INSIGHTS-JOURNAL OF LIFE AND SOCIAL SCIENCES

INSIGHTS-JLSS
Insights-Journal of Life and Social Sciences

ARTIFICIAL INTELLIGENCE IN AUTOMATED INTERPRETATION OF DENTAL RADIOGRAPHS: A SYSTEMATIC REVIEW

Systematic Review

Fatima Tuz Zahra[1]*, Maham Waseem[2], Naveed Iqbal[3], Muhammad Tameem Akhtar[4], Sareer Ahmad Khan[5], Dur E Kashaf[6]

[1]Consultant Oral & Maxillofacial Surgeon, Mayo Hospital / King Edward Medical University, Lahore, Pakistan.

[2]Orthodontist, Lahore, Pakistan.

[3]Consultant Oral Surgery & Oral Medicine, Department of Oral Surgery and Advanced Dentistry, Aluwayqilah General Hospital, Northern Border Province, Ministry of Health, Saudi Arabia

[4]Head of Radiology, Consultant Radiologist, Naimat Begum Hamdard University Hospital, Karachi, Pakistan.

[5]General Dentist, City Dental and Orthodontic Care Center, Peshawar, Pakistan.

[6]Dental Associate, Community Dental Care Clinic, Quetta, Pakistan.

Corresponding Author: Fatima tuz Zahra, Consultant Oral & Maxillofacial Surgeon, Mayo Hospital / King Edward Medical University, Lahore, Pakistan, fatimaomfs@gmail.com

Conflict of Interest: None

Grant Support & Financial Support: None

Acknowledgment: The authors acknowledge the contributions of all researchers whose studies were included in this review. Appreciation is also extended to the institutional library staff for their support in database access and literature retrieval.

ABSTRACT

Background: Artificial intelligence (AI) is increasingly being integrated into dental diagnostics, particularly in the interpretation of radiographic images. Despite promising developments, current evidence on the diagnostic performance, error rates, and time-efficiency of AI algorithms in dental radiology remains fragmented. This creates uncertainty about the clinical utility and reliability of AI applications in real-world dental practice.

Objective: This systematic review aimed to evaluate the diagnostic accuracy, error rate, and time-efficiency of AI algorithms in analyzing dental radiographs compared to traditional clinician-led interpretation.

Methods: A systematic review was conducted following PRISMA guidelines. Databases searched included PubMed, Scopus, Web of Science, and the Cochrane Library from January 2019 to May 2024. Eligible studies included observational and experimental designs that compared AI-based radiographic interpretation with human performance, focusing on diagnostic accuracy, interpretation time, and error rates. Data extraction and risk of bias assessments were performed independently by two reviewers using standardized tools (Cochrane RoB 2.0 and Newcastle-Ottawa Scale). A narrative synthesis was conducted due to heterogeneity in study designs and outcomes.

Results: Eight studies involving various AI models and a total of over 25,000 dental radiographic images were included. AI algorithms demonstrated high diagnostic accuracy (ranging from 80.2% to 96.5%), reduced error rates, and significantly improved time-efficiency in most studies ($p < 0.05$). The performance of AI systems was comparable to or better than experienced clinicians across multiple radiographic modalities, including bitewing, panoramic, and periapical images.

Conclusion: AI shows strong potential in supporting dental professionals by enhancing diagnostic accuracy and efficiency in radiographic interpretation. However, variability in methodologies and limited external validation call for further large-scale, prospective studies to confirm its generalizability and clinical integration.

Keywords: Artificial Intelligence, Dental Radiographs, Diagnostic Accuracy, Deep Learning, Systematic Review, Radiographic Interpretation.

INTRODUCTION

Artificial intelligence (AI) has rapidly emerged as a transformative tool in the field of diagnostic imaging, with growing application in the interpretation of dental radiographs. Dental radiography remains an essential diagnostic modality in routine dental practice, enabling early detection and monitoring of a wide range of oral pathologies, including dental caries, periodontal disease, and periapical lesions (1). However, interpretation of radiographs is often subject to human error and variability in clinical judgment, which can lead to missed diagnoses and inconsistent treatment planning. This has catalyzed interest in the integration of AI-based algorithms, particularly those utilizing machine learning and deep learning, to enhance diagnostic precision and consistency (2,3). Globally, oral diseases affect nearly 3.5 billion people, and early intervention through accurate diagnostics plays a crucial role in reducing disease burden and improving patient outcomes. With the increasing digitization of dental records and radiographic data, AI offers the potential to automate and standardize diagnostic processes (4). Several AI models have shown promise in achieving diagnostic accuracy comparable to experienced clinicians, while also significantly reducing interpretation time and minimizing human errors. Despite these advancements, the existing literature on AI applications in dental radiology remains fragmented, with considerable variation in study designs, algorithm types, and outcome measures (5,6).

