



Automatic and Interpretable Model for Periodontitis Diagnosis in Panoramic Radiographs

Haoyang Li^{1,2,3}, Juexiao Zhou^{1,4}, Yi Zhou⁵, Jieyu Chen^{6(✉)}, Feng Gao^{6(✉)}, Ying Xu^{2,5(✉)}, and Xin Gao^{1(✉)}

¹ Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
`xin.gao@kaust.edu.sa`

² Cancer Systems Biology Center, China-Japan Union Hospital, Jilin University, Changchun, China

³ College of Computer Science and Technology, Jilin University, Changchun, China
⁴ Southern University of Science and Technology, Shenzhen, China

⁵ Department of Biochemistry and Molecular Biology and Institute of Bioinformatics, University of Georgia, Athens, