

Convolutional neural networks for automated tooth numbering on panoramic radiographs: A scoping review

Ramadhan Hardani Putra^{1,*}, Eha Renwi Astuti¹, Aga Satria Nurrachman¹, Dina Karimah Putri^{1,2}, Ahmad Badruddin Ghazali³, Tjio Andrinanti Pradini⁴, Dhinda Tiara Prabaningtyas⁴

¹Department of Dentomaxillofacial Radiology, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

²Division of Dental Informatics and Radiology, Tohoku University Graduate School of Dentistry, Sendai, Japan

³Oral Radiology Unit, Department of Oral Maxillofacial Surgery and Oral Diagnosis, Kulliyah of Dentistry, International Islamic University Malaysia, Malaysia

⁴Undergraduate Program, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

ABSTRACT

Purpose: The objective of this scoping review was to investigate the applicability and performance of various convolutional neural network (CNN) models in tooth numbering on panoramic radiographs, achieved through classification, detection, and segmentation tasks.

Materials and Methods: An online search was performed of the PubMed, Science Direct, and Scopus databases. Based on the selection process, 12 studies were included in this review.

Results: Eleven studies utilized a CNN model for detection tasks, 5 for classification tasks, and 3 for segmentation tasks in the context of tooth numbering on panoramic radiographs. Most of these studies revealed high performance of various CNN models in automating tooth numbering. However, several studies also highlighted limitations of CNNs, such as the presence of false positives and false negatives in identifying decayed teeth, teeth with crown prosthetics, teeth adjacent to edentulous areas, dental implants, root remnants, wisdom teeth, and root canal-treated teeth. These limitations can be overcome by ensuring both the quality and quantity of datasets, as well as optimizing the CNN architecture.

Conclusion: CNNs have demonstrated high performance in automated tooth numbering on panoramic radiographs. Future development of CNN-based models for this purpose should also consider different stages of dentition, such as the primary and mixed dentition stages, as well as the presence of various tooth conditions. Ultimately, an optimized CNN architecture can serve as the foundation for an automated tooth numbering system and for further artificial intelligence research on panoramic radiographs for a variety of purposes. (*Imaging Sci Dent* 2023; 53: 271-81)

KEY WORDS: Artificial Intelligence; Technology Transfer; Deep Learning; Dentition; Radiography, Panoramic

Introduction

In the field of dentistry, panoramic radiography is a standard examination procedure used in clinical practice to capture an image of the dental and maxillofacial region. This technique is instrumental in making diagnoses, planning

treatments, and evaluating treatment outcomes. Panoramic radiography allows the assessment of various factors, including the condition of the teeth, the presence of lesions and trauma, the structure of the jawbone, the status of edentulous patients, and the location of the third molar.^{1,2} One advantage of panoramic techniques is that they provide an image of the maxillary and mandibular regions with a relatively low radiation dose. Specifically, the radiation dose a patient receives from 1 panoramic radiograph is equivalent to the dose from 4 intraoral radiographs.³

Prior to creating a radiodiagnostic report based on panoramic radiography, tooth numbering serves as a guide for

Received March 8, 2023; Revised June 14, 2023; Accepted July 14, 2023

Published online September 4, 2023

*Correspondence to : Prof. Ramadhan Hardani Putra

Department of Dentomaxillofacial Radiology, Faculty of Dental Medicine, Universitas Airlangga, Jalan Prof. Dr. Moestopo No. 47, Surabaya 60132, East Java, Indonesia

Tel) 62-31-503-0255, E-mail) ramadhan.hardani@fkg.unair.ac.id

Copyright © 2023 by Korean Academy of Oral and Maxillofacial Radiology

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Imaging Science in Dentistry · pISSN 2233-7822 eISSN 2233-7830

dental and periodontal charting, based on tooth anatomy and location. A numbering system is employed in dental radiological reports, particularly those involving panoramic radiographs.⁴ Various tooth numbering methods are routinely implemented in dental practice, with popular systems including the Fédération Dentaire Internationale (FDI) system, the Universal Numbering System, and the Zsigmondy-Palmer system.⁵ The FDI system is the most commonly utilized and is widely used in diagnostic reports. Tooth numbering is also essential for interpreting radiographs, recording patient dental medical history, and performing forensic tasks.⁶ As an important diagnostic imaging tool, panoramic radiographs are taken in large numbers every day. However, the interpretation of these radiographs can be subjective, relying heavily on the dentist's experience and knowledge. This can potentially lead to misinterpretation, especially in the context of a heavy daily workload.^{7,8} The application of an artificial intelligence (AI)-based radiographic diagnostic tool is anticipated to address these issues by reducing errors, shortening the overall treatment duration, and enhancing the quality of patient dental care.⁹

Over the past decade, convolutional neural networks (CNNs) have garnered considerable attention due to their high performance in image recognition and computer vision.¹⁰ In 2012, a CNN outperformed traditional machine learning techniques in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC).¹¹ Essentially, a CNN employs a deep learning approach to independently determine the most effective features for image representation. A CNN comprises several layers, including convolution, pooling, nonlinearity, and fully connected layers. These allow the CNN to perform feature extraction and selection for classification within these hidden layers, thereby automatically discerning the relevant features of an image.¹² Due to these advantages, numerous studies have sought to apply CNNs in the realm of digital radiography.¹³ To recognize an object on a digital radiograph, a CNN performs 3 primary tasks: classification, detection, and segmentation. Classification involves categorizing or labeling an image into a specific class. Detection facilitates the identification of the location of lesions, organs, or other objects of interest. Finally, segmentation is employed to delineate the precise pixel-wise boundaries of an organ or pathological feature.^{12,14} In dentomaxillofacial radiology, CNNs have been utilized for a variety of purposes, including the identification and recognition of dental caries, periodontal bone loss, intraosseous lesions, dental implants, and osteoporosis.¹⁴⁻¹⁶

