Systematic Review

Artificial Intelligence-Based Dental Caries Detection in Cone Beam Computed Tomography Images: A Systematic Review

Mitra Montazerlotf [1], PhD; Mehrdad Hosseini Shakib [2], PhD; Mohamad Javad Kharazifard [3], PhD; Reza Radfar [4], PhD;

[1] Dept. of Information Technology Management, Faculty of Management and Economics, SR.C., Islamic Azad University, Tehran, Iran.
[2] Dept. of Industrial Management, Ka.C., Islamic Azad University, Karaj, Iran.
[3] Dental Research Center, Dentistry Research Institute, Tehran University of Medical Sciences, Tehran, Iran.
[4] Dept. of Industrial Management, SR.C., Islamic Azad University, Tehran, Iran.

KEY WORDS
Artificial Intelligence;
Cone Beam Computed Tomography;
Dental Caries;
Machine Learning;
Deep Learning;

ABSTRACT

Background: Dental caries remains one of the most prevalent oral diseases worldwide. Although cone beam computed tomography (CBCT) is not routinely indicated for caries detection due to radiation dose considerations, it provides three-dimensional (3D) information that may support caries assessment when CBCT scans are already clinically justified. Artificial intelligence (AI), particularly machine learning (ML) and deep learning approaches, has shown potential in improving diagnostic consistency and efficiency in dental imaging.

Purpose: This systematic review aimed to evaluate current evidence on the application of AI-based methods for dental caries detection and classification using CBCT images.

Materials and Method: A systematic literature search was conducted in PubMed, Scopus, Web of Science, IEEE Xplore, and Google Scholar for studies published between 2018 and 2024. Eligible studies applied AI techniques to CBCT images for caries detection and reported quantitative diagnostic performance metrics. Study selection, data extraction, and quality assessment were performed according to preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines.

Results: Five studies meeting the inclusion criteria were analyzed. Reported AI models, predominantly convolutional neural networks (CNNs), demonstrated sensitivity ranging from 86.67% to 96.5% and specificity from 91.3% to 99.46% in controlled study settings. In limited clinical evaluation studies, AI assistance improved clinician sensitivity while maintaining specificity and reducing assessment time.

Conclusion: AI-based approaches applied to CBCT images show promising preliminary diagnostic performance for caries detection in controlled research settings. However, the limited number of available studies, heterogeneity in study design, predominantly retrospective single-center designs, and lack of external validation restrict the generalizability and clinical applicability of these findings. Further large-scale, multi-center prospective investigations with external validation are required before routine clinical implementation can be recommended.

Received: 16 December 2025;
Revised: 31 January 2026;
Accepted: 10 May 2026;

Corresponding Author: Hosseini Shakib M, Dept. of Industrial Management, Ka.C., Islamic Azad University, Karaj, Iran. Email: mehrdad.shakib@kiau.ac.ir

Cite this article as:

Introduction

Dental caries remains one of the most prevalent chronic oral diseases worldwide, affecting more than 2.5 billion individuals and posing a substantial burden on public health systems [1-2]. Early and accurate detection of carious lesions is essential to prevent disease progression, reduce the need for invasive treatments, and preserve tooth structure [3-4]. In routine clinical practice, caries diagnosis is primarily based on visual–tactile examination combined with two-dimensional (2D) radiographic modalities, particularly bitewing and periapical radiographs, which are considered the current standard of care [5-6].

Despite their widespread use, conventional diagnostic approaches have inherent limitations. Visual inspection is subjective and highly dependent on clinician experience, while 2D radiography is affected by anatomical superimposition and limited sensitivity for early or non-cavitated lesions [6]. Previous studies have reported that up to 30–40% of proximal carious lesions may remain undetected in their early stages using conventional radiography alone [6-7]. These limitations have driven ongoing research into adjunctive diagnostic tools aimed at improving diagnostic accuracy and consistency.

Cone beam computed tomography (CBCT) provides three-dimensional (3D) volumetric imaging that eliminat-es anatomical overlap and allows multiplanar visualiz-ation of dental structures [7-8]. CBCT has demonstrated improved diagnostic capability for certain dental conditi-ons, including assessment of complex root morphology, periapical lesions, and anatomical variations [9-11]. Seve-ral studies have also explored its diagnostic performance in detecting interproximal, occlusal, and root caries [12-13].

However, CBCT is not routinely indicated for caries detection due to its higher radiation dose, increased cost, and the principles of radiation protection, particularly the principles of as low as reasonably achievable (ALARA) and as low as diagnostically acceptable (ALADA), which mandate justification and optimization of exposure [14-15] .Consequently, CBCT should not be acquired solely for the purpose of caries diagnosis, and its role in this context is limited to situations in which scans are already clinically justified for other diagnostic reasons [5,14].

In recent years, artificial intelligence (AI) including machine learning (ML) and deep learning approaches such as convolutional neural networks (CNNs), has shown substantial potential in medical image analysis by enabling automated feature extraction and pattern recognition [16-17]. In dentistry, AI-based systems have been successfully applied to various diagnostic tasks, including periapical lesion detection, tooth segmentation, and mandibular canal identification [18-20].

Most AI research on dental caries detection to date has focused on 2D imaging modalities, particularly bitewing and periapical radiographs, where AI systems have achieved diagnostic performance comparable to experienced clinicians while reducing interpretation time and inter-observer variability [21-26]. In contrast, the application of AI to CBCT-based caries detection remains relatively underexplored, partly due to the increased complexity of volumetric data and the limited availability of annotated CBCT datasets.

Recent methodological advances, including transfer learning, self-supervised learning, and interpretable AI models, have begun to address some of these challenges by improving performance with limited labeled data and enhancing model transparency [27-31]. These developments raise important questions regarding the potential role of AI as a decision-support tool for caries assessment on CBCT images that are already available from clinically justified scans.

Although several narrative and systematic reviews have examined AI applications in dental radiography more broadly, most have predominantly focused on 2D imaging modalities and have not specifically addressed CBCT-based caries detection [24-26]. Therefore, a focused synthesis of the existing evidence is needed to clarify the current state of AI applications in CBCT caries detection, identify methodological limitations, and outline directions for future research.

