The Automatic Detection of Caries in Third Molars on Panoramic Radiographs Using Deep Learning: A Pilot Study

Shankeeth Vinayahalingam

Radboud University Medical Center

Steven Kempers

Radboud University Medical Center

Lorenzo Limon

Radboud University Medical Center

Dionne Deibel

Radboud University Medical Center

Thomas Maal

Radboud University Medical Center

Stefaan Bergé

Radboud University Medical Center

Tong Xi (✉ Tong.Xi@radboudumc.nl)

Radboud University Medical Center

Marcel Hanisch

Universitätsklinikum Münster

Research Article

Keywords: panoramic radiographs, convolutional neural network, MobileNet V2, deep-learning

Posted Date: April 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-379636/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The objective of this study is to assess the diagnostic accuracy of dental caries on panoramic radiographs using deep-learning algorithms. A convolutional neural network (CNN) was trained on a reference data set consisted of 400 cropped panoramic images in the detection of carious lesions in mandibular and maxillary third molars, based on the CNN MobileNet V2. For this pilot study, the trained MobileNet V2 was applied on a test set consisting of 100 cropped OPG(s). The detection accuracy and the area-under-the-curve (AUC) were calculated. The proposed method achieved an accuracy of 0.87, a sensitivity of 0.87, a specificity of 0.86 and an AUC of 0.90 for the detection of carious lesions of third molars on OPG(s). A high diagnostic accuracy was achieved in caries detection in third molars based on the MobileNet V2 algorithm as presented. This is beneficial for the further development of a deep-learning based automated third molar removal assessment in future.

Introduction

The removal of third molars is one of the most commonly performed surgical procedures in oral surgery. Recent guidelines recommend the removal of pathologically erupting third molars in order to prevent future complications[1, 2]. The second molar is frequently disrupting the eruption path of the third molar, evoking it to only erupt partially or not at all, which can adversely affect periodontal health of the second molar. Impacted or partially erupted third molars are often the cause for various pathology such as pericoronitis, cysts, periodontal disease, damage to the adjacent tooth and carious lesions[3]. The prevalence of carious lesions in third molars is reported to range between 2,5% and 86%[4].

Orthopantomograms (OPG(s)) are generally used in the decision making process whether to remove a third molar or not. An OPG is an excellent diagnostic tool for planning the third molar removals, however, they are not suitable for the detection of dental caries. Previous studies have reported a lower caries diagnosis accuracy of OPG(s) compared to conventional bitewings and periapical radiographs[5]. To reduce the radiation exposure and to avoid taking additional radiographs, enhanced caries diagnostics on OPG(s) is useful.

In recent years, deep learning models like convolutional neural networks (CNNs) have been used to analyse medical images and to support the diagnostic procedure[6].

In the field of dentistry, CNNs have been applied for the detection of carious lesions on different image modalities such as periapical radiographs[7], bitewings[8], near-infrared light transillumination images[9, 10] and clinical photos[11, 12]. However, none of the studies have explored automated caries detection on OPG(s). The aim of this study is to train a CNN-based deep learning model for the detection of caries on third molars on OPG(s) and to assess the diagnostic accuracy.

Results

Table 1 summarizes the classification performance of MobileNet V2 on the test set. The diagnostic accuracy was 87%. The model achieved an AUC of 0.90 (Fig. 1). The confusion matrix is presented in Fig. 2.

Table 1
The Accuracy, precision, F1-score, recall and specificity for the detection of dental caries in third molars on OPG(s)

Accuracy	Precision	F1-score	Recall	Specificity
0.87	0.87	0.87	0.87	0.86

The class activation heatmaps (CAM) for carious third molar and non-carious third molar are illustrated in Fig. 3–4. These heatmaps visualize the discriminative regions used by the MobileNet V2 to classify the third molar (M3) as carious and non-carious. Optical inspection indicates a broader region of interest in non-carious M3(s). A more centered and focused region of interest is found for carious M3(s).

Discussion

A daily dilemma in dentistry and oral surgery is to determine whether a third molar should be removed or not. In cases of diseased third molar, where pain or pathology are obvious, there is a general consensus that surgical removal is indicated[3]. Improved diagnostics, e.g. on OPG(s), might improve the selection process whether to remove or not. A more stringent indication pathway may reduce millions of unnecessary third molar removals every year, so reducing comorbidity and health costs[13].

