

Evaluation of an Artificial Intelligence web-based software to detect and classify dental structures and treatments in panoramic radiographs

Monica Bonfanti-Gris^a, Angel Garcia-Cañas^a, Raul Alonso-Calvo^b, Maria Paz Salido Rodriguez-Manzaneque^{a,*}, Guillermo Pradies Ramiro^a

^a Department of Conservative and Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid. Plaza Ramón y Cajal, s/n. 28040 Madrid, Spain

^b Department of Informatics Systems and Languages, Faculty of Software Engineering, Polytechnic University of Madrid. Campus Montegancedo s/n, Boadilla del Monte. 28660 Madrid, Spain

ARTICLE INFO

Keywords:

Artificial intelligence
Convolutional neural network
Dental radiology
Diagnostic imaging
Machine learning
Panoramic radiography

ABSTRACT

Objectives: To evaluate the diagnostic reliability of a web-based Artificial Intelligence program on the detection and classification of dental structures and treatments present on panoramic radiographs.

Methods: A total of 300 orthopantomographies (OPG) were randomly selected for this study. First, the images were visually evaluated by two calibrated operators with radiodiagnosis experience that, after consensus, established the “ground truth”. Operators’ findings on the radiographs were collected and classified as follows: metal restorations (MR), resin-based restorations (RR), endodontic treatment (ET), Crowns (C) and Implants (I). The orthopantomographies were then anonymously uploaded and automatically analyzed by the web-based software (Denti.Ai). Results were then stored, and a statistical analysis was performed by comparing them with the ground truth in terms of Sensitivity (S), Specificity (E), Positive Predictive Value (PPV) Negative Predictive Value (NPV) and its later representation in the area under (AUC) the Receiver Operating Characteristic (ROC) Curve.

Results: Diagnostic metrics obtained for each study variable were as follows: (MR) S=85.48%, E=87.50%, PPV=82.8%, NPV=42.51%, AUC=0.869; (PR) S=41.11%, E=93.30%, PPV=90.24%, NPV=87.50%, AUC=0.672; (ET) S=91.9%, E=100%, PPV=100%, NPV=94.62%, AUC=0.960; (C) S=89.53%, E=95.79%, PPV=89.53%, NPV=95.79%, AUC=0.927; (I) S, E, PPV, NPV=100%, AUC=1.000.

Conclusions: Findings suggest that the web-based Artificial intelligence software provides a good performance on the detection of implants, crowns, metal fillings and endodontic treatments, not being so accurate on the classification of dental structures or resin-based restorations.

Clinical Significance: General diagnostic and treatment decisions using orthopantomographies can be improved by using web-based artificial intelligence tools, avoiding subjectivity and lapses from the clinician.

1. Introduction

Making a correct diagnosis is one of the most important procedures carried out in the dental office. This should be assessed not only from the professionals’ dental experience, but also from complementary techniques which provide essential information [1]. Orthopantomographies (OPGs) can be considered the preferred general diagnostic method due to the visualization of all the oro-facial structures in a single image, as well as the comfort with which it is achieved and the low radiation that is absorbed by the patient [2]. However, although they are undoubtedly helpful, it must be considered that their interpretation will always be

subjected to the level of knowledge and expertise of the operator who analyzes it. It has been stated that more experienced examiners show an improved diagnostic accuracy compared with less experienced ones [3]. Yet, any qualified clinician can suffer from both mental and eye fatigue, which could lead to ignoring important radiographic signs that can interfere on their final diagnostic decisions, leading to an incorrect or misinterpreted diagnosis [4].

Recently, technologies such as Artificial Intelligence (AI) have revolutionized the health field. These systems were mainly created to ease and improve the diagnostic and therapeutic capacity of professionals. AI was introduced by John McCarthy in 1995 and it refers to

* Corresponding author at: Universidad Complutense De Madrid, Plaza De Ramón Y Cajal S/N. 28040 Madrid. Spain.

E-mail address: mpsalido@ucm.es (M.P. Salido Rodriguez-Manzaneque).

<https://doi.org/10.1016/j.jdent.2022.104301>

Received 23 April 2022; Received in revised form 13 September 2022; Accepted 15 September 2022

Available online 21 September 2022

0300-5712/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the capacity of certain machines to imitate both human knowledge and behavior based on sequences of algorithms [5]. It is important to mention that Machine Learning (ML) – an AI subfield - has proven to be a reliable and effective way to train convolutional neural networks that analyze radiographic images due to their ability to predict results on unseen data [3]. Therefore, automated assistant systems for dental radiographic diagnosis could overcome the above-mentioned shortcomings, as well as dispose of discrepancies produced by the subjectivity of different clinicians [5].

This new technology has demonstrated an excellent performance in some medical fields [6], as well as in general dentistry imagery (both in 2D and 3D) [7–9]. Most of the studies carried out on these types of image-testing base their main objective on the performance evaluation of convolutional neural networks (CNN) to detect anatomical structures, treatments and pathologies – such as dental caries [10–12]. Nonetheless, other CNNs were reported to detect different mandibular tumors [13], root fracture, or even to perform human identification and age determination from the teeth appearing in these radiographic images - with forensic purposes [14–16]. Although its role in general dentistry diagnostics is determining, AI-based software's have been noted to have an important turning point on the detection and classification of dental implants - which can sometimes be a challenging task even for operators. Several studies have already been published addressing this problematic with both panoramic and periapical radiographs [17–19]. Researches referring to 3D imagery have evaluated the performance analysis of several CNN to automatically detect both anatomical and dental structures [20–22] such as the mandible [23], mandibular nerve canal [24] or the pharyngeal airway [25].

