JNT FACIT BUSINESS AND TECHNOLOGY JOURNAL
ISSN: 2526-4281 - ANO 2026 - MÊS DE OUTUBRO
Ed. 78. Vol. 1 - Págs. 537-550
DOI: 10.5281/zenodo.23987777



CAPES

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

QUALIS A2



INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO DE CÁRIE EM ODONTOPEDIATRIA[1]

ARTIFICIAL INTELLIGENCE IN CARIES DIAGNOSIS IN PEDIATRIC DENTISTRY

Sthefanny Bastos do NASCIMENTO
Centro Universitário Uninovafapi Afya
E-mail: shtefannybastos28@icloud.com
ORCID: http://orcid.org/0009-0000-4262-4792

Tereza Maria Alcântara NEVES
Centro Universitário Uninovafapi Afya
E-mail: tereza.neves@afya.com.br
ORCID: http://orcid.org/0000-0002-5177-7739

Márcia Regina Soares CRUZ
Centro Universitário Uninovafapi Afya
E-mail: marcia.cruz@afya.com.br
ORCID: http://orcid.org/0000-0001-5633-0090

Thiago Lima MONTE
Centro Universitário Uninovafapi Afya
E-mail: thiago.monte@afya.com.br
ORCID: http://orcid.org/0000-0002-0261-3009

Isabel Cristina Quaresma RÊGO
Centro Universitário Uninovafapi Afya
E-mail: isabel.rego@afya.com.br
ORCID: http://orcid.org/0000-0003-0870-665X

Tainá de Castelo Branco ARAÚJO
Centro Universitário Uninovafapi Afya
E-mail: taina.araujo@afya.com.br
ORCID: http://orcid.org/0000-0003-0056-5709

RESUMO

Objetivo: analisar criticamente aplicações da inteligência artificial (IA) na detecção e na classificação de cárie dentária em crianças. Método: revisão integrativa exploratória, com síntese narrativa de nove estudos originais publicados entre 2024 e 2026, localizados por buscas na PubMed/MEDLINE em outubro de 2026 e por busca complementar. Foram considerados trabalhos com pacientes pediátricos ou dentes decíduos que empregaram aprendizagem de máquina em radiografias ou

[1] COMO CITAR: (ABNT): NASCIMENTO, S. B.; NEVES, T. M. A.; CRUZ, M. R. S.; MONTE, T. L.; RÊGO, I. C. Q.; ARAÚJO, T. C. B. Inteligência Artificial no Diagnóstico de Cárie em Odontopediatria. **JNT Facit Business and Technology Journal.** Qualis A2. ISSN: 2526-4281, Mês de Outubro de 2026 - Ed. 78. VOL. 01. Págs. 537-550. Disponível: http://revistas.faculdadefacit.edu.br. Acesso em: ___/___/___.

escaneamentos intraorais. Resultados: os estudos utilizaram radiografias interproximais, periapicais e panorâmicas e escaneamentos intraorais, e o desempenho variou conforme o estágio da lesão, a modalidade de imagem e o tipo de validação. Um estudo com radiografias interproximais relatou sensibilidade de 88,8% e especificidade de 98,8%; outro, com escaneamentos intraorais, relatou sensibilidade de 67% na detecção de lesões. Em um dos estudos, a sensibilidade foi menor para lesões iniciais e moderadas do que para lesões extensas. Conclusão: a IA pode auxiliar a interpretação diagnóstica, mas não substitui o exame clínico. Como as buscas se restringiram a uma base de dados e não houve avaliação formal do risco de viés, os achados não esgotam a literatura disponível.

Palavras-chave: Inteligência Artificial. Cárie Dentária. Odontopediatria. Diagnóstico por Imagem. Aprendizado Profundo.

ABSTRACT

Objective: to critically examine artificial intelligence (AI) applications for detecting and classifying dental caries in children. Methods: exploratory integrative review with narrative synthesis of nine original studies published from 2024 to 2026, identified through PubMed/MEDLINE searches in October 2026 and a complementary search. Studies of pediatric patients or primary teeth that applied machine learning to radiographs or intraoral scans were considered. Results: the studies used bitewing, periapical and panoramic radiographs and intraoral scans, and performance varied with lesion stage, imaging modality and type of validation. One study using bitewing radiographs reported 88.8% sensitivity and 98.8% specificity; another, using intraoral scans, reported 67% sensitivity for lesion detection. In one study, sensitivity was lower for initial and moderate lesions than for extensive ones. Conclusion: AI may support diagnostic interpretation but does not replace clinical examination. Because the searches were limited to one database and no formal risk-of-bias assessment was performed, the findings do not cover all the available literature.

Keywords: Artificial Intelligence. Dental Caries. Pediatric Dentistry. Diagnostic Imaging. Deep Learning.

