Journal of Imaging Informatics in Medicine (2024) 37:1824–1845
https://doi.org/10.1007/s10278-024-01054-5

A Systematic Review on Caries Detection, Classification, and Segmentation from X-Ray Images: Methods, Datasets, Evaluation, and Open Opportunities

Luiz Guilherme Kasputis Zanini[1] · Izabel Regina Fischer Rubira-Bullen[2] · Fátima de Lourdes dos Santos Nunes[1,3]

Received: 7 September 2023 / Revised: 19 December 2023 / Accepted: 2 January 2024 / Published online: 1 March 2024

Abstract
Dental caries occurs from the interaction between oral bacteria and sugars, generating acids that damage teeth over time. The importance of X-ray images for detecting oral problems is undeniable in dentistry. With technological advances, it is feasible to identify these lesions using techniques such as deep learning, machine learning, and image processing. Therefore, the survey and systematization of these methods are essential to determining the main computational approaches for identifying caries in X-ray images. In this systematic review, we investigated the primary computational methods used for classifying, detecting, and segmenting caries in X-ray images. Following the PRISMA methodology, we selected relevant studies and analyzed their methods, strengths, limitations, imaging modalities, evaluation metrics, datasets, and classification techniques. The review encompassed 42 studies retrieved from the Science Direct, IEEExplore, ACM Digital, and PubMed databases from the Computer Science and Health areas. The results indicate that 12% of the included articles utilized public datasets, with deep learning being the predominant approach, accounting for 69% of the studies. The majority of these studies (76%) focused on classifying dental caries, either in binary or multiclass classification. Panoramic imaging was the most commonly used radiographic modality, representing 29% of the cases studied. Overall, our systematic review provides a comprehensive overview of the computational methods employed in identifying caries in radiographic images and highlights trends, patterns, and challenges in this research field.

Keywords Caries · Dental caries · Cavities · Radiography · Machine learning · Deep learning · Image processing · Systematic review

Introduction

The deterioration of enamel and dentin caused by bacteria in dental plaque is a disease known as dental caries, which impacts oral health. In the absence of treatment, the disease can progress to the inner part of the tooth, known as the dental pulp, where nerves and blood vessels are present, causing inflammation and tooth loss [1]. Researchers have been developing computational methods for diagnosing various dental abnormal conditions, such as periodontal disease, dental abscess, and lesions in dental canals, by using different imaging modalities [2–4].

In particular, X-ray methods are especially useful for diagnosing interproximal caries, which occur between the contact areas of the teeth and where the field of view is limited [5]. Images generated for these methods are frequently used as inputs for computer methods that diagnose caries.

Caries diagnosis has been addressed in previous systematic reviews (SRs), with a focus on methods using

Izabel Regina Fischer Rubira-Bullen and Fátima de Lourdes dos Santos Nunes contributed equally to this work.

✉ Luiz Guilherme Kasputis Zanini
luiz.kasputis@usp.br

Izabel Regina Fischer Rubira-Bullen
izabullen@usp.br

Fátima de Lourdes dos Santos Nunes
fatima.nunes@usp.br

1 Department of Computer Engineering and Digital Systems, University of São Paulo, Av. Prof. Luciano Gualberto 158, São Paulo 05508-010, São Paulo, Brazil

2 School of Dentistry at Bauru, University of São Paulo, Alameda Doutor Octávio Pinheiro Brisolla, Bauru 17012-191, São Paulo, Brazil

3 School of Arts, Sciences and Humanities, University of São Paulo, Rua Arlindo Béttio, 1000, São Paulo 03828-000, São Paulo, Brazil

deep learning techniques and different types of radiographic modalities. In [6], the goal was to identify different approaches for diagnosing caries in periapical images. In contrast, some studies [7, 8] only focused on investigating deep learning methods. Additionally, the systematic review presented in [9] identified studies that employed machine learning algorithms but did not address the differentiation of exam acquisition for caries diagnosis. However, none of the previous studies has comprehensively addressed the comparison between different radiographic modalities, incorporating a variety of diagnostic objectives for dental caries, including the comparison of classical machine learning, image processing, and deep learning.

In this SR, we investigated the influence of different methods and computational approaches on the diagnostic performance of caries in X-ray images. We review the limitations of these approaches, including challenges with clinical practices. We also highlight the strengths of these methods, and explore the types of imaging modalities used, ranging from periapical radiographs to computer tomography, as well as how different caries classifications are approached. Furthermore, we discuss the main evaluation criteria used and describe the main image datasets used. As a result, this study provides a comprehensive view of current advances and challenges in applying computational methods for caries diagnosis in X-ray images.

- Identification of key trends and research gaps
- Discussion about the challenges that computational methods face in accurately diagnosing dental caries
- Exploration of the strengths and limitations associated with computer methods in diagnosis of dental caries
- Discussion about the challenges that computational methods face in accurately diagnosing dental caries
- Identification of the main image datasets to test the approaches and their characteristics
- Influences of methods and approaches on diagnostic performance
- Analysis of the main metrics to evaluate the approaches

This review intends to establish a solid basis for future studies and research to improve the accuracy and effectiveness of caries diagnosis through radiographic images.

In this article, the background is addressed in "Background", presenting the main concepts necessary to understand the different approaches. The research questions, the methods, and the systematization process of the review are described in "Methods". In "Overview of the Included Studies", an overview of the included studies is presented, as well as an analysis of the data extracted from the included studies. In "Computational Methods", the studies are detailed and divided into specific objectives, presenting the limitations and the use of technologies in each study. A discussion of the results and answers to the research questions set at the beginning of the review is presented in "Discussion". Section "Open Gaps and Research Possibilities" addresses research opportunities and gaps that need to be filled, followed by the conclusion presented in "Final Remarks"

Background

This section presents relevant concepts to understand tooth structure and computational aspects used during the SR process.

Dental Structures, Caries, and Imaging Modalities

The tooth is a hard and mineralized anatomical structure responsible for cutting, crushing, and grinding food, consisting of three different parts (Fig. 1): enamel, dentin, and dental pulp [10]. Dental caries is a disease caused by the activity of bacteria that metabolize sugar-producing acids that can destroy dental tissues [11]. The progression of caries can cause pain, sensitivity, and even tooth loss [12]. Caries is the most predominant oral disease in the world, affecting approximately 2 billion people worldwide [10].

Three main strategies are used to diagnose dental caries [13]. The first technique is visual diagnosis, where the dentist examines the teeth with a mirror in the areas affected by caries. The second technique is laser technology, which allows early detection of cavities, even before they become visible. Finally, the third technique is the use of X-ray exams to acquire radiographic images to identify regions injured

Fig. 1 The tooth is composed of three main elements: enamel, dentin, and root. Enamel is the white outer layer, dentin is the beige structure, and the root is the red part

by dental caries. The four radiographic methods used for diagnosing dental caries are provided in this review:

- **Periapical radiographs** provide an expanded field of view that encompasses the root and crown of the entire tooth, in addition to the surrounding bone and periodontal tissues [14].
- **Interproximal radiographs or bitewing** provide a two-dimensional image that allows the visualization of the crowns of two or three adjacent teeth, facilitating the evaluation of the presence of caries between teeth [14].
- **Panoramic X-ray** offers a broad view of the dental arch, including the maxillary and mandibular bones, paranasal sinuses, and temporomandibular joints [14].
- **Cone beam computed tomography (CBCT)** provides detailed three-dimensional (3D) images of the teeth, bones, and soft tissues of the head, allowing the dentist to visualize the presence of caries and other purposes, such as dental canal, low bone formation, and surgical treatment planning [15].

In the dental context, the choice of the imaging modality depends on the treatment and diagnosis that the dentist expects to perform. The X-rays (periapical, interproximal, and panoramic) are part of the two-dimensional (2D) imaging technique used in dentistry. Each of these X-rays provides a distinct view of dental and periodontal structures and a comprehensive overview of the patient's oral health. Three-dimensional (3D) radiographic modalities allow visualization of the maxillofacial structure from different angles, making them useful for orthodontic treatment planning, implant placement, and diagnosis of various diseases [16].

Radiographic modalities with a reduced field of view, such as interproximal and periapical, favor lower radiation emissions and are often used for detecting cavities and visualizing dental structures. In contrast, 3D imaging modalities have a wider field of view but require a higher radiation dose. A large number of radiographic modalities are available for diagnosing cavities, and this is one of the aspects analyzed in this SR.

Computer Concepts

Computer vision is a research field that employs computational approaches to enable computers to process and interpret images and videos in a manner similar to humans [17]. It encompasses a variety of methods and algorithms to perform a range of tasks, including classification, segmentation, and object detection in images or videos (Fig. 2), the categories used in the present SR.

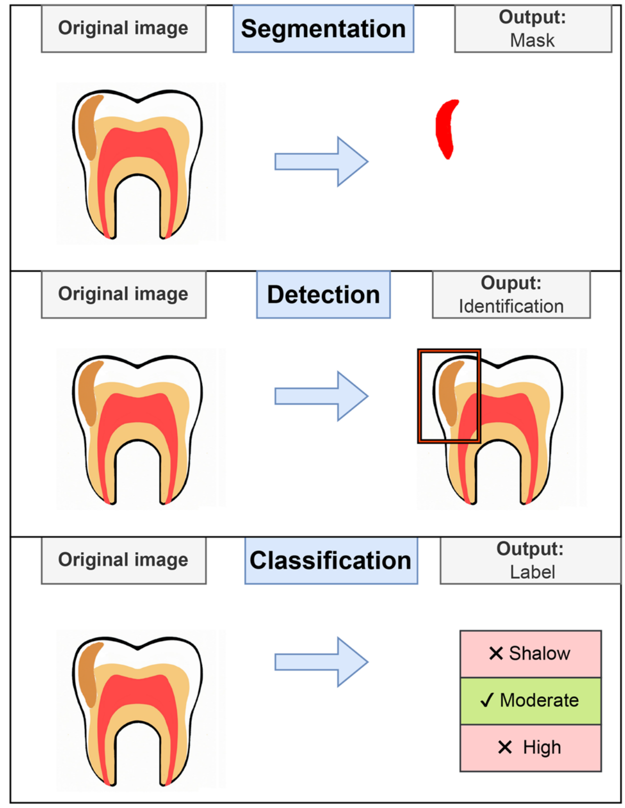


Fig. 2 Three main objective categories analyzed in this SR: classification, detection, and segmentation in computer vision

Segmentation is used to divide the image into distinct regions or segments based on the characteristics of the object of interest present in the image. **Detection** identifies and precisely locates objects within the image, generating an output in the form of bounding rectangles that surround the region of interest. **Classification** is a technique used to assign categories to an image, resulting in an output that indicates the association of the image with a specific category.

Diagnosing cavities through radiographic modalities can be challenging due to several factors. The most challenging aspect is diagnosing cavities in the early stages when the demineralization of tooth tissue is not very perceptive in the image without some processing. Additionally, distinguishing active and inactive caries lesions is challenging for dentists since cavity activity is related to factors such as saliva pH and the presence of sugar [18]. Thus, diagnosis through images is challenging since the disease's interpretation is made by radiologists and dentists, who must have experience in this type of analysis. Therefore, using computational methods may be a factor to improve disease diagnosis.