One potential application of this technology is in the num-

bering of teeth on panoramic radiographs. The task of tooth numbering presents a challenge, as the CNN must learn to identify and classify multiple objects, in this case teeth, within a single radiographic image. CNN-based architectures have demonstrated the capacity to recognize and classify types of teeth based on dental anatomy and location.^{17,18} The use of an automatic CNN approach for tooth numbering can reduce the time spent by dentists on identifying teeth shown on panoramic radiographs. The automation of tooth numbering is crucial, as the system must be capable of identifying the number of the tooth for subsequent analysis during the initial stage of interpretation. For instance, this functionality is important in identifying and diagnosing radiopathology in a specific tooth. The findings from these workflows are then consolidated in the diagnostic report or dental filling chart of the panoramic radiographs.

This scoping review was designed to gather and examine published studies regarding the performance of various CNN architectures in the context of tooth numbering systems for panoramic radiographs. The review incorporated studies employing a structured exploration to address a central question: how can CNNs be utilized to automate tooth numbering on panoramic radiographs? The applicability, performance, and limitations of the various CNN architectures reported in existing studies were analyzed and discussed. Ultimately, the findings of this review may serve as a basis for future advancements in AI research, with the aim of enhancing the overall oral health care system.

Materials and Methods

Search strategy

This study was designed as a scoping review, which is used to map the literature on a specific topic or research area. This approach provides an opportunity to identify key concepts, detect gaps in the research, and pinpoint sources of evidence.^{19,20} A systematic literature search was performed for studies published between January 2012 and September 2022, with the process conducted by 2 reviewers. The publications were limited to those published in 2012 or later because CNNs have demonstrated significant performance improvements since that time.¹¹ A systematic search was conducted of 3 databases: PubMed, Science Direct, and Scopus. The article search strategy employed Boolean operators, including AND, OR, and NOT. As each database has a unique procedure for article searches, the search terms were adjusted in accordance with the guidelines of each respective database (Table 1).

Table 1. Article search

Database	Keywords
PubMed	“panoramic radiograph” AND “convolutional neural network” OR “deep learning” AND “tooth detection” OR “tooth numbering”
ScienceDirect	“panoramic radiograph” AND “convolutional neural network” OR “CNN” AND “tooth numbering” AND “classification” AND “detection” AND “segmentation”
Scopus	“panoramic radiograph” AND “convolutional neural network” AND “tooth numbering” AND “classification” OR “detection” OR “segmentation”

Study selection

Articles from 3 online databases were screened based on the pertinence of their titles and abstracts to the topic of the review. Two reviewers independently screened these titles and abstracts for potential eligibility, then conducted a full-text review in accordance with the selection criteria. All studies demonstrating the use of CNNs for tooth numbering on panoramic radiographs, including those utilizing classification, detection, and segmentation, were included. The inclusion criteria for this study were articles or journal publications that used a CNN approach for tooth numbering, were written in English, and provided full-text access. The exclusion criteria were articles that employed machine learning methods other than CNN for tooth numbering on periapical, bitewing, or 3-dimensional imaging. This review focused on the applicability of CNN for tooth numbering, given the increased popularity of CNN over the past decade due to its superior performance in computer vision tasks.

Data extraction

One reviewer performed the data extraction, which was subsequently discussed in depth with a second reviewer. The data extracted in this study included the author, year of publication, CNN architecture, application of the transfer learning method, tooth numbering method (classification, detection, or segmentation), number of datasets, sources of errors in tooth numbering, performance, and primary findings. The tooth numbering methods examined in this study included classification, detection, and segmentation. The performance metrics extracted for this review were accuracy, precision, sensitivity, specificity, and F1 score.

Results

A total of 540 studies were retrieved from 3 online databases for screening. Following an initial review based on the titles and abstracts, 28 studies were identified as potentially suitable for further consideration. Upon full-text eval-

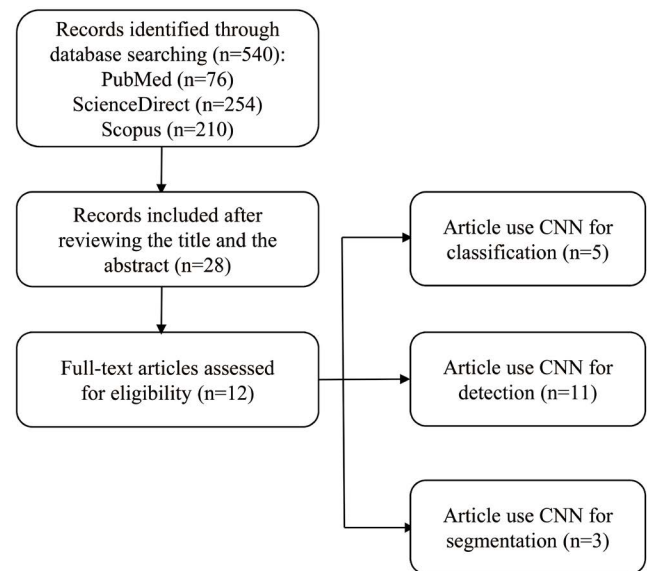


Fig. 1. Flowchart illustrating the article search process. CNN: convolutional neural network.

uation, 12 studies fulfilled the inclusion criteria and were incorporated into this review (Fig. 1). The results derived from the data extraction of all included studies are displayed in Table 2.