The aim of this systematic review was to critically evaluate the available evidence on AI-based detection and classification of dental caries using CBCT images. Specifically, this review assesses the AI architectures and methodologies employed, diagnostic performance metrics, comparison with human expert performance where available, and key technical and clinical challenges. Importantly, this review considers AI-assisted CBCT caries detection as an adjunctive approach rather than a justification for CBCT acquisition, in accordance with current clinical guidelines and radiation protection principles.

Materials and Method

Study Design

This systematic review was conducted according to the

Table 1: Search Strategy Concept Groups

Concept Group	Search Terms
Concept 1- Dental Caries	"dental caries" OR "tooth decay" OR "dental cavity" OR "cavities" OR "carious lesion" OR "enamel caries" OR "dentin caries" OR "root caries"
Concept 2- CBCT Imaging	"cone beam computed tomography" OR "CBCT" OR "cone-beam CT" OR "volumetric tomography" OR "digital volume tomography"
Concept 3- AI	"artificial intelligence" OR "AI" OR "machine learning" OR "deep learning" OR "neural network" OR "convolutional neural network" OR "CNN" OR "automated detection" OR "computer-aided diagnosis"
Combined Search Formula	(Concept 1) AND (Concept 2) AND (Concept 3)

Abbreviations: cone beam computed tomography (CBCT); artificial intelligence (AI); convolutional neural network (CNN)

PRISMA guidelines. The review protocol focused on studies applying AI and ML techniques to dental caries detection in CBCT images.

Search Strategy

A comprehensive literature search was performed across five major electronic databases: PubMed/ MEDLINE, Scopus, Web of Science, IEEE Xplore, and Google Scholar. To ensure the comprehensiveness of the search, multiple biomedical, technical, and interdisciplinary databases were systematically queried, and reference lists of included studies were manually screened. No additional eligible studies meeting the predefined inclusion criteria were identified through these supplementary searches.

The search period spanned from January 2018 to September 2024. The year 2018 was selected as the starting point of the search because the application of deep learning architectures, particularly CNNs, to dental imaging and CBCT analysis became methodologically mature and more widely adopted from this period onward. Earlier studies were largely exploratory or based on traditional ML approaches with limited clinical relevance.

The search strategy combined three concept groups using Boolean operators: dental caries terminology, CBCT imaging terminology, and AI terminology. The detailed search terms for each concept group and their combinations are presented in Table 1. Database-specific search strings with applied filters are detailed in Table 2. This comprehensive search strategy ensured broad coverage of relevant literature while maintaining specificity to the research question.

Eligibility Criteria

Studies were included in this systematic review if they met the following criteria. First, only original research articles published in peer-reviewed journals were considered eligible. Second, studies had to be published in English language to ensure accurate interpretation and analysis. Third, the research focus needed to be specifically on dental caries detection or classification using CBCT images. Fourth, studies had to apply AI or ML methods for image analysis. Fifth, quantitative performance metrics such as accuracy, sensitivity, specificity, or F1-score had to be provided. Finally, studies had to involve either human subjects or *ex vivo* human teeth specimens to ensure clinical relevance. Studies were excluded from this review under several conditions. Review articles, conference abstracts, letters, editorials, and case reports were not considered as they do not pre-

Table 2: Database-Specific Search Strings

Database	Search String	Filters Applied
PubMed	(("dental caries"[MeSH] OR "tooth decay"[Title/Abstract] OR "cavity"[Title/Abstract]) AND ("cone beam computed tomography"[MeSH] OR "CBCT"[Title/Abstract]) AND ("artificial intelligence"[MeSH] OR "machine learning"[Title/Abstract] OR "deep learning"[Title/Abstract]))	Date: 2018-2024; Language: English; Article types: Journal Article
Scopus	TITLE-ABS-KEY((("dental caries" OR "tooth decay") AND ("CBCT" OR "cone beam computed tomography") AND ("artificial intelligence" OR "machine learning" OR "deep learning")))	Year: 2018-2024; Document type: Article; Language: English
Web of Science	TS=(("dental caries" OR "tooth decay") AND ("CBCT" OR "cone beam CT") AND ("artificial intelligence" OR "machine learning" OR "neural network"))	Timespan: 2018-2024; Document types: Article; Language: English
IEEE Xplore	(("dental caries" OR "tooth decay") AND ("CBCT" OR "cone beam computed tomography") AND ("artificial intelligence" OR "machine learning" OR "deep learning"))	Content Type: Journals; Year: 2018-2024
Google Scholar	"dental caries" AND "CBCT" AND ("artificial intelligence" OR "machine learning")	Date: 2018-2024 (sorted by relevance)

Abbreviations: cone beam computed tomography (CBCT); artificial intelligence (AI); Medical Subject Headings (MeSH); title-abstract-keywords (TITLE-ABS-KEY); topic search (TS) sent original research data. Studies using imaging modalities other than CBCT were excluded to maintain clouded to maintain focus on the specific imaging technique of interest. Research not focused on caries detection, even if utilizing CBCT and AI, was excluded. Studies with insufficient quantitative results for performance evaluation were also excluded, as objective assessment of diagnostic performance was essential to this review. Additionally, studies employing non-AI diagnostic methods were excluded as they fell outside the scope of this systematic review. Finally, studies for which full text was not available in English were excluded to ensure accurate data extraction and interpretation.

Study Selection Process

Two independent reviewers conducted the screening of titles and abstracts of all retrieved articles. The first reviewer was a specialist in oral and maxillofacial radiology with expertise in CBCT interpretation and dental diagnosis. The second reviewer was a specialist in information technology management with expertise in AI and ML applications in healthcare. Following the initial screening, full-text articles of potentially eligible studies were assessed against the predetermined inclusion and exclusion criteria by both reviewers independently. Any disagreements between the two reviewers were resolved through discussion and consensus. In cases where consensus could not be reached, a third reviewer with expertise in dental research methodology was consulted. The entire selection process was documented using a PRISMA flow diagram that showed the number of records identified, screened, and included or excluded at each stage, along with the specific reasons for exclusion.