Bearing in mind all the above-mentioned considerations and the existence on the internet of web-based radiological diagnostic programs with scarce documentation about their reliability, we considered to set up this analytical, observational and cross-sectional study. The aim of this research was to evaluate the diagnostic capacity in terms of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of a web-based convolutional neural network software on the detection and classification of dental structures and treatments (including both metallic and resin-based fillings, root canal treatments, crowns and implants). The null hypotheses were established as there were no differences between what was detected by the operators - considered as Ground Truth - and the results offered by the artificial intelligence-based software.

2. Material and methods

The present study compared the performance of dental operators and a convolutional neural network on the detection and classification of different findings in panoramic radiographs. These two tasks are very commonly studied when reporting results from CNNs. Thus, it has to be cleared up the differences between them – detection refers to the process of object identification, whether classification refers to object categorization. Reporting this study followed STARD (Standards for Reporting of Diagnostic Accuracy Studies) and the CLAIM (Checklist for Artificial Intelligence in Medical Imaging) checklists [26–28].

2.1. Image dataset protection and sampling

This study was conducted by the Department of Conservative and Prosthetic Dentistry at Complutense's University in Madrid and was approved by the San Carlos Clinic Hospital (approval n° 21/297E). No collaboration nor working condition was established between university researchers and AI company at any time. 300 anonymized panoramic radiographs exported in .jpeg file format were randomly acquired using Kodak 9000 3D Extraoral Imaging System (Carestream, Carestream Health, Inc., New York, USA) with an image sizing and resolution of 2706×1536 and 323×323 pixels, respectively. The exposure conditions were set at 73 kVp, 12 mA and an exposure time of 12.5 seconds,

according to the manufacturer's instructions. The radiographs consisted of patients with permanent dentition (with or without any dental absence or treatments). They were selected without considering any sex or race differences. All images containing deciduous and mixed dentition were excluded.

2.2. Image dataset processing and reference test

Sample size was calculated setting the expected average prevalence (P) of any event at 0,8 and the precision (d) at 0,05. For Z value of 1,96 the formula $n = \frac{z^2 P(1-P)}{d^2}$ offered a sample size of 245,86; finally a dataset of 300 panoramic radiographs were selected. An exhaustive analysis of the selected images was performed by two independent clinical evaluators (M.B-G. and A.G-C.) with diagnostic experience who were previously trained and calibrated. For this purpose, operators were presented with 20 panoramic radiographs where all study variables were present and, after analysis, a carefully-driven diagnostic criteria for each one of them was carried out. Interclass Correlation Coefficient (ICC) was calculated to assure a correct reliability in both experts analyzing outcomes, obtaining results between 0.993 and 1.000 for each variable. Thus, the diagnostic consensus established between both dental professionals formed the basis from which the web based software analysis results would later be compared. In the event that there was no agreement between the two main operators, a third party (G.P-R.) acted evaluating the case and achieving consensus among the evaluators. If harmony was not accomplished, the case was excluded. The existing dental structures were labeled with Federation Dentaire International (FDI) nomenclature. Dental treatments found were identified and labeled according to the following categories: (MR) for metallic restorations, (RR) for resin-based restorations, (ET) for endodontic treatments, (C) for both dental and implant supported crowns and pontics, (I) for osseointegrated implants. Afterwards, the images were uploaded and analyzed by the software. Although this AI-based program offered the possibility to modify the brightness and contrast of the radiographs, no modification was made in order to ensure both consistency and equality terms between the operators' and the programs' evaluation. To establish the diagnostic capacity of the web-based software for each study variable, the following given categories were marked: crown, endodontic treatment, implant, pontic and filling. In addition, the detection threshold of the program was adjusted to 0% - being this the percentage of trustworthiness with which the software specifies the presence of any dental structure or treatment. This way, we were able to evaluate everything that was detected by it (regardless of the percentage of reliability obtained).

2.3. Architecture of the deep convolutional neural network

In this study, a pre-trained Convolutional Neural Network (Denti.Ai, Denti.Ai Technology Inc., Toronto ON, Canada) was used to analyze all 300 panoramic radiographs. Denti.Ai® is an automated online software based on artificial intelligence that was released online in 2017 and that is capable of generating an automatic diagnosis of radiological images using Deep Learning technologies. To carry out the task of detecting and classifying dental structures found in the OPG's, this program uses different detection modules: first, the image is processed to establish the limits of each tooth present and then marks the structures found with bounding boxes. The tooth numbering module then classifies each delimited tooth region according to the FDI dental nomenclature. This way, the system generates coordinates in the bounding boxes and associates them with dentition numbers [29] (Fig. 1). Regarding the detection and classification of both pathologies and treatments, the program is run by a coding system based on a gray scale [29]. Outcomes can be audited and verified by a dental practitioner, who can modify them at any time according to his experience and professional opinion. This allows to nurture the neural network in case of an incorrect