Datasets

The sizes of the datasets ranged from 160 to 8,000 panoramic images. Ten studies utilized a distinct test dataset. In 2 studies,^{26,27} the total number of images was the only information provided, with no details regarding the partitioning of the dataset. In 6 of the 10 studies with a distinct test set,^{24,25,28,30-32} data were divided into 3 datasets (training, validation, and testing), while 4 studies^{21-23,29} reported division into 2 datasets (training and testing).

CNN architecture

The CNN architecture employed varied considerably among the included studies. Eight studies^{21-25,27-30} reported

Table 2. Results of data extraction from the included studies

Author	CNN architecture	Transfer learning method	Numeration method	Number of datasets	Sources of errors in tooth numbering	Performance	Conclusion
Tuzoff et al. (2019) ²¹	Faster R-CNN (detection) VGG-16 (classification)	Yes	Detection, classification	1,574 (training: 1,352, test: 222)	Severely decayed tooth, crown prosthetic, missing neighboring tooth, dental implant, root remnant	Detection Precision: 99.45% Sensitivity: 99.41% Classification Sensitivity: 98% Specificity: 99.94%	The proposed CNN-based system can be used to detect and number permanent teeth.
Kim et al. (2020) ²²	Faster R-CNN Inception-v3 (detection) Heuristic algorithm (classification)	No	Detection, classification	303 (training: 253, test: 50)	Implant fixtures, crown prosthetic	Detection Precision: 84.5% Sensitivity: 75.5% Specificity: 80.4% Classification Accuracy: 84.5% Sensitivity: 84.2% Specificity: 75.5%	The proposed CNN-based system can be used to detect and number permanent teeth.
Mahdi et al. (2020) ²³	Faster R-CNN with ResNet	Yes	Detection	1,000 (training: 900, test: 100)	Not reported	ResNet-50 Precision: 94.2% F1 score: 96.5% ResNet-101 Precision: 96.4% F1 score: 97.8%	The proposed CNN-based system can be used to detect and number permanent teeth.
Bilgic et al. (2021) ²⁴	Faster R-CNN Inception-v2	Yes	Detection	2,482 (training: 1,984, validation: 249, test: 249)	Not reported	Precision: 96.52% Sensitivity: 95.59% F1 score: 96.06%	The proposed CNN-based system can be used to detect and number permanent teeth.
Kilic et al. (2021) ²⁵	Faster R-CNN Inception-v2	Yes	Detection	421 (training: 329, validation: 46, test: 46)	Not reported	Precision: 95.71% Sensitivity: 98.04% F1 score: 96.86%	The proposed CNN-based system can be used to detect and number deciduous teeth.
Lin and Chang (2021) ²⁶	AlexNet, VGGNet, ResNet, Xception, GoogLeNet	No	Classification	895 (dataset partition not reported)	Not reported	Accuracy AlexNet: 84.68% VGGNet: 85.62% ResNet: 90.93% Xception: 89.68% GoogLeNet: 85%	ResNet had the best performance for tooth classification.

Table 2. Continued

Author	CNN architecture	Transfer learning method	Numeration method	Number of datasets	Sources of errors in tooth numbering	Performance	Conclusion
Prados-Privado et al. (2021) ²⁷	Detection Mask R-CNN Classification ResNet-101	Yes	Detection, classification	8,000 (dataset partition not reported)	Missing neighboring tooth, root remnant, wisdom tooth, pontic	Accuracy Detection: 99.24% Classification: 93.83%	The proposed CNN-based system can be used to detect and number permanent teeth.
Vinayahalingam et al. (2021) ²⁸	Mask R-CNN with ResNet-50	No	Detection, segmentation, classification	2,000 (training: 1,600, validation: 200, test: 200)	Root remnant, dental implant, dental filling, root canal treatment, crown prosthetic	F1 score Detection: 99.3% Segmentation: 95.2% Classification: 97%	The proposed CNN-based system showed promising results for automatic chart filling.
Yükseket al. (2021) ²⁹	DENTECT	No	Segmentation, detection	1,005 (training: 855, test: 150)	Not reported	Segmentation Pixel accuracy: 96.4% Detection Precision: 89.1%	The proposed CNN-based system can be used to number teeth and classify multiple diseases.
Chandrashekar et al. (2022) ³⁰	Collaborative learning (combination between Faster R-CNN and YOLO-v5 + Mask R-CNN and U-Net)	Yes	Detection, segmentation	1,500 (various dataset partition)	Not reported	Segmentation Accuracy: 98.77% F1 score: 98.83% mAP: 97.30% Detection Accuracy: 98.44% F1 score: 98.75% mAP: 97.75%	The collaborative model is significantly more effective than the individual model in detecting and segmenting teeth.
Choi et al. (2022) ³¹	EfficientDet-D3	No	Detection	1,638 (training: 971, validation: 359, test: 308)	Not reported	Precision: 99.1% Sensitivity: 99.6%	The proposed CNN-based system can be used to detect and number permanent teeth and to classify dental treatment.
Mima et al. (2022) ³²	Faster R-CNN	Yes	Detection	160 (K-fold validation method)	Wisdom tooth, metal restoration	Accuracy Detection: 98.9% Classification: 91.7%	The performance of CNN for tooth detection and numbering can be improved by dividing the panoramic radiograph into 6 areas.

CNN: convolutional neural network, mAP: mean average precision.

using a combination of 2 to 4 different CNN architectures, while 4 studies^{26,29,31,32} utilized individually constructed CNN architectures for the enumeration of teeth on panoramic radiographs. The most frequently used architecture was Faster R-CNN, as reported in 7 studies, followed by ResNet in 4 studies. Inception and Mask R-CNN were each used in 3 studies, VGG in 2, and AlexNet, Xception, GoogLeNet, DENTECT, YOLO-v5, U-Net, and EfficientDet in 1 study each.