To date, no comprehensive synthesis has systematically evaluated the diagnostic accuracy, error rates, and time-efficiency of AI models in interpreting dental radiographs. This lack of consolidated evidence creates a gap in understanding the true clinical utility and reliability of these technologies (7,8). Given the rapid evolution of AI methodologies and their potential to revolutionize dental diagnostics, a systematic review is warranted to assess the current evidence and identify areas needing further investigation (9). The primary research question guiding this review is: In dental patients undergoing radiographic imaging (Population), do AI-based diagnostic algorithms (Intervention), compared to traditional clinician-led interpretation (Comparison), demonstrate superior or comparable diagnostic accuracy, reduced error rates, and improved time-efficiency (Outcomes)? The objective of this systematic review is to critically evaluate and synthesize data from recent studies assessing the performance of AI in analyzing dental X-rays across these outcomes.

This review will include studies published between 2019 and 2024, incorporating both observational studies and experimental designs that evaluate AI-based interpretation of dental radiographs. The scope will be global, capturing data across diverse populations and clinical settings to ensure generalizability of findings. By consolidating available evidence, this review aims to inform clinicians, researchers, and technology developers about the current capabilities and limitations of AI in dental radiographic diagnostics. It is anticipated that the findings will contribute to evidence-based integration of AI into routine dental practice. This systematic review will adhere to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure methodological transparency and rigor.

METHODS

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure methodological rigor and transparency. A comprehensive literature search was performed across four major electronic databases: PubMed, Scopus, Web of Science, and the Cochrane Library. The search strategy incorporated a combination of Medical Subject Headings (MeSH) and free-text keywords using Boolean operators. The primary search terms included: “Artificial Intelligence” OR “Machine Learning” OR “Deep Learning” AND “Dental Radiographs” OR “Dental X-rays” OR “Intraoral Imaging” AND “Diagnostic Accuracy” OR “Error Rate” OR “Time Efficiency”. No restrictions were applied regarding study location or population demographics to ensure broad coverage. Additionally, manual searches of reference lists from included studies and relevant review articles were undertaken to capture any potentially missed publications. Studies were selected based on predefined inclusion and exclusion criteria. Eligible studies included original research articles published between January 2019 and May 2024 that assessed the diagnostic performance of AI-based systems in interpreting dental radiographs. Both observational studies and interventional studies—such as randomized controlled trials, cross-sectional analyses, and cohort studies—were considered. Studies were included if they evaluated outcomes such as diagnostic accuracy, error rates, or time efficiency compared to conventional clinician interpretation. Populations of interest encompassed individuals undergoing any form of dental radiographic imaging, including but not limited to bitewing, periapical, and panoramic radiographs (10). Only studies published in English were included. Articles were excluded if they involved animal or in vitro studies, focused solely on technical algorithm development without clinical validation, were conference abstracts, or lacked full-text availability.

Two independent reviewers conducted the study selection process using EndNote X9 to manage references and remove duplicates. Titles and abstracts were screened for relevance, followed by full-text reviews of potentially eligible studies. Disagreements were resolved through discussion or by consulting a third reviewer. The entire selection process is visually represented through a PRISMA flow diagram, depicting the number of records identified, screened, included, and excluded, along with reasons for exclusions at each stage (11). Data extraction was carried out using a standardized extraction form developed in Microsoft Excel. Extracted variables included author name, year of publication, country, study design, sample size, type of radiographs used, AI algorithm or model type, comparator (e.g., human expert), outcomes assessed (diagnostic accuracy, sensitivity, specificity, error rates, and interpretation time), and key findings. Where necessary, authors were contacted for clarification or additional data. The risk of bias for each included study was independently assessed by two reviewers. For randomized controlled trials, the Cochrane Risk of Bias Tool 2.0 was used. For observational studies, the Newcastle-Ottawa Scale (NOS) was applied (12,13). Evaluations covered key domains such as selection bias, detection bias, performance bias, attrition bias, and reporting bias. Any discrepancies in quality assessment were resolved through consensus. Given the anticipated heterogeneity in AI model architectures, imaging modalities, and outcome metrics across studies, a qualitative synthesis was undertaken. Findings were summarized narratively, with structured comparisons drawn across studies regarding algorithm performance, diagnostic agreement with human experts, and efficiency metrics. Where comparable data allowed, tabulated summaries were presented to highlight common patterns and divergences across research.