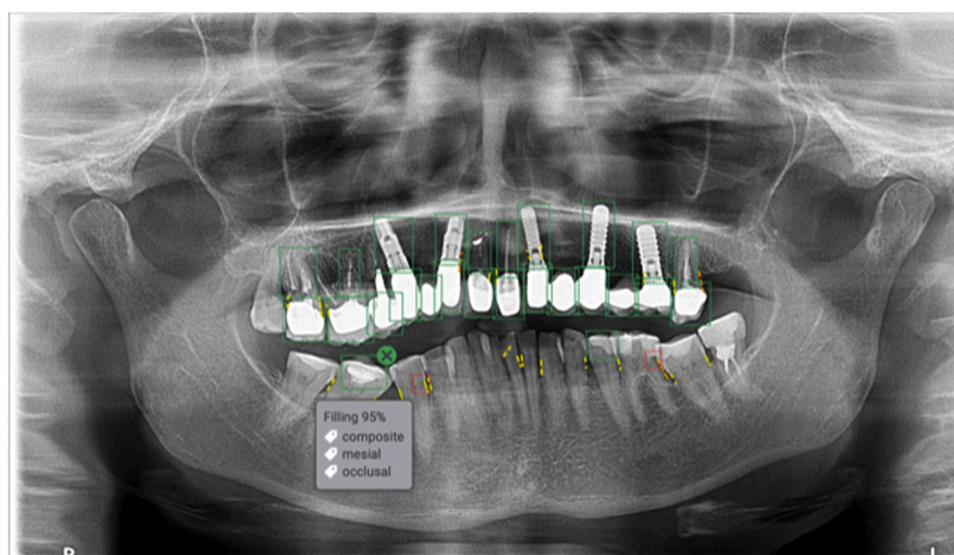


Fig. 1. Resulting Analysis made by the Convolutional Neural Network where bounding boxes of all present dental structures and treatments can be observed.

diagnosis given by it.

2.4. Validation metrics and statistical analysis

The performance of the automatic detection stage was evaluated on the basis of detection sensitivity, which was calculated as the number of teeth/treatments automatically detected by the software out of the number of teeth/treatments identified by the operators (ground truth). The diagnostic sensitivity (S), specificity (E), positive predictive value (PPV), negative predictive value (NPV), receiver operating characteristic curve (ROC) and area under the ROC curve (AUC) of the tested dataset were calculated by using IBM SPSS Statistics version 26 (SPSS, Chicago, IL, USA) for every study variable. Due to the existence of certain diagnostic errors when the task of detection and classification of dental structures was performed, a descriptive analysis was carried out. The observed detection mistakes were classified as follows: Error Type 0 (ET0), correct numbering classification; Error Type 1 (ET1), mistaken detection of implant-supported crowns with natural dental structures; Error Type 2 (ET2), mistaken detection of pontics with natural dental structures; Error Type 3 (ET3), mistaken numbering classification between dental pieces present in radiographic image; Error Type 4 (ET4), no detection of dental structures present in radiographic image; and Error Type 5 (ET5), no detection of supernumeraries and included teeth.

3. Results

A descriptive analysis was carried out for each variable included in the present study. From the 300 subjects included in the study, 44% were male and 56% were female patients, aged with a mean of 42 years old. Average detection numbers of both operator and the web-based convolutional neural network were compared, showing very close-related results between both diagnostic sources (Table 1).

The number of teeth detected on each panoramic radiograph by the program was 28.120 ± 4.126 , being the one detected by Gold Standard 28.110 ± 4.139 . Descriptive analysis on the CNNs' detection and classification of teeth was performed due to existing classification mistakes, from which was observed that there was a 50% chance that the AI based software could miss the correct detection/classification of dental structures in panoramic radiographs (ET1, 8.8%; ET2, 11.3%; ET3, 20.7%; ET4, 1.8%; ET5, 7.3%) (Table 2, Fig. 2).

Once evaluator agreement and results offered by the software were collected, Interclass Correlation Coefficient (ICC) was calculated to analyze the level of agreement between both parties for all variables

Table 1

Operators and Convolutional Neural Network mean and standard deviation for detection capacity of each study variable included in the study.

Variable	Operator's Average (Mean)	CNN Average (Mean)
Total N Teeth Detected	28.110 ± 4.139	28.120 ± 4.126
Total N Restorations Detected	3.460 ± 3.352	3.420 ± 3.322
Total N Metallic Restorations Detected	1.250 ± 2.114	1.260 ± 2.115
Total N Resin-Based Restorations Detected	2.210 ± 2.763	2.160 ± 2.741
Total N Root Canal Treatments Detected	0.880 ± 1.488	0.880 ± 1.489
Total N Crowns Detected	1.320 ± 3.029	1.320 ± 3.030
Total N Implants Detected	0.300 ± 1.001	0.300 ± 1.001

Table 2

List of types of errors made by the software in the detection and classification of dental structures (ET0, correct numbering classification; ET1, mistaken detection of implant-supported crowns with natural dental structures; ET2, mistaken detection of pontics with natural dental structures; ET3, mistaken numbering classification between dental pieces present in radiographic image; ET4, No detection of dental structures present in radiographic image; and ET5, No detection of supernumeraries and included teeth.).

		RESULTS	
		N	Percentage
ERROR TYPE	ET0	164	50,0%
	ET1	29	8,8%
	ET2	37	11,3%
	ET3	68	20,7%
	ET4	6	1,8%
	ET5	24	7,3%
Total		328	100,0%

included in the study. Results obtained were always higher than 0.98, indicating high performance of the AI based software (Table 3).