Table 1 Database and search strings

Database	Search strings
IEEExplore	(dental caries OR dental cavity OR caries) AND (classification OR detection OR Identification OR segmentation OR localization) AND (CBCT OR computer tomography OR X-ray OR panoramic OR bitewing OR radiography OR radiographic)
ACM Digital	(dental caries OR dental cavity OR caries) AND (classification OR detection OR Identification OR segmentation) AND (CBCT OR computer tomography OR X-ray OR panoramic OR bitewing OR radiography OR radiographic)
Science Direct	(dental caries OR dental cavity OR caries) AND (classification OR detection OR Identification OR segmentation OR localization)
PubMed	(dental caries OR dental cavity OR caries) AND (classification OR detection OR Identification OR segmentation OR localization) AND (CBCT OR computer tomography OR X-ray OR panoramic OR bitewing OR radiography OR radiographic)

Methods

This SR study was based on the PRISMA methodology [19]. This methodology consists of a sequential series of steps, starting with the formulation of research questions, followed by literature searches, screening and selection of pertinent studies, extraction of data, and ultimately, analysis and synthesis of the findings.

Before starting the SR, we conducted an exploratory analysis of articles of interest within the theme we decided to explore to define databases and keywords to fill a protocol. Next, this SR was divided into four main stages: (i) planning, (ii) selection of primary studies, (iii) data extraction, and (iv) interpretation of the results, as presented below.

Planning

In the first step of the SR, a protocol was described and discussed among the authors; the main topics are presented below.

The following databases and keywords were based on a previous exploratory analysis. They are *ScienceDirect*,[1] *ACM*,[2] *PubMed*,[3] and *IEEE*.[4]

To reach the main goal of the SR, we defined the following research questions:

1. What computer methods have been used to diagnose dental caries?
2. What is the influence of the approaches and methods on diagnostic performance?
3. What are the limitations of the approaches?
4. What are the strengths of the approaches?
5. What types of imaging modalities were used?
6. What types of caries classification were considered?
7. What were the main evaluation metrics used?
8. What were the main image datasets used?

Table 1 presents the search strings used, which were selected based on studies on detecting and segmenting dental caries discovered in a previous exploratory analysis. The search period was limited from 2010 to February 2023. The search for studies considered occurrences of the searched terms in the keywords title and abstract of the studies. However, the search in ScienceDirect was more extensive due to limitations in using Boolean operators.

To select relevant studies, we established a set of criteria for inclusion (I) and exclusion (E) (Table 2). These criteria served as guidelines to determine which studies should be considered for the systematic analysis.

After including the studies, we applied quality criteria. The criteria were evaluated by using a scale ranging from 0.0, indicating noncompliance, to 0.5, indicating partial compliance, and finally to 1.0, indicating full compliance with the criterion (Table 3).

Study Selection

Figure 3 presents a PRISMA flow diagram with the quantitative results obtained after each step executed in this SR, as well as the reasons for the exclusion of articles, according to the criteria presented in Table 2.

The first step was identifying the main information for electronic and manual searches. In the screening stage, we applied the inclusion and exclusion criteria to a total of 1125 records. During this phase, two exclusion criteria (6 and 7) were applied using a semiautomatic method. This method involved assessing the language of the article and determining if the study was a review or survey. A total of 857 articles remained for the next step.

The next step was to read the titles of the articles and apply exclusion criteria, resulting in a set of 165 records. Subsequently, the abstracts of these articles were read, and further exclusion criteria were applied again, leaving 53 articles.

Finally, the 53 remaining articles were read in their entirety, and the criteria were applied once again. A total of 42 articles met all the specified criteria.

1 http://www.sciencedirect.com
2 http://dl.acm.org
3 https://www.ncbi.nlm.nih.gov/pubmed
4 https://ieeexplore.ieee.org/

Table 2 Criteria for inclusion and exclusion and description of the criteria

Criteria	Description
I1	The study proposed methods for the classification, detection, identification, or segmentation of caries in radiographic images
E1	The study did not propose methods for classification, detection, identification, or segmentation of caries
E2	The study did not use radiological images
E3	The study considered other diseases or tooth structure (periapical disease, tooth root, tooth canal, etc.)
E4	The study considered only deciduous dentition
E5	The study had less than five pages (work in progress)
E6	The study was not primary (e.g., reviews and surveys)
E7	The study was not written in English

Data Extraction and Interpretation

In this stage, we systematically extracted information from each included study to address the initial research questions. Some attributes to be extracted vary depending on the specific research objective, emphasizing various computational methods utilized in the studies. In general, the following attributes were extracted: image dataset, including whether they were publicly available and their size; type of radiographic image being used; main project objectives, categorized into segmentation, classification, and detection; and computer techniques employed, categorized into deep learning, image processing, or classical machine learning.

The data extracted from the included 42 can be seen in Table 2. After reading the full papers, we carefully analyzed the relevant contributions. The findings were then categorized and discussed to identify patterns and trends in their applications.

Overview of the Included Studies

This section summarizes the studies, focusing on quality criteria and extracted attributes. Figure 4 offers a snapshot of how the studies align with the quality criteria outlined in Table 3. Additionally, Table 4 outlines the key information drawn from the studies included in this analysis.

Table 3 Quality criteria and description of the criteria

Criteria	Description
Q1	The classification or detection technique is clearly presented, indicating the objective of the work
Q2	The study differentiates levels of caries lesions
Q3	The limitations of the method are presented and discussed
Q4	The results are compared with other approaches
Q5	The evaluation method is clearly presented, for example, the definition of evaluation metrics

Overview Quality Criteria

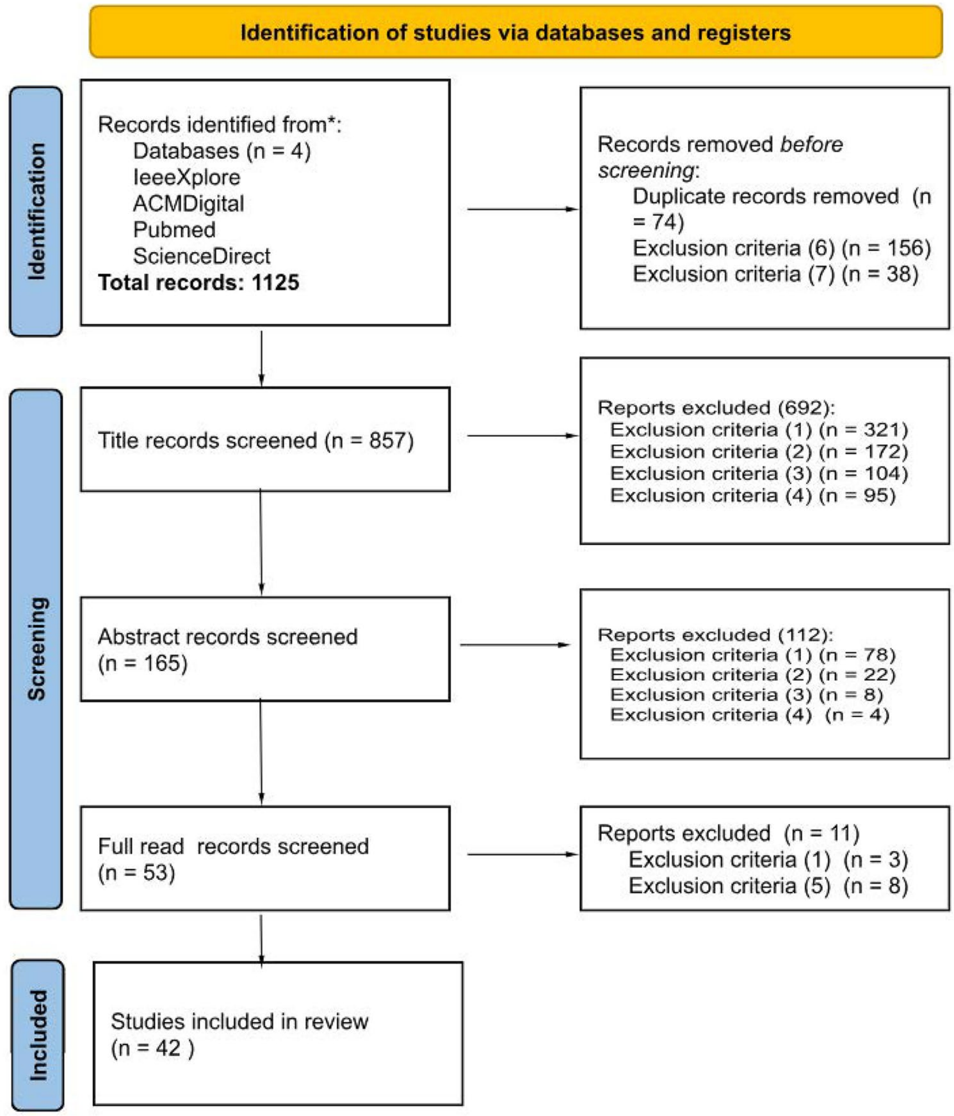
The overall criterion of the included studies is shown in Fig. 4. The first criterion measures how clearly the methods are presented. It occurs that 45% of the studies meet this criterion, and half partially attended. A study was considered complete when it clearly articulated its contributions and objectives. Studies that mentioned their goals but did not clearly describe how they contributed to the overall objective were classified as partially completed (50%). The lack of clarity of objectives is problematic since some studies do not provide information about the purpose of new research or even their contribution to the area.

The criterion Q2 evaluates the ability of studies to distinguish between different levels of caries. Notably, a significant majority (71%) of the studies fail to satisfy this criterion, underscoring a critical area that warrants further focus. Meanwhile, 12% of the studies achieve only a partial fulfillment of this criterion, which typically involves merely stating the number of classification categories without providing detailed methodologies for such classification. This shortfall is primarily due to the complexity of accurately determining caries levels. Additionally, the high percentage of studies could be attributed to the lack of advanced imaging technologies or specialized clinical expertise in most cases.

The Q3 criterion is related to the discussion of a method's limitations. It shows that 52% of the papers discuss this aspect. A total of 7% of studies only mention future studies and do not address limitations in the actual approach, and thus, they were considered partially completed. Failing to address limitations can lead to an incomplete grasp of the research's applicability, highlighting the necessity for detailed and thorough critical analysis in all studies.

In criterion Q4, referring to comparison with other approaches, with 45% of articles achieving completeness. However, 31% of the studies only partially met this criterion, mainly due to their limited scope in comparing various methods within the study itself instead of including comparisons with other researches. These comparisons are essential

Fig. 3 PRISMA flow diagram that represents the study selection and inclusion process



to contextualize new approaches in the existing scenario and highlight advances in the area.

Finally, criterion Q5, which assesses clarity in the presentation of evaluation metrics, showed that most works (60%) met this requirement. There are 30% of studies that attend partially, showing only common metrics but not indicating how the classification, detection, and segmentation are measured. It is necessary to provide clearer metrics presentations, whether through tables, graphs, or figures, to better understand and allow replicate results.