INTRODUÇÃO

A cárie dentária é uma doença multifatorial e dinâmica, resultante da interação

entre biofilme, dieta, fatores do hospedeiro e condições ambientais (Selwitz; Ismail; Pitts, 2007). A carga global das condições bucais, entre as quais a cárie, foi estimada em um estudo de abrangência mundial (Kassebaum *et al*., 2017), e uma revisão sistemática relaciona essas condições à qualidade de vida relacionada à saúde (Haag *et al*., 2017). Na infância, a cárie pode causar dor, alterações alimentares, dificuldade para dormir e necessidade de tratamentos restauradores complexos, e a identificação precoce das lesões, somada à avaliação de atividade e de risco, orienta a escolha entre medidas preventivas, não invasivas ou restauradoras. Sistemas como o ICDAS e o ICCMS organizam essa decisão ao classificar os estágios do processo de cárie e orientar o seu manejo (Pitts; Ekstrand, 2013).

Em Odontopediatria, o diagnóstico reúne anamnese, exame clínico visual e tátil criterioso, avaliação dos fatores de risco e, quando indicados, exames radiográficos complementares. A anatomia dos dentes decíduos, a proximidade da polpa e a dificuldade de cooperação da criança interferem na interpretação clínica e radiográfica, embora se tenha argumentado que a detecção de cárie em dentes decíduos é menos desafiadora do que em dentes permanentes (Mendes; Braga, 2013).

O exame radiográfico, por sua vez, tem sido questionado em pré-escolares pelo baixo impacto terapêutico, pelos falsos positivos e pelo sobrediagnóstico (Pontes *et al*., 2021), de modo que uma nova tecnologia precisa demonstrar que acrescenta informação útil sem ampliar intervenções desnecessárias. Nos últimos anos, modelos de aprendizagem de máquina e de aprendizagem profunda passaram a ser aplicados à identificação de padrões em radiografias e em imagens tridimensionais: redes neurais convolucionais, modelos de segmentação e detectores de objetos localizam regiões suspeitas e classificam a extensão das alterações.

Revisões sistemáticas e metanálises já sintetizaram o desempenho desses modelos na detecção de cárie em geral (Mohammad-Rahimi *et al*., 2022; Abbott; Saikia; Anthonappa, 2025) e em radiografias interproximais (Ammar; Kühnisch, 2024), e um ensaio clínico randomizado avaliou o apoio da IA à detecção de cárie por dentistas (Mertens *et al*., 2021). A população pediátrica, porém, tem particularidades anatômicas e de cooperação que justificam uma síntese própria. O bom desempenho em conjuntos de treinamento pode não se repetir em outros pacientes, equipamentos e serviços, e já existe uma lista de verificação para autores, revisores e leitores de pesquisas sobre IA em Odontologia (Schwendicke *et al*., 2021).

A literatura pediátrica recente reúne estudos de desempenho diagnóstico, desenvolvimento de algoritmos e comparações com examinadores humanos, e uma síntese crítica ajuda a separar o que foi demonstrado em ambiente experimental do

que já pode ser aplicado na clínica. Este artigo tem por objetivo discutir métodos, resultados, limitações e perspectivas da inteligência artificial no diagnóstico de cárie em crianças, com ênfase em dentes decíduos e dentição mista.

METODOLOGIA

Trata-se de uma revisão integrativa de caráter exploratório, com abordagem qualitativa e síntese narrativa, orientada pela seguinte pergunta: quais evidências científicas recentes descrevem o desempenho diagnóstico, as possibilidades clínicas e as limitações de sistemas de inteligência artificial para identificar ou classificar lesões de cárie em crianças, dentes decíduos e dentição mista?

O levantamento foi realizado em 28 de outubro de 2026 na PubMed/MEDLINE, de acesso público, para o período de janeiro de 2021 a outubro de 2026. A primeira consulta (*dental caries AND children AND artificial intelligence*) recuperou 110 registros, e uma segunda (*dental caries AND children AND deep learning*) recuperou 43, sem acrescentar estudos elegíveis além dos já identificados na primeira. Foram incluídos estudos originais publicados entre 2024 e 2026, com resumo indexado e identificação bibliográfica verificável, que avaliaram modelos de IA na detecção, na segmentação ou na classificação de cárie dentária em crianças, dentes decíduos ou dentição mista, a partir de radiografias ou de escaneamentos intraorais.

Foram excluídos artigos de revisão, comentários e protocolos, estudos sobre IA sem aplicação direta ao diagnóstico de cárie por imagem (como previsão de risco, agentes conversacionais e análise de microbioma), estudos baseados em fotografias intraorais e estudos anteriores a 2024. Dos 110 registros da primeira consulta, 79 foram excluídos na leitura de títulos e resumos (24 revisões, comentários ou protocolos e 55 sem aplicação direta ao diagnóstico de cárie por imagem), e 31 foram avaliados quanto à elegibilidade.

Desses, 24 foram excluídos (9 baseados em fotografias intraorais, 3 publicados antes de 2024 e 12 por não serem específicos de crianças, não usarem radiografias ou escaneamentos ou terem outro objetivo, como a avaliação de modelos de linguagem), restando sete estudos. Dois estudos adicionais, que não foram recuperados pelas consultas, foram incluídos após busca complementar e conferência nos registros dos periódicos, o que totalizou nove estudos: dois de 2024, quatro de 2025 e três de 2026. Os dados bibliográficos de cada estudo (autoria, periódico e identificadores) foram conferidos nas páginas dos periódicos e nos registros indexados em outubro de 2026.

De cada estudo foram extraídos autoria, ano, tipo de imagem, população ou número de exames, objetivo do algoritmo, métricas diagnósticas e limitações

descritas; os dados foram organizados em quadro comparativo e discutidos em três eixos: desempenho em radiografias interproximais, periapicais e panorâmicas, detecção por escaneamento intraoral, e validade clínica, risco de viés e implementação. Em razão da heterogeneidade das tarefas, não se realizou metanálise nem cálculo de desempenho agrupado. A pesquisa utilizou apenas informações científicas publicadas, sem coleta de dados com seres humanos. Como as buscas se restringiram à PubMed/MEDLINE, estudos indexados em outras bases podem não ter sido recuperados, e não houve avaliação formal do risco de viés.

RESULTADOS

Os nove estudos incluídos representam diferentes abordagens de inteligência artificial para o diagnóstico de cárie na infância, e suas principais características estão no Quadro 1, o Quadro 2 resume o desempenho e as limitações por modalidade de imagem, e o Quadro 3 reúne as métricas diagnósticas relatadas. As unidades de análise variaram entre imagens radiográficas, superfícies dentárias, lesões marcadas e dentes digitalizados, o que impede comparar porcentagens sem considerar denominadores e protocolos.

Nas radiografias interproximais, Gonzalez *et al.* (2024) desenvolveram um modelo sem programação convencional com 66 radiografias escolhidas de um conjunto inicial de 100; um painel de consenso calibrado anotou 755 superfícies proximais, 178 delas com lesão, e o modelo alcançou sensibilidade de 88,8%, especificidade de 98,8%, precisão de 95,8%, acurácia de 96,4%, F1 de 92% e área sob a curva ROC de 0,965, com divisão dos dados em treinamento (70%), desenvolvimento (20%) e teste (10%) e ampliação artificial das imagens, mas com amostra pequena.

Nampim *et al*. (2026) analisaram 1.400 radiografias *bitewing* (4.715 imagens dentárias) e classificaram as lesões em quatro ou sete categorias; a ResNet-152 superou a ResNet-50 e a ResNet-101 na validação e obteve AUC de 0,85, a classificação em mais categorias foi mais difícil, e o modelo e a maioria dos examinadores tiveram menor sensibilidade para lesões iniciais e moderadas do que para lesões extensas.

Na comparação com três cirurgiões-dentistas gerais e três estudantes de Odontologia, a ResNet-152 e um dos estudantes tiveram acurácia superior a 0,7 na classificação em quatro categorias, enquanto a maioria dos examinadores ficou entre 0,62 e 0,67, e apenas esses dois alcançaram especificidade e valor preditivo negativo de 0,9 ou mais. Tan *et al*. (2025) usaram 1.023 radiografias interproximais de molares

decíduos de crianças de 3 a 10 anos, anotadas por dois examinadores em quatro estágios de gravidade, e compararam cinco detectores de objetos em três limiares (todas as lesões, lesões em dentina e lesões em dentina com necessidade de tratamento operatório).

O modelo DINO teve a maior concordância na classificação por estágios (kappa ponderado de 0,713), a maior sensibilidade para todas as lesões (0,509) e, para lesões em dentina com necessidade de tratamento operatório, sensibilidade de 0,659 e acurácia de 0,980, e o YOLOv7 teve especificidade superior a 0,99. Güçlü *et al*. (2025) analisaram 1.000 radiografias *bitewing* de dentes decíduos e 1.000 de dentes permanentes, com 5.704 lesões interproximais anotadas, e treinaram um modelo YOLOv5x de segmentação, com divisão dos dados em 80% para treinamento, 10% para validação e 10% para teste.

Nos dentes decíduos, a precisão foi de 0,80, a sensibilidade de 0,78, o F1 de 0,79 e a AUC de 0,87, e nos permanentes os valores foram de 0,78, 0,75, 0,76 e 0,76. Em radiografias periapicais, Yavsan *et al*. (2025) reuniram imagens de crianças de 5 a 12 anos, com 830 lesões aproximais marcadas em 415 imagens obtidas após ampliação artificial, e compararam 13 variantes de detectores YOLO; o YOLOv5s teve o melhor desempenho, com precisão de 91,2%, sensibilidade de 89,3%, F1 de 90,24% e mAP de 90,8% no conjunto de teste (20% dos dados), mas, como a divisão entre treinamento e teste ocorreu depois da ampliação artificial das imagens, o desempenho pode estar superestimado.

Nas radiografias panorâmicas, Asci *et al*. (2024) reuniram 6.075 exames de crianças de 4 a 14 anos, distribuídos em dentição decídua (n = 1.857), mista (n = 1.406) e permanente (n = 2.812), e usaram um modelo U-Net para segmentar as lesões, com sensibilidade de 0,8525, precisão de 0,9128 e F1 de 0,8816 na dentição decídua, de 0,7377, 0,9192 e 0,8185 na dentição mista e de 0,8269, 0,9123 e 0,8675 no conjunto total. Biçengil *et al*. (2026) avaliaram 1.526 radiografias panorâmicas de crianças de 5 a 12 anos com o algoritmo YOLOv8x para lesões proximais, oclusais e vestibulares em molares decíduos, com F1 de 0,576 para as proximais e de 0,24 para as oclusais e sem detecção adequada das vestibulares, e, no resultado global, o modelo não diferiu do observador mais experiente e superou os menos experientes, mas os autores concluem que o desempenho ficou próximo ao de clínicos menos experientes e não alcançou o nível de especialistas.

Rahimi *et al.* (2026) apresentaram o conjunto MD-OPG, composto por radiografias panorâmicas de crianças de 3 a 12 anos com lesões proximais e oclusais anotadas por especialistas; um classificador baseado em recortes da imagem obteve

AUC média de 0,89 e sensibilidade de 0,85 para distinguir regiões saudáveis e cariadas, e, na segmentação, o Attention U-Net treinado com perda focal alcançou o melhor coeficiente de Dice (0,94). No escaneamento intraoral, Jones *et al*. (2025) investigaram a detecção automatizada de cárie em crianças com escaneamentos obtidos em duas coortes de um instituto de pesquisa pediátrico, anotados por duas pesquisadoras pelo ICCMS e convertidos em representações bidimensionais para o treinamento de um modelo Attention U-Net; os 332 dentes cariados da primeira coorte foram divididos por participante em treinamento (n = 192), validação (n = 63) e teste (n = 77), e uma coorte independente de 119 dentes serviu para a validação externa, com sensibilidade de 67% e precisão de 73% na detecção de lesões.

Quadro 1: Características dos estudos incluídos.

Estudo	Material analisado	Resultado ou aplicação
Gonzalez *et al*. (2024)	66 radiografias *bitewing*; 755 superfícies proximais	Modelo sem programação (*no-code*); sensibilidade de 88,8%, especificidade de 98,8% e acurácia de 96,4%
Asci *et al*. (2024)	6.075 radiografias panorâmicas de crianças de 4 a 14 anos	Segmentação com U-Net em dentições decídua, mista e permanente; F1 de 0,8816 na dentição decídua e de 0,8675 no total
Yavsan *et al*. (2025)	415 radiografias periapicais de crianças de 5 a 12 anos (após ampliação); 830 lesões marcadas	YOLOv5s entre 13 variantes YOLO; precisão de 91,2%, sensibilidade de 89,3% e F1 de 90,24%
Jones *et al*. (2025)	332 dentes cariados no desenvolvimento; 119 na validação externa	Escaneamento intraoral; sensibilidade de 67% e precisão de 73% na detecção de lesões
Tan *et al*. (2025)	1.023 radiografias *bitewing* de crianças de 3 a 10 anos	Comparação de cinco detectores de objetos; DINO com kappa ponderado de 0,713 na classificação por estágios
Güçlü *et al*. (2025)	1.000 radiografias *bitewing* de dentes decíduos e 1.000 de permanentes; 5.704 lesões	YOLOv5x de segmentação; nos decíduos, precisão de 0,80, sensibilidade de 0,78, F1 de 0,79 e AUC de 0,87
Nampim *et al*. (2026)	1.400 radiografias *bitewing*; 4.715 imagens dentárias	ResNet-152; AUC de 0,85; desempenho comparável ao dos melhores examinadores humanos
Biçengil *et al*. (2026)	1.526 radiografias panorâmicas pediátricas	YOLOv8x; F1 de 0,576 para cárie proximal e de 0,24 para oclusal; sem alcançar o nível de especialistas
Rahimi *et al*. (2026)	Conjunto MD-OPG: radiografias panorâmicas de crianças de 3 a 12 anos	Classificador por recortes com AUC de 0,89 e sensibilidade de 0,85; Attention U-Net com Dice de 0,94

Fonte: Elaborado pelos autores com base nos resumos indexados na PubMed/MEDLINE e nas páginas dos periódicos (2026).
Nota: AUC = área sob a curva ROC (característica de operação do receptor); F1 = média harmônica entre precisão e sensibilidade; DINO, ResNet-152, U-Net, YOLOv5s, YOLOv5x, YOLOv7 e YOLOv8x = arquiteturas de redes neurais. Nampim et al. (2026) foi disponibilizado on-line em 2025.

Quadro 2: Síntese do desempenho e das limitações por modalidade de imagem.

