Critical Review
Imaginology

Demystifying artificial intelligence and deep learning in dentistry

Jonas Almeida RODRIGUES[(a)]
Joachim KROIS[(b)]
Falk SCHWENDICKE[(b)]

[(a)]Universidade Federal do Rio Grande do Sul – UFRGS, School of Dentistry, Department of Surgery and Orthopedics, Porto Alegre, RS, Brazil.

[(b)]Charité – Universitätsmedizin Berlin, Digital Health and Health Services Research, Department of Oral Diagnostics, Berlin, Germany

Declaration of Interests: The authors certify that they have no commercial or associative interest that represents a conflict of interest in connection with the manuscript.

Corresponding Author:
Jonas Almeida Rodrigues
E-mail: jorodrigues@ufrgs.br

https://doi.org/10.1590/1807-3107bor-2021.vol35.0094

Submitted: April 25, 2021
Accepted for publication: July 5, 2021
Last revision: July 7, 2021

Abstract: Artificial intelligence (AI) is a general term used to describe the development of computer systems which can perform tasks that normally require human cognition. Machine learning (ML) is one subfield of AI, where computers learn rules from data, capturing its intrinsic statistical patterns and structures. Neural networks (NNs) have been increasingly employed for ML complex data. The application of multilayered NN is referred to as "deep learning", which has been recently investigated in dentistry. Convolutional neural networks (CNNs) are mainly used for processing large and complex imagery data, as they are able to extract image features like edges, corners, shapes, and macroscopic patterns using layers of filters. CNN algorithms allow to perform tasks like image classification, object detection and segmentation. The literature involving AI in dentistry has increased rapidly, so a methodological guidance for designing, conducting and reporting studies must be rigorously followed, including the improvement of datasets. The limited interaction between the dental field and the technical disciplines, however, remains a hurdle for applicable dental AI. Similarly, dental users must understand why and how AI applications work and decide to appraise their decisions critically. Generalizable and robust AI applications will eventually prove helpful for clinicians and patients alike.

Keywords: Artificial Intelligence; Deep Learning; Neural Networks, Computer; Diagnostic Imaging; Dentistry.

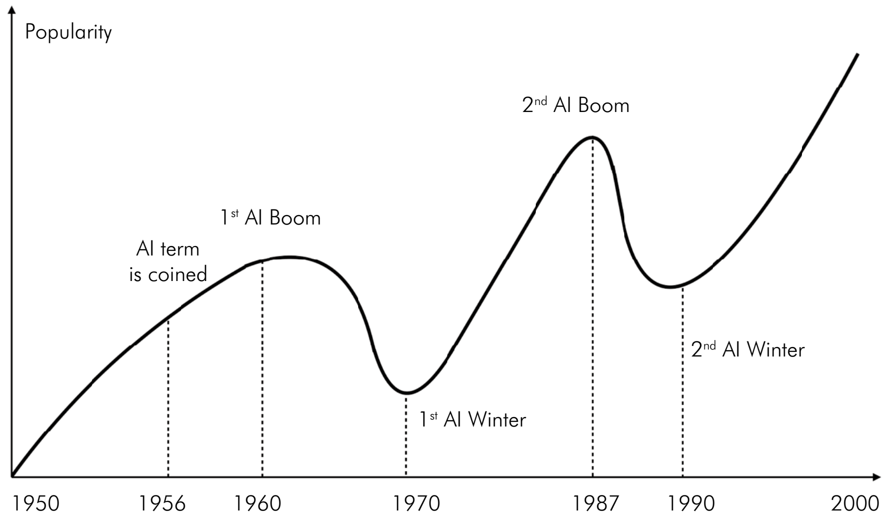
Introduction

Artificial intelligence (AI) has recently attracted significant public interest and is impacting many industries worldwide. Especially in healthcare it promises to be truly transformative. AI has a potential to transfer time-consuming human tasks to machines, improving patient outcomes[1] at higher safety and efficiency and relieving burdened healthcare providers and systems.

The excitement around AI is not a new one. The term was first used in 1956; since then, AI has lived through phases of excitement and disappointment ("AI winters") (Figure 1).

Recently, the advances in data availability, for example via electronic health records and digital imaging, the growth in computational power and the development of software approaches allowing to employ big

A



B

1938 - 1946: Golden age of science fiction

1950 - 1960: First AI Boom: Can machines think? Turing test, method to determine the intelligence of a computer by A. Turing; Logic Theorists, the fist AI programme is created; Dartmouth summer conference on AI, the term AI is used for the first time;

1965: ELIZA, a natural language programme is created, which simulates a conversation with a psychotherapist;

1970s: First AI winter;

1980s - 1990s: Second AI Boom: E. Feigenbraum introduces expert systems, which emulates decisions of human experts;

1990s: Second AI winter;

1997: Deep Blue, a chess-playing computer beats the world champion Gary Kasparov;

2006: Geoffrey Hinton from the University of Toronto develops Deep Learning and publishes "A Fast Learning Algorithm for Deep Belief Nets";

2011 - 2014: Watson (IBM) wins jeopardy; Apple integrates Siri, an assistant with voice; Alexa, virtual assistant from Amazon;

2018: European Union establishes guidelines for dealing with ethics in AI;

Figure 1. Timeline illustration of AI development

data fuel an unseen optimism: In healthcare, AI has been successfully employed for automated evaluation of imagery such as mammograms for breast cancer detection,[2] photographs for skin cancer screening[3] and eye examinations for assessing diabetic retinopathy.[4] In dentistry, AI has been used to aid the detection of carious lesions on bitewing radiographs,[5] for periodontal bone loss detection in periapical radiographs[6] and panoramics,[7] to assess periapical lesions on panoramics[8] and peri-apicals,[9] to detect carious lesions on near-infrared light transillumination imagery,[10] to diagnose tumors in the mandible in panoramic radiographs,[11] or for the classification of restorations in periapical radiographs.[12] Schwendicke et al.,[13] through a scoping review, discussed the literature on AI in dentistry and found that AI is being widely used for imaging diagnosis, but that the usefulness, safety and generalizability of many AI applications remained unclear at present.

Definitions of AI

AI is a general term used to describe the theory and development of computer systems which can perform tasks that normally require human cognition (perception, language, understanding, reasoning, learning, planning, and problem solving).[14] In healthcare "narrow AI" applications are currently in the focus, referring to AI systems that are specified to handle a singular or limited task, with human cognition remaining needed. Narrow AI systems lack the self-awareness, consciousness, and genuine intelligence to match human intelligence, and do not encompass wider skill sets required for complex decision-making (i.e. therapy decisions integrating a wide range of data sources, but also patients' expectations and the clinician's experience, among others). In informatics, there is a theorem which states that a person with a computer is better than a person alone,[14] and the same may apply to AI; it "augments" human intelligence rather than replacing it.

Machine learning (ML) is one prolific subfield of AI, where computers learn rules from data (rather than humans providing these rules). This learning from data, capturing its intrinsic statistical patterns and structures,[15] can be done in various ways, the three most popular ML domains are:

Supervised learning

The most common learning strategy. Data and data labels (output) are provided, and the ML model is iteratively optimized towards representing this data-label pair. For example, pictures of dogs labeled "dog" allow the machine to develop an algorithm which can eventually classify new, unseen images (dog present yes/no). Supervised learning is resource-intensive; especially in medicine establishing a large number of labels is challenging.

Unsupervised learning

It is used to understand the structure and relationships among input features rather than try to predict an outcome label from them. The data given to the learning algorithm is unlabeled, and the algorithm is asked to identify patterns in the input data. Examples are a recommendation system of an e-commerce website where the learning algorithm discovers items often bought together or clustering genetic patterns to analyze evolutionary biology.

Reinforcement learning

It is used for the computer to learn how to make decisions on its own, with the consequences of those decisions potentially appearing much later after the decisions were made. It differs from other forms of supervised learning because the sample dataset does not train the machine. Instead, it learns by trial and error. Therefore, a series of right decisions would strengthen the method as it better solves the problem. Reinforcement learning is used to solve different games and achieve superhuman performance.