Figure 5 presents a boxplot graph that shows each category's median and quartiles, illustrating the distribution difference between articles published in conferences and journals based on quality criteria. A notable observation is that some

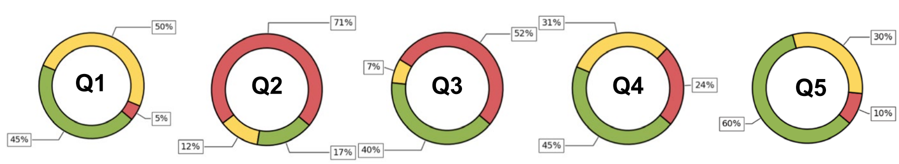


Fig. 4 Application of the quality criteria applied to the 42 studies included. Red represents studies that did not meet the criteria, yellow indicates studies that partially met the criteria, and green represents studies that fully met the criteria

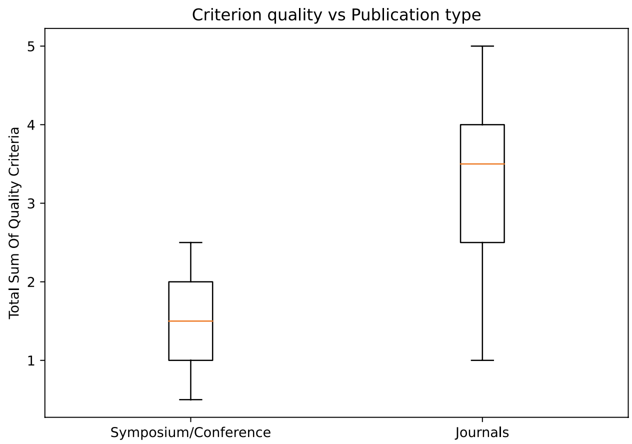


Fig. 5 The relationship between publication type and criterion quality can be observed through the boxplot, which displays each category's maximum and minimum values

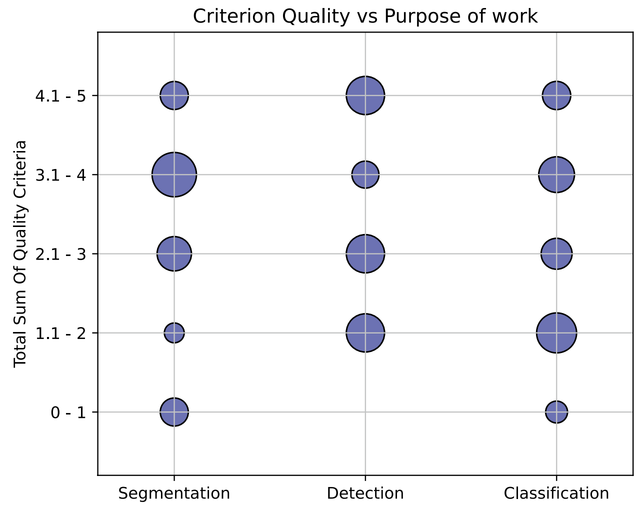


Fig. 6 Relationship between the purpose of the study and quality criteria. The scatter points were normalized by dividing the number of studies present in a specific group by the total number of studies in the task. This normalization establishes a relationship between the individual studies and their distribution within the task

studies published in journals fully met the criteria, obtaining the maximum value (five), while articles published in conferences and symposiums obtained a maximum value of 2.5. Furthermore, there is a substantial range between the minimum and maximum values for studies published in journals, indicating that not all papers fully meet the established criteria.

Additionally, the studies were analyzed according to their objective categories, as shown in Fig. 6. Each bubble within a category represents the total percentage of studies falling within a specific interval. The studies with lower scores are related to the classification objective, with the maximum score ranging between 1.1 and 2 because some of the studies were missing how caries images were classified. Regarding segmentation, the score distributions appear to be quite similar, with a higher concentration in the interval [3.1,4]; however, segmentation studies require a detailed analysis of the pixels and an understanding of the structures within the image, which is reflected in the study in question. In the detection category, no study obtained a score in the lower interval, and similar distributions were in the interval [1.1, 5], suggesting that detection studies have a better distribution.

Overview Attributes

Table 4 shows the attributes extracted from each included article. These attributes represent relevant information found in each article. These data are crucial for the analysis and understanding of the results obtained in this research. In the following sections, we describe these attributes, followed by a comprehensive analysis of the collected data.

Figure 7 indicates the distribution of the included studies over the years. It is important to note that, in recent years, there has been notable growth in the number of studies aimed at diagnosing dental caries. This increase is largely due to improvements in deep learning techniques and increased data across multiple radiographic modalities. Furthermore, it is important to highlight that the exclusive use of image processing did not have such a significant increase, as this approach is being used as a preliminary step in the deep learning process.

The datasets used in each study were classified as public or private. Datasets that offer open access were classified as public, while datasets without this access were categorized as private. Datasets in papers that intended to make data available, such as through contacting the author via mail but did not provide an external link for open access, were categorized as private.

The dataset size can differ significantly depending on the study. The sample generally consists of images, but in some cases, studies divide the data by teeth or specimens. In this category, we separated the studies into total images used, which corresponds to the value of the training and testing images provided by the authors.

The computational techniques presented in the articles were divided into three categories, considering the computational areas in the studies surveyed. The first category is image processing (IP), which corresponds to applying specific techniques for image analysis and manipulation. The second category is classical supervised learning (CSL), which uses feature extraction to apply conventional machine learning algorithms, such as naive bayes (NB), support vector machine (SVM), random forest (RF), and AdaBoost. Finally, the third category comprises studies that employed deep learning (DL) approaches, which utilize a

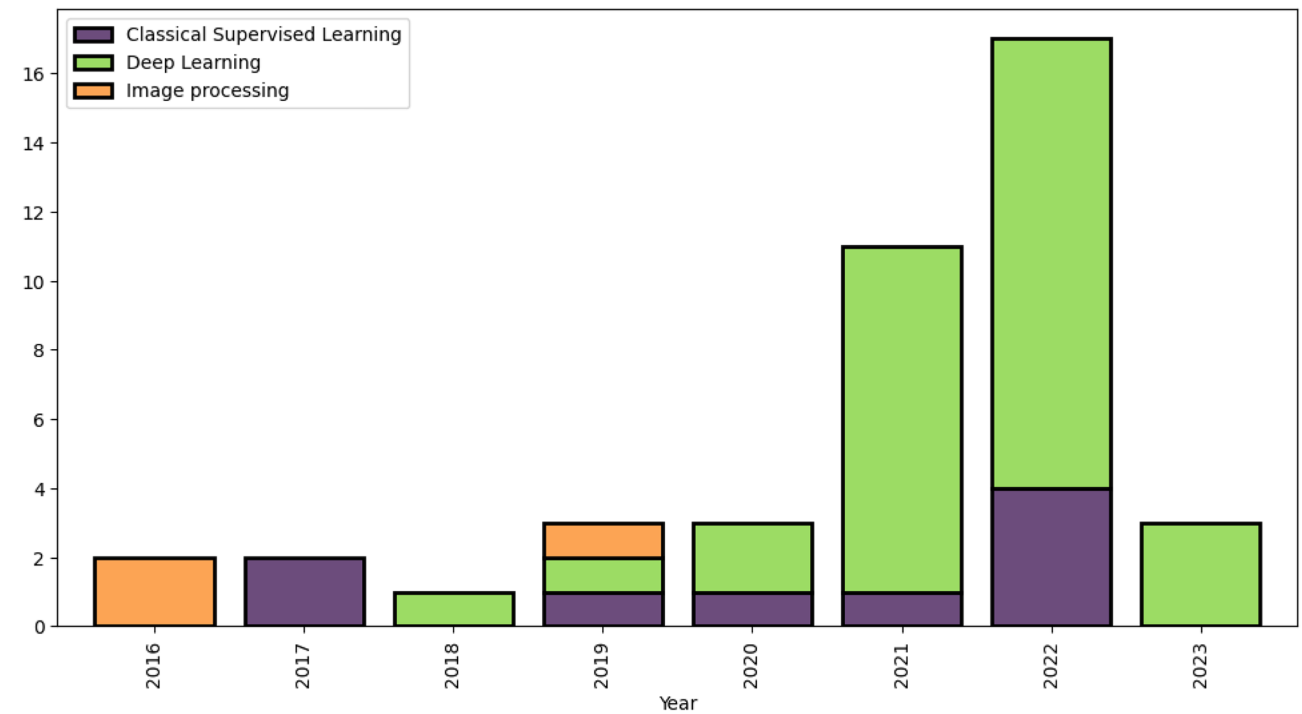


Fig. 7 Distribution of studies according to years and relation to the use of classical supervised learning, deep learning, and image processing techniques

convolutional neural network architecture for extracting features through convolution layers and neurons to make predictions.

The study objectives were categorized into three main classes: classification, detection, and segmentation. Some studies involve multiple objectives simultaneously. Section "Background" of the study describes these objectives in a computer vision context.

The image type attribute refers to the imaging modality presented in "Background". Based on this, it was possible to identify the main types of radiographic modalities and provide an overview of the state of the art in applying computational diagnosis in caries. However, some studies did not provide detailed information on the radiographic modality used, limiting themselves to mentioning only "X-rays" as the modality.

Computational Methods

To provide a detailed view of the computational methods presented in the included articles, the sections below are divided into three categories, whose definitions are presented in "Background": classification, detection, and segmentation. Within each category, the techniques and metrics used are related, as well as the strengths and limitations of the tasks related to each objective. Figure 8 shows the occurrence of techniques separated into three groups.

Segmentation

Table 5 presents studies that concentrate on segmenting different regions of the tooth. The main objective of these studies is to identify and distinguish various regions by using a radiographic image. As can be observed, all the computational technique categories were found in segmentation studies. Different techniques and metrics were used within each category.

The approaches described in [20, 26, 27] utilized image processing techniques to segment specific regions of interest. The study [20] employed a combination of Otsu thresholding, median filtering, and morphological operations to generate a mask of the desired region. In the study by [26], a pipeline was developed to enhance image quality, focusing on tasks such as tooth separation, lesion localization, and caries region segmentation. The authors of the study used the geodesic active contour technique to accurately segment caries regions by leveraging contour gradients and image intensities. The study [27] compared the analysis of various methods to segment regions of interest in teeth to determine the most effective approach for segmenting these areas in dental images.

Table 4 Summary of the 42 included studies

Index	Ref.	Year	Dataset category	Dataset size	Technique category	Objective	Image type
1	[20]	2016	Public	80	IP	Segmentation	Bitewing
2	[21]	2016	Private	-	IP	Detection	X-ray
3	[22]	2017	Public	120	CSL	Classification	Periapical
4	[23]	2017	Private	93	CSL	Classification	X-ray
5	[24]	2018	Private	3000	DL	Classification	Periapical
6	[25]	2019	Private	120	CSL	Classification	X-ray
7	[26]	2019	Public	152	IP	Segmentation	Periapical
8	[27]	2019	Private	1500	DL	Classification segmentation	Panoramic
9	[28]	2020	Private	250	DL	Classification	Panoramic
10	[29]	2020	Private	1900	DL	Classification	Panoramic
11	[30]	2020	Private	105	CSL	Classification	Periapical
12	[31]	2021	Private	396	CSL	Classification	X-ray
13	[32]	2021	Private	470	DL	Classification	Panoramic
14	[33]	2021	Private	1160	DL	Classification segmentation	Panoramic
15	[34]	2021	Private	654	DL	Classification segmentation	Bitewing
16	[35]	2021	Private	4398	DL	Segmentation	CBCT
17	[36]	2021	Private	2325	DL	Detection segmentation	Bitewing
18	[37]	2021	Private	112	DL	Classification	Bitewing
19	[38]	2021	Private	278	DL	Classification	Bitewing
20	[39]	2021	Private	400	DL	Classification	Panoramic
21	[40]	2021	Public	480	DL	Classification	X-ray
22	[41]	2021	Private	206	DL	Segmentation	Periapical
23	[42]	2022	Private	476	DL	Classification	Panoramic
24	[43]	2022	Private	340	DL	Classification	Periapical
25	[44]	2022	Private	220	CSL	Classification	X-ray
26	[45]	2022	Private	198	CSL	Classification	X-ray
27	[46]	2022	Private	1037	DL	Segmentation	Periapical
28	[47]	2022	Private	978	DL	Classification detection	Bitewing
29	[48]	2022	Private	188	DL	Classification	Periapical
30	[49]	2022	Private	38,000	DL	Classification	Bitewing
31	[50]	2022	Private	175	DL	Classification detection	Bitewing
32	[51]	2022	Private	10,000	DL	Classification detection	Panoramic
33	[52]	2022	Private	1159	DL	Classification segmentation	Panoramic
34	[53]	2022	Private	200	DL	Detection	Periapical
35	[54]	2022	Private	4129	DL	Classification	Periapical
36	[55]	2022	Private	2758	DL	Classification	Bitewing
37	[56]	2022	Private	153	DL	Segmentation	X-ray
38	[57]	2022	Public	120	DL	Classification	X-ray
39	[58]	2022	Private	120	DL	Classification detection	Bitewing
40	[59]	2023	Private	562	DL	Classification	Panoramic
41	[60]	2023	Private	10,000	DL	Classification segmentation	Panoramic
42	[61]	2023	Private	500	DL	Segmentation	Bitewing