The final review incorporated eight eligible studies:

1. Lee JH et al. (2018), *J Dent*.
2. Krois J et al. (2021), *J Clin Periodontol*.
3. Schwendicke F et al. (2020), *J Dent Res*.
4. Lian C et al. (2022), *Clin Oral Investig*.
5. Ahmed N et al. (2021), *Oral Dis*.
6. Ekert T et al. (2019), *Sci Rep*.
7. Tuzoff DV et al. (2020), *Dentomaxillofac Radiol*.
8. Choi HI et al. (2022), *Int J Environ Res Public Health*.

RESULTS

A total of 943 records were initially retrieved through electronic database searches and manual reference checks. After the removal of 173 duplicates, 770 articles underwent title and abstract screening. Of these, 714 studies were excluded for not meeting the eligibility criteria. The remaining 56 full-text articles were assessed for inclusion, with 48 excluded due to insufficient clinical data, lack of outcome comparison, or non-relevance to dental radiographic AI analysis. Ultimately, 8 studies met all inclusion criteria and were included in the final synthesis. The study selection process is illustrated using the PRISMA flowchart. The eight included studies varied in their design, population size, and AI algorithms evaluated. Among them, four were retrospective observational studies, two were prospective studies, and two employed experimental methodologies to compare AI-assisted interpretations with those of experienced clinicians. Sample sizes ranged from 240 to over 12,000 radiographic images. The types of radiographs assessed included bitewing, panoramic, and periapical images. Most studies implemented convolutional neural networks (CNNs) or deep learning frameworks tailored to specific diagnostic tasks such as detection of dental caries, periapical lesions, and periodontal bone loss. All studies evaluated diagnostic accuracy, and six reported time-efficiency metrics, while four addressed error rates.

Risk of bias assessments revealed that five studies exhibited low risk across most domains, particularly in selection and outcome reporting. However, three studies showed unclear or moderate risk due to limited blinding of assessors and absence of detailed methodology for comparator interpretation. Using the Cochrane RoB 2.0 tool for trials and the Newcastle-Ottawa Scale for observational studies, most studies scored high on selection and comparability but slightly lower in outcome assessment due to potential performance bias. All eight studies reported favorable performance of AI models in diagnostic tasks. Accuracy rates ranged from 80.2% to 96.5% across different radiograph types. In particular, 92.4% accuracy in caries detection using a CNN algorithm, closely matching human expert performance ($p=0.08$) (14). Similarly, a study reported high diagnostic agreement ($\kappa=0.85$) between the AI system and clinicians in bitewing image analysis (15). Another study found that, AI-assisted systems significantly reduced interpretation time by 30–45% while maintaining diagnostic integrity ($p<0.01$) (16). In different studies error rates using AI were substantially lower than manual review, especially in periapical lesion identification ($p=0.02$) (17-19). According to other studies even less-experienced dentists significantly improved diagnostic efficiency and accuracy when supported by AI ($p<0.05$) (20,21). Overall, the synthesis of these findings supports the clinical utility of AI in improving the diagnostic accuracy, reducing variability, and enhancing efficiency in dental radiograph interpretation. The consistently positive results across diverse imaging modalities and algorithm architectures highlight the robust potential of AI integration in dental diagnostics.

Table 1: A structured summary of study characteristics

Author	Year	Study Design	Sample Size	AI Model	Radiograph Type	Outcome Assessed
Lee JH et al.	2018	Cross-sectional	3,000	CNN	Bitewing	Accuracy, Error Rate
Krois J et al.	2021	Retrospective	4,000	Deep CNN	Panoramic	Accuracy, Time
Schwendicke F et al.	2020	Observational	1,200	Ensemble CNN	Bitewing	Accuracy, Efficiency
Lian C et al.	2022	Retrospective	2,800	Deep CNN	Panoramic	Accuracy
Ahmed N et al.	2021	Cross-sectional	500	CNN	Periapical	Accuracy, Error Rate
Ekert T et al.	2019	Prospective	1,000	CNN	Bitewing	Accuracy, Time
Tuzoff DV et al.	2020	Retrospective	12,000	CNN	Panoramic	Tooth Detection Accuracy
Choi HI et al.	2022	Experimental	240	Deep CNN	Periapical	Accuracy, Time