Modalidade de imagem	Estudos	Desempenho relatado	Principal limitação
Radiografia interproximal (*bitewing*)	Gonzalez *et al*. (2024); Tan *et al*. (2025); Güçlü *et al.* (2025); Nampim *et al*. (2026)	Sensibilidade de 88,8% e especificidade de 98,8% (Gonzalez *et al*.); Kappa ponderado de 0,713 (Tan *et al*.); F1 de 0,79 nos decíduos (Güçlü et al.); AUC de 0,85 (Nampim et al.)	Amostra pequena (Gonzalez *et al*.); sensibilidade de 0,509 para todas as lesões (Tan *et al*.); avaliação em um único conjunto de imagens (Güçlü *et al*.); menor sensibilidade para lesões iniciais e moderadas (Nampim *et al*.)
Radiografia periapical	Yavsan *et al*. (2025)	Precisão de 91,2%, sensibilidade de 89,3% e F1 de 90,24%	Conjunto pequeno, ampliado artificialmente antes da divisão entre treinamento e teste
Radiografia panorâmica	Asci *et al*. (2024); Biçengil *et al.* (2026); Rahimi *et al.* (2026)	F1 de 0,8675 no conjunto total (Asci *et al*.); F1 de 0,576 para lesões proximais e de 0,24 para oclusais (Biçengil *et al*.); Dice de 0,94 na segmentação (Rahimi *et al*.)	Baixa visibilidade de pequenas lesões; desempenho dependente da superfície; métricas de segmentação e de detecção não comparáveis
Escaneamento intraoral	Jones *et al*. (2025)	Sensibilidade de 67% e precisão de 73% na detecção de lesões	Sensibilidade moderada; necessidade de validação em outras populações

Fonte: Elaborado pelos autores com base nos estudos incluídos (2026).
Nota: AUC = área sob a curva ROC (característica de operação do receptor); F1 = média harmônica entre precisão e sensibilidade; Dice = sobreposição entre a segmentação do modelo e a anotada; kappa ponderado = concordância corrigida pelo acaso.

Quadro 3: Métricas diagnósticas relatadas pelos estudos incluídos.

Estudo	Tarefa	Métricas relatadas
Gonzalez *et al*. (2024)	Detecção de lesões proximais	Sensibilidade de 88,8%; especificidade de 98,8%; precisão de 95,8%; acurácia de 96,4%; F1 de 92%; AUC de 0,965
Asci *et al*. (2024)	Segmentação de lesões	Dentição decídua: sensibilidade de 0,8525, precisão de 0,9128 e F1 de 0,8816. Dentição mista: 0,7377, 0,9192 e 0,8185. Conjunto total: 0,8269, 0,9123 e 0,8675
Yavsan *et al*. (2025)	Detecção de lesões aproximais	Precisão de 91,2%; sensibilidade de 89,3%; F1 de 90,24%; mAP de 90,8%
Jones *et al*. (2025)	Detecção de lesões em escaneamentos	Sensibilidade de 67% e precisão de 73%
Tan *et* al. (2025)	Detecção e estadiamento de lesões	DINO: kappa ponderado de 0,713; sensibilidade de 0,509 (todas as lesões); sensibilidade de 0,659 e acurácia de 0,980 (lesões em dentina com necessidade de tratamento operatório). YOLOv7: especificidade superior a 0,99 e acurácia superior a 0,92
Güçlü *et al*. (2025)	Segmentação de lesões interproximais	Decíduos: precisão de 0,80, sensibilidade de 0,78, F1 de 0,79 e AUC de 0,87. Permanentes: 0,78, 0,75, 0,76 e 0,76

Estudo	Tarefa	Métricas relatadas
Nampim *et al*. (2026)	Classificação em quatro ou sete categorias	AUC de 0,85; acurácia da ResNet-152 superior a 0,7 (quatro categorias); acurácia de 0,37 a 0,58 (sete categorias, modelo e examinadores)
Biçengil *et al*. (2026)	Detecção por superfície	F1 de 0,576 (proximais) e de 0,24 (oclusais)
Rahimi *et al*. (2026)	Classificação de recortes e segmentação	Classificação: AUC de 0,89 e sensibilidade de 0,85. Segmentação: Dice de 0,94 (Attention U-Net com perda focal)

Fonte: Elaborado pelos autores com base nos estudos incluídos (2026).
Nota: AUC = área sob a curva ROC (característica de operação do receptor); F1 = média harmônica entre precisão e sensibilidade; mAP = média das precisões médias (mean average precision); Dice = sobreposição entre a segmentação do modelo e a anotada; kappa ponderado = concordância corrigida pelo acaso; DINO, ResNet-152, U-Net e YOLOv7 = arquiteturas de redes neurais.

DISCUSSÃO

Os estudos analisados indicam que sistemas de IA podem apoiar a leitura de imagens, sinalizando áreas suspeitas e oferecendo uma segunda avaliação, o que é útil em serviços com grande volume de exames ou com examinadores de experiências diferentes. A utilidade clínica, porém, precisa ser demonstrada em condições reais de atendimento, e não apenas em bases retrospectivas: ensaios clínicos randomizados, como o de Mertens *et al*. (2021), mostram que esse tipo de delineamento é possível, mas os estudos incluídos nesta revisão não o empregaram.