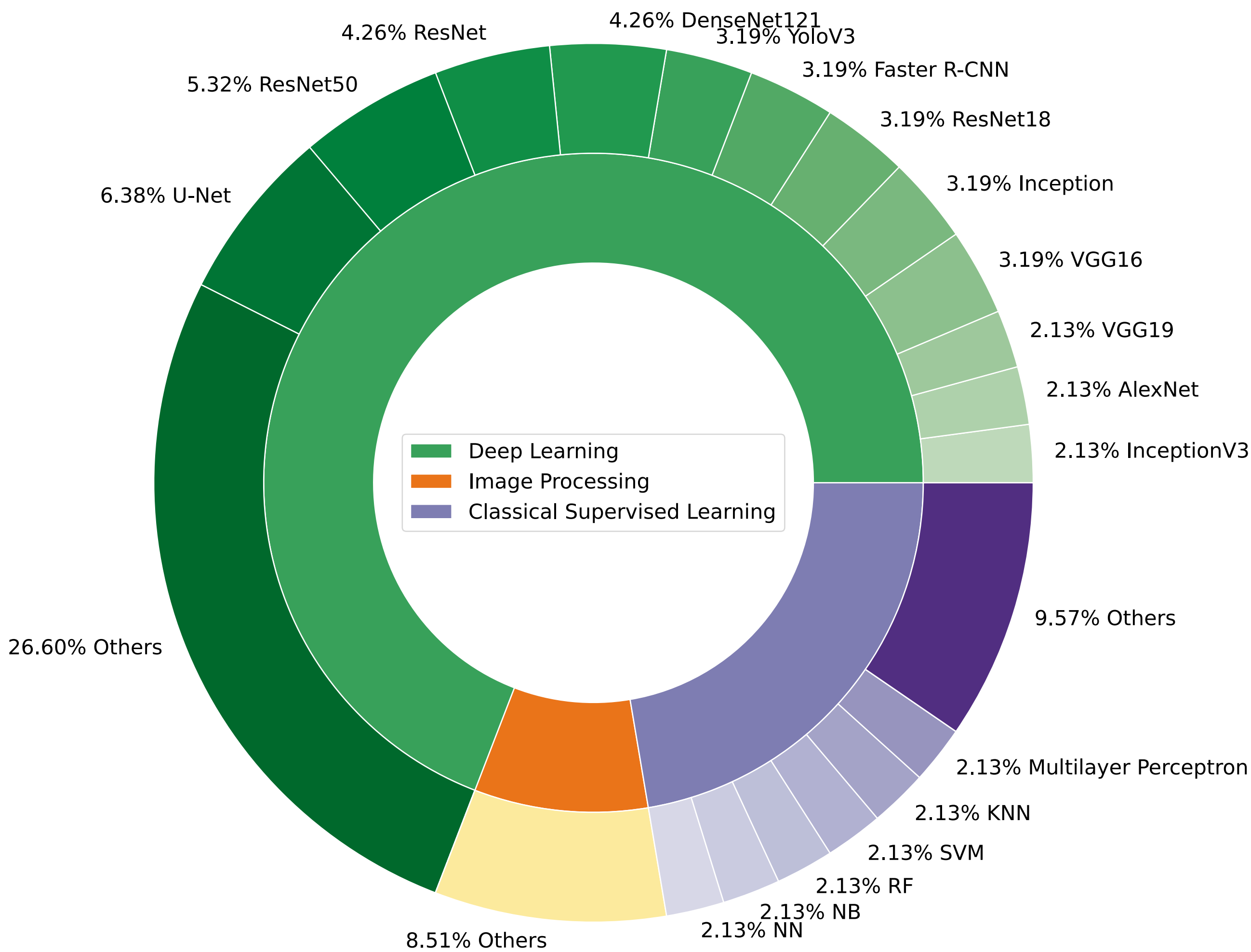


Fig. 8 Analysis of the frequency of techniques employed in the analysis of 42 works, category use of image processing, deep learning, and classical supervised learning. Techniques grouped in the "other" category are those that were identified only once

Table 5 Studies that present segmentation as the main objective

Index	Ref.	Techniques	Metrics
1	[20]	Preprocessing; edge recognition; thresholding; connected component labeling	Not informed
7	[26]	Segmented region filtration; modified k-means clustering; geodesic active counter	Accuracy; Dice; specificity; sensitivity
8	[27]	Hybrid graph-cut segmentation	Accuracy; F1-score; specificity; precision; recall
14	[33]	nnU-Net	Dice; IoU
15	[34]	U-Net	F1-score; sensitivity; precision
16	[35]	U-Net	Sensitivity; specificity
17	[36]	U-Net	Confusion matrix
22	[41]	U-Net; XNet; SegNet; U-Net + DenseNet-121	Dice; IoU
27	[46]	ResNet-101 and Mask R-CNN	Confusion matrix
33	[52]	CariesNet	Dice
37	[56]	Neural Network combined U-Net; trans-Unet; ResNet	Dice; precision
41	[60]	ResNet-50-DCDNet	F1-score
42	[61]	U-Net	F1-score; sensitivity; precision

Several deep learning segmentation approaches have been proposed for different types of radiography; the most chosen deep learning architecture was U-Net [34–36, 61]. In [34], the U-Net neural network was used to segment different regions to distinguish different lesion levels of lesions. In contrast, [35] used CBCT first to individualize the regions of each tooth and segmented the carious region. Finally, the segmentation proposed by [61] highlighted restorations, dental pulp, dental crowns, and caries.

In [33], a deep learning framework for medical image segmentation (nnU-Net) was utilized to automatically configure the preprocessing, network architecture, training, and postprocessing steps. The study [41] compared different segmentation architectures, including U-Net, SegNet, XNet, and U-Net + DenseNet-121, and found that U-Net architectures achieved better results. The approaches [52, 56] proposed convolutional architectures with modifications. In the case of CariesNet [52], the architecture is similar to U-Net but incorporates an additional module (partial encoder) to enhance the aggregation of high-level features and improve overall performance. The study [56] built upon the U-Net implementation, introducing a new approach with an encoder layer that is a variant of ResNet for feature extraction and integrated a visual transformation module in the backbone layer to enhance the incorporation of contextual information.

The study [46] used the Mask R-CNN ResNet-101 architecture, a neural network combining the R-CNN network with a segmentation mask to detect and segment objects in images, which is recognized for its accuracy in image classification. The architecture Dental Caries Detection Network (DCDNet) proposed in the study [60] presented a segmentation structure that contains a multipredicted output (MPO), where the final feature map is divided into three distinct paths to detect occlusal, proximal, and cervical caries.

Strengths and Limitations of Segmentation Tasks

Segmentation is an area that faces considerable challenges, especially concerning different dimensions, scales, resolutions, and imaging inputs [62]. Multiobjective segmentation is important for simultaneously targeting all relevant structures in an image, in contrast to single-object segmentation, such as targeting different tooth structures or degrees of caries. The studies [34, 52, 60] used multiple-objective approaches to segment caries, maxillofacial structure, or even segmenting the gravity of the caries.

Furthermore, segmentation with different radiographic modalities can be challenging, as different modalities may present different anatomical information. For example, segmentation in conventional radiography can be limited owing to low spatial resolution, while the presence of image artifacts can compromise segmentation in computed tomography [63].

Different approaches have been proposed to address segmentation challenges, including the combination of various convolutional neural networks with specific architectures for multiple-objective segmentation, as well as image preprocessing techniques to enhance image quality. Furthermore, developing a well-annotated image base and conducting rigorous evaluations of segmentation methods are crucial for achieving improved outcomes in this research area. Additionally, it is essential to recognize the importance of data balancing in segmentation tasks in machine learning models. Effective data balancing strategies can significantly improve the performance of such models, ensuring more accurate and reliable results.

Image processing can be a solution with small datasets. This category of approaches requires specialized knowledge of the problem and the application of different techniques. A limitation of this approach is that the quality of the input image can negatively affect the accuracy and reliability of the segmentation results, especially if the image contains noise or artifacts. Moreover, image processing yields results without requiring a training step.

The U-Net architecture is widely used in biomedical image segmentation due to its advantages [64]. First, U-Net uses an encoder-decoder architecture that makes input data more efficient, minimizing the need for a large dataset. Second, it features "skip connections" between the encoder and decoder, helping preserve low-level information and improve segmentation accuracy, especially for smaller, well-defined objects. Finally, U-Net can easily adapt to different image sizes, resolutions and segmentation issues. References that attest to these advantages include [34–36, 61].

The ResNet architecture was employed in the studies [46, 56, 60]. ResNet's strength lies in its ability to prevent the vanishing gradient problem in deep neural networks [65], enabling the construction of models that can learn complex and deep hierarchical features. This capability significantly improves the accuracy in identifying various types of dental caries and other dental conditions, leading to more precise diagnoses and more effective treatments in dentistry. The integration of ResNet into deep learning models underscores the potential of this architecture to improve medical image analysis, especially in dentistry.

The limitations of segmentation methods can be summarized as follows:

- **Sensitivity to noise**: The noise can lead to inaccurate segmentations if the method does not have a good generalization.
- **Dependence on image quality**: The quality of the image influences the segmentation, especially in images with the lowest quality.

Table 6 Studies whose main objective was the detection of regions with caries

Index	Ref.	Techniques	Metrics
2	[21]	Particle swarm optimization	Error rate
17	[36]	VGG16	Confusion matrix
28	[47]	Faster R-CNN with ResNet-50	AUC; Confusion matrix; RoC curve
31	[50]	YOLOv3	IoU
32	[51]	Faster R-CNN; ResNet; Inception	Precision; sensitivity; specificity
34	[53]	Faster R-CNN; YOLOv3	Average precision; F1-score
39	[58]	Deep gradient–based LeNet	Error rate

- **Computational limitations**: The methods for certain types of images do not work well in other images.
- **Understanding of the problem domain**: To employ the right techniques, a comprehensive understanding of the problem domain is necessary

Detection

Table 6 highlights the studies of detecting carious regions. These studies identified and located areas within an image that indicate the presence of caries. As evident from the findings, detection studies incorporated various categories of computational techniques. Each category utilized a diverse set of techniques and metrics.

In this study [21], the particle swarm optimization (PSO) algorithm was employed to divide the image into multiple regions or segments. Therefore, the proposed approach identified the intersection of two lines, thus detecting restorations and caries.