It should be noted that the final number of included studies (n=5) reflects the strict eligibility criteria and the current scarcity of original research specifically addressing AI-based dental caries detection using CBCT imaging. A substantial number of potentially relevant articles were excluded because they relied on 2D radiographic modalities or did not apply AI-based methods, underscoring the limited and emerging nature of this research field rather than limitations of the search strategy.

Data Extraction

For each included study, data were systematically extracted using a standardized data extraction form. Study characteristics including authors, year of publication, country of origin, and study design were recorded. Sample characteristics encompassed sample size, CBCT devices used, and patient population demographics when available. AI methodology details were comprehensively extracted, including preprocessing techniques applied to CBCT images, specific AI architectures employed, and training strategies utilized. Performance metrics were carefully documented, including accuracy, sensitivity, specificity, and F1-score when reported. Comparison with human expert performance was extracted when available in the original studies. Finally, challenges and limitations reported by the study authors were systematically recorded to identify common obstacles and areas requiring further research.

Quality Assessment

Study quality was assessed using adapted versions of quality assessment of diagnostic accuracy studies for AI (QUADAS-AI) and the checklist for AI in medical imaging (CLAIM checklist). The quality assessment focused on several key domains. Patient selection and reference standard were evaluated to ensure appropriate study populations and gold standard comparisons. AI model development and validation methodology were assessed to determine the rigor of the ML approaches. Performance evaluation and reporting standards were examined to ensure transparent and comprehensive reporting of results. Finally, risk of bias in study design and conduct was systematically evaluated across all included studies to provide an overall assessment of the evidence quality. CBCT imaging quality assessment protocols similar to those employed in recent maxillofacial studies [10, 25] were considered when evaluating the technical adequacy of included investigations.

Given the methodological heterogeneity and limited number of included studies, a formal GRADE assessment was not applied. However, the risk of bias and reporting quality were systematically evaluated using QUADAS-AI and CLAIM checklists to provide a structured appraisal of the strength and limitations of the available evidence.

A meta-analysis was not performed due to the limited number of eligible studies, substantial heterogeneity in study design, AI architectures, outcome definitions, and performance metrics, as well as differences in CBCT devices and validation strategies. Under these conditions, quantitative synthesis was considered methodologically inappropriate and potentially misleading.

Table 3: QUADAS-AI Risk of Bias Assessment

Study	Patient Selection	Index Test (AI)	Reference Standard	Flow & Timing	Overall Risk
Ezhov *et al.* [33]	Low	Low	Low	Low	Low
Esmaeilyfard *et al.* [3]	Unclear	Low	Unclear	Low	Moderate
Chen *et al.* [34]	Low	Low	Low	Low	Low
Amasya *et al.* [4]	Low	Low	Low	Low	Low
Zanini *et al.* [29]	Unclear	Low	Unclear	Low	Moderate

Abbreviations: artificial intelligence (AI);

Therefore, a qualitative systematic synthesis was undertaken in accordance with PRISMA recommendations.

Results

Study Selection

The systematic literature search and selection process is illustrated in the PRISMA flow diagram (Figure 1). The initial database search identified 847 records across all five electronic databases. After removing duplicates and screening by title and abstract, 89 articles underwent full-text assessment. Finally, five studies met all inclusion criteria and were included in this systematic review [3-4, 29, 33-34]. Quality assessment using adapted QUADAS-AI criteria revealed generally low risk of bias across most domains (Table 3). Three studies [3, 33, 34] demonstrated low risk across all assessed domains. Two studies [4, 29] showed unclear risk in patient selection or reference standard domains due to insufficient reporting of consecutive enrollment and blinding procedures. Among the included studies, three employed a multi-center design, while two were conducted in single-center settings, as detailed in Table 4.

The primary reasons for exclusion at the full-text stage were non-use of CBCT imaging (n=32), absence of AI/ML methods (n=26), insufficient performance data (n=15), and lack of focus on caries detection (n=11).

Characteristics of Included Studies

The five included studies were published between 2021 and 2024, representing the most recent advances in AI-based caries detection using CBCT [3-4,29,33-34]. Three of the included studies employed a multi-center design, while the remaining two were conducted as single-center studies. Sample sizes varied considerably, ranging from 204 to 1,346 CBCT scans, reflecting different research designs and objectives. Table 4 presents a comprehensive overview of the characteristics of all included studies, including publication year, country of origin, sample size, CBCT devices used, types of caries investigated, AI architectures employed, and main findings.

CBCT Image Preprocessing Techniques

All included studies employed preprocessing steps to prepare CBCT images for AI model training, recognizing the critical importance of image standardization for optimal performance [3-4, 29, 33-34].

Table 4: Characteristics of Included Studies

Study	Year	Country	Sample Size	CBCT Device(s)	Caries Types	AI Architecture	Main Findings
Ezhov *et al.* [33]	2021	Multi-center (3 sites)	1,346 training scans + 30 clinical evaluation	Multiple devices	Overall caries detection	Modified 3D *U-Net* + ResNeXt/ DenseNet (Diagnocat system)	Sensitivity: 92.39%, Specificity: 98.99%
Esmaeilyfard *et al.*[3]	2024	Iran	785 molars (382 carious, 403 healthy)	Not specified	Type I-IV (location), D1-D3 (depth)	Multi-input CNN (axial, sagittal, coronal)	Accuracy: 93-97%, Sensitivity: 87-96%, Specificity: 91-97%
Chen *et al.* [34]	2024	China	2,713 slices from 204 teeth	Not specified	Proximity to pulp classification	ResNet50_vd, MobileNetV3 + interpretable model with *U-Net*/ DeepLabv3 +/ OCR-Net	Interpretable model: >91.7% all metrics vs. Direct model: 70% accuracy
Amasya *et al.* [4]	2023	Multi-center	500 CBCT volumes (clinical evaluation: 30 scans, 24 dentists)	Multiple devices	Overall caries detection	Diagnocat system (same as Ezhov)	AI-assisted: Sensitivity 85.37% vs. unassisted 76.72%; 6.78% time reduction
Zanini *et al.*[29]	2024	Brazil	Unlabeled + labeled CBCT data	Not specified	ICDAS classification	*ResNet*-18 + SimCLR (self-supervised learning)	F1-score: 88.42%, 5.5% improvement with SSL

Abbreviations: cone beam computed tomography (CBCT); artificial intelligence (AI); convolutional neural network (CNN); International Caries Detection and Assessment System (ICDAS); self-supervised learning (SSL).