Among those domains mentioned above the supervised learning approach is the most widely applied in healthcare. The performance of a supervised learning algorithm is measured on a different and independent set of examples called a test set. This serves to test the generalization ability of the machine — its ability to produce sensible answers on new inputs that it has never seen during training.[16] On this test set, performance metrics can be calculated, such as sensitivity, specificity, area under the ROC curve.

Data employed for ML can be simple or complex; for the latter, neural network (NNs) have been increasingly employed.[15] NNs build on the idea of artificial neurons, which are semi-parametric mathematical non-linear models. When these neurons are organized in layers of different form and size and connected using mathematical operations, classification and regression tasks might be performed.[17] The application of multilayered neural networks is referred to as "deep learning" (DL).[18] DL is especially suitable complex data, like imagery, and its application has been investigated in dentistry. DL will be now discussed in more details. Figure 2 shows the hierarchy of AI.

Deep learning (DL) in dentistry

One particular type of DL involves Convolutional Neural Networks (CNNs). CNNs are mainly used for processing large and complex imagery data, as they are able to extract image features like edges, corners, shapes, and macroscopic patterns[16] using layers of filters. CNNs allow to perform tasks like image classification ("does this image contain a caries lesion?"), object detection ("where on this image is a tooth or a caries lesion?"), and segmentation ("which pixels are affected by caries?").[13]