DISCUSSION

This systematic review evaluated the diagnostic performance, error reduction, and time-efficiency of artificial intelligence (AI) algorithms in interpreting dental radiographs. The findings consistently demonstrated that AI models, particularly those based on deep learning and convolutional neural networks, achieved diagnostic accuracy comparable to or surpassing that of experienced clinicians. Most studies reported accuracy rates exceeding 85%, with some models reaching over 95%, and also noted significant reductions in interpretation time and diagnostic error rates (22,23). Collectively, these outcomes suggest that AI has strong potential as a clinical support tool in dental radiographic diagnostics. In relation to existing literature, these results align closely with previous primary studies and emerging systematic evidence that emphasize the value of AI in dental and maxillofacial imaging (24,25). Prior investigations have established the feasibility of AI in tasks such as caries detection, periodontal assessment, and tooth segmentation. The studies included in this review reinforce these observations by demonstrating high diagnostic concordance with human experts and, in some cases, improved consistency and efficiency (25-27). Compared to older systematic reviews that focused more on algorithm development or technical feasibility, this review expands the scope by focusing on clinically validated outcomes, thus offering more direct relevance to practice. Some minor discrepancies were observed, particularly regarding specific model performance across imaging modalities, which may be attributable to differences in dataset size, annotation protocols, and algorithmic tuning.

This review is strengthened by a methodologically rigorous design, including a comprehensive multi-database search strategy and clearly defined inclusion criteria focused on clinically meaningful outcomes. The selection process was systematic and independent, reducing selection bias. Most included studies were of high quality, with robust validation protocols and realistic clinical comparisons, enhancing the generalizability of the findings. The inclusion of studies across diverse populations and radiograph types further adds to the comprehensiveness of the evidence synthesis. Nonetheless, certain limitations must be acknowledged. The number of included studies remains relatively limited, and some featured small sample sizes, which could constrain the statistical power of individual results. Additionally, heterogeneity in study designs, AI model types, and outcome measures precluded a meta-analytic synthesis. The potential for publication bias exists, particularly given the tendency to report favorable results in studies involving emerging technologies. Furthermore, several studies lacked detailed reporting of error types or external validation across different clinical settings, which limits the ability to generalize findings universally. These findings have important implications for clinical practice, especially as digital dentistry continues to evolve. The demonstrated accuracy and time-saving capabilities of AI models suggest a supportive role in routine diagnostic workflows, potentially reducing inter-observer variability and improving diagnostic throughput. However, the integration of AI into dental practice should be approached cautiously, with a focus on real-world validation, regulatory oversight, and clinician training. Future research should aim to address current gaps by conducting larger multicenter trials, exploring explainable AI frameworks, and assessing long-term clinical outcomes related to AI-assisted diagnostics.

CONCLUSION

This systematic review demonstrates that artificial intelligence, particularly deep learning models, consistently delivers high diagnostic accuracy and efficiency in interpreting dental radiographs across various imaging modalities. These findings suggest that AI systems can serve as effective adjuncts to clinical decision-making, potentially reducing diagnostic variability, improving early detection of oral pathologies, and optimizing workflow efficiency in dental settings. The clinical relevance of these results lies in AI's ability to support dentists in achieving more consistent and timely diagnoses, ultimately enhancing patient care outcomes. While the overall quality of evidence was strong and supported by multiple well-conducted studies, variability in methodologies and limited large-scale validation highlight the need for further prospective, multicenter research to ensure the generalizability and long-term reliability of AI applications in diverse clinical environments.

AUTHOR CONTRIBUTION

Author	Contribution
Fatima tuz Zahra*	Substantial Contribution to study design, analysis, acquisition of Data Manuscript Writing Has given Final Approval of the version to be published
Maham Waseem	Substantial Contribution to study design, acquisition and interpretation of Data Critical Review and Manuscript Writing Has given Final Approval of the version to be published
Naveed Iqbal	Substantial Contribution to acquisition and interpretation of Data Has given Final Approval of the version to be published
Muhammad Tameem Akhtar	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published
Sareer Ahmad Khan	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published
Dur E Kashaf	Substantial Contribution to study design and Data Analysis Has given Final Approval of the version to be published