Table 5: Performance Metrics of AI Systems

Study	Accuracy	Sensitivity	Specificity	F1-Score	Additional Metrics
Ezhov *et al.* [33] (Diagnocat)	-	92.39%	98.99%	-	Comparable to radiologists (93.18-94.38% sensitivity)
Esmaeilyfard *et al.* [3] - Type I (Occlusal)	93.3%	92.3%	94.6%	92.8%	-
Esmaeilyfard *et al.* [3] - Type II (Proximal)	93.7%	95.2%	94.2%	94.4%	-
Esmaeilyfard *et al.* [3] - Depth D1	89.7%	87.7%	90.6%	89.0%	-
Esmaeilyfard *et al.* [3] - Depth D2	93.3%	95.0%	94.2%	93.6%	-
Esmaeilyfard *et al.* [3] - Depth D3	96.2%	96.5%	97.5%	97.3%	-
Chen *et al.* [34] - Direct model	70.0%	60.6%	-	61.6%	Precision: 78.6%
Chen *et al.* [34] - Interpretable model	>91.7%	>91.7%	-	>91.7%	All metrics >91.7%
Amasya *et al.* [4] - AI-assisted	-	85.37%	96.72%	-	Time: 17.55 min
Amasya *et al.* [4] - Unassisted	-	76.72%	96.16%	-	Time: 18.74 min
Zanini *et al.* [29] (with SSL)	-	86.67%	-	88.42%	Precision: 90.44%

Abbreviations: artificial intelligence (AI); self-supervised learning (SSL)

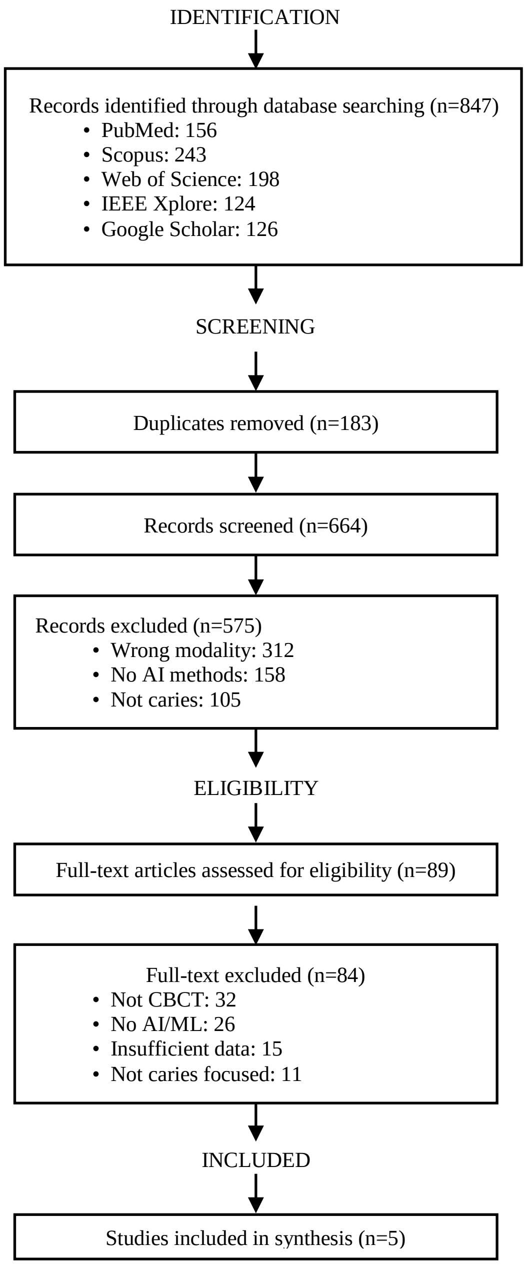

Figure 1: PRISMA Flow Diagram (Abbreviations**:** artificial intelligence (AI); cone beam computed tomography (CBCT); machine learning (ML))

Ezhov *et al.* [33] implemented comprehensive normalization of Hounsfield unit values, including clipping values below -1000 to remove air density artifacts, removing outliers below the 5th and above the 95th percentile, and standardizing the remaining values by subtracting the mean and dividing by the standard deviation to achieve consistent intensity distributions across different CBCT devices.

Data augmentation techniques were utilized to increase dataset size and improve model generalizability. Esmaeilyfard *et al.* [3] applied vertical and horizontal flipping, rotation within ±20 degrees, and zooming up to two times magnification, achieving a tenfold increase in training data. Zanini *et al.* [29] combined traditional image processing with self-supervised learning to extract valuable information from unlabeled CBCT data, addressing the challenge of limited annotated medical imaging datasets.

Region of interest extraction was implemented differently across studies. Chen *et al.* [34] employed center-slice sampling with 60×60 pixel cropping and random shifts of ±30 pixels to introduce positional diversity. Ezhov *et al.* [33] developed a multi-stage processing pipeline that progressively refined analysis from coarse 3mm³ to fine 0.15mm³ voxel resolution, balancing computational efficiency with diagnostic precision. Amasya *et al.* [4] applied device-specific software for artifact removal to standardize image quality across different CBCT devices in their multi-center study.