Table 4 summarizes Sensitivity, Specificity, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) calculated for each variable included in the present study. The neural network showed an overall accuracy of $> 80\%$. Values dropped under 50% only for the detection of resin-based restorations (41.11%) and the total detection of restorative treatments (both metallic and resin-based).

Likewise, values obtained from each variable were represented in

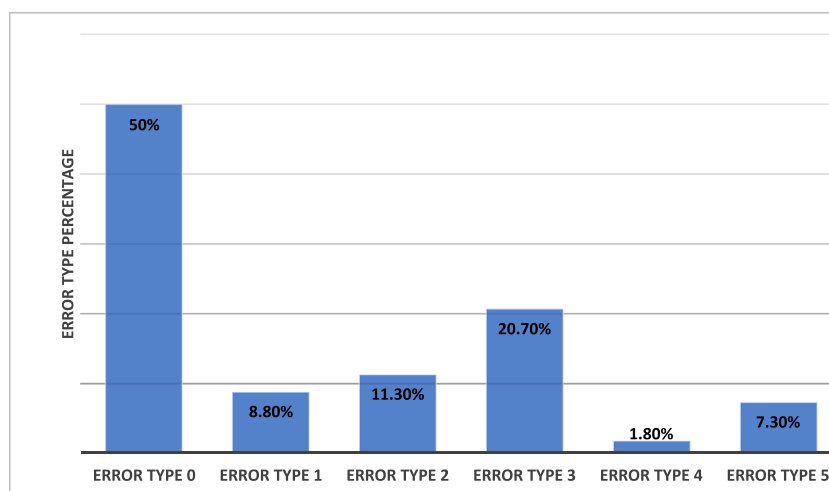


Fig. 2. Graphic representation of the types of errors made by the software in the detection and classification of dental structures.

Table 3

Interclass Correlation Coefficient obtained for each variable included in the study.

Variable	Interclass Correlation Coefficient (ICC) ($p < 0.05$)
Total N Teeth Detected	0.999
Total N Restorations Detected	0.996
Total N Metallic Restorations Detected	0.995
Total N Resin-Based Restorations Detected	0.985
Total N Root Canal Treatments Detected	0.998
Total N Crowns Detected	1.000
Total N Implants Detected	1.000

Table 4

Representation of Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value for each variable included in the study.

Variable	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Total N Restorations Detected	56.95%	92.20%	95.48%	42.51%
Total N Metallic Restorations Detected	85.48%	87.50%	82.8%	89.53%
Total N Resin-Based Restorations Detected	41.11%	93.30%	90.24%	87.50%
Total N Root Canal Treatments Detected	91.9%	100%	100%	94.62%
Total N Crowns Detected	89.53%	95.79%	89.53%	97.79%
Total N Implants Detected	100%	100%	100%	100%

ROC Curves offering their corresponding areas under the curve (AUCs) higher than 0.80 except for resin-based restorations (0.672) (Fig. 3).

4. Discussion

The aim of this study was to evaluate the clinical efficacy in terms of accuracy of an AI web-based software for helping professionals in the automatic detection and classification of dental structures and treatments present in panoramic radiographs. According to the results, the

present study found that the AI-based software could be considered as a helpful tool to increase the efficiency on the detection of dental structures present in panoramic radiographs (Gold Standard Detection Average 28.110 ± 4.139 ; CNNs Detection Average 28.120 ± 4.126). However, certain errors were observed when the program proceeded to their classification (FDI nomenclature), being the most repetitive one the confusion of natural teeth with pontics or implant-supported crowns. Also, the software failed to detect and classify included teeth, supernumeraries or teeth present in the image that were associated with artifacts or blurred areas. Hence, the null hypothesis was partially rejected.

Nowadays, the application of Artificial Intelligence in the medical field can be considered as one of the most remarkable advances achieved in the last decades. This novelty has already been applied in multiple dental specialties, such as orthodontics or general dentistry [7–9]. Regarding dental radiodiagnosis, several AI-based systems have been developed to focus their attention in 2D images – such as Denti.Ai, Overjet or Pearl – and numerous studies have attempted the automation of diagnosis for both pathologies and treatments in intraoral radiographs [30], but not too many studies targeted this new technology in orthopantomographies [30,31]. Despite this, other AI-based softwares have been developed to analyze and automatically detect anatomical structures in 3D imagery, such as Diagnocat, or Relu.

As detailed in several AI research articles, the use of several neural networks with different detection tasks can improve the output of the CNN. This way, results obtained from their synergy have proven to have a faster learning stage and more precise findings [4,29]. Leite et al. carried out an investigation in which positive results were obtained in the detection, numbering and dental segmentation in panoramic radiographs. Specifically, authors pointed out that the best performance was achieved by the combination of CNN DeepLab-v3 and ResNet 101 CNNs [32]. In the present study, the used CNN also functions with two different neural networks that ultimately help to achieve better performance: latest version of the Faster R-CNN (Region Based Convolutional Neural Network) model is used for radiographic detection of teeth, and VGG-16 computer architecture is utilized for achieving the FDI numbering of the dental structures identified [29]. A recent systematic review reported that these Deep Learning models are considered the best to perform object detection precision [33].