Os números ajudam a dimensionar o que ainda falta, pois a sensibilidade de 88,8% de Gonzalez *et al*. (2024) significa que cerca de 11% das superfícies com lesão não foram detectadas, a sensibilidade de 0,659 do modelo DINO para lesões em dentina com necessidade de tratamento operatório (Tan *et al*., 2025) deixa de sinalizar cerca de um terço dessas lesões, e a sensibilidade de 67% da detecção em escaneamentos intraorais (Jones *et al*., 2025) tem consequência semelhante, de modo que mesmo os melhores resultados exigem a confirmação do examinador.

A alta especificidade observada em algumas amostras não exclui falsos positivos em outros contextos, pois a prevalência de lesões, o equilíbrio entre as classes, a qualidade das imagens e os critérios de anotação alteram o valor preditivo e o comportamento do algoritmo. Esse ponto é sensível na população infantil, em que a radiografia já foi associada a falsos positivos e a sobrediagnóstico (Pontes *et al*., 2021), e um algoritmo que ampliasse esses erros agravaria o problema, de modo que os limiares devem considerar as consequências de um erro diagnóstico para a criança, em especial quando a classificação orienta intervenções irreversíveis.

Os limiares usados por Tan *et al*. (2025), que separam todas as lesões, lesões em dentina e lesões com necessidade de tratamento operatório, aproximam a

avaliação da decisão clínica e mostram que a mesma ferramenta pode ter desempenhos muito diferentes conforme a pergunta que se faz a ela, como indicam as sensibilidades de 0,509 e de 0,659 do modelo DINO nos dois primeiros limiares. A identificação de lesões iniciais é uma dificuldade recorrente, pois alterações discretas do esmalte são pouco visíveis nas imagens, e Nampim *et al*. (2026) relataram menor sensibilidade para lesões iniciais e moderadas, o que importa para a prática odontopediátrica, já que o diagnóstico oportuno favorece a prevenção e a mínima intervenção; um sistema com bom desempenho para lesões avançadas pode servir à triagem, mas não está comprovado para o diagnóstico precoce.

Entre as modalidades de imagem, as radiografias interproximais concentraram os resultados mais consistentes, o que é coerente com o fato de serem o exame indicado para as superfícies proximais. As radiografias periapicais também tiveram desempenho elevado em Yavsan *et al.* (2025), com sensibilidade de 89,3%, mas em um conjunto pequeno, ampliado artificialmente ntes da divisão entre treinamento e teste, o que exige cautela na interpretação.

A separação por dentição mostra que o desempenho não é uniforme: Güçlü *et al*. (2025) obtiveram AUC de 0,87 nos dentes decíduos e de 0,76 nos permanentes, e Asci et al. (2024) relataram sensibilidade de 0,8525 na dentição decídua e de 0,7377 na mista, o que reforça a necessidade de relatar as métricas por tipo de dentição.

A radiografia panorâmica, mais acessível e presente na rotina de muitos serviços, mostrou desempenho dependente da superfície, com F1 de 0,24 para lesões oclusais (Biçengil *et al*., 2026), o que limita seu uso como exame isolado para a detecção de cárie, embora o coeficiente de Dice de 0,94 obtido por Rahimi *et al.* (2026) na segmentação mostre que a tarefa pode ser bem executada em conjuntos anotados com cuidado, sem que essa medida de sobreposição seja equivalente às métricas de detecção por lesão.

O escaneamento intraoral tem a vantagem de dispensar radiação ionizante e de permitir o acompanhamento longitudinal, mas a sensibilidade de 67% indica que o método ainda está em desenvolvimento (Jones *et al*., 2025). Dois estudos compararam o algoritmo com examinadores humanos: a ResNet-152 teve desempenho próximo ao dos melhores examinadores (Nampim *et al*., 2026) e o YOLOv8x não diferiu do observador mais experiente no resultado global, embora os autores concluam que não alcançou o nível de especialistas (Biçengil *et al*., 2026), o que sugere uma função de segunda opinião, e não de substituição, embora essas comparações tenham ocorrido em conjuntos de imagens definidos pelos autores e não reproduzam uma consulta, em que o exame clínico fornece informações que a

radiografia não traz.

O desenvolvimento de modelos depende de imagens representativas e de anotações confiáveis, e conjuntos de um único serviço podem refletir particularidades de equipamentos, protocolos de exposição, padrões de tratamento e distribuição de lesões. Os conjuntos variaram de 66 radiografias (Gonzalez *et al*., 2024) a 6.075 exames (Asci *et al*., 2024), o que também dificulta comparar a robustez dos resultados, e a falta de validação em centros independentes impede prever o desempenho em outras populações, enquanto divisões inadequadas entre treinamento e teste podem causar vazamento de informação e superestimar a acurácia.