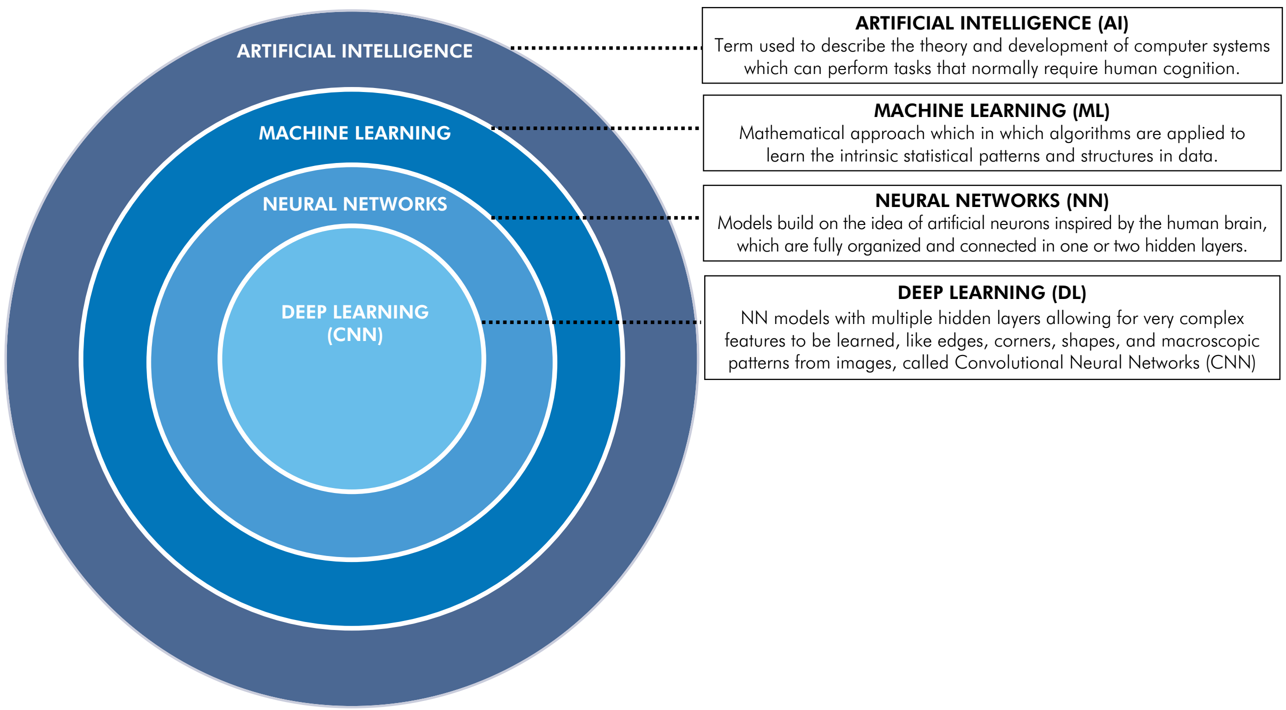


Figure 2. Hierarchy of AI

For image analysis (a field of AI termed as "Computer vision") via CNNs, it is usually (at least when employing supervised learning) necessary to provide labelled imagery. One major difficulty here is establishing a truthful label; in many instances, a hard ground truth (gold standard), for example via histologic assessment, cannot be established. Hence, labelling usually involves multiple human experts who classify the same images (for example, presence or absence of an apical lesion). Having multiple experts allows to overcome the individual limitations of each expert but results in a "fuzzy" ground truth.[19] For example, five experts may have assessed an image and four classified it as showing an apical lesion, while one deviated and thought that no lesion was present. Various ways of unifying these fuzzy labels (majority votes etc.) are available, while admittedly, all come with limitations: In our example, we cannot know if the single, deviating expert is nevertheless right. Triangulation from other data, for example electronic health records or clinical assessments, may help here.

CNNs showed high performance for tooth segmentation and identification,[20-24] for dental implants planning;[25] for biofilm classification on fluorescence images;[26] for diagnosing maxillary sinusitis on panoramic radiography,[27] for cephalometric landmarks detection,[28] or for root morphological classification.[29]

For caries detection, CNNs have also shown good performance on periapical[30] and bitewing[31,32] imagery. Recently, a large sample of 3686 bitewings images were assessed by four experienced dentists and a CNN algorithm was trained and validated for approximal caries detection. The neural network showed an accuracy of 0.80 compared to 0.70 obtained by the experienced dentists, being more sensitive than those.[32]

Challenges and key considerations on AI

The number of studies involving AI in healthcare increased rapidly, although many concerns remain as to the applied methods to develop, validate, test and eventually deploy them. Therefore, it is recommended that studies on AI follow as rigorously as possible a methodological guidance.[33]

Studies on AI in dentistry should be planned, conducted and reported keeping in mind that the desirable endpoint might be a clinical application, even if this distance seems to be long and hard to achieve. Another important point to be considered is having a clear definition of the study aim and of the datasets that will be used. Most recently published studies used small, imbalanced or homogenous datasets stemming from one specific population. Increasing the size of training datasets is desirable for clinical applications, as robustness increases; expanding beyond data from one population also strengthens generalizability.[34]

The definition of a reference test is also an important step during planning and conducting the study. Depending on the study design, several independent or joint annotators might be necessary in order to label the data to be learnt and tested on. In this case, they would constitute the gold standard for the model. This should be also clearly reported.

An independent dataset sample should be used for testing the model, and this should be planned and designed in advance. This sample should not have been used for training the model. In order to assure a certain generalizability or robustness it is suggested that a completely external dataset be used at this stage.

Involving dental expertise in the rather technical process of training and testing AI is essential: The interaction between the dental field and the technical disciplines is currently limited, but dental domain knowledge seems crucial to develop an applicable, useful AI, but also to assess AI for its inherent decision logic (explainability, transparency, see below).

Efforts are essential in order to develop a public dataset, such as in the medical field to build algorithms that can be used in clinical applications. It would be ideal that researchers release the data used in their studies with removal of personal information. However, this is challenging as legal and institutional support from each stakeholder would be necessary. Specific regulatory laws should be encouraged in order to allow the use of data for AI purposes. The development of a common and free repository that can reliably collect, catalog, and archive publicly available data would be valuable in the dental field. This repository should reflect the wealth of conditions, populations, data sources (*e.g.* image generation machines) and usecases.