Diagnostic Performance

AI systems across all included studies demonstrated consistently high diagnostic performance that met or exceeded clinically relevant thresholds. The detailed performance metrics including accuracy, sensitivity,

Table 6: Performance by Caries Location (Esmaeilyfard *et al.* [3])

Caries Type	Location	Accuracy	Sensitivity	Specificity	F1-Score
Type I	Occlusal	93.3%	92.3%	94.6%	92.8%
Type II	Proximal	93.7%	95.2%	94.2%	94.4%
Type III	Cervical	91.6%	88.7%	91.3%	89.4%
Type IV	Multiple sites	97.2%	95.1%	96.4%	94.0%

Table 7: Performance by Caries Depth (Esmaeilyfard *et al.* [3])

Depth	Description	Accuracy	Sensitivity	Specificity	F1-Score
D1	Enamel only	89.7%	87.7%	90.6%	89.0%
D2	Into dentin	93.3%	95.0%	94.2%	93.6%
D3	Near/into pulp	96.2%	96.5%	97.5%	97.3%

specificity, and F1-score for all included studies are presented in Table 5.

Several key findings emerge from the performance analysis. Performance varied by caries location, with Esmaeilyfard *et al.* [3] achieving highest accuracy for Type IV (multiple sites, 97.2%) and Type II (proximal, 93.7%) caries as detailed in Table 6. AI systems demonstrated high overall performance with sensitivity ranging from 86.67% to 96.5% and specificity from 91.3% to 99.46%, meeting clinical diagnostic accuracy requirements. Caries depth detection showed improved performance with increasing lesion depth, achieving 89.7%, 93.3%, and 96.2% accuracy for D1, D2, and D3 respectively [3], likely due to more distinct radiographic features in advanced caries. Table 7 presents comprehensive performance metrics across all depth categories, demonstrating progressive improvement in sensitivity and specificity with increasing lesion severity. Interpret ability provided substantial benefits, with Chen *et al.* [34] demonstrating performance improvement exceeding 20 percentage points when incorporating interpretable components, effectively addressing the "black box" limitation. Self-supervised learning showed clear advantages, with Zanini *et al.* [29] demonstrating 5.5% F1-score improvement, particularly valuable when labeled data is limited.

Comparison with Human Expert Performance

Three studies provided direct comparison between AI systems and human expert performance. Ezhov *et al.* [33] demonstrated that the Diagnocat system achieved performance comparable to experienced oral and maxillofacial radiologists, with AI sensitivity of 92.39% nearly matching human radiologist sensitivity of 93.18% to 94.38%, and AI specificity of 98.99% matching human specificity of 98.99% to 99.46%. This comparable performance suggested that AI systems have matured to match specialist expertise in caries detection.

Amasya *et al.* [4] conducted the most comprehensive clinical evaluation with 24 dentists performing caries detection with and without AI assistance. Sensitivity improved significantly from 76.72% to 85.37% with AI assistance, representing an 8.65% increase in detecting true carious lesions. Specificity remained stable at 96.16% to 96.72%, indicating AI assistance did not increase false positive diagnoses. Assessment time decreased from 18.74 to 17.55 minutes, representing a 6.78% time reduction. Additionally, inter-observer agreement improved with AI assistance, suggesting AI systems can help standardize diagnostic criteria and reduce subjective variability among clinicians.

Chen *et al.* [34] validated their AI-derived classifica-

Table 8: CNN Architecture Specifications

Study	Architecture	Input Dimensions	Key Components	Training Details
Ezhov *et al.* [33]	Modified 3D *U-Net* + ResNeXt/DenseNet	3D volumetric (varying voxel sizes: 3mm³ to 0.15mm³)	- 3D convolutions - Squeeze-excite blocks - Multi-stage pipeline	- Multi-center data - Rare case mining - K-fold validation
Esmaeilyfard *et al.* [3]	Multi-input CNN	96×160×1 per view (axial, sagittal, coronal)	- Separate conv paths - Concatenation layer - Fully connected layers	- Data augmentation (10× increase) - 100 epochs - Adam optimizer
Chen *et al.* [34]	*ResNet*50_vd, MobileNetV3, + segmentation networks	2D slices 60×60	- Direct classification - Interpretable with *U-Net*/Deep Labv 3+/OCR-Net - LIME visualization	- Transfer learning - Pre-segmentation - ImageNet pretrained
Amasya *et al.* [4]	Diagnocat (same as Ezhov)	3D volumetric	- 5-module system - Hierarchical processing	- Commercial system - Multi-device validation
Zanini *et al.* [29]	*ResNet*-18 + SimCLR	2D slices	- Self-supervised pretraining - Contrastive learning - Fine-tuning	- Unlabeled data extraction - SSL pretraining - Supervised fine-tuning

Abbreviations: convolutional neural network (CNN); self-supervised learning (SSL); local interpretable model-agnostic explanations (LIME) tion system based on minimum distance between caries and dental pulp with clinical specialists. Treatment decisions showed significant correlation with AI-generated distance measurements, with correlation coefficients ranging from -0.46 to -0.50 ($p< 0.0001$), demonstrating that AI-derived metrics were clinically relevant and aligned with expert clinical decision-making.

Discussion

This systematic review of five studies with generally low to moderate risk of bias demonstrates that AI-based systems show promising accuracy in detecting dental caries from CBCT images [3-4, 29, 33-34]. The evidence reveals several key preliminary findings. First, AI systems demonstrated diagnostic performance in controlled study settings with sensitivity ranging from 86.67% to 96.5% and specificity from 91.3% to 99.46%. In two studies that included direct comparison [4, 33], AI performance approached that of experienced radiologists under specific study conditions.

Second, CNN-based architectures, particularly *U-Net* and *ResNet* variants, proved most effective [3, 29, 33-34], with multi-module hierarchical systems successfully balancing efficiency and accuracy [4, 33]. Third, advanced techniques including interpretable models and self-supervised learning provided significant performance improvements [29, 34] addressing key limitations in traditional deep learning approaches. Fourth, clinical evaluation demonstrated that AI assistance improved clinician sensitivity by 8.65% while reducing assessment time by 6.78% [4], supporting AI as an effective decision-support tool.

Technical Considerations

The predominance of CNN-based architectures, particularly *U-Net* and *ResNet* variants [3-4, 29, 33-34], reflects their proven effectiveness in medical image analysis. Detailed architectural specifications including input dimensions, key components, and training strategies for all evaluated systems are summarized in Table 8. *U-Net*'s encoder-decoder structure excels at preserving spatial information crucial for lesion localization [33], while *ResNet*'s skip connections enable training of deeper networks [3, 29]. The Diagnocat system's hierarchical coarse-to-fine processing from 3mm³ to 0.15mm³ voxel resolution [4, 33] successfully balanced computational efficiency with diagnostic precision, an important de sign principle for resource-constrained clinical settings.