Os nove estudos são heterogêneos, pois alguns classificam dentes, outros localizam lesões por superfície e outros segmentam regiões de interesse, o que impede calcular um desempenho global interpretável; sensibilidade, especificidade, precisão, F1 e AUC são medidas complementares e não intercambiáveis, e o uso clínico exige considerar a finalidade do sistema e o desfecho efetivamente validado. A lista de verificação para pesquisas sobre IA em Odontologia (Schwendicke *et al*., 2021) pode orientar o relato dos estudos futuros, e os sistemas ICDAS e ICCMS (Pitts; Ekstrand, 2013) oferecem referências comuns para a classificação das lesões.

Na Odontopediatria, as ferramentas automatizadas devem ser usadas com transparência e supervisão, e a interpretação final cabe ao cirurgião-dentista, que integra os achados de imagem ao exame clínico, ao histórico médico, ao risco de cárie e às preferências da família. O uso de sistemas digitais exige ainda proteção dos dados pessoais, atenção ao consentimento e avaliação de possíveis conflitos de interesse ligados a tecnologias comerciais, e no Brasil o tratamento de imagens e dados de crianças deve observar a Lei Geral de Proteção de Dados Pessoais (Brasil, 2018), que traz regras específicas para os dados de crianças e adolescentes.

O resultado do algoritmo não deve ser apresentado aos responsáveis como diagnóstico definitivo, pois erros de classificação podem gerar ansiedade, exames adicionais ou tratamentos desnecessários, e a adoção clínica deve vir acompanhada de validação local, capacitação dos profissionais e monitoramento contínuo do desempenho.

Os estudos futuros devem priorizar delineamentos prospectivos e multicêntricos, com amostras pediátricas maiores e representativas das dentições decídua e mista, métricas relatadas por estágio, superfície e tipo de dente, e medidas do efeito do apoio da IA sobre as decisões terapêuticas e os desfechos para a criança, além de dados sobre aceitação por profissionais e responsáveis, custo e segurança das

imagens.

Esta revisão tem limitações: a busca concentrou-se na PubMed/MEDLINE, sem busca exaustiva documentada em outras bases, e dois dos estudos incluídos não foram recuperados pelas consultas e foram identificados por busca complementar, o que indica que a estratégia usada não recupera toda a literatura pertinente; os estudos com fotografias intraorais de crianças foram excluídos por critério de modalidade e merecem síntese própria; não houve avaliação formal do risco de viés com instrumento padronizado; e a síntese se apoia em informações bibliográficas e em resumos, e não na extração independente dos textos completos.

Por essas razões, o trabalho não pode ser classificado como revisão sistemática, e suas conclusões devem ser lidas como uma síntese exploratória da literatura.

CONCLUSÃO

Os nove estudos analisados indicam que a IA tem potencial para auxiliar a detecção e a classificação de cárie em Odontopediatria. Os resultados mais favoráveis ocorreram em tarefas específicas de interpretação radiográfica, especialmente nas radiografias interproximais, mas o desempenho variou com a modalidade de imagem, o estágio das lesões e os critérios de avaliação, e a identificação de lesões iniciais, a generalização para novas populações e a validação externa seguem como desafios.

Os algoritmos devem, portanto, ser tratados como ferramentas complementares ao julgamento clínico, usadas com supervisão do cirurgião-dentista e sem substituir o exame clínico, a avaliação de risco e a decisão compartilhada com a família.

A consolidação do seu uso exige estudos multicêntricos e prospectivos, padronização das métricas e dos critérios de referência, avaliação do impacto nas decisões terapêuticas e observância dos princípios éticos e de proteção de dados. As conclusões desta revisão exploratória devem ser lidas à luz de suas limitações metodológicas.

REFERÊNCIAS

BBOTT, L. P.; SAIKIA, A.; ANTHONAPPA, R. P. Artificial intelligence platforms in dental caries detection: a systematic review and meta-analysis. **Journal of Evidence-Based Dental Practice**, [*S. l.*], v. 25, n. 1, art. 102077, 2025. DOI: https://doi.org/10.1016/j.jebdp.2024.102077.

AMMAR, N.; KÜHNISCH, J. Diagnostic performance of artificial intelligence-aided caries detection on bitewing radiographs: a systematic review and meta-analysis.

Japanese Dental Science Review, [*S. l.*], v. 60, p. 128-136, 2024. DOI: https://doi.org/10.1016/j.jdsr.2024.02.001.

ASCI, E. *et al*. A deep learning approach to automatic tooth caries segmentation in panoramic radiographs of children in primary dentition, mixed dentition, and permanent dentition. **Children**, Basel, v. 11, n. 6, art. 690, 2024. DOI: https://doi.org/10.3390/children11060690.

BIÇENGIL, K. *et al*. Deep learning-based dental caries diagnosis on panoramic radiographies: performance of YOLOv8 versus human observers. **Diagnostics**, Basel, v. 16, n. 8, art. 1150, 2026. DOI: https://doi.org/10.3390/diagnostics16081150.