Modern approaches to object detection in images utilize the faster region-based convolutional neural network (R-CNN) architecture, which is known for being fast and accurate [47, 51, 53]. This model is an extension of R-CNN and consists of three main parts. The first part is a convolutional neural network (CNN), which extracts features from the input image. The second part is a region proposition network (RPN) that generates object suggestions based on features extracted from the CNN. Finally, R-CNN uses the suggestions generated by RPN to classify and identify objects in the suggestions using a set of fully connected layers. The approach is related to the type of X-ray modality used and different convolutional architectures used to extract features [47]. In addition, the model's performance when using different neural networks, such as YOLOv3, ResNet, and Inception [51, 53], was also compared.

The study [36] used the VGG16 neural network, which is used as a base network (backbone) for object detection tasks, such as object detection in images or videos, having one of the advantages in obtaining a performance very interesting compared to more recent and complex architectures [66].

In addition, the study [50] used YOLOv3, which can detect, classify, and locate multiple objects in a single image. The model can detect objects in a single step instead of performing several steps as other models do. Finally, the study [58] presented the deep gradient–based learning neural network, which is an extension of the LeNet architecture using gradient optimization with deep learning, allowing the network to learn more optimally.

Strengths and Limitations of Detection Tasks

Detection has distinct strengths and weaknesses. It is ideal for real-time applications such as self-driving cars and security cameras because the approach does not have to process and identify individual pixels for each instance [67]. However, detection only provides a bounding box around the object, while segmentation provides a more precise region by highlighting the pixels corresponding to the region of interest.

Data annotation is crucial for detection models to train and compare results. The images must annotate the areas of interest, meaning that each object's location and category must be identified. The annotated images must have experienced professionals to perform this task, making the process expensive and time-consuming.

In addition, other limitations encountered in the caries detection process include the following:

- **Data availability**: It is often difficult to find enough annotated data to train an accurate detection model.
- **Variation in the appearance of the teeth**: The appearance of the teeth can change depending on the patient's age, oral health, and personal hygiene. Thus, the variation between the images makes it difficult to detect cavities.
- **Computational limitations**: Detection models may have difficulty detecting caries in uneven tooth surface areas, which may affect their accuracy.

Classification

Table 7 displays the different studies aimed at classifying image types. As evident from the observations, classification studies encompass a variety of computational techniques. Within each category, diverse methods and metrics are

Table 7 Works with the objective of classifying a table with information on the type of techniques, classification, and metrics

Index	Ref.	Techniques	Classification	Metrics
3	[22]	NN	Binary	Accuracy
4	[23]	AdaBoost; NB; SMO; decision tree; RF; decision stump; RBF; NN	Binary	Accuracy; precision; sensitivity, specificity; PPV
5	[24]	Inception V3	Binary	Accuracy; sensitivity; specificity; precision; RoC curve
6	[25]	Adaptive neural network	Binary	F1-score; precision; recall; specificity
8	[27]	CNN with graph-cut technique	Binary	Accuracy; F1-score; specificity; precision; recall
9	[28]	CNN-SVM	Binary	Accuracy; specificity; sensitivity
10	[29]	AlexNet	Binary	Accuracy; F1-score; precision; recall
11	[30]	BPNN; SVM; KNN; NB; Bagging; RF; XGBoost	Binary	AUC; F1-score; precision; recall; TP rate; FP rate; PRC
12	[31]	KNN; SVM	4 classes	Confusion matrix
13	[32]	VGG19; VGG16	Binary	AUC; accuracy; F1-score; precision
14	[33]	DenseNet-121	Caries only in enamel; caries in less than one-third of dentin; caries in more than one-third of dentin	Accuracy; F1-score; sensitivity; specificity; precision; NPV
15	[34]	U-Net	Initial; moderate; extensive	F1-score; sensitivity; precision
18	[37]	Inception; ResNet	Normal; incipient; advanced	AUC; precision; recall; specificity; RoC curve
19	[38]	AlexNet; GoogLeNet; VGG19; ResNet-50	Binary	Accuracy
20	[39]	MobileNet V2	Binary	AUC; confusion matrix; Roc curve
21	[40]	Hybrid neural network: Sparse AutoEncoder; logistic regression	Enamel; pulp; root injury	Accuracy
23	[42]	3-layer CNN; ResNet-18; ResNet-50	Binary	Confusion matrix
24	[43]	Network MI-DCNNE	Binary	AUC; confusion matrix
25	[44]	Multilayer perceptron	4 classes	AUC; confusion matrix; RoC curve
26	[45]	Multilayer perceptron	4 classes	Confusion matrix
28	[47]	Faster R-CNN with ResNet-50	Caries only in enamel; caries in less than one-third of dentin; caries in more than one-third of dentin	AUC; confusion matrix; RoC curve
29	[48]	VGG16; Inception V3; ResNet-50; DenseNet-121	Binary	Confusion matrix
30	[49]	ResNet-18	Binary	AUC; sensitivity; specificity; RoC curve
31	[50]	YOLOv3	ICMS: 2, 4, and 7 classes	F1-score; precision; recall
32	[51]	Faster R-CNN; ResNet; Inception	Coronal; proximal; cervical	Precision; sensitivity; specificity
33	[52]	CariesNet	High; moderate; shallow	Accuracy; F1-score; precision; recall
35	[54]	Modified ResNet-18	Binary	AUC; F1-score; sensitivity; specificity; precision; NPV; F1-score; RoC curve
36	[55]	ResNet-18; ResNet-50; ResNet-101; ResNet-152	ICMS: 4 and 7 classes	AUC, accuracy; sensitivity; specificity; RoC curve
38	[57]	Image processing techniques with ResNeXt-RNN	Binary	Accuracy; F1-score; sensitivity; precision
39	[58]	Deep gradient–based LeNet	Overall; premolar; molar	Precision; sensitivity; error rate
40	[59]	EfficientNet-B0; DenseNet-121; ResNet-50	Binary	AUC; confusion matrix; RoC curve
41	[60]	ResNet-50-DCDNet	Occlusal; proximal; cervical	Precision; sensitivity; specificity; F1-score

utilized. In the following subsection, each individual work is presented and discussed.

Classical Supervised Learning

Several studies have been conducted using classical machine learning approaches to classify X-ray images [23, 31, 44, 45]. In one of the studies [31], the authors employed the gray-level co-occurrence matrix (GLCM) to extract meaningful texture features from the gray levels and then used K-nearest neighbors (KNN) and support vector machine (SVM) algorithms to classify the images. Both studies [44, 45] used a multilayer perceptron for caries classification. The study [44] employed the GLCM to extract texture features from gray levels, while [23] used Hu moments to extract shape information from the images. In [23], the radon transformation and discrete cosine transformation were applied to extract patterns and capture global information about pixel intensities in different directions. They also reduced dimensionality using principal components from the original features and trained machine learning algorithms such as SVM, decision tree, and AdaBoost.

Studies [22, 25, 30] focused on image classification using neural networks (NNs). However, these approaches differ from deep learning, as conventional neural networks typically consist of a limited number of layers, whereas deep learning utilizes multiple layers. Another study [22] employed the GLCM to extract features from grayscale images and applied linearly adaptive particle swarm optimization in conjunction with an NN. This optimization technique specifically targeted the learning rate parameter. The study [25] utilized a different methodology: after preprocessing the images, they employed the morphological principal component analysis (MPCA) for feature extraction, and the classification task was then performed using an NN. The study conducted by [30] employed both GLCM and the gray-level difference method for feature extraction. The authors evaluated SVM, KNN, and NN to compare their performance, showing better accuracy using the NN approach. These works showcase the diverse approaches to utilizing NNs for caries classification tasks.

Deep Learning

The studies [27, 28, 40, 57] proposed a hybrid approach using a modified neural network or hybrid neural network. In [40], a hybrid neural network with a sparse encoder was proposed to differentiate lesions into types: present in the root, enamel, and dentin. The study [57] used a modified ResNeXt-RNN neural network to classify the images into carie and noncarie. A combination of different techniques with convolutional layers was performed in [27, 28]. In [28], the SVM algorithm was used and [27] combined the hybrid graph-cut segmentation technique, which divides the images into different regions and segments the various parts of the teeth.

In the studies [42, 59], the authors presented comparisons between different deep neural networks: ResNet-50, ResNet-18, DenseNet-121, and EfficientNet-B0 with binary classification in panoramic radiography. The studies [29, 32, 39] compared different convolutional architectures (AlexNet, VGG16, and MobileV2) for executing binary classification of caries in panoramic radiography.

Different classifications were used to evaluate lesions [33, 51, 52, 60]. In [33], segmentation was performed in the previous stage and was used as an input in lesion classification using DenseNet-121, classifying into three classes. In [52, 60], the studies presented a new architecture of deep neural networks based on modern implementations capable of differentiating several levels of caries. Finally, [51] used a neural network to locate the disease and evaluate the severity of caries in coronal, proximal, and cervical regions, obtaining better results in the Faster R-CNN approach.

Several studies have been conducted to compare or even introduce a new approach to classify caries [24, 43, 46, 48, 54]. The study [43] employed a multi-input CNN ensemble approach that incorporated two types of inputs: a raw image and an “image enhanced" version. The CNNs processed these two images, and the resulting outputs were combined using a score-based fusion technique to classify the images as either caries or noncaries. The study [54] developed a deep learning model using two cascaded ResNet-18 backbones capable of classifying caries and periapical periodontitis. Meanwhile, studies [24, 46, 48] compared deep neural network architectures such as ReNest101, LeNetV3, VGG16, ResNet-50, and DenseNet-121.

Multiple studies have explored lesion detection and classification [37, 47, 50, 55, 58]. These studies primarily employed deep neural networks to accurately detect and localize lesions by generating bounding boxes around them. Additionally, the detection process involved classifying the identified lesions often represented as a percentage indicating the confidence level of the classification.

In [38], alternative strategies to enhance the classification process were employed, using preprocessing techniques to separate the tooth from the radiographic images. Subsequently, the model was trained using individual images of the teeth. The results showed improved classification task accuracy. In [49], the authors explored the utilization of unlabeled data and applied a self-supervised learning technique to enhance the performance of deep learning models. The results obtained in this study demonstrated the benefits and gains achieved through the application of this technique. In addition to using of techniques to improve the classification, the study [34] used the U-Net neural network to differentiate lesions into initial, moderate, or extensive classifications.

Strengths and Limitations of Classification Tasks

The results in Table 7 did not show a consensus on using different classifications. Most studies used binary classification, and few proposed lesion identification at different stages, making it difficult to compare approaches.

Furthermore, the performance metrics proposed by computer approaches are often not easy to compare since the datasets are different. The lack of public image bases often impacts the comparisons of the methods; for example, extensive lesions are easier to diagnose, which can be extended to computational methods that offer better metrics only in extensive lesions. The use of a dataset can result in overestimating computer method performance, which can be a problem in a computer approach applied to real clinical situations. Therefore, it is crucial that studies use an image base representative of clinical practice and make the data publicly available to make comparisons more fairly and reliably.

Data balance is crucial for enhancing machine learning models. When there is an imbalance in the classes, the model tends to become biased toward the majority class, leading to possible underestimation or loss of precision in classifying the minority class. The most common method for dealing with class balance was using data augmentation techniques to avoid it.

