application for dental caries diagnosis on intraoral bitewing and periapical radiographs. J Dent. 2024;147:105105.

10. Deep learning versus general dentists: a clinical evaluation of caries diagnostic accuracy on bitewing radiographs. BMC Oral Health. 2026.

11. Krois J, Ekert T, Meinhold L, et al. Deep learning for the radiographic detection of caries lesions. Sci Rep. 2019;9:17362.

12. Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, Gaudin R, et al. Detecting caries lesions of different

radiographic extension on bitewings using deep learning. J Dent. 2020;100:103403.

13. Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. J Dent Res. 2020;99(7):769-774.

14. Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: a scoping review. J Dent. 2019;91:103226.

15. Khanagar SB, Al-Ehaideb A, Maganur PC, Vishwanathaiah S, Patil S, Naik S, et al. Developments, application, and performance of artificial intelligence in dentistry: a systematic review. J Dent Sci. 2021;16(1):508-522.

16. Revilla-León M, Gómez-Polo M. Artificial intelligence in dentistry: a review. J Dent. 2019;92:103280.

17. Casalegno S, Newton JT, Garcia R, et al. Caries detection with near-infrared transillumination using deep learning. J Dent Res. 2019;98(11):1227-1233.

18. Javed A, et al. Artificial intelligence applications in dental radiology: diagnostic performance and clinical implementation. Dentomaxillofac Radiol. 2022;51:20210278.

19. Moutselos K, Berdouses E, Oulis C, Maglogiannis I. Recognition of occlusal caries in dental intraoral images using deep learning. Annu Int Conf IEEE Eng Med Biol Soc. 2019;2019:6302-6305.

20. Abbott PV, Saikia A, Anthonappa RP. Artificial intelligence platforms in dental caries detection: a systematic review and meta-analysis. J Evid Based Dent Pract. 2025;25(1):102077.