This pilot study assesses the capability of a deep learning model (MobileNet V2) to detect carious third molars on OPG(s) and is therefore a mosaic stone in the picture of automation of M3 removal diagnostics. Caries detection on third molars using OPG(s) is flawed by limited and varying accuracy of individual examiners leading to inconsistent decisions and consequently suboptimal care[8]. The use of deep neural networks might bring us a more reliable, faster and reproducible way of diagnosing pathology, and can therefore reduce the number of unnecessary third molar removals [14]. As previously stated, the assessment of third molars on OPG(s) is additionally interesting, as radiation exposure can be reduced when additional radiographs are avoided.

In dental radiology, previous studies applied deep learning models for caries detection on different image modalities[7–12, 15]. Lee JH et al. applied a pre-trained GoogLeNet Inception v3 CNN network on periapical radiographs achieving accuracies up to 0.89[7]. Casalegno et al. applied U-net with VGG-16 as an encoder, on near-infrared transillumination images (NITS)[9] with a reported AUC between 0.836 and 0.856. ResNet-18 and ResNext-50 were applied by Schwendicke et al. on NITS[10]. The reported AUC ranged from 0.730 to 0.856 in these studies. Two other studies explored the caries detection on clinical photos using Mask R-CNN with ResNet, reporting an accuracy of 0.870[12] and a F1-score of 0.889[11]. Finally, U-net with EfficientNet-B5 as an encoder was used to segment caries on bitewings with an accuracy of 0.8[8]. It is important to note that the performance of the deep learning models are highly

dependent on the dataset, the hyperparameters, the image modality and the architecture itself[6, 15]. As these parameters differed between the studies, a direct comparison of these studies would be misleading.

In this study, an accuracy of 0.87 and an AUC of 0.90 was achieved for caries detection on third molars on OPG(s). In comparison, previous studies have stated an AUC of 0.768 for caries detection on OPG(s) by clinicians[16]. Several factors are associated with the model performance. Firstly, the use of depthwise separable convolutions and the inverted residual with linear bottleneck reduced the number of parameters and the amount of memory constraint while retaining a high accuracy[17]. These characteristics make the MobileNet V2 less prone to overfitting. Overfitting is a modelling error that occurs when a good fit is achieved on the training data, while the generalization of the model on unseen data is unreliable. Secondly, a histogram equalization was applied on the OPG(s) as a pre-processing step. Histogram equalization is a method for adjusting image intensities to enhance the contrast and this can increase the prediction accuracy[18]. Lastly, transfer learning was used to prevent overfitting. Transfer learning is a technique that pre-trains very deep networks on large datasets in order to learn the generic and low-level features in the early layers of the network. By reusing these learned weights on other tasks, the need to re-learn these low-level features in new data sets is eliminated, which greatly reduces the amount of data and time required to converge such a deep neural network[19].

A limitation of the present study is that only cropped images of third molars were included. Training and testing the model with cropped premolars, incisors and canines might further increase the robustness and the generalizability to assess all caries on OPG(s). Secondly, the clinical and radiological assessment by surgeons is not the gold standard in detection of caries. Histological confirmations of caries and further extension of labeled data are required, to overcome the model's limits in this present study.

To the best of our knowledge, this is the first publication to rely deep learning using solely OPG(s) for caries detection on third molars. Furthermore, class activation maps are generated to increase the interpretability of the model predictions. Considering the encouraging results, future work should reside on the detection of other pathologies associated with third molars such as pericoronitis, periapical lesions, root resorption or cysts. Also, the potential bias in these algorithms with possible risks of limited robustness, generalizability and reproducibility has to be assessed in future studies using external datasets and is a necessary step to further implement deep learning successfully in daily clinical practice.

In conclusion, a convolutional neural network (CNN) was developed that achieved a F1 score of 0.87 for caries detection on third molars using panoramic radiographs. This forms a promising foundation for the further development of automatic third molar removal assessment.

Material And Methods

Data selection

Random preoperative OPG(s) of patients who underwent third molar removal were retrospectively selected from the Department of Oral and Maxillofacial Surgery of Radboud University Nijmegen Medical

Centre, Netherlands. The accumulated OPG(s) were acquired with a Cranex Novus e device (Soredex, Helsinki, Finland), operated at 90 kV and 10 mA, using a CCD sensor. The inclusion criteria were a minimum age of 16 and the presence of at least one third molar (M3). Blurred and incomplete OPG(s) were excluded from further analysis. This study has been conducted in accordance with the code of ethics of the world medical association (Declaration of Helsinki). The approval of this study was granted by the Institutional Review Board (Commissie Mensgebonden Onderzoek regio Arnhem-Nijmegen) and informed consent were not required as all image data were anonymized and de-identified prior to analysis (decision no. 2019–5232).