Esmaeilyfard *et al.*'s [3] multi-input CNN architecture, integrating axial, sagittal, and coronal views, effectively leveraged CBCT's 3D nature, achieving 93-97% accuracy. This multi-view integration addresses limitations of applying 2D architectures to volumetric data where single-slice analysis may miss critical diagnostic information. The superior performance compared to single-view approaches in earlier studies [21-22] suggests future architectures should explicitly incorporate 3D spatial relationships.

The dramatic performance improvement achieved by Chen *et al.*'s [34] interpretable model- from 70% to greater than 91.7% accuracy- represents the most significant architectural finding. Incorporating pre-segmentation of caries and pulp regions with local interpretable model-agnostic explanations (LIME) visualization [35] improved both interpretability and performance, challenging the assumption that interpretability trades off against accuracy. This demonstrates that forcing models to explicitly identify relevant anatomical structures before classification creates transparent pipelines where each stage can be clinically validated [30-31]. Future architectures should prioritize inherent interpretability aligned with clinical reasoning.

Zanini *et al.*'s [29] SimCLR-based self-supervised learning, achieving 5.5% F1-score improvement, addresses labeled data scarcity by learning from abundant unlabeled CBCT data. This contrastive learning approach is particularly valuable in dentistry where expert annotation is expensive and time-consuming [29]. Comprehensive preprocessing, particularly Hounsfield unit normalization and artifact management [33], proved essential for robust performance across different CBCT devices, emphasizing the need for device-agnostic normalization protocols [15, 33]. Artifact management remains particularly crucial given CBCT's known limitations in dental material visualization, as demonstrated in endodontic studies evaluating filling material integrity [32-33].

Clinical Implications and Challenges

AI systems achieved diagnostic accuracy comparable to specialists [33] and improved clinician sensitivity by 8.65% while maintaining specificity [4], suggesting potential for standardized diagnostic criteria, reduced inter-observer variability, and improved detection of subtle lesions. The 6.78% time reduction [4] could translate to increased patient throughput and reduced radiologist fatigue. These clinical benefits, however, must be contextualized within appropriate clinical indications. AI-assisted CBCT caries detection should be viewed exclusively as an adjunctive tool for opportunistic assessment when CBCT scans are already justified and acquired for other diagnostic purposes, rather than as justification for CBCT acquisition solely for caries diagnosis [14-15]. This distinction is essential to prevent inappropriate radiation exposure and maintain adherence to evidence-based imaging guidelines. Chen *et al.*'s [34] caries-pulp distance classification system provides actionable treatment planning information, with AI-derived measurements showing significant correlation with clinical decisions ($p< 0.0001$).

The clinical significance of CBCT in caries detection must be carefully evaluated within the framework of radiation protection principles. The ALARA and ALADA principles are fundamental to medical imaging and mandate that radiation exposure be justified, optimized, and kept to the minimum necessary to achieve diagnostic objectives [14-15]. CBCT delivers substantially higher effective radiation doses (5-1073 μSv depending on field of view and device settings) compared to conventional intraoral radiography (1-8 μSv for bitewing radiographs), representing a 100-1000 fold difference in patient exposure [14]. Additionally, CBCT imaging involves higher cost, longer acquisition time, and increased technical complexity compared to standard dental radiography. Consequently, professional guidelines from the American Academy of Oral and Maxillofacial Radiology, European Commission, and International Commission on Radiological Protection explicitly state that CBCT should not be used routinely for caries detection when conventional radiography is adequate [14-15]. The appropriate clinical context for AI-assisted CBCT caries detection is therefore limited to opportunistic assessment when CBCT scans have already been justified and acquired for other indications such as implant treatment planning, evaluation of impacted teeth, assessment of pathological lesions, or complex endodontic cases. In these scenarios, AI systems can provide added diagnostic value by systematically analyzing existing volumetric data for caries without requiring additional radiation exposure. This adjunctive role aligns with radiation protection principles while maximizing the diagnostic utility of clinically indicated CBCT examinations.

Despite high performance, several critical barriers limit clinical implementation. Data-related challenges include small, single-center datasets [3, 29, 34] limiting diversity of imaging protocols and patient populations, class imbalance under representing rare caries types [3], and label quality variability from differences in annotation protocols [4, 33]. Technical challenges include significant image quality variability across different CBCT devices and manufacturers [15, 33], metallic restoration artifacts obscuring tooth structures [15], and substantial computational requirements for volumetric data processing [26, 33].

Device-related variability represents a particularly critical challenge for AI model generalizability, as CBCT scanners differ substantially in voxel size (ranging from 0.076 mm to 0.4mm), field of view options, tube voltage and current settings, reconstruction algorithms, and artifact reduction protocols [15, 33]. These technical differences directly affect image quality parameters including spatial resolution, contrast resolution, noise levels, and Hounsfield unit consistency, which in turn influence AI model performance [33]. Ezhov *et al.* [33] specifically addressed this challenge through comprehensive Hounsfield unit normalization protocols including outlier removal and standardization across different CBCT devices in their multi-center validation. Amasya *et al.* [4] similarly implemented device-specific preprocessing and artifact reduction to standardize image quality across multiple CBCT systems in their clinical evaluation. However, most included studies [3, 29, 34] utilized single CBCT devices or did not specify device-agnostic validation strategies, raising concerns about model transferability to different clinical settings with varied imaging equipment. Future AI development must prioritize robust preprocessing pipelines and multi-device validation to ensure reliable performance across the diverse CBCT systems encountered in clinical practice.

The "black box" nature of many deep learning models limits clinical trust [34], though interpretability techniques [29, 31, 35] show promise. Generalizability concerns arise as models trained on single-center data may underperform on external datasets [3-4]. Only two studies used multi-center data [4, 33], and none performed external validation on independent datasets. Practical challenges related to workflow integration and user interface design [4] require attention to ensure AI enhances rather than disrupts clinical practice.