As presented in the article published by Tuzoff et al., results for the dental detection task of the AI software were very positive. Such is, that their sensitivity and specificity values reached 99.41% and 99.45% respectively. Another research where a similar deep learning model was used – R-CNN – also achieved high detection accuracy on dental structures (F1 Score 0.95) [34]. Our results noted a sensitivity of 69.3% for

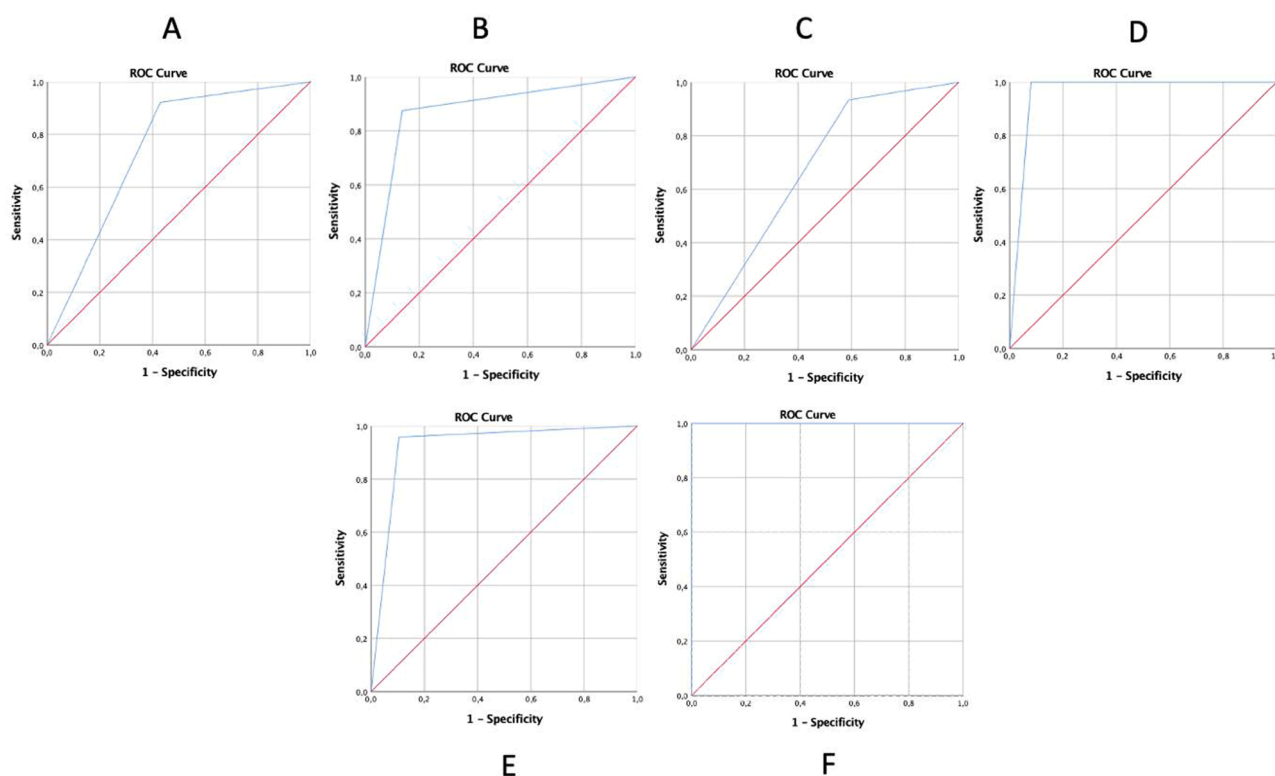


Fig. 3. ROC Curve representations for each variable. A) Total of Restorations, B) Total of Metallic Restorations, C) Total of Resin-Based Restorations, D) Endodontic Treatments, E) Dental and Implant supported Crowns and Pontics and F) Osseointegrated Implants.

this task, being this the percentage of cases in which the software detected the same structures as the Ground Truth did. Regarding the task of numbering tooth structures, Tuzoff et al. reported sensitivity and specificity values of 98% and 99.94%, respectively [29]. Our results showed that the program would only succeed in 50% of the cases, being the remaining ones attributed to previously mentioned errors.

Results obtained by Bilgir et al. in 2021 also demonstrated that an AI based-software could be considered as a successful way to detect and number the teeth present in panoramic radiographs, showing sensitivity and a PPV value of 0.9559 and 0.9652, respectively. This results were also proved by Leite et al., where both the sensitivity and the precision (PPV) obtained were greater than 0.98 [4,10].

Other studies where the same AI-based software was used describe good results in the detection of teeth in good quality radiographic images both with normal dental positions and in cases with overlapping dental structures. The authors also reported successful results in cases of blurry or lower quality images, as well as differentiating natural pieces from the ones replaced with pontics and implants, contrary to what was found in our study. It was mentioned that the cases in which the program was unable to detect tooth structures (false negatives) were associated with small root remnants - difficult to detect even by clinicians, also supported by Vinayahalingam et al. - presence of orthopedic objects, or large dental crowding [34]. Regarding the numbering provided by the program, authors reflected that the main errors were associated with the absence of teeth adjacent to the dental structures present in the radiograph. A higher frequency of this error was observed being related to molar regions, also supported by findings in our research [4,10,29,35].