Regarding changes over the years studied, Faster R-CNN was initially introduced in 2019 for the purpose of tooth detection, and this was followed by the use of VGG-16 for numbering teeth on panoramic radiographs.²¹ In the following year, modifications were made to Faster R-CNN using other CNN architectures such as ResNet²³ and Inception-v2^{24,25} to evaluate the performance of tooth numbering systems. Starting in 2021, a number of studies began developing tooth segmentation-based architectures using Mask R-CNN, which was also modified with other CNN architectures like ResNet^{27,28} and Faster R-CNN.³⁰ In the same year, Yüksel et al.²⁹ developed a CNN-based architecture, termed DENTECT, that involved segmentation into 4 different quadrants before performing tooth detection and numbering on panoramic radiographs. Most recently, a collaborative

learning approach was introduced that combined Faster R-CNN and YOLO-v5 for tooth detection and Mask R-CNN and U-Net for tooth segmentation. This approach has been demonstrated effective in a variety of complex situations.³⁰

Types of tooth numbering methods

The primary functions of CNN in recognizing dental objects on panoramic radiographs include 3 tasks: classification, detection, and segmentation. A schematic representation of these tasks is provided in Figure 2. For tooth numbering, 11 studies employed the detection method, 5 utilized the classification method, and 3 adopted the segmentation method. Table 3 summarizes a comparative analysis of these 3 distinct approaches to tooth numbering on panoramic radiographs.

Detection enables the localization and numerical identification of a specific tooth by drawing a bounding box around it. Related studies have most commonly applied Faster R-CNN for tooth detection, occasionally modifying it with other CNN architectures to enhance the classification of tooth numbers. The references and/or data annotations for this task were supplied by various professionals, including radiologists,^{21,24,31} dentists,^{22,23} pedodontists,²⁵ endodontists,²⁹ and experts whose qualifications were not de-

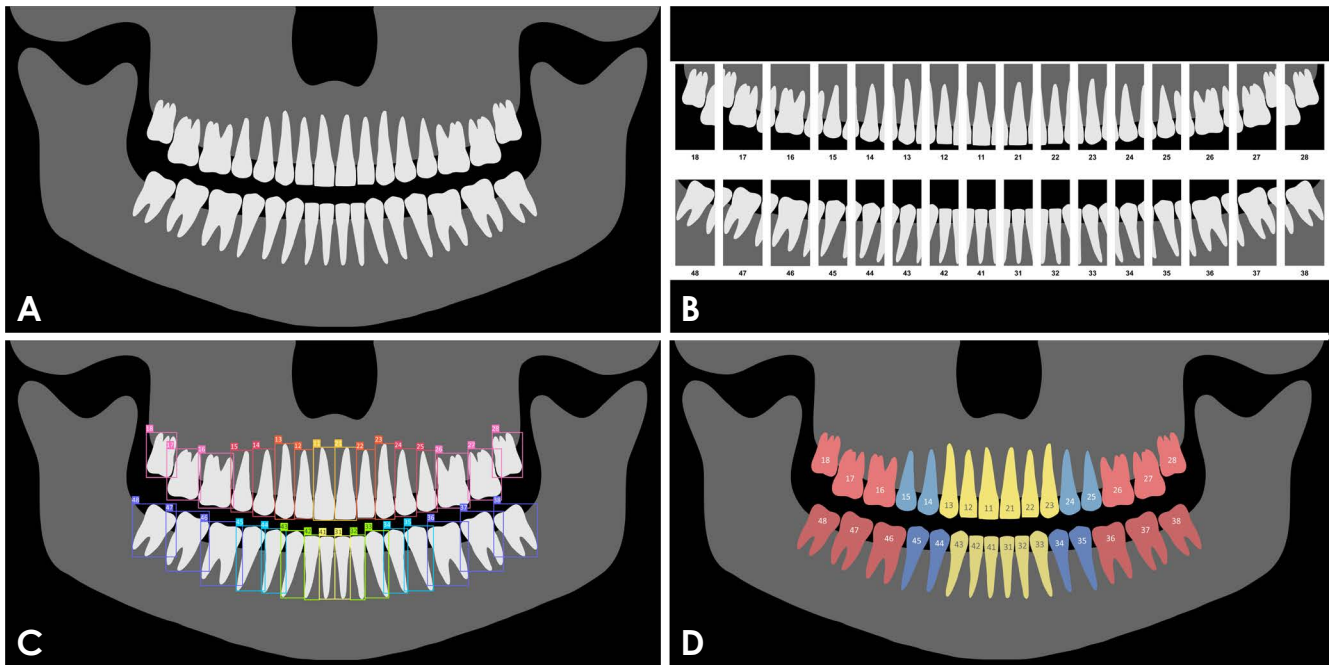


Fig. 2. Illustration of the various tasks involved in automated tooth numbering on panoramic radiographs. A. The pre-processed panoramic image. B. The classification task, which requires a labeled dataset or cropped tooth images from panoramic images, is employed for tooth numbering on each image. C. The detection task, which requires a panoramic radiograph with a marked region of interest, facilitates the localization and identification of the tooth object by drawing a bounding box around it. D. The segmentation task is carried out to delineate precise boundaries and identify the tooth objects on the panoramic radiograph.