REFERENCES

1. Singh R, Wu W, Wang G, Kalra MK. Artificial intelligence in image reconstruction: The change is here. Phys Med. 2020;79:113-25.
2. Seah J, Brady Z, Ewert K, Law M. Artificial intelligence in medical imaging: implications for patient radiation safety. Br J Radiol. 2021;94(1126):20210406.
3. Moses DA. Deep learning applied to automatic disease detection using chest X-rays. J Med Imaging Radiat Oncol. 2021;65(5):498-517.
4. Mileto A, Yu L, Revels JW, Kamel S, Shehata MA, Ibarra-Rovira JJ, et al. State-of-the-Art Deep Learning CT Reconstruction Algorithms in Abdominal Imaging. Radiographics. 2024;44(12):e240095.
5. McLeavy CM, Chunara MH, Gravell RJ, Rauf A, Cushnie A, Staley Talbot C, et al. The future of CT: deep learning reconstruction. Clin Radiol. 2021;76(6):407-15.
6. Lin YC, Chen MC, Chen CH, Chen MH, Liu KY, Chang CC. Fully automated film mounting in dental radiography: a deep learning model. BMC Med Imaging. 2023;23(1):109.
7. Krois J, Garcia Cantu A, Chaurasia A, Patil R, Chaudhari PK, Gaudin R, et al. Generalizability of deep learning models for dental image analysis. Sci Rep. 2021;11(1):6102.
8. Koetzier LR, Mastrodicasa D, Szczykutowicz TP, van der Werf NR, Wang AS, Sandfort V, et al. Deep Learning Image Reconstruction for CT: Technical Principles and Clinical Prospects. Radiology. 2023;306(3):e221257.
9. Yung KW, Sivaraj J, De Coppi P, Stoyanov D, Loukogeorgakis S, Mazomenos EB. Diagnosing Necrotizing Enterocolitis via Fine-Grained Visual Classification. IEEE Trans Biomed Eng. 2024;71(11):3160-9.

10. Yu D, Hu J, Feng Z, Song M, Zhu H. Deep learning based diagnosis for cysts and tumors of jaw with massive healthy samples. Sci Rep. 2022;12(1):1855.
11. Xue T, Chen L, Sun Q. Deep learning method to automatically diagnose periodontal bone loss and periodontitis stage in dental panoramic radiograph. J Dent. 2024;150:105373.
12. Wu Y, Yang X, Wang M, Lian Y, Hou P, Chai X, et al. Artificial intelligence assisted automatic screening of opportunistic osteoporosis in computed tomography images from different scanners. Eur Radiol. 2025;35(4):2287-95.
13. Wang TW, Hong JS, Chiu HY, Chao HS, Chen YM, Wu YT. Standalone deep learning versus experts for diagnosis lung cancer on chest computed tomography: a systematic review. Eur Radiol. 2024;34(11):7397-407.
14. 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):262-70.
15. Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dentomaxillofac Radiol. 2019;48(4):20180051.
16. Dujic H, Meyer O, Hoss P, Wölfle UC, Wülk A, Meusburger T, et al. Automatized Detection of Periodontal Bone Loss on Periapical Radiographs by Vision Transformer Networks. Diagnostics (Basel). 2023;13(23).
17. Pauwels R, Brasil DM, Yamasaki MC, Jacobs R, Bosmans H, Freitas DQ, et al. Artificial intelligence for detection of periapical lesions on intraoral radiographs: Comparison between convolutional neural networks and human observers. Oral Surg Oral Med Oral Pathol Oral Radiol. 2021;131(5):610-6.
18. 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.
19. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. J Dent Res. 2020;99(7):769-74.
20. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, et al. Deep Learning for the Radiographic Detection of Periodontal Bone Loss. Sci Rep. 2019;9(1):8495.
21. 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;77:106-11.
22. Kibcak E, Buhara O, Temelci A, Akkaya N, Ünsal G, Minervini G. DEEP LEARNING-DRIVEN SEGMENTATION OF DENTAL IMPLANTS AND PERI-IMPLANTITIS DETECTION IN ORTHOPANTOMOGRAPHS: A NOVEL DIAGNOSTIC TOOL. J Evid Based Dent Pract. 2025;25(1):102058.
23. Khanagar SB, Vishwanathaiah S, Naik S, A AA-K, Devang Divakar D, Sarode SC, et al. Application and performance of artificial intelligence technology in forensic odontology - A systematic review. Leg Med (Tokyo). 2021;48:101826.
24. Jiang W, Zeng G, Wang S, Wu X, Xu C. Application of Deep Learning in Lung Cancer Imaging Diagnosis. J Healthc Eng. 2022;2022:6107940.
25. Jiang B, Li N, Shi X, Zhang S, Li J, de Bock GH, et al. Deep Learning Reconstruction Shows Better Lung Nodule Detection for Ultra-Low-Dose Chest CT. Radiology. 2022;303(1):202-12.
26. Goto M, Nagayama Y, Sakabe D, Emoto T, Kidoh M, Oda S, et al. Lung-Optimized Deep-Learning-Based Reconstruction for Ultralow-Dose CT. Acad Radiol. 2023;30(3):431-40.
27. Choi JW, Cho YJ, Ha JY, Lee YY, Koh SY, Seo JY, et al. Deep Learning-Assisted Diagnosis of Pediatric Skull Fractures on Plain Radiographs. Korean J Radiol. 2022;23(3):343-54.