Critical limitations in the current evidence base must be acknowledged when interpreting these findings. All included studies were retrospective in design, limiting the ability to assess prospective real-world performance [3-4, 29, 33-34]. Most studies [3, 29, 34] were conducted in single centers with limited patient diversity and imaging protocol variation. Crucially, none of the included studies performed external validation on independent datasets from different institutions, raising substantial concerns about model generalizability to diverse clinical settings with different patient populations, CBCT devices, and imaging protocols. The reported performance metrics therefore reflect controlled research conditions rather than validated clinical performance. Additionally, sample sizes were generally modest (204-1346 scans), and validation strategies varied considerably across studies, precluding meta-analysis and limiting confidence in pooled estimates. These methodological limitations significantly constrain the strength of evidence and prevent definitive conclusions regarding clinical readiness or equivalence to expert human performance.

Contextualization and Future Directions

The limited evidence base reveals important constraints on the strength and applicability of current findings. The small number of included studies (n=5) reflects the nascent state of AI research in CBCT-based caries detection rather than inadequate search methodology. Critical methodological limitations include retrospective study designs in all included studies precluding assessment of prospective real-world performance [3-4, 29, 33-34], predominantly single-center investigations [3, 29, 34] limiting patient and protocol diversity, complete absence of external validation on independent datasets raising substantial generalizability concerns, modest sample sizes of 200-800 specimens in most studies [3, 29, 34], focus on common caries presentations with underrepresentation of atypical or early lesions [3], and heterogeneity in AI architectures, validation strategies, and outcome definitions preventing meta-analysis. These limitations restrict this review to qualitative synthesis and prevent definitive conclusions regarding clinical equivalence to expert performance or readiness for routine clinical implementation. The evidence demonstrates proof-of-concept and technical feasibility but does not yet support claims of clinical validation or deployment readiness.

Future research should prioritize large-scale multi-center studies exceeding 5,000 cases with standardized annotation protocols, advanced AI architectures including 3D attention mechanisms and vision transformers, prospective randomized controlled trials comparing AI-assisted versus conventional diagnosis, health economic analyses of cost-effectiveness [4, 34], explainable AI development with advanced interpretability techniques [30-31], multimodal integration of CBCT with intraoral scans and clinical data, and standardization efforts including public benchmark datasets and evaluation protocols.

Before clinical implementation can be responsibly recommended, AI systems for CBCT caries detection must demonstrate robust performance through (1) prospective validation studies comparing AI-assisted diagnosis with unassisted expert diagnosis in real clinical workflows; (2) external validation on independent multi-center datasets to establish generalizability across diverse patient populations and imaging protocols; (3) clinical impact studies demonstrating improved patient outcomes, not merely diagnostic metrics; and (4) regulatory approval through appropriate pathways ensuring safety and effectiveness standards. Current evidence establishes technical feasibility and promising preliminary performance but does not constitute sufficient validation for clinical deployment.

AI-based caries detection systems have progressed to feasible clinical implementation levels. However, these systems should augment rather than replace human expertise [4]. With appropriate multi-center validation, regulatory oversight, and thoughtful implementation respecting clinical judgment, AI systems have genuine potential to enhance diagnostic accuracy, improve workflow efficiency, and contribute to better oral health outcomes globally. The evidence provides a solid foundation for responsible development and clinical translation of AI-based caries detection systems.

Conclusion

This systematic review indicates that AI-based systems, particularly those using CNN architectures, demonstrate promising preliminary diagnostic performance for detecting dental caries in CBCT images under controlled research conditions. Reported sensitivity and specificity values are high, and limited comparative evidence suggests that AI performance may approach that of experienced radiologists in specific scenarios. In addition, AI assistance appears to enhance clinician sensitivity and reduce assessment time, supporting its potential as a decision-support tool. Recent technical advancements, including interpretable models and self-supervised learning approaches, further contribute to improved accuracy and transparency, which are essential for clinical trust and adoption.

Despite these encouraging findings, significant methodological limitations restrict immediate clinical implementation. The existing evidence is based entirely on retrospective studies, predominantly conducted in single-center settings, with no external validation on independent datasets. These limitations, along with challenges such as limited training data, variability in CBCT image quality, and the lack of prospective clinical impact studies, constrain generalizability and real-world applicability. Current evidence supports proof-of-concept and technical feasibility rather than clinical validation. Robust prospective, multi-center studies with external validation, standardized evaluation protocols, and demonstrated clinical impact are essential before these systems can be safely and effectively integrated into routine clinical practice.

Acknowledgments

The authors declare that no financial or personal assistance was received for this study.

Conflict of Interest

The authors declare no competing interests.

References

[1] Paknahad M, Dokohaki S, Khojastepour L, Shahidi S, Haghnegahdar A. A radio-odontometric analysis of sexual dimorphism in first molars using cone-beam computed tomography. Am J Forensic Med Pathol. 2022; 43: 46-51.

[2] Selwitz RH, Ismail AI, Pitts NB. Dental caries. Lancet. 2007; 369(9555): 51-59.

[3] Esmaeilyfard R, Bonyadifard H, Paknahad M. Dental caries detection and classification in CBCT images using deep learning. Int Dent J. 2024; 74: 328-334.

[4] Amasya H, Alkhader M, Serindere G, Futyma-Gąbka K, Aktuna Belgin C, Gusarev M, et al. Evaluation of a decision support system developed with deep learning approach for detecting dental caries with cone-beam computed tomography imaging. Diagnostics (Basel). 2023; 13: 3471.

[5] Park YS, Ahn JS, Kwon HB, Lee SP. Current status of dental caries diagnosis using cone beam computed tomography. Imaging Sci Dent. 2011; 41: 43-51.

[6] Young S, Lee J, Hodges R, Chang TL, Elashoff D, White S. A comparative study of high-resolution cone beam computed tomography and charge-coupled device sensors for detecting caries. Dentomaxillofac Radiol. 2009; 38: 445-451.

[7] Scarfe WC, Farman AG. What is cone-beam CT and how does it work? Dent Clin North Am. 2008; 52: 707-730.