Regarding the detection and classification of dental restorations, Abdalla-Aslan et al. obtained satisfactory results in their study. Their article describes a high diagnostic performance by the AI-based system, obtaining an overall sensitivity for the task of 94.6% and an overall specificity of 99.4%. In this case, the used CNN based its decision-making on different gray threshold values and the position they had within the image. As reported by the authors, this mechanism works

well in regions where objects of two different densities are clearly differentiated, but not in those where gray balance tends to be similar. This is the reason why lower values in the detection and classification of treatments such as composite fillings or root canal treatments were obtained (83.1% and 83.2%, respectively). This was also described in Vinayahalingam et al. findings, where lowest F1 score was attributed to root canal fillings (0.873) [34]. The results obtained in our study gave a lower detection sensitivity to treatments like resin-based restorations (41.11%). However, it should be noted that the cases in which the AI-based software presented higher resin-based restoration detection error were associated with cases where even the operators struggled identifying [1].

In the same way as described in Abdalla-Aslan et al. results, satisfactory performance was achieved in the detection of metal fillings (85.58%) and implants (100%) in our study [1], although our CNN sometimes confused crowns (89.53%) with extensive metal restorations. This was contrary to what was found on other articles, where implant detection statistical charts obtained a F1 score of 0.809 [34].

The present study has certain limitations. Nowadays the applications of the AI-based tools are being widely studied in 3D imagery, as it is thought to be its near future application. Also, it has to be pointed that a single radiodiagnosis device was used, so the diagnostic capacity of both the software and the human operators may have been influenced by the quality characteristics of the images recruited [33]. Likewise, the diagnostic capabilities of the software could have been better tested by performing certain modifications in each image (exposure, contrast, etc). By not undergoing them, the software was only trained in the identification of structures and treatments on good quality images [4]. Another limitation of this study was not utilizing two separate groups of imaginary sets for software's training and testing phases - only testing images were used to carry out this research. Most of the published articles about these AI-based technologies used two sets of radiographs for both above-mentioned purposes. This was intended to correct system failures so that it would "learn" to correctly detect dental

structures/treatments in the final testing images [33]. Additionally, in the present study, erroneous annotations made by the program were not corrected by the operators. This fact assumes that the program understood that what it detected was correct - when it certainly was not - and, as a result, learning capacity of the software could not be measured. It should also be observed that the number of treatments per variable included were not equally present in the total number of radiographs (e. g. total number of implants detected was not the same as the total number of resin-based restorations or root canal treatments). Thus, for a better understanding of both CNNs' detection and classification capacities, the same number of treatments for each variable should be included in later studies (assuring equity in results analysis).

Suggestions to optimize the AI algorithm and increase the neural networks' accuracy might include more heterogeneous imagery - avoiding underrepresented pathologies or treatments - to make the results obtained more generalizable. Also, more independent dental operators experts should contribute to determine gold standard values and datasets should include images that were not previously uploaded (same patient years later, for instance) to avoid "data snooping bias" [26]. It also needs to be considered the importance to have an ethical board approval and to follow AI ethics-related guidelines when approaching to develop a research concerning this topic. As Mörch et al. stated on their recently published article, this observation underlines the importance for future professionals to receive formal academic training on the ethical and social impacts of AI technologies [36]. In this study, both considerations were taken into account and successfully achieved. Also, literature concerning new AI-based software's have insisted on the usefulness of these tools to dental clinicians, since subjectivity and self-doubt play an important role on their daily working routine. Allowing an objective diagnosis of both present pathologies and treatments may be a starting point to what may come next in digital dentistry.

5. Conclusions

Findings from the present study suggest that the use of a web-based convolutional neural network can provide an overall reliable performance in the detection and classification of dental structures and treatments in panoramic radiographs with some specific limitations. General outcomes for both variables - dental structures and treatments - was higher than 70%, despite resin-based restorations, which detection accuracy was inferior to 50%. However, due to the deep learning characteristics of the AI-based software and the limitations explored in this study, a greater quantity of data has to be uploaded and analyzed to obtain better and more reliable results in the near future.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or non-for-profit sectors.

CRediT authorship contribution statement

Monica Bonfanti-Gris: Project administration, Methodology, Data curation, Writing - original draft. **Angel Garcia-Cañas:** Methodology, Writing - review & editing. **Raul Alonso-Calvo:** Methodology, Data curation, Supervision. **Maria Paz Salido Rodriguez-Manzaneque:** Project administration, Methodology, Data curation, Supervision. **Guillermo Pradies Ramiro:** Conceptualization, Project administration, Methodology, Data curation, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare no conflict of interest.