Table 3. Comparison of the use of convolutional neural networks (CNNs) for classification, detection, and segmentation for tooth numbering on panoramic radiographs

Task	Number of studies	Most common CNN architecture	Most common outcome metric	Findings
Classification	5	ResNet (n = 3)	Accuracy (n = 3)	ResNet showed the best performance for the classification task.
Detection	11	Faster R-CNN (n = 7)	Precision (n = 7)	A combination of Faster R-CNN and another CNN model can improve the detection task.
Segmentation	3	Mask R-CNN (n = 2)	F1 score (n = 2)	A combination of Mask R-CNN and another CNN model can potentially be used for automated dental chart filling.

scribed.^{27,28,32} For detection tasks, the models were assessed using several metrics: precision (n = 7), F1 score (n = 5), sensitivity (n = 5), specificity (n = 1), accuracy (n = 3), and mean average precision (n = 1). In a comparative test, 1 study contrasted the performance of tooth detection using Faster R-CNN and YOLO v5 architectures within a collaborative learning framework.³⁰

Classification is employed to categorize and enumerate isolated or cropped tooth images derived from panoramic radiographs. ResNet architectures are predominantly utilized for this task and demonstrate high performance. One study involved a comparison of 5 distinct CNN architectures, namely AlexNet, VGGNet, ResNet, Xception, and GoogLeNet, to determine the most effective option for classification. The reference standard for data annotation and performance evaluation was supplied by a radiologist,²¹ a dentist,²² a medical expert,²⁶ and an expert whose qualifications were not described.^{27,28} The performance of the CNNs was assessed using several metrics, including accuracy (n = 3), F1 score (n = 1), and specificity and sensitivity (n = 2).

Segmentation is carried out to delineate the pixel-wise boundaries of the tooth object on panoramic radiographs. Mask R-CNN was primarily employed for teeth segmentation, followed by other CNN architectures to identify pathologies or specific conditions affecting each tooth. This method can also be utilized to facilitate automated dental chart filling. The standard references for the segmentation task were an endodontist and a medical expert.^{28,29} The performance of the segmentation task was evaluated using F1 score (n = 2), accuracy (n = 1), and pixel accuracy (n = 1). One study compared the performance of tooth segmentation using Mask R-CNN and U-Net architectures in a collaborative learning setting.³⁰

Findings

Assessing the sources of errors in tooth numbering is key

for the advancement of this system. As indicated in Table 2, only 5 studies reported an evaluation of the mistakes made by the system in tooth numbering. The presence of severely decayed teeth, crown prosthetics (including metal restorations, dental fillings, and pontics), missing adjacent teeth, dental implants, root remnants, wisdom teeth, and root canal-treated teeth can impair the performance of CNNs in tooth numbering. This is due to the potential for these factors to generate false positives or false negatives. The analysis of error sources is vital to address the limitations of CNNs in tooth numbering. Consequently, studies should include error analysis reports to aid in the development of optimized CNN architecture and dataset preparation for future research.

Although the aforementioned tooth conditions can decrease the performance of CNNs in tooth numbering, numerous studies have demonstrated that CNNs can extend beyond this purpose. They can identify the presence of dental implants, missing teeth, teeth treated with root canals, root remnants, and teeth with crown restorations and fillings. This capability can be further refined for automated dental chart filling in panoramic radiographs.²⁸⁻³¹ CNNs can also be utilized for numbering deciduous teeth on panoramic radiographs.²⁵ However, employing CNNs for numbering in other stages of dentition presents a challenge. This is due to the more intricate process required to perform tooth recognition tasks, particularly in the mixed dentition stage, because of the presence of tooth germs.

Discussion

All of the included studies reported that CNNs, irrespective of the specific architecture employed, demonstrated good performance in the automatic numbering of teeth on panoramic radiographs. Vinayahalingam et al.²⁸ applied the CNN methodology to the automation of dental chart filling on panoramic radiographs. The objective of automating den-

tal chart filling extends beyond merely facilitating tooth numbering; it can also be utilized to automatically identify the presence of crowns, fillings, root canals, implants, and residual roots on panoramic radiographs. The findings of this study indicated that CNNs achieved an F1 score exceeding 95% for tasks related to classification, detection, and segmentation. CNNs have been shown to be effective for automatic tooth numbering, and their application can be further expanded to automate dental chart filling using panoramic radiographs.

Faster R-CNN is the architecture most commonly employed for tooth numbering on panoramic radiographs.³³ This model is an enhancement of R-CNN and Fast R-CNN.³⁴ Faster R-CNN utilizes a region proposal network (RPN), a type of neural network that supplants the function of selective search. This substitution reduces the excessive computational demands on the computer, thereby accelerating object detection based on deep learning.^{35,36} The RPN shares full-image convolutional features with the detection network, which allows for nearly cost-free region proposals. It is a fully convolutional network that concurrently predicts object boundaries and objectness scores at each position.³⁷ One benefit of using an RPN is the increase in speed during both training and prediction. Given that an RPN is a straightforward network that only employs convolutional layers, the prediction time can be faster than when using the classification base network.^{38,39}

The primary functions of a CNN in identifying objects, including tooth numbers, on digital radiography images are classification, detection, and segmentation. Most of the studies examined^{21-25,27-32} involved the use of detection as a numeration method. Detection is an important step in pinpointing the location of an object. The most frequently used CNN architecture for performing detection tasks is Faster R-CNN. This architecture was employed in 7 studies for detection purposes. Bilgir et al.²⁴ and Kılıc et al.²⁵ chose Faster R-CNN for tooth numbering due to its design, which simplifies the complexity of CNNs with an architecture constructed to be comparatively deep and wide. CNNs can also be employed to number primary teeth using the detection task. Kılıc et al.²⁵ utilized a dataset comprised of 421 anonymous panoramic radiographic images of pediatric patients aged between 5 and 7 years. The performance achieved included a precision of 95.71%, a sensitivity of 98.04%, and an F1 score of 96.86%. This demonstrates that CNNs can be used for numbering of both permanent and primary teeth. Further advancements can be made in the use of CNNs for primary tooth numbering, such as the application of tooth germ numeration, age estimation, and

digital forensic purposes. Following successful detection, the identified tooth can undergo further analysis using other machine learning methods to detect pathologies or anomalies. The architecture of the CNN can be modified or expanded to detect various pathologies in the dental area of panoramic radiographs.