[8] Mozzo P, Procacci C, Tacconi A, Martini PT, Andreis IA. A new volumetric CT machine for dental imaging based on the cone-beam technique: preliminary results. Eur Radiol. 1998; 8: 1558-1564.

[9] Davvaz B, Hosseini S, Ghahramani Y, Haghnegahdar A. Exploring taurodontism using cone beam computed tomography: a retrospective study of its morphology and association with C-shaped canal. J Dent (Shiraz). 2025; 26: 371-378.

[10] Movahhedian N, Ghahremani L, Akbarizadeh F, Khojastepour L. The impact of age on the accuracy of sexual dimorphism using anthropometric indices of maxillary sinus. J Dent (Shiraz). 2025. doi: 10.30476/dentjods.2025.104945.2561.

[11] Miracle AC, Mukherji SK. Conebeam CT of the head and neck, part 1: physical principles. AJNR Am J Neuroradiol. 2009; 30: 1088-1095.

[12] Ramis-Alario A, Llena MC, Castillo-Oliva JL, Paredes-Gallardo V, Ibañez-Vera AJ. Comparison of diagnostic accuracy between periapical and panoramic radiographs and cone beam computed tomography in measuring the periapical area of teeth scheduled for periapical surgery: a cross-sectional study. J Clin Exp Dent. 2019; 11: e732-e738.

[13] Liang YH, Li G, Wesselink PR, Wu MK. Detection and measurement of artificial periapical lesions by cone-beam computed tomography. Int Endod J. 2014; 47: 332-338.

[14] Price JB. Caries detection with dental cone beam computed tomography. In: Zandona AF, Longbottom C, edi-

tors. Detection and assessment of dental caries: a clinical guide. 1st ed. Springer, Cham: Switzerland; 2019. p. 127-138.

[15] Isman O, Aktan AM, Ertas ET. Evaluating the effects of orthodontic materials, field of view, and artifact reducution mode on accuracy of CBCT-based caries detection. Clin Oral Investig. 2020; 24: 2487-2496.

[16] Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O. Deep learning with convolutional neural network in radiology. Jpn J Radiol. 2018; 36: 257-272.

[17] Gao H, Xiao J, Yin Y, Liu T, Shi J. A mutually supervised graph attention network for few-shot segmentation: The perspective of fully utilizing limited samples. Available at: https://ieeexplore.ieee.org/abstract/document/ 9733-415

[18] Bayrakdar IS, Orhan K, Akarsu S, Celik O, Atasoy S, Pekince A, et al. A U-Net approach to apical lesion segmentation on panoramic radiographs. BioMed Res Int. 2022; 2022: 7035367.

[19] Jang TJ, Kim KC, Cho HC, Seo JK. A fully automated method for 3D individual tooth identification and segmentation in dental CBCT. IEEE Trans Pattern Anal Mach Intell. 2022; 44: 6562-6568.

[20] Cipriano M, Allegretti S, Bolelli F, Bartolomeo MD, Pollastri F, Pellacani A, et al. Deep segmentation of the mandibular canal: a new 3D annotated dataset of CBCT volumes. IEEE Access. 2022; 10: 11500-11510.

[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: 106-111.

[22] Ying S, Wang B, Zhu H, Liu W, Huang F. Caries segmentation on tooth X-ray images with a deep network. J Dent. 2022; 119: 104076.

[23] Lin X, Hong D, Zhang D, Huang M, Yu H. Detecting proximal caries on periapical radiographs using convolutional neural networks with different training strategies on small datasets. Diagnostics (Basel). 2022; 12: 1047.

[24] Khanagar SB, Alfouzan K, Awawdeh M, Alkadi L, Albalawi F, Alfadley A. Application and performance of artificial intelligence technology in detection, diagnosis and prediction of dental caries (DC): a systematic review. Diagnostics (Basel). 2022; 12: 1083.

[25] Hung KF, Ai QYH, Wong LM, Yeung AWK, Li DTS, Leung YY. Current applications of deep learning and radiomics on CT and CBCT for maxillofacial diseases. Diagnostics (Basel). 2023; 13: 110.

[26] Fan W, Zhang J, Wang N, Li J, Hu L. The application of deep learning on CBCT in dentistry. Diagnostics (Basel). 2023; 13: 2056.

[27] Azizi S, Mustafa B, Ryan F, Beaver Z, Freyberg J, Deaton J, et al. Big self-supervised models advance medical image classification. Proc IEEE Int Conf Comput Vis. 2021: 3478-3488.

[28] Islam MM, Hossain MB, Akhtar MN, Moni MA, Hasan KF. CNN based on transfer learning models using data augmentation and transformation for detection of concrete crack. Algorithms. 2022; 15: 287.

[29] Zanini LGK, Rubira-Bullen IRF, Nunes FLS. Enhancing dental caries classification in CBCT images by using image processing and self-supervised learning. Comput Biol Med. 2024; 183: 109221.

[30] Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. Available at: https://openaccess.thecvf.com/content_cvpr_ 2016/papers/Zhou_Learning_Deep_Features_CVPR_2016 _ paper.pdf

[31] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. Available at: https://arxiv.org/abs/1610.02391

[32] Eskandarinezhad M, Razi S, Pirzadeh T, Bagheri Sabzevar S, Nazari L, Akhondian S. Investigation of voids in the apical plug of MTA using cone-beam computed tomography, digital radiography, and analog radiography. J Dent (Shiraz). 2026; 27: 27-32.

[33] Ezhov M, Gusarev M, Golitsyna M, Yates JM, Kushnerev E, Tamimi D, et al. Clinically applicable AI system for dental diagnosis with CBCT. Sci Rep. 2021; 11: 15006.

[34] Chen S, Yang Y, Wu W, Wei R, Wang Z, Tay FR. Classification of caries based on CBCT: a deep learning network interpretability study. J Imaging Inform Med. 2024; 37: 3160-3173.

[35] Tulio Ribeiro M, Singh S, Guestrin C. Why should I trust you? Explaining the predictions of any classifier. Available at: https://dl.acm.org/doi/10.1145/2939672.2939778