Last but not least, humans must understand why and how important decisions or predictions are made by AI for health applications. The principles of transparency, interpretability and explainability must be ensured. Explainable AI (XAI) provides interpretable explanations in natural language or easier-to-understand presentations, allowing dentists, patients, and other stakeholders to understand why a decision is made by the AI applications, and thereby to question its validity to avoid undesirable consequences.

Conclusions

AI in dentistry is rapidly growing; many studies and applications lack, however, robustness and generalizability. Methodological guidance and rigour are needed when developing dental AI, and close interaction between the dental field and the technical disciplines is needed. Dental users must understand why and how AI applications work and decide to appraise their decisions critically. Generalizable and robust AI applications will eventually prove helpful for clinicians and patients alike.

Acknowledgement

To CAPES/PrInt - UFRGS funding program for the Visiting Senior Professorship (PVE Senior), grant 88887.194845/2018-00.

References

1. Forbes. AI and healthcare: a giant opportunity. 2019 [cited 2021 Apr 23]. Available from: https://www.forbes.com/sites/insights-intelai/2019/02/11/ai-and-healthcare-a-giant-opportunity/?sh=512423eb4c68

2. Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T, Boss A. Deep learning in mammography diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. Invest Radiol. 2017 Jul;52(7):434-40. https://doi.org/10.1097/RLI.0000000000000358
3. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 Feb;542(7639):115-8. https://doi.org/10.1038/nature21056
4. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016 Dec;316(22):2402-10. https://doi.org/10.1001/jama.2016.17216
5. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent. 2018 Oct;77:106-11. https://doi.org/10.1016/j.jdent.2018.07.015
6. Lee JH, Kim DH, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci. 2018 Apr;48(2):114-23. https://doi.org/10.5051/jpis.2018.48.2.114
7. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, et al. Deep learning for the radiographic detection of periodontal bone loss. Sci Rep. 2019 Jun;9(1):8495. https://doi.org/10.1038/s41598-019-44839-3
8. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, et al. Deep learning for the radiographic detection of apical lesions. J Endod. 2019 Jul;45(7):917-922.e5. https://doi.org/10.1016/j.joen.2019.03.016
9. Endres MG, Hillen F, Salloumis M, Sedaghat AR, Niehues SM, Quatela O, et al. Development of a deep learning algorithm for periapical disease detection in dental radiographs. Diagnostics (Basel). 2020 Jun;10(6):430. https://doi.org/10.3390/diagnostics10060430
10. Schwendicke F, Elhennawy K, Paris S, Friebertshäuser P, Krois J. Deep learning for caries lesion detection in near-infrared light transillumination images: a pilot study. J Dent. 2020 Jan;92:103260. https://doi.org/10.1016/j.jdent.2019.103260
11. Poedjiastoeti W, Suebnukarn S. Application of convolutional neural network in the diagnosis of Jaw tumors. Healthc Inform Res. 2018 Jul;24(3):236-41. https://doi.org/10.4258/hir.2018.24.3.236
12. 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 Mar;9(1):3840. https://doi.org/10.1038/s41598-019-40414-y
13. Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: A scoping review. J Dent. 2019 Dec;91:103226. https://doi.org/10.1016/j.jdent.2019.103226
14. Friedman CP. A "fundamental theorem" of biomedical informatics. J Am Med Inform Assoc. 2009 Mar-Apr;16(2):169-70. https://doi.org/10.1197/jamia.M3092
15. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: chances and Challenges. J Dent Res. 2020 Jul;99(7):769-74. https://doi.org/10.1177/0022034520915714
16. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May;521(7553):436-44. https://doi.org/10.1038/nature14539
17. Niño Sandoval TC, Guevara Pérez SV, González FA, Jaque RA, Infante Contreras C. [Use of artificial neural networks for mandibular morphology prediction through craniomaxillar variables in a póstero-anterior view]. Univ Odontol. 2016 Jun;35(74):21-8. Spanish. https://doi.org/10.11144/Javeriana.uo35-74.urna
18. Burt JR, Torosdagli N, Khosravan N, RaviPrakash H, Mortazi A, Tissavirasingham F, et al. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. Br J Radiol. 2018 Sep;91(1089):20170545. https://doi.org/10.1259/bjr.20170545
19. Walsh T. Fuzzy gold standards: approaches to handling an imperfect reference standard. J Dent. 2018 Jul;74 Suppl 1:S47-9. https://doi.org/10.1016/j.jdent.2018.04.022
20. Eun H, Kim C. Oriented tooth localization for periapical dental X-ray images via convolutional neural network. 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA) 2016. Institute of Electrical and Electronics Engineers; 2017. p. 1-7.
21. Wirtz A, Mirashi SG. Wesarg S Automatic teeth segmentation in panoramic x-ray images using a coupled shape model in combination with a neural network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. Springer; 2018. p. 712-9. (Lecture Notes in Computer Science, v. 11073)
22. Jader G, Fontineli J, Ruiz M, Abdalla K, Pithon M. Oliveira L Deep instance segmentation of teeth in panoramic X-Ray images. Proceedings. In: 31st Conference on Graphics, Patterns and Images, Institute of Electrical and Electronics Engineers; 2019. p. 400-7.
23. Singh P, Sehgal P. Numbering and Classification of panoramic dental images using 6-layer convolutional neural network. Pattern Recognit Image Anal. 2020 Jan;30(1):125-33. https://doi.org/10.1134/S1054661820010149
24. Zhu G, Piao Z. Kim SC Tooth detection and segmentation with mask R-CNN. 2020 International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2020. Institute of Electrical and Electronics Engineers; 2020. p. 70-2.
25. Jae-Hong L. Identification and classification of dental implant systems using various deep learning-based convolutional neural network architectures. Clin Oral Implants Res. 2019 Sep;30 S19:217. https://doi.org/10.1111/clr.175_13509