During the analysis of the studies, the limitations identified were generally related to image availability. The main limitations were highlighted:

- **Small number of images**: The dataset used in the study was relatively small, which may limit the generalization of the results.
- **Limited to high-quality radiographic images**: The study only used high-quality radiographic images, which may not reflect real-world scenarios where lower-quality images are often encountered. This may limit the applicability of the model in clinical settings.
- **Lack of multiple examiners**: The study may have been limited because only one examiner was used to interpret the images. Having multiple examiners would increase the reliability of the results and ensure that they are not biased by a single person's interpretation.
- **Restricted to a single imaging device**: The study only used one imaging device, which may not represent the range of devices used in clinical practice. This may limit the generalization of the results and the model's applicability in different settings.

Image Characteristics and Evaluation

The distribution of studies among different radiographic modalities is quite proportional (Fig. 9). Each exam type presents a contribution of approximately 20%, except for CBCT. This can be explained because it is not recommended for the exclusive diagnosis of dental caries and is used with a complementary exam [68], with the smallest percentage being 2%.

The radiographic modalities used to diagnose caries have distinct advantages. Periapical radiographs offer a detailed visualization of the internal structure of a tooth, enabling the detection of caries that may not be apparent clinically. Interproximal X-rays allow for the early identification of cavities located between teeth. Panoramic X-rays provide a comprehensive overview of the teeth and neighboring structures, facilitating the simultaneous detection of cavities in multiple teeth. CBCT images offer three-dimensional images of teeth, bones, and adjacent structures, enabling a more precise and comprehensive assessment of caries extent and location [69].

Datasets

Over the years, deep learning techniques have grown significantly, as shown in Fig. 7. However, using these techniques requires a large quantity of data. Additionally, comparing approaches that utilize private data is difficult since datasets can have biases.

Good practice should make datasets available, with the correct treatment to make the data anonymous. It could provide more transparency in the results. For example, studies [22, 26, 57] used the same public data [70]. This later

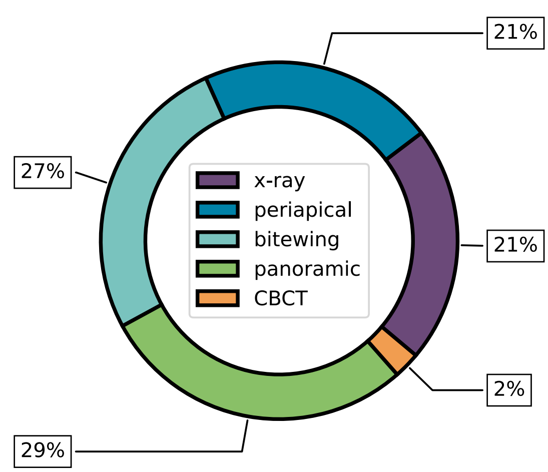


Fig. 9 Distribution of works according to imaging modalities

Table 8 Performance evaluation equations used to compare and analyze the performance of several different methods

Metric	Equation	Used in
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	Segmentation/detection/classification
Precision/PPV	$\frac{TP}{TP+FP}$	Segmentation/detection/classification
NPV	$\frac{TN}{TN+FN}$	Classification
Sensitivity/recall/TPR	$\frac{TP}{TP+FN}$	Segmentation/detection/classification
Specificity	$\frac{TN}{FP+TN}$	Segmentation/detection/classification
Error rate	$\frac{FP+FN}{n}$	Detection/classification
F1-score	$\frac{2*TP}{2*TP+FP+FN}$	Segmentation/detection/classification
Dice	$\frac{2\lvert S\cap G\rvert}{\lvert S\rvert+\lvert G\rvert}$	Segmentation
IoU/Jaccard	$\frac{\lvert S\cap G\rvert}{\lvert S\cup G\rvert}$	Segmentation/detection

study provided a set of 120 periapical radiograph images and a file with 152 teeth and their respective caries marked. Some images contain annotations, but no mask indicates the injured region. In contrast, an image base available in the International Symposium on BIOMEDICAL IMAGING 2015 competition [71] contains 120 bitewing radiographs. However, the link to access these images is unavailable. The lack of public image bases for different radiography modalities is evident and could be an important contribution to researchers in the area.

Metrics

As shown in Tables 5, 6, and 7, the metrics are different for each category of computational techniques used in the approaches. Segmentation, detection, and classification metrics are employed to evaluate the performance of computational methods in different tasks, such as predicting the input image class, locating and categorizing objects in an image, and segmenting an image into similar regions. Table 8 summarizes the main metrics cited in the included articles. In general, classification metrics use the relationship between true-positive values (TP), true-negative values (TN), false-negative values (FN), and false-positive values (FP). Segmentation and detection metrics used the relation where S corresponds to the manual segmentation by the expert and G corresponds to the segmentation by the method.

In addition to the metrics presented in Table 7, the confusion matrix is frequently used to evaluate the performance of approaches, showing the graphical relationship between classification metrics such as accuracy, recall, F1 score, sensitivity, and specificity. The area under the curve (AUC) is a metric that obtains an overall view of the algorithm's performance across all classes, which assesses the model's ability to discriminate between different probability thresholds of classification. In addition to the common metrics, the error rate is often used to evaluate the efficiency of the model used by work [58]. The study [30] used the precision-recall curve (PRC), which is a measure for evaluating classification performance, and compared the precision and recall, mainly in situations with class imbalance.

The studies conducted by [33, 41, 52, 56] were the only studies to propose evaluating segmentation using specific metrics such as the Dice similarity coefficient or intersection over union (IoU). Using segmentation metrics is essential to evaluating segmentation algorithm quality against an original segmentation. However, these metrics did not consider the relationship between the segmentation produced by the algorithm and the ground truth segmentation. In addition to pixel-based metrics, the main metrics used to evaluate segmentation performance, as detailed in Table 8, include accuracy, precision, recall, F1-score, and specificity. These later metrics are essential for assessing the degree of alignment between the segmentation results produced by the algorithm and the ground truth. However, these metrics do not quantify the region segmented by the computational method compared to the ground truth. Thus, specific segmentation metrics must be used to provide this detailed quantification in order to precisely locate topics where the algorithms should be improved.

To evaluate the performance of caries detection models, metrics such as precision, recall, F1-score, accuracy, and AUC are commonly used, similar to those used in classification models. However, a study [50] proposed using the IoU metric to evaluate the overlap between the regions demarcated by the model and the annotated area of a professional, providing a more accurate measure of performance.

Discussion

Upon analyzing the findings presented in the 42 articles, it becomes possible to provide comprehensive responses to the questions described in "Methods".

What Computer Methods Have Been Used to Diagnose Dental Caries?

The methods varied depending on the specific objectives of each technique. Different methods were designed to extract features and perform classification tasks. Additionally, there were approaches dedicated to segmenting regions of interest, while another set of methods focused on detecting and returning bonding regions

In classification tasks, supervised learning using machine learning algorithms was commonly employed, first extracting features from images and then classifying them. One prevalent method for feature extraction used the GLCM [22, 30, 31, 44], which captures features at different pixel intensity levels and combines them with machine learning algorithms. In the deep learning domain, several popular architectures, such as DenseNet, Inception, VGG, and ResNet, have been extensively explored. Comparative studies often focused on evaluating the performance of these architectures [29, 32, 39, 42, 59]. Additionally, some studies proposed hybrid architectures or modifications to existing neural networks to improve classification accuracy [27, 28, 40, 57]. The studies [33, 38] suggested that separating the tooth region from the original image produces better results compared to using the raw image.

In segmentation tasks, the U-Net architecture appeared to be the explored deep learning approach. Many studies utilized this deep learning architecture due to its capacity to handle images of different sizes and its effectiveness in scenarios with limited training samples. Furthermore, some works have extended the U-Net format by incorporating additional neural networks into the backbone structure [52, 56]. This modification allows the segmentation of injured regions across different radiographic modalities, enhancing the versatility and applicability of the segmentation models. In the studies that employ image processing techniques for segmentation, a wide range of techniques are utilized to extract and isolate the regions of interest. Typically, these studies began by enhancing the image quality and applying a thresholding technique to distinguish the target region from the background [20, 26, 27]. Therefore, different techniques were employed to separate and isolate the specific regions of interest, allowing further analysis or processing. The choice of techniques may change depending on the characteristics of the image and the specific objectives of the study.

Finally, in the detection process, deep learning architectures such as Faster R-CNN [47, 51, 53], and neural networks like YOLO [50, 53], have shown promising approaches. These methods are capable of detecting multiple objects simultaneously and defining regions of interest by surrounding them with bounding boxes. In many research studies, these techniques are commonly used to generate bounding boxes around detected objects, accompanied by a percentage indicating the confidence or probability of belonging to a specific class. The wide range of available detection approaches offers versatility and adaptability to different contexts.

What Is the Influence of the Approaches and Methods on Diagnostic Performance?

In studies employing bitewing radiograph images, diverse computational methods were compared with human evaluations. The study [47] assessed the effectiveness of a CNN in detecting proximal caries, comparing it with the performance of graduate students. The study revealed a significant difference in sensitivity between the computational model and the students ($p < 0.05$), particularly in early lesions, where the model demonstrated superior sensitivity. In [36], the authors compared models with five experienced observers on a dataset. The results show that computational models outperformed the experts. The study [34] analyzed the influence of a CNN model on the diagnostic sensitivity of dentists examining bitewing radiographs. The results show that dentists using this model, the sensitivity was significantly increased for identifying early and moderate caries ($p < 0.05$). However, the model also increased the number of false-positive rates. These studies demonstrated that computational methods can enhance or, at minimum, maintain the performance of health professionals. Yet, a crucial consideration, as noted in [34], is the inherent trade-off in these models. While there is a significant increase in sensitivity for detecting caries, there is also a corresponding rise in false-positive rates.

In the context of panoramic radiographs, the study [33] compared the performance of a computational method with the performance of experienced dentists, finding similar results in terms of precision and recall (0.986 and 0.821, respectively) between the computational approach and dentists. In [35], 24 dentists assessed 30 CBCT scans, half with the assistance of the system and the other half without. A statistically significant difference was observed between the groups ($p = 0.032$), with the aid system demonstrating improved accuracy in classifying different dental diseases. However, no significant difference was noted in caries detection ($p > 0.05$), indicating the challenges in diagnosing caries with CBCT, even when using computational methods.

In summary, computational methods can potentially enhance diagnostic performance in dentistry, but significant limitations remain to address. The balance between increased sensitivity and the possibility of false positives, coupled with the need for more research involving a considerable number of professionals, is crucial in assessing the real effectiveness of these systems in clinical applications. Thus, while computational methods show promise, a definitive conclusion regarding notable improvement across all dental imaging modalities cannot yet be reached.

What Are the Limitations of the Approaches?

The limitations of the approaches discussed in this SR vary according to the specific tasks, as detailed in sections "Strengths and Limitations of Segmentation Tasks", "Strengths and Limitations of Detection Tasks", and "Strengths and Limitations of Classification Tasks". We have identified and highlighted several common limitations across different methodologies, which are presented following

A notable limitation is the relatively small dataset of images used in many studies. This small sample size can significantly limit the generalizability of the results, as models trained on limited data may perform less well in diverse or unexpected clinical scenarios. Additionally, the reliance on only high-quality images in the study presents another limitation, potentially reducing the applicability of the findings in real-world clinical settings where image quality can vary greatly.

Furthermore, the testing of models on images acquired with a single type of imaging device, which occurs in most of the included studies, constrains the broad applicability of the results. This limitation is crucial as different imaging devices may produce images with varying qualities and characteristics, affecting the model's performance in diverse clinical environments.