Five studies employed a classification method for numbering teeth on panoramic radiographs.^{21,22,26-28} This task is essential for categorizing tooth numbers based on their features. ResNet, one of the most common CNN architectures, is frequently chosen for this classification task. The popularity of ResNet (short for Residual Network) is largely due to its residual connection mechanism, which addresses the vanishing gradient problem. This problem arises when the gradient results analyzed by the model fail to reach the first layer, thereby preventing CNNs from studying the calculated errors.^{40,41} The operational principle of ResNet involves constructing a network that is relatively complex, while concurrently determining the optimal number of layers to overcome the vanishing gradient problem.⁴² In 2015, ResNet emerged as the winner of the ILSVRC and the Common Objects in Context (COCO) competitions. Specifically, it won in the categories of ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation.⁴³ Regarding CNN architecture, depth is an important element in developing high-performing CNN models. ResNet offers several variations of layer types, including 18, 34, 50, 101, and 152 layers.⁴⁴ Mahdi et al. utilized ResNet-50, comprising 50 layers, and ResNet-101, comprising 101 layers, to compare the precision and F1 score performance for permanent tooth numbering on panoramic radiographs.²³ ResNet-101, being a deeper network than ResNet-50, delivered superior performance. Consequently, from a computational perspective, ResNet-101 demands greater computing power than ResNet-50. Both architectures are applicable in the field of dentistry.^{45,46} Vinayahalingam et al.²⁸ employed Mask R-CNN with ResNet-50 to automate dental chart filling based on panoramic radiographs. The F1 score performance achieved for detection was 99.3%. This indicates that the detection method is not solely useful for tooth numbering, but can also be applied to automated dental chart filling, a feature that is beneficial for forensic identification.

Only 3 studies employed segmentation for tooth numbering on panoramic radiographs.²⁸⁻³⁰ While most of the included studies utilized Mask R-CNN, U-Net is among the most frequently used architectures in medical imaging segmentation due to its effectiveness and superior performance. Image segmentation is a semantic process, dividing images

into objects and non-objects.^{47,48} Chandrashekar et al.³⁰ implemented a collaborative technique, combining 2 architectures for segmentation and detection tasks in panoramic radiography to improve results. The collaborative CNN architecture used was a combination of Mask R-CNN and U-Net for segmentation, which was then modified into a detection CNN architecture model, namely Faster R-CNN and YOLO-v5. The accuracy achieved for segmentation was 98.44%, while the detection performance was 98.77%. The findings of this study suggest that the collaborative model is significantly more effective than the individual learning model. CNNs can also be utilized for automated dental chart filling on panoramic radiographs using segmentation methods.²⁸ Further advancements are possible, particularly in using the segmentation method for tooth numbering on panoramic radiographs, as only 3 articles have utilized this method to date. Segmentation plays a key role in computer vision and image processing, providing an effective process to facilitate analysis by dividing the image into 2 parts: the object and the background.⁴⁹⁻⁵¹

The studies included in this review utilized a range of dataset sizes to develop their neural network models. The smallest dataset, used by Mima et al., contained 160 images,³² while the largest, used by Prados-Privado et al., contained 8,000.²⁷ Eight of 10 studies included in this review used datasets of at least 1,000 images. A substantial quantity of data is necessary for the development of a CNN model. Studies involving fewer than 1,000 images can be considered to include a small dataset, which may result in a less accurate output model.⁵² Since gathering, processing, and labeling data is exhaustive, several methods have been employed to enhance output model performance. These methods include transfer learning and data augmentation. Transfer learning is a technique that can expedite the learning process by transferring the pre-trained base layer of the CNN model using readily available datasets, then training the remaining layers on a smaller, local dataset. Seven of the studies included in this review (as shown in Table 2) demonstrated that this technique can be effective for tooth recognition, although the transfer learning was performed using datasets other than panoramic radiographs, such as ImageNet and COCO. Data augmentation is another technique that can expand the training dataset through the application of image transformations such as flipping, color manipulation, cropping, magnification, rotation, and translation.⁵³

The studies included in this review highlighted the limitations of CNNs in tooth numbering for panoramic radiographs. Numerous studies reported various errors made by

the system during the testing phase. The system may fail to identify a tooth due to the presence of severely decayed teeth, crown prosthetics (including metal restorations, dental fillings, and pontics), missing adjacent teeth, dental implants, root remnants, wisdom teeth, and root canal treatments on panoramic images. The most frequent errors, resulting in either false negatives or false positives, were associated with the presence of root remnants. The system also demonstrated errors in detecting dental implants and teeth with restoration materials.^{21,22,27,28} For the further advancement of CNNs in tooth numbering, it is crucial to include a variety of potential tooth forms in the training process. These forms may include primary teeth, tooth germs, residual roots, tooth anomalies, or teeth with crown restorations. The presence of an edentulous area should also be considered in the training dataset, as it can disrupt the learning process of the CNN model, leading to false positives on adjacent teeth.^{21,27} These considerations should be addressed during the dataset preparation stage of CNN system development for tooth numbering on panoramic radiographs.²⁸

In future studies, CNN performance could be optimized by considering dataset quality as well as quantity, based on the requirements of the development process.³¹ Merely increasing dataset size may not necessarily improve performance, as the diversity of samples and the quality of the dataset are also crucial for optimizing results. For an effective learning system, it is essential that the sample is diverse and evenly distributed. This is particularly important given the challenging task of identifying multiple tooth objects with similar anatomical features in various locations. Once an appropriate dataset is assembled, the selection of appropriate CNN architecture is another important step prior to architecture modification. Collaborative or hybrid learning can be employed by integrating the CNN architecture based on the study's objectives, as each architecture has specific advantages for certain tasks. This optimization depends on the computational power and the researcher's knowledge, as optimal performance can be achieved with smaller datasets and lower computational power when the study is well-designed. The optimal CNN architectures for tooth numeration on panoramic radiographs can be utilized in various AI research projects for further advancements in the field of dentistry.