Data annotation

The present third molars (M3(s)) on OPG(s) were classified and labeled as carious M3 and non-carious M3 based on electronic medical records (EMR). Subsequently, a crop of 256 by 256 pixels around the M3 was created. The cropped data consisting of carious M3(s) and non-carious M3(s) were de-identified and anonymized prior to further analysis. All anonymized data were revalidated by two clinicians (SV, MH). In cases of disagreement, the cropped OPG(s) were excluded. The final dataset consisted of 250 carious M3(s) and 250 non-carious M3(s).

The model

The MobileNet V2 was used in this study. This model is characterized by depthwise separable convolutions and an inverted residual with linear bottleneck. The low-dimensional compressed representation is expanded to a higher dimension with a lightweight depthwise convolution. Subsequently, features are projected back to low-dimensional representation with a linear convolution. [17] The applied model structure is shown in Fig. 5.

Model training

The total dataset was randomly divided into 3 sets, 320 for training, 80 for validation and 100 for testing. All datasets had an equal class distribution of carious and non-carious third molars. Histogram equalization and data augmentation techniques were employed on the training dataset, in order to improve the model generalization. The MobileNet V2 was pretrained on the 2012 ILSVRC ImageNet dataset[20]. During the training process, hyperparameters and optimization operations were empirically determined, so that a maximum model performance was achieved on the validation set. Subsequently, the best model was used to perform predictions on the test set.

The optimization algorithm employed was the Adam optimizer, at a learning rate of 0.0001, with a batch-size of 32 and batch normalization. The training and optimization process were carried out using the Keras library in the Colaboratory Jupyter Notebook environment[21].

Statistical analysis

The diagnostic accuracy for caries detection was assessed based on the true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). Classification metrics are reported as

follows for the test set: accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$, precision = $\frac{TP}{TP+FP}$, dice = $\frac{2TP}{2TP+FP+FN}$ (also known as the F1-score), recall = $\frac{TP}{TP+FN}$ (also known as sensitivity), specificity = $\frac{TN}{TN+FP}$. Furthermore, the area-under-the-curve-receiver-operating-characteristics-curve (AUC) and confusion matrix are presented. Gradient-weighted Class Activation Mapping (Grad-CAM), a class-discriminative localization technique was applied, in order to generate visual explanations highlighting the important regions in the cropped image for detecting carious lesions[22].

Data availability

The data used in this study can be made available if needed within the regulation boundaries for data protection.

Declarations

Data availability

The data used in this study can be made available if needed within the regulation boundaries for data protection.

Competing interests

The author(s) declare no competing interests.

Author contributions statement

Shankeeth Vinayahalingam: Study design, data collection, statistical analysis, writing the article. Steven Kempers: Data collection, statistical analysis, writing the article. Lorenzo Limon: Study design, statistical analysis, writing the article. Dionne Deibel: Study design, statistical analysis, writing the article. Thomas Maal: Study design, article review, supervision. Stefaan Bergé: Supervision, writing article, article review. Marcel Hanisch: Statistical analysis, supervision, writing article, article review. Tong Xi: Study design, data collection, statistical analysis, writing the article.

Sources of support

This research is partially funded by the Radboud AI for Health collaboration between Radboud University and the Innovation Center for Artificial Intelligence (ICAI) of the Radboud University Nijmegen Medical Centre.

References

1. Venta, I. Current care guidelines for third molar teeth. *J Oral Maxillofac Surg.* **73**, 804–805 <https://doi.org/10.1016/j.joms.2014.12.039> (2015).