References

- [1] R. Abdalla-Aslan, T. Yeshua, D. Kabla, I. Leichter, C. Nadler, An artificial intelligence system using machine-learning for automatic detection and classification of dental restorations in panoramic radiography, *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* (2020), <https://doi.org/10.1016/j.oooo.2020.05.012>.
- [2] F. Kunz, A. Stellzig-Eisenhauer, F. Zeman, J. Boldt, Artificial intelligence in orthodontics: Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network, *J. Orofac. Orthop.* 81 (2020) 52–68, <https://doi.org/10.1007/s00056-019-00203-8>.
- [3] A.G. Cantu, S. Gehrung, J. Krois, A. Chaurasia, J.G. Rossi, R. Gaudin, K. Elhennawy, F. Schwendicke, Detecting caries lesions of different radiographic extension on bitewings using deep learning, *J. Dent.* 100 (2020), 103425, <https://doi.org/10.1016/j.jdent.2020.103425>.
- [4] A.F. Leite, A. Van Gerven, H. Willems, T. Beznik, P. Lahoud, H. Gaëta-Araujo, M. Vranckx, R. Jacobs, Artificial intelligence-driven novel tool for tooth detection and segmentation on panoramic radiographs, *Clin. Oral Investig.* (2020) 1–11, <https://doi.org/10.1007/s00784-020-03544-6>.
- [5] F. Schwendicke, W. Samek, J. Krois, Artificial intelligence in dentistry: chances and challenges, *J. Dent. Res.* 99 (2020) 769–774, <https://doi.org/10.1177/0022034520915714>.
- [6] A. Gomolin, E. Netchiporouk, R. Gniadecki, I.V. Litvinov, Artificial intelligence applications in dermatology: where do we stand? *Front. Med.* 7 (2020) 1–7, <https://doi.org/10.3389/fmed.2020.00100>.
- [7] F. Kunz, A. Stellzig-Eisenhauer, F. Zeman, J. Boldt, Artificial intelligence in orthodontics. Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network, *J. Orofac. Orthop.* (2020) 52–68, <https://doi.org/10.25259/ajos.123.2019>.
- [8] T. Takahashi, K. Nozaki, T. Gonda, T. Mamenno, M. Wada, K. Ikebe, Identification of dental implants using deep learning—pilot study, *Int. J. Implant Dent.* 6 (2020) 53, <https://doi.org/10.1186/s40729-020-00250-6>.
- [9] B. Erpaçal, Ö. Adıgüzel, S. Cangül, The use of micro-computed tomography in dental applications, *Int. Dent. Res.* 9 (2019) 78–91, <https://doi.org/10.5577/intdntres.2019.vol9.no2.7>.
- [10] E. Bilgir, I.Ş. Bayraktar, Ö. Çelik, K. Orhan, F. Akkoca, H. Sağlam, A. Odaş, A. F. Aslan, C. Özçetin, M. Kılı, I. Rozylo-Kalinowska, An artificial intelligence approach to automatic tooth detection and numbering in panoramic radiographs, *BMC Med. Imaging.* 21 (2021) 1–9, <https://doi.org/10.1186/s12880-021-00656-7>.
- [11] S. Mertens, J. Krois, A.G. Cantu, L.T. Arsiwala, F. Schwendicke, Artificial intelligence for caries detection: Randomized trial, *J. Dent.* 115 (2021), <https://doi.org/10.1016/j.jdent.2021.103849>.
- [12] H. Mohammad-Rahimi, S.R. Motamedian, M.H. Rohban, J. Krois, S.E. Uribe, E. Mahmoudinia, R. Rokhsad, M. Nadimi, F. Schwendicke, Deep learning for caries detection: a systematic review, *J. Dent.* 122 (2022), 104115, <https://doi.org/10.1016/j.jdent.2022.104115>.
- [13] Z. Liu, J. Liu, Z. Zhou, Q. Zhang, H. Wu, G. Zhai, J. Han, Differential diagnosis of ameloblastoma and odontogenic keratocyst by machine learning of panoramic radiographs, *Int. J. Comput. Assist. Radiol. Surg.* 16 (2021) 415–422, <https://doi.org/10.1007/s11548-021-02309-0>.
- [14] J. Kim, H.S. Lee, I.S. Song, K.H. Jung, DeNTNet: Deep Neural Transfer Network for the detection of periodontal bone loss using panoramic dental radiographs, *Sci. Rep.* 9 (2019) 1–10, <https://doi.org/10.1038/s41598-019-53758-2>.
- [15] Y. Cheng Guo, M. Han, Y. Chi, H. Long, D. Zhang, J. Yang, Y. Yang, T. Chen, S. Du, Accurate age classification using manual method and deep convolutional neural network based on orthopantomogram images, *Int. J. Legal Med.* 135 (2021) 1589–1597, <https://doi.org/10.1007/s00414-021-02542-x>.
- [16] S. Kim, Y.H. Lee, Y.K. Noh, F.C. Park, Q.S. Auh, Age-group determination of living individuals using first molar images based on artificial intelligence, *Sci. Rep.* 11 (2021) 1–11, <https://doi.org/10.1038/s41598-020-80182-8>.
- [17] M. Hadj Saïd, M.-K. Le Roux, J.-H. Catherine, R. Lan, Development of an artificial intelligence model to identify a dental implant from a radiograph, *Int. J. Oral Maxillofac. Implants.* 35 (2020) 1077–1082, <https://doi.org/10.11607/jomi.8060>.
- [18] S. Sukeyawa, K. Yoshii, T. Hara, K. Yamashita, K. Nakano, N. Yamamoto, H. Nagatsuka, Y. Furuki, Deep neural networks for dental implant system classification, *Biomolecules* 10 (2020) 1–13, <https://doi.org/10.3390/biom10070984>.
- [19] S. Sukeyawa, K. Yoshii, T. Hara, T. Matsuyama, K. Yamashita, K. Nakano, K. Takabatake, H. Kawai, H. Nagatsuka, Y. Furuki, Multi-task deep learning model for classification of dental implant brand and treatment stage using dental panoramic radiograph images, *Biomolecules* (2021) 11, <https://doi.org/10.3390/biom11060815>.
- [20] E. Shaheen, A. Leite, K.A. Alqahtani, A. Smolders, A. Van Gerven, H. Willems, R. Jacobs, A novel deep learning system for multi-class tooth segmentation and classification on cone beam computed tomography. A validation study: deep learning for teeth segmentation and classification, *J. Dent.* (2021) 115, <https://doi.org/10.1016/j.jdent.2021.103865>.
- [21] R.C. Fontenele, M. do N. Gerhardt, J.C. Pinto, A. Van Gerven, H. Willems, R. Jacobs, D.Q. Freitas, Influence of dental fillings and tooth type on the performance of a novel artificial intelligence-driven tool for automatic tooth segmentation on CBCT images - a validation study, *J. Dent.* (2022) 119, <https://doi.org/10.1016/j.jdent.2022.104069>.
- [22] M. do N. Gerhardt, R.C. Fontenele, A.F. Leite, P. Lahoud, A. Van Gerven, H. Willems, A. Smolders, T. Beznik, R. Jacobs, Automated detection and labelling of teeth and small edentulous regions on cone-beam computed tomography using