In conclusion, CNNs can be employed to automate the process of tooth numbering on panoramic radiographs through classification, detection, and segmentation. Enhancing and optimizing the CNN architecture based on the task at hand can improve the performance and results of the automated tooth numbering system. Future developments

in CNN-based models should focus on various stages of dentition, such as the primary and mixed dentition stages. These models should also consider a range of tooth conditions, including teeth with restorations, dental anomalies, residual teeth, and edentulous areas, to enhance the recognition of tooth number. An optimized CNN architecture is anticipated to be useful for automated tooth numbering on panoramic radiographs, irrespective of the tooth condition. Ultimately, it could serve as a foundational element for AI research on panoramic radiography for a variety of purposes.

Conflicts of Interest: None

References

- Ossowska A, Kusiak A, Świetlik D. Artificial intelligence in dentistry - narrative review. *Int J Environ Res Public Health* 2022; 19: 3449.
- Izzetti R, Nisi M, Aringhieri G, Crocetti L, Graziani F, Nardi C. Basic knowledge and new advances in panoramic radiography imaging techniques: a narrative review on what dentists and radiologists should know. *Appl Sci (Basel)* 2021; 11: 7858.
- Rozylo-Kalinowska I. Panoramic radiography in dentistry. In: *Imaging techniques in dental radiology*. Cham; Springer: 2020. p. 43-56.
- Thanathornwong B, Suebnukam S. Automatic detection of periodontal compromised teeth in digital panoramic radiographs using faster regional convolutional neural networks. *Imaging Sci Dent* 2020; 50: 169-74.
- Havale R, Sheetal BS, Patil R, Hemant Kumar R, Anegundi RT, Inushekar KR. Dental notation for primary teeth: a review and suggestion of a novel system. *Eur J Paediatr Dent* 2015; 16: 163-6.
- Forrest A. Forensic odontology in DVI: current practice and recent advances. *Forensic Sci Res* 2019; 4: 316-30.
- Müller A, Mertens SM, Göstemeyer G, Krois J, Schwendicke F. Barriers and enablers for artificial intelligence in dental diagnostics: a qualitative study. *J Clin Med* 2021; 10: 1612.
- Mosadeghrad AM. Factors influencing healthcare service quality. *Int J Health Policy Manag* 2014; 3: 77-89.
- Joda T, Yeung AW, Hung K, Zitzmann NU, Bornstein MM. Disruptive innovation in dentistry: what it is and what could be next. *J Dent Res* 2021; 100: 448-53.
- Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: current applications and future directions. *PLoS Med* 2018; 15: e1002707.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional. In: *Pereira F, Burges CJ, Bottou L, Weinberger KQ. Advances in Neural Information Processing Systems 25 (NIPS 2012)*. San Mateo: Morgan Kaufmann Publishers; 2012.
- Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional neural networks for radiologic images: a radiologist's guide. *Radiology* 2019; 290: 590-606.
- Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: a scoping review. *J Dent* 2019; 91: 103226.
- Putra RH, Doi C, Yoda N, Astuti ER, Sasaki K. Current applications and development of artificial intelligence for digital dental radiography. *Dentomaxillofac Radiol* 2022; 51: 20210197.
- Roosanty A, Widyaningrum R, Diba SF. Artificial intelligence based on Convolutional Neural Network for detecting dental caries on bitewing and periapical radiographs. *J Radiol Dentomaxillofac Indones* 2022; 6: 89-94.
- Musri N, Christie B, Ichwan SJ, 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* 2021; 51: 237-42.
- Miki Y, Muramatsu C, Hayashi T, Zhou X, Hara T, Katsumata A, et al. Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput Biol Med* 2017; 80: 24-9.
- Chen H, Zhang K, Lyu P, Li H, Zhang L, Wu J, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci Rep* 2019; 9: 3840.
- Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. *J Med Internet Res* 2020; 22: e17256.
- Daudt HM, van Mossel C, Scott SJ. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC Med Res Methodol* 2013; 13: 48.
- Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac Radiol* 2019; 48: 20180051.
- Kim C, Kim D, Jeong H, Yoon SJ, Youm S. Automatic tooth detection and numbering using a combination of a CNN and heuristic algorithm. *Appl Sci (Basel)* 2020; 10: 5624.
- Mahdi FP, Motoki K, Kobashi S. Optimization technique combined with deep learning method for teeth recognition in dental panoramic radiographs. *Sci Rep* 2020; 10: 19261.
- Bilgir E, Bayrakdar İŞ, Çelik Ö, Orhan K, Akkoca F, Sağlam H, et al. An artificial intelligence approach to automatic tooth detection and numbering in panoramic radiographs. *BMC Med Imaging* 2021; 21: 124.
- Kılıc MC, Bayrakdar İS, Çelik Ö, Bilgir E, Orhan K, Aydın OB, et al. Artificial intelligence system for automatic deciduous tooth detection and numbering in panoramic radiographs. *Dentomaxillofac Radiol* 2021; 50: 20200172.
- Lin SY, Chang HY. Tooth numbering and condition recognition on dental panoramic radiograph images using CNNs. *IEEE Access* 2021; 9: 166008-26.
- Prados-Privado M, García Villalón J, Blázquez Torres A, Martínez-Martínez CH, Ivorra C. A convolutional neural network for automatic tooth numbering in panoramic images. *Biomed Res Int* 2021; 2021: 3625386.
- Vinayahalingam S, Goey RS, Kempers S, Schoep J, Cherici T, Moin DA, et al. Automated chart filing on panoramic radiographs using deep learning. *J Dent* 2021; 115: 103864.
- Yüksel AE, Gültekin S, Simsar E, Özdemir ŞD, Gündoğar M, Tokgöz SB, et al. Dental enumeration and multiple treatment