Finally, a critical limitation is the image's resolution in diagnosing minor caries. The current approaches may not adequately account for the spatial nuances in images, which can be essential for accurately identifying and diagnosing smaller carious lesions. This inadequacy in resolution can lead to missed diagnoses or misinterpretations, especially in minor caries requiring a high level of detail and precision for accurate identification.

What Are the Strengths of the Approaches?

The strengths of the approaches discussed in this SR are as varied as the tasks themselves, as detailed in sections "Strengths and Limitations of Segmentation Tasks", "Strengths and Limitations of Detection Tasks", and "Strengths and Limitations of Classification Tasks".

The use of IP has three main strengths for caries detection and segmentation. The first is a shorter processing time, an essential feature of IP, which does not require a training step like methods such as CSL and DL. Second, preprocessing can improve image quality, an essential aspect of enhancing complex images, which usually are not performed in methods of other categories. A third advantage is working effectively with limited data samples, an essential factor in the available data. For affected regions, Otsu thresholding and median filtering techniques are often used, where gray-level features of the image are used to detect caries. In addition, PSO optimization is used to accurately detect cavities and dental fillings, which improves the IP in caries diagnosis.

In caries classification, methods using CSL have some strengths. The range of available libraries and inductive algorithms provides a lot of options for choosing the most appropriate strategy for each image type and extent of caries. Furthermore, CSL algorithms can adjust to different data, making them useful tools in caries classification. Finally, an advantage of using CSL is their efficiency with smaller data sets, differentiating them from DL, which usually requires large amounts of data for training.

There are some strengths in using DL in classifying, segmenting, and detecting caries. The first advantage is the capacity to learn patterns without prior knowledge of the data. The second is that DL tends to outperform standard methods in complex tasks because it can analyze and learn from datasets when the amount of images is diverse and provided on a large scale. The third strength is DL's adaptability and flexibility, which enable application in different clinical and diagnostic contexts. In detection, DL architectures such as Faster R-CNN and YOLO can identify regions of interest to provide fast insights even in complex imaging scenarios. In segmentation, architectures such as U-Net are very useful with limited training samples, segmenting the injured region. In classification, architectures such as DenseNet, Inception, VGG, and ResNet can extract features from dental images and offer the capacity of distinguish levels and, with some adaptations, could explain the results using Grad-CAM.

What Types of Imaging Modalities Were Used?

Figure 9 shows the distribution of studies across different radiographic modalities. The most common choice was panoramic radiography, which has wide application in surgical procedures and aids in caries detection. In contrast, there were relatively few studies utilizing CBCT due to the challenges associated with obtaining images from this type of tomography.

The field of segmentation, classification, and detection in 3D models presents distinct challenges due to its reliance on a series of interconnected 2D images. Notably, panoramic radiography stands out as a valuable approach, as it offers a comprehensive view of all teeth, enabling enhanced analysis and detection capabilities.

What Types of Caries Classification Were Considered?

Analyzing Table 7 reveals no standardization in the classification used, varying according to each study. Some studies proposed classifying caries into different levels of severity, such as initial, moderate, and extensive. It is important to highlight that caries can present a wide variety of lesions, making comparing results problematic and reinforcing the importance of a consistent classification.

A significant contribution to dental caries classification was presented in the works of [50, 55], the authors proposed using the International Caries Classification and Management System, internationally recognized by dentists. Using recognized classification in dentistry can facilitate radiographic image annotation, providing a consensus classification among studies and helping to improve disease diagnosis.

What Were the Main Evaluation Metrics Used?

The metrics employed in each study can vary significantly, but some commonly used metrics include AUC, F1-score, precision, recall, sensitivity, and accuracy. These metrics are often utilized in works involving binary classification or other types of classification tasks. They provide valuable insights into the performance and effectiveness of the classification models being evaluated. Some works have proposed using classic techniques such as Dice and IoU in the segmentation stage. Finally, in the detection stage, the metrics used were presented in values such as the confusion matrix and IoU (Tables 5, 6, and 7).

However, in regard to assessing model performance in classification tasks, there is a lack of consensus on the most suitable metrics. This lack of consensus makes it challenging to compare results and make advancements in the literature regarding this particular application.

What Were the Main Image Datasets Used?

As highlighted in "Datasets", most research into dental caries diagnosis uses particular data sets. This approach has its limitations, especially when we consider the accessibility and diversity of this data. In this section, we will discuss in more detail the characteristics of public datasets available for caries diagnosis research.

The Periapical Dental X-ray dataset [70] contains a collection of 120 digital periapical radiographs. Dental experts from University Teknologi Malaysia's Dental Clinic evaluated and annotated this dataset. The dataset does not provide specific segmentation or identification of regions of interest. It contains only images that show the presence of caries. The availability of such specialized information is crucial for applying computational methods, especially for automatic diagnosis and image analysis. The lack of detailed segmentation or explicit dataset identification limits its immediate usefulness in advanced computing techniques, requiring further processing or annotation for computer methods.

The [71] study was limited in its availability because it was conducted as a competition in 2015, and data access has been restricted since then. The dataset included 400 cephalometric radiographs collected from different groups of patients aged 6 to 60 years and 120 bitewing radiographs, which included diversity in age and provided a wide scope for algorithm testing.

The SR indicates an increase in the number of studies that have used computational methods for diagnosing dental caries in recent years. With the progress of deep learning, obtaining more efficient and accurate diagnostic aid systems is possible. However, challenges can be overcome, such as lesion appearance variation, difficulties obtaining high-quality images, and lack of consensus on lesion classification. Therefore, a joint effort is needed to improve techniques and standardize study methodologies.

Open Gaps and Research Possibilities

There are several open gaps and research possibilities in dental classification, detection, and segmentation.

1. **Public databases**: An opportunity for research and advancement in the literature is the availability of public datasets with precise annotations to assist in comparisons of classification, detection, and segmentation tasks of dental caries. Adopting an international classification is fundamental to ensure standardization in the classification of methods and to assure that the methods that use this set of public data have a good generalization in caries of different levels of severity. This effort can significantly contribute to developing more effective and accurate computational methods for diagnosing dental caries.
2. **Exploration of the use of CBCT**: CBCT is a relatively new X-ray modality in dentistry, and its use for diagnosing dental caries has been little explored. One of the reasons for this may be the complexity of analyzing the data generated by the three-dimensional images obtained with CBCT. Furthermore, interpreting these images requires specific skills and advanced knowledge of medical image analysis. Therefore, considering using CBCT to diagnose caries presents a challenging opportunity to combine one of the techniques of this 3D modality.

3. **Standardization of classification, segmentation, and detection metrics**: One of the major difficulties in diagnosing caries is the standardization of classification, segmentation, and detection indicators. Since there are multiple techniques and algorithms for performing these tasks, it is important to define concise and well-defined metrics to assess the effectiveness and accuracy of these methods. The lack of standardization can make studies difficult to compare and limit the generalizability of results. Therefore, establishing clear criteria for evaluating the quality and efficacy of algorithms used for caries diagnosis is important to ensure the reliability and safety of analyses by experts in the field.
4. **Hybrid methods using more than one type of image**: This SR did not find an approach to diagnosing caries by combining different types of radiographs. Multiple types of X-rays, such as panoramic and CBCT, are often necessary to obtain a complete view of the region of interest. Using a hybrid approach that combines different types of radiographs could be an interesting research opportunity to assess the diagnostic efficacy of caries. Therefore, studying this approach can contribute to developing new diagnostic techniques and improving the accuracy of the caries diagnostic process.

Final Remarks

Despite the contributions provided in this work, our method is limited, as it did not address methods that use approaches with nonradiographic images, such as optical coherence tomography images and laser fluorescence images. Although these technologies have the potential to offer greater sensitivity and specificity for detecting early lesions, they do not offer a field of view for the diagnosis of interproximal carie [5].

The contribution of this SR was separated into items which can be seen below.

- **Computational methods**: The SR identified studies that used several computational techniques to classify, detect, and segment dental caries, such as deep learning, machine learning, and image processing.
- **Radiographic modalities**: During the analysis of the SR, we evaluated and discussed studies that used different radiographic modalities, including panoramic radiography, cone beam computed tomography (CBCT), and periapical radiography, for the purpose of classifying, segmenting, and detecting dental caries.
- **Datasets**: The SR identified the datasets most commonly used in studies focusing on the diagnosis of dental caries. Additionally, the SR highlighted potential research opportunities that could provide valuable guidance for future investigations.
- **Commonly used metrics**: The SR metrics are frequently used to assess the effectiveness of methods in the diagnosis of dental caries, such as sensitivity, specificity, precision, AUC, ROC curve, Dice, and IoU. Identifying these metrics can standardize the evaluation of the methods and simplify the comparison between different studies.

The SR identified studies that explored the diagnoses of dental caries across a wide range of lesion types and severities, spanning from initial caries to advanced lesions. This information enables the development of specific approaches for each type of lesion and severity, thereby enhancing the accuracy of diagnosis.

Approximately 12% of the analyzed studies used public datasets while using deep learning accounted for 69% of the total works. The majority of these studies aimed to classify dental caries, either in binary or multiclass classification, comprising 76% of the total. Among the radiographic modalities, panoramic imaging was the most commonly used, representing 29% of cases, whereas CBCT had the lowest representation, with only 2%. The dataset sizes varied greatly, ranging from a total of 38,000 images to as few as 80 images. This wide fluctuation in the amount of data used reflects the diversity and complexity of machine learning studies applied to dentistry.

In summary, the SR carried out systematized the main methods used in the diagnosis of dental caries, presenting specific objectives, techniques, X-ray modalities, datasets used, limitations, and research possibilities. The review allowed an overview of existing knowledge and consolidated the available information to guide future research.

Funding This work is supported in part by the Brazilian National Council for Scientific and Technological Development (CNPq grant number 307710/2022-0), in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, Brazil, and Itaú Unibanco S.A. through the PBI program of the Centro de Ciência de Dados (C2D) of Escola Politécnica at Universidade de São Paulo, and São Paulo Research Foundation (FAPESP) – National Institute of Science and Technology – Medicine Assisted by Scientific Computing (INCT-MACC) – grant 2014/50889-7.

Data Availability The data in this systematic review, like the search protocol, are available upon reasonable request to the corresponding author.

References

1. Rathee, M. & Sapra, A. in *Dental Caries* (StatPearls Publishing, Treasure Island (FL), 2023).
2. Pan, Y.-C., Chan, H.-L., Kong, X., Hadjiiski, L. M. & Kripfgans, O. D. Multi-class deep learning segmentation and automated measurements in periodontal sonograms of a porcine model. *Dentomaxillofacial Radiology* **51**, 20210363 (2022).