- convolutional neural networks, *J. Dent.* (2022) 122, <https://doi.org/10.1016/j.jdent.2022.104139>.
- [23] P.J. Verhelst, A. Smolders, T. Beznik, J. Meewis, A. Vandemeulebroucke, E. Shaheen, A. Van Gerven, H. Willems, C. Politis, R. Jacobs, Layered deep learning for automatic mandibular segmentation in cone-beam computed tomography, *J. Dent.* 114 (2021), <https://doi.org/10.1016/j.jdent.2021.103786>.
- [24] P. Lahoud, S. Diels, L. Niclaes, S. Van Aelst, H. Willems, A. Van Gerven, M. Quirynen, R. Jacobs, Development and validation of a novel artificial intelligence driven tool for accurate mandibular canal segmentation on CBCT, *J. Dent.* 116 (2022), 103891, <https://doi.org/10.1016/j.jdent.2021.103891>.
- [25] S. Shujaat, O. Jazil, H. Willems, A. Van Gerven, E. Shaheen, C. Politis, R. Jacobs, Automatic segmentation of the pharyngeal airway space with convolutional neural network, *J. Dent.* 111 (2021), 103705, <https://doi.org/10.1016/j.jdent.2021.103705>.
- [26] F. Schwendicke, T. Singh, J.H. Lee, R. Gaudin, A. Chaurasia, T. Wiegand, S. Uribe, J. Krois, Artificial intelligence in dental research: checklist for authors, reviewers, readers, *J. Dent.* (2021) 107, <https://doi.org/10.1016/j.jdent.2021.103610>.
- [27] P.M. Bossuyt, J.B. Reitsma, D.E. Bruns, C.A. Gatsonis, P.P. Glasziou, L. Irwig, J. G. Lijmer, D. Moher, D. Rennie, H.C.W. De Vet, H.Y. Kressel, N. Rifai, R.M. Golub, D.G. Altman, L. Hooft, D.A. Korevaar, J.F. Cohen, STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies, *BMJ* 351 (2015) 1–9, <https://doi.org/10.1136/bmj.h5527>.
- [28] J. Mongan, L. Moy, C.E. Kahn, Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers, *Radiol. Artif. Intell.* 2 (2020), e200029, <https://doi.org/10.1148/ryai.2020200029>.
- [29] D.V. Tuzoff, L.N. Tuzova, M.M. Bornstein, A.S. Krasnov, M.A. Kharchenko, S. I. Nikolenko, M.M. Sveshnikov, G.B. Bednenko, Tooth detection and numbering in panoramic radiographs using convolutional neural networks, *Dentomaxillofac. Radiol.* 48 (2019) 1–10, <https://doi.org/10.1259/dmfr.20180051>.
- [30] J.H. Lee, S.S. Han, Y.H. Kim, C. Lee, I. Kim, Application of a fully deep convolutional neural network to the automation of tooth segmentation on panoramic radiographs, *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* 129 (2020) 635–642, <https://doi.org/10.1016/j.oooo.2019.11.007>.
- [31] K. Moutselos, E. Berdouses, C. Oulis, I. Maglogiannis, Recognizing occlusal caries in dental intraoral images using deep learning, in: *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, 2019, pp. 1617–1620, <https://doi.org/10.1109/EMBC.2019.8856553>.
- [32] T. Hiraiwa, Y. Arij, M. Fukuda, Y. Kise, K. Nakata, A. Katsumata, H. Fujita, E. Arij, A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography, *Dentomaxillofac. Radiol.* 48 (2019) 1–7, <https://doi.org/10.1259/dmfr.20180218>.
- [33] F. Umer, S. Habib, N. Adnan, Application of deep learning in teeth identification tasks on panoramic radiographs, *Dentomaxillofac. Radiol.* (2022), <https://doi.org/10.1259/dmfr.20210504>.
- [34] S. Vinayahalingam, R. Shan Goey, S. Kempers, J. Schoep, T. Cheric, D.A. Moin, M. Hanisch, Automated chart filing on panoramic radiographs using deep learning, *J. Dent.* (2021) 115, <https://doi.org/10.1016/j.jdent.2021.103864>.
- [35] S.-N. Lee, Jae-hong; Jeong, Efficacy of deep convolutional neural network algorithm for the identification and classification of dental implant systems, using panoramic and periapical radiographs, *Open Med.* 99 (2020) 1–7.
- [36] C.M. Mörch, S. Atsu, W. Cai, X. Li, S.A. Madathil, X. Liu, V. Mai, F. Tamimi, M. A. Dilhac, M. Ducret, Artificial intelligence and ethics in dentistry: a scoping review, *J. Dent. Res.* 100 (2021) 1452–1460, <https://doi.org/10.1177/00220345211013808>.