2. Kaye, E. *et al.* Third-Molar Status and Risk of Loss of Adjacent Second Molars. *J Dent Res.* **22034521990653**, <https://doi.org/10.1177/0022034521990653> (2021).
3. Ghaeminia, H. *et al.* Surgical removal versus retention for the management of asymptomatic disease-free impacted wisdom teeth. *Cochrane Database Syst Rev.* **5**, CD003879 <https://doi.org/10.1002/14651858.CD003879.pub5> (2020).
4. Shugars, D. A. *et al.* Incidence of occlusal dental caries in asymptomatic third molars. *J Oral Maxillofac Surg.* **63**, 341–346 <https://doi.org/10.1016/j.joms.2004.11.009> (2005).
5. Akarslan, Z. Z., Akdevelioglu, M., Gungor, K. & Erten, H. A comparison of the diagnostic accuracy of bitewing, periapical, unfiltered and filtered digital panoramic images for approximal caries detection in posterior teeth. *Dentomaxillofac Radiol.* **37**, 458–463 <https://doi.org/10.1259/dmfr/84698143> (2008).
6. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Med Image Anal.* **42**, 60–88 <https://doi.org/10.1016/j.media.2017.07.005> (2017).
7. Lee, J. H., Kim, D. H., Jeong, S. N. & Choi, S. H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *Journal of dentistry.* **77**, 106–111 <https://doi.org/10.1016/j.jdent.2018.07.015> (2018).
8. Cantu, A. G. *et al.* Detecting caries lesions of different radiographic extension on bitewings using deep learning. *J Dent.* **100**, 103425 <https://doi.org/10.1016/j.jdent.2020.103425> (2020).
9. Casalegno, F. *et al.* Caries Detection with Near-Infrared Transillumination Using Deep Learning. *Journal of dental research.* **98**, 1227–1233 <https://doi.org/10.1177/0022034519871884> (2019).
10. Schwendicke, F., Elhennawy, K., Paris, S., Friebertshuser, P. & Krois, J. Deep learning for caries lesion detection in near-infrared light transillumination images: A pilot study. *Journal of dentistry.* **92**, 103260 <https://doi.org/10.1016/j.jdent.2019.103260> (2020).
11. Moutselos, K., Berdouses, E., Oulis, C. & Maglogiannis, I. Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2019, 1617–1620, doi:10.1109/EMBC.2019.8856553 (2019).
12. Liu, L. Z. *et al.* A Smart Dental Health-IoT Platform Based on Intelligent Hardware, Deep Learning, and Mobile Terminal. *Ieee Journal of Biomedical and Health Informatics.* **24**, 898–906 <https://doi.org/10.1109/Jbhi.2019.2919916> (2020).
13. Friedman, J. W. The prophylactic extraction of third molars: a public health hazard. *Am J Public Health.* **97**, 1554–1559 <https://doi.org/10.2105/AJPH.2006.100271> (2007).
14. Vinayahalingam, S., Xi, T. & Berg Automated detection of third molars and mandibular nerve by deep learning. *Scientific reports.* **9**, 9007 <https://doi.org/10.1038/s41598-019-45487-3> (2019).
15. Schwendicke, F., Golla, T., Dreher, M. & Krois, J. Convolutional neural networks for dental image diagnostics: A scoping review. *J Dent.* **91**, 103226 <https://doi.org/10.1016/j.jdent.2019.103226> (2019).

16. Abdinian, M., Razavi, S. M., Faghihian, R., Samety, A. A. & Faghihian, E. Accuracy of Digital Bitewing Radiography versus Different Views of Digital Panoramic Radiography for Detection of Proximal Caries. *J Dent (Tehran)*. **12**, 290–297 (2015).
17. Sandler, M., Howard, A., Zhu, M. L., Zhmoginov, A. & Chen, L. C. MobileNetV2: Inverted Residuals and Linear Bottlenecks 2018 *IEEE/Cvf Conference on Computer Vision and Pattern Recognition (Cvpr)*, 4510–4520, doi:10.1109/Cvpr.2018.00474 (2018).
18. Radhiyah, A., Harsono, T. & Sigit, R. Comparison Study of Gaussian and Histogram Equalization Filter on Dental Radiograph Segmentation For Labelling Dental Radiograph 2016 *International Conference on Knowledge Creation and Intelligent Computing (Kcic)*, 253–258 (2016).
19. Shao, L., Zhu, F. & Li, X. Transfer learning for visual categorization: a survey. *IEEE Trans Neural Netw Learn Syst*. **26**, 1019–1034 <https://doi.org/10.1109/TNNLS.2014.2330900> (2015).
20. Deng, J. *et al.* ImageNet: A Large-Scale Hierarchical Image Database. *Proc Cvpr Ieee*. 248–255 <https://doi.org/10.1109/cvpr.2009.5206848> (2009).
21. Lee, H. & Song, J. Introduction to convolutional neural network using Keras; an understanding from a statistician. *Commun Stat Appl Met* **26**, 591–610, doi:10.29220/Csam.2019.26.6.591 (2019).
22. Selvaraju, R. R. *et al.* Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *Ieee I Conf Comp Vis*. 618–626 <https://doi.org/10.1109/lccv.2017.74> (2017).

Figures

Figure 1

Area-under-the-curve-receiver-operating-characteristics-curve (AUC)

Figure 2

Confusion matrix showing the classification results

Figure 3

Class activation map for carious third molars. The left column shows the cropped carious M3s. The middle column represents the class activation map. The right column illustrates the overlay.

Figure 4

Class activation map for non-carious third molars. The left column shows the cropped non-carious M3s. The middle column represents the class activation map. The right column illustrates the overlay.