26. Imangaliyev S, Veen MH, Volgenant CM, Keijser BJ, Crielaard W, Levin E. Deep learning for classification of dental plaque images. Lecture Notes in Computer Science. Springer; 2016. p. 407-10. (Artificial Intelligence and Lecture Notes in Bioinformatics, v. 10122). https://doi.org/10.1007/978-3-319-51469-7_34
27. Murata M, Ariji Y, Ohashi Y, Kawai T, Fukuda M, Funakoshi T, et al. Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. Oral Radiol. 2019 Sep;35(3):301-7. https://doi.org/10.1007/s11282-018-0363-7
28. Arık SÖ, Ibragimov B, Xing L. Fully automated quantitative cephalometry using convolutional neural networks. J Med Imaging (Bellingham). 2017 Jan;4(1):014501. https://doi.org/10.1117/1.JMI.4.1.014501
29. Hiraiwa T, Ariji Y, Fukuda M, Kise Y, Nakata K, Katsumata A, et al. A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. Dentomaxillofac Radiol. 2019 Mar;48(3):20180218. https://doi.org/10.1259/dmfr.20180218
30. Choi J, Eun H, Kim C. Boosting proximal dental caries detection via combination of variational methods and convolutional neural network. J Signal Process Syst Signal Image Video Technol. 2018 Jan;90(1):87-97. https://doi.org/10.1007/s11265-016-1214-6
31. Srivastava MM, Kumar P, Pradhan L, Varadarajan S. Detection of tooth caries in bitewing radiographs using deep learning. ArXiv; 2017.
32. Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, Gaudin R, et al. Detecting caries lesions of different radiographic extension on bitewings using deep learning. J Dent. 2020 Sep;100:103425. https://doi.org/10.1016/j.jdent.2020.103425
33. Schwendicke F, Singh T, Lee JH, Gaudin R, Chaurasia A, Wiegand T, et al. Artificial intelligence in dental research: checklist for authors, reviewers, readers. J Dent. 2021 Apr;107:103610. https://doi.org/10.1016/j.jdent.2021.103610
34. Weese J, Lorenz C. Four challenges in medical image analysis from an industrial perspective. Med Image Anal. 2016 Oct;33:44-9. https://doi.org/10.1016/j.media.2016.06.023