- detection on panoramic X-rays using deep learning. *Sci Rep* 2021; 11: 12342.
30. Chandrashekar G, AlQarni S, Bumann EE, Lee Y. Collaborative deep learning model for tooth segmentation and identification using panoramic radiographs. *Comput Biol Med* 2022; 148: 105829.
 31. Choi HR, Siadari TS, Kim JE, Huh KH, Yi WJ, Lee SS, et al. Automatic detection of teeth and dental treatment patterns on dental panoramic radiographs using deep neural networks. *Forensic Sci Res* 2022; 7: 456-66.
 32. Mima Y, Nakayama R, Hizukuri A, Murata K. Tooth detection for each tooth type by application of faster R-CNNs to divided analysis areas of dental panoramic X-ray images. *Radiol Phys Technol* 2022; 15: 170-6.
 33. Shirsat S, Abraham S. Tooth detection from panoramic radiographs using deep learning. In: Srirama SN, Lin JC, Bhatnagar R, Agarwal S, Reddy PK. *Big Data Analytics: 9th International Conference on Big Data Analytics Proceedings*. Berlin: Springer-Verlag; 2021. p. 54-63.
 34. Su Y, Li D, Chen X. Lung nodule detection based on faster R-CNN framework. *Comput Methods Programs Biomed* 2021; 200: 105866.
 35. Gavrilesco R, Zet C, Foşalău C, Skoczylas M, Cotovanu D. Faster R-CNN: an approach to real-time object detection. In: 2018 International Conference and Exposition on Electrical and Power Engineering (EPE). Iasi, Romania: IEEE; 2018. p. 0165-8. Available from <https://ieeexplore.ieee.org/document/8559776>.
 36. Shih KH, Chiu CT, Lin JA, Bu YY. Real-time object detection with reduced region proposal network via multi-feature concatenation. *IEEE Trans Neural Netw Learn Syst* 2020; 31: 2164-73.
 37. Wang C, Peng Z. Design and implementation of an object detection system using faster R-CNN. In: 2019 International Conference on Robots & Intelligent System (ICRIS). Haikou, China: IEEE; 2019. p. 204-6. Available from <https://ieeexplore.ieee.org/abstract/document/8806291>.
 38. Bharati P, Pramanik A. Deep learning techniques - R-CNN to mask R-CNN: a survey. In: *Computational Intelligence in Pattern Recognition*. Springer; 2020. p. 657-68.
 39. Zhang Y, Chu J, Leng L, Miao J. Mask-refined R-CNN: a network for refining object details in instance segmentation. *Sensors (Basel)* 2020; 20: 1010.
 40. Hu Y, Huber A, Anumula J, Liu SC. Overcoming the vanishing gradient problem in plain recurrent networks [Internet]. *arXiv* 2018; 1801.06105 [cited 2023 Mar 8]. Available from <https://arxiv.org/abs/1801.06105v3>.
 41. Zhang K, Sun M, Han TX, Yuan X, Guo L, Liu T. Residual networks of residual networks: Multilevel residual networks. *IEEE Trans Circuits Syst Video Technol* 2018; 28: 1303-14.
 42. Wickramasinghe CS, Marino DL, Manic M. ResNet autoencoders for unsupervised feature learning from high-dimensional data: deep models resistant to performance degradation. *IEEE Access* 2021; 9: 40511-20.
 43. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition [Internet]. *arXiv* 2015; 1512.03385v1 [cited 2023 Mar 8]. Available from <https://arxiv.org/abs/1512.03385>.
 44. Mukhometzianov R, Carrillo J. CapsNet comparative performance evaluation for image classification [Internet]. *arXiv* 2018; 1805.11195. Available from <https://arxiv.org/abs/1805.11195>.
 45. Hu J, Shen L, Albanie S, Sun G, Vedaldi A. Gather-Excite: exploiting feature context in convolutional neural networks [Internet]. *arXiv* 2019; 1810.12348 [cited 2023 Mar 8]. Available from <https://arxiv.org/abs/1810.12348>.
 46. Ding X, Zhang X, Ma N, Han J, Ding G, Sun J. RepVGG: making VGG-style convnets great again [Internet]. *arXiv* 2021; 2101.03697 [cited 2023 Mar 8]. Available from <https://arxiv.org/abs/2101.03697>.
 47. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: a nested u-net architecture for medical image segmentation [Internet]. *arXiv* 2018; 1807.10165 [cited 2023 Mar 8]. Available from <https://arxiv.org/abs/1807.10165>.
 48. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation [Internet]. *arXiv* 2015; 1505.04597 [cited 2023 Mar 8]. Available from <https://arxiv.org/abs/1505.04597>.
 49. Chen X, Williams BM, Vallabhaneni SR, Czanner G, Williams R, Zheng Y. Learning active contour models for medical image segmentation. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA; 2019. p. 11624-32.
 50. Patil DD, Deore SG. Medical image segmentation: a review. *Int J Comput Sci Mob Comput* 213; 2: 22-7.
 51. Fan J, Zhang Z, Song C, Tan T. Learning integral objects with intra-class discriminator for weakly-supervised semantic segmentation. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020. p. 4282-91.
 52. Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. *Imaging Sci Dent*. 2019; 49: 1-7.
 53. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 2021; 8: 53.