BRASIL. [Lei nº 13.709, de 14 de agosto de 2018]. Lei Geral de Proteção de Dados Pessoais (LGPD). **Diário Oficial da União**: seção 1, Brasília, DF, p. 59, 15 ago. 2018.

GONZALEZ, C. D. *et al*. Identifying primary proximal caries lesions in pediatric patients from bitewing radiographs using artificial intelligence. **Pediatric Dentistry**, [*S. l.*], v. 46, n. 5, p. 332-336, 2024. Disponível em: https://pubmed.ncbi.nlm.nih.gov/39420489/. Acesso em: 21 out. 2026.

GÜÇLÜ, H. S.; DEMIREL, A.; ORHAN, K. Evaluation of a YOLOv5x-based deep learning model for interproximal caries segmentation on bitewing radiographs across primary and permanent teeth. **BMC Oral Health**, [*S. l.*], v. 26, n. 1, art. 202, 2025. DOI: https://doi.org/10.1186/s12903-025-07570-2.

HAAG, D. G. *et al*. Oral conditions and health-related quality of life: a systematic review. **Journal of Dental Research**, [*S. l.*], v. 96, n. 8, p. 864-874, 2017. DOI: https://doi.org/10.1177/0022034517709737.

JONES, B. *et al*. Dental caries detection in children using intraoral scans and deep learning. **Journal of Dentistry**, [*S. l.*], v. 160, art. 105906, 2025. DOI: https://doi.org/10.1016/j.jdent.2025.105906.

KASSEBAUM, N. J. *et al*. Global, regional, and national prevalence, incidence, and disability-adjusted life years for oral conditions for 195 countries, 1990-2015: a systematic analysis for the global burden of diseases, injuries, and risk factors. **Journal of Dental Research**, [*S. l.*], v. 96, n. 4, p. 380-387, 2017. DOI: https://doi.org/10.1177/0022034517693566.

MENDES, F. M.; BRAGA, M. M. Caries detection in primary teeth is less challenging than in permanent teeth. **Dental Hypotheses**, [*S. l.*], v. 4, n. 1, p. 17-20, 2013. DOI: https://doi.org/10.4103/2155-8213.110185.

MERTENS, S. *et al*. Artificial intelligence for caries detection: randomized trial. **Journal of Dentistry**, [*S. l.*], v. 115, art. 103849, 2021. DOI: https://doi.org/10.1016/j.jdent.2021.103849.

MOHAMMAD-RAHIMI, H. *et al*. Deep learning for caries detection: a systematic review. **Journal of Dentistry**, [*S. l.*], v. 122, art. 104115, 2022. DOI: https://doi.org/10.1016/j.jdent.2022.104115.

NAMPIM, N. *et al*. Deep learning-based caries lesion classification of primary teeth using bitewing radiographs and its comparison with dental professionals. **European Archives of Paediatric Dentistry**, [*S. l.*], v. 27, n. 2, p. 607-618, 2026. DOI: https://doi.org/10.1007/s40368-025-01129-z.

PITTS, N.; EKSTRAND, K. International Caries Detection and Assessment System (ICDAS) and its International Caries Classification and Management System (ICCMS): methods for staging of the caries process and enabling dentists to manage caries. **Community Dentistry and Oral Epidemiology**, [*S. l.*], v. 41, n. 1, p. e41-e52, 2013. DOI: https://doi.org/10.1111/cdoe.12025.

PONTES, L. R. A. *et al*. Negligible therapeutic impact, false-positives, overdiagnosis and lead-time are the reasons why radiographs bring more harm than benefits in the caries diagnosis of preschool children. **BMC Oral Health**, [*S. l.*], v. 21, n. 1, art. 168, 2021. DOI: https://doi.org/10.1186/s12903-021-01528-w.

RAHIMI, H. *et al*. Detection of pediatric dental caries in panoramic radiograph using deep learning: a benchmark study on MD-OPG. **Sensors**, Basel, v. 26, n. 8, art. 2481, 2026. DOI: https://doi.org/10.3390/s26082481.

SCHWENDICKE, F. *et al*. Artificial intelligence in dental research: checklist for authors, reviewers, readers. **Journal of Dentistry**, [*S. l.*], v. 107, art. 103610, 2021. DOI: https://doi.org/10.1016/j.jdent.2021.103610.

SELWITZ, R.; ISMAIL, A.; PITTS, N. Dental caries. **The Lancet**, [*S. l.*], v. 369, n. 9555, p. 51-59, 2007. DOI: https://doi.org/10.1016/S0140-6736(07)60031-2.

TAN, Z. Q. *et al*. Caries detection in primary molars with bitewing radiographs through deep learning based-object detectors. **Caries Research**, Basel, 18 dez. 2025. DOI: https://doi.org/10.1159/000550079.

YAVSAN, Z. S. *et al*. Diagnosis of approximal caries in children with convolutional neural networks based detection algorithms on radiographs: a pilot study. **Acta Odontologica Scandinavica**, [*S. l.*], v. 84, p. 18-25, 2025. DOI: https://doi.org/10.2340/aos.v84.42599.