3. Rana, A. *et al.* Automated segmentation of gingival diseases from oral images. *IEEE Healthcare Innovations* 144–147 (2017).
4. Rimi, I. F. *et al.* Machine learning techniques for dental disease prediction. *Iran J Comput Sci* **5**, 187–195 (2022).
5. Felemban, O. M., Loo, C. Y. & Ramesh, A. Accuracy of cone-beam computed tomography and extraoral bitewings compared to intraoral bitewings in detection of interproximal caries. *J Contemp Dent Pract* **21**, 1361–1367 (2020).
6. Musri, N., Christie, B., Ichwan, S. J. A. & Cahyanto, A. Deep learning convolutional neural network algorithms for the early detection and diagnosis of dental caries on periapical radiographs: A systematic review. *Imaging Sci Dent* **51**, 237–242 (2021).
7. Prados-Privado, M., García Villalón, J., Martínez-Martínez, C. H., Ivorra, C. & Prados-Frutos, J. C. Dental Caries Diagnosis and Detection Using Neural Networks: A Systematic Review. *J Clin Med* **9**, 3579 (2020).
8. Mohammad-Rahimi, H. *et al.* Deep learning for caries detection: A systematic review. *J Dent* **122**, 104115 (2022).
9. Reyes, L. T., Knorst, J. K., Ortiz, F. R. & Ardenghi, T. M. Machine Learning in the Diagnosis and Prognostic Prediction of Dental Caries: A Systematic Review. *Caries Res* **56**, 161–170 (2022).
10. Morris, A. L. & Tadi, P. in *Anatomy, Head and Neck, Teeth* (StatPearls Publishing, Treasure Island (FL), 2022).
11. Marsh, P. D. Dental plaque as a microbial biofilm. *Caries Res* **38**, 204–211 (2004).
12. Featherstone, J. D. The science and practice of caries prevention. *J Am Dent Assoc* **131**, 887–899 (2000).
13. Braga, M. M., Mendes, F. M. & Ekstrand, K. R. Detection Activity Assessment and Diagnosis of Dental Caries Lesions. *Dental Clinics* **54**, 479–493 (2010).
14. White, S. C. & Pharoah, M. J. *Oral radiology: Principles and interpretation* (Elsevier Health Sciences, 2014).
15. Setzer, F. C., Hinckley, N., Kohli, M. R. & Karabucak, B. A Survey of Cone-beam Computed Tomographic Use among Endodontic Practitioners in the United States. *J Endod* **43**, 699–704 (2017).
16. Wang, S. & Ford, B. Imaging in Oral and Maxillofacial Surgery. *Dental Clinics* **65**, 487–507 (2021).
17. Ikeuchi, K. *Computer vision: A reference guide* (Springer, 2021).
18. Guo, L. & Wenyuan, S. Salivary biomarkers for caries risk assessment. *Journal of the California Dental Association* **41**, 107–118 (2013).
19. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* n71 (2021).
20. Bhan, A., Goyal, A., Harsh, Chauhan, N. & Wang, C.-W. Feature Line Profile Based Automatic Detection of Dental Caries in Bitewing Radiography. *ICMETE* 635–640 (2016).
21. Naebi, M. *et al.* Detection of Carious Lesions and Restorations Using Particle Swarm Optimization Algorithm. *Int J Dent* **2016**, 3264545 (2016).
22. Sornam, M. & Prabhakaran, M. A new linear adaptive swarm intelligence approach using back propagation neural network for dental caries classification. *ICPCSI* 2698–2703 (2017).
23. Singh, P. & Sehgal, P. Automated caries detection based on Radon transformation and DCT. *ICCCNT* 1–6 (2017).
24. Lee, J.-H., Kim, D.-H., Jeong, S.-N. & Choi, S.-H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent* **77**, 106–111 (2018).
25. Patil, S., Kulkarni, V. & Bhise, A. Algorithmic analysis for dental caries detection using an adaptive neural network architecture. *Heliyon* **5**, e01579 (2019).
26. Datta, S., Chaki, N. & Modak, B. A Novel Technique to Detect Caries Lesion Using Isophote Concepts. *IRBM* **40**, 174–182 (2019).
27. Al Kheraif, A. A., Wahba, A. A. & Fouad, H. Detection of dental diseases from radiographic 2d dental image using hybrid graph-cut technique and convolutional neural network. *Measurement* **146**, 333–342 (2019).
28. Verma, D., Puri, S., Prabhu, S. & Smriti, K. Anomaly detection in panoramic dental x-rays using a hybrid Deep Learning and Machine Learning approach. *TENCON* 263–268 (2020).
29. Lakshmi, M. M. & Chitra, P. Classification of Dental Cavities from X-ray images using Deep CNN algorithm. *ICOEI* 774–779 (2020).
30. Geetha, V., Aprameya, K. S. & Hinduja, D. M. Dental caries diagnosis in digital radiographs using back-propagation neural network. *Health Inf Sci Syst* **8**, 8 (2020).
31. Jusman, Y. *et al.* Comparison of Dental Caries Level Images Classification Performance using KNN and SVM Methods 167–172 (2021).
32. Choudhary, A. *et al.* An Effective Approach for Classification of Dental Caries using Convolutional Neural Networks. *SMART* 204–209 (2021).
33. Lian, L., Zhu, T., Zhu, F. & Zhu, H. Deep Learning for Caries Detection and Classification. *Diagnostics (Basel)* **11**, 1672 (2021).
34. Lee, S. *et al.* Deep learning for early dental caries detection in bitewing radiographs. *Sci Rep* **11**, 16807 (2021).
35. Ezhov, M. *et al.* Clinically applicable artificial intelligence system for dental diagnosis with CBCT. *Sci Rep* **11**, 15006 (2021).
36. Bayrakdar, I. S. *et al.* Deep-learning approach for caries detection and segmentation on dental bitewing radiographs. *Oral Radiol* **38**, 468–479 (2022).
37. Moran, M. *et al.* Classification of Approximal Caries in Bitewing Radiographs Using Convolutional Neural Networks. *Sensors (Basel)* **21**, 5192 (2021).
38. Mao, Y.-C. *et al.* Caries and Restoration Detection Using Bitewing Film Based on Transfer Learning with CNNs. *Sensors* **21**, 4613 (2021).
39. Vinayahalingam, S. *et al.* Classification of caries in third molars on panoramic radiographs using deep learning. *Sci Rep* **11**, 12609 (2021).
40. Megalan Leo, L. & Kalapalatha Reddy, T. Learning compact and discriminative hybrid neural network for dental caries classification. *Microprocessors and Microsystems* **82**, 103836 (2021).
41. Khan, H. A. *et al.* Automated feature detection in dental periapical radiographs by using deep learning. *Oral Surgery, Oral Medicine, Oral Pathology* **131**, 711–720 (2021).
42. Evaluation of Convolutional Neural Network for Automatic Caries Detection in Digital Radiograph Panoramic on Small Dataset. *ICoDSE.*
43. Imak, A. *et al.* Dental Caries Detection Using Score-Based Multi-Input Deep Convolutional Neural Network. *IEEE Access* **10**, 18320–18329 (2022).
44. Jusman, Y., Widyaningrum, A. & Puspita, S. Algorithm of Caries Level Image Classification Using Multilayer Perceptron Based Texture Features. *CyberneticsCom* 168–173 (2022).
45. Jusman, Y., Widyaningrum, A., Tyassari, W., Puspita, S. & Saleh, E. Classification of Caries X-Ray Images using Multilayer Perceptron Models Based Shape Features. *ICITDA* 1–6 (2022).
46. Jayasinghe, H. *et al.* Effectiveness of Using Radiology Images and Mask R-CNN for Stomatology. *ICAC* 60–65 (2022).
47. Chen, X., Guo, J., Ye, J., Zhang, M. & Liang, Y. Detection of Proximal Caries Lesions on Bitewing Radiographs Using Deep Learning Method. *Caries Res* **56**, 455–463 (2022).
48. Liu, F. *et al.* Recognition of Digital Dental X-ray Images Using a Convolutional Neural Network. *J Digit Imaging* **36**, 73–79 (2023).
49. Taleb, A. *et al.* Self-Supervised Learning Methods for Label-Efficient Dental Caries Classification. *Diagnostics (Basel)* **12**, 1237 (2022).
50. Panyarak, W., Suttapak, W., Wantanajittikul, K., Charuakkra, A. & Prapayasatok, S. Assessment of YOLOv3 for caries detection in bitewing radiographs based on the ICCMS™radiographic scoring system. *Clin Oral Investig* (2022).
51. Kim, C., Jeong, H., Park, W. & Kim, D. Tooth-Related Disease Detection System Based on Panoramic Images and Optimization

Through Automation: Development Study. *JMIR Med Inform* **10**, e38640 (2022).
52. Zhu, H. *et al.* CariesNet: a deep learning approach for segmentation of multi-stage caries lesion from oral panoramic X-ray image. *Neural Comput Appl* 1–9 (2022).
53. Zhu, Y. *et al.* Faster-RCNN based intelligent detection and localization of dental caries. *Displays* **74**, 102201 (2022).
54. Li, S. *et al.* Artificial intelligence for caries and periapical periodontitis detection. *Journal of Dentistry* **122**, 104107 (2022).
55. Panyarak, W., Wantanajittikul, K., Suttapak, W., Charuakkra, A. & Prapayasatok, S. Feasibility of deep learning for dental caries classification in bitewing radiographs based on the ICCMS™radiographic scoring system. *Oral Surg Oral Med Oral Pathol Oral Radiol* **135**, 272–281 (2023).
56. Ying, S., Wang, B., Zhu, H., Liu, W. & Huang, F. Caries segmentation on tooth X-ray images with a deep network. *Journal of Dentistry* **119**, 104076 (2022).
57. Ramana Kumari, A., Nagaraja Rao, S. & Ramana Reddy, P. Design of hybrid dental caries segmentation and caries detection with meta-heuristic-based ResneXt-RNN. *Biomedical Signal Processing and Control* **78**, 103961 (2022).
58. Vimalarani, G. & Ramachandraiah, U. Automatic diagnosis and detection of dental caries in bitewing radiographs using pervasive deep gradient based LeNet classifier model. *Microprocessors and Microsystems* **94**, 104654 (2022).
59. Oztekin, F. *et al.* An Explainable Deep Learning Model to Prediction Dental Caries Using Panoramic Radiograph Images. *Diagnostics* **13**, 226 (2023).
60. Dayı, B., Üzen, H., Çiçek, p. B. & Duman, U. B. A Novel Deep Learning-Based Approach for Segmentation of Different Type Caries Lesions on Panoramic Radiographs. *Diagnostics (Basel)* **13**, 202 (2023).
61. Baydar, O., Różyło-Kalinowska, I., Futyma-Gabka, K. & Sağlam, H. The U-Net Approaches to Evaluation of Dental Bite-Wing Radiographs: An Artificial Intelligence Study. *Diagnostics* **13**, 453 (2023).
62. Yu, Y. *et al.* Techniques and challenges of image segmentation: A review. *Electronics* **12** (2023).
63. Sultana, F., Sufian, A. & Dutta, P. Evolution of Image Segmentation using Deep Convolutional Neural Network: A Survey. *Knowledge-Based Systems* **201-202**, 106062 (2020).
64. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation 234–241 (2015).
65. He, K., Zhang, X., Ren, S. & Sun, J. *Deep residual learning for image recognition*, 770–778 (2016).
66. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).
67. Sanil, N., venkat, P. A. N., Rakesh, V., Mallapur, R. & Ahmed, M. R. Deep Learning Techniques for Obstacle Detection and Avoidance in Driverless Cars 1–4 (2020).
68. Scarfe, W. C., Farman, A. G., Levin, M. D. & Gane, D. Essentials of Maxillofacial Cone Beam Computed Tomography. *Alpha Omegan* **103**, 62–67 (2010).
69. Wenzel, A. Radiographic modalities for diagnosis of caries in a historical perspective: from film to machine-intelligence supported systems. *Dentomaxillofacial Radiology* **50**, 20210010 (2021). PMID: 33661697.
70. Rad, A. E., Rahim, M. S., Rehman, A. & Saba, T. Digital Dental X-ray Database for Caries Screening. *3D Res.* **7**, 96:1–96:5 (2016).
71. Wang, C.-W. *et al.* A benchmark for comparison of dental radiography analysis algorithms. *Medical Image Analysis* **31**, 63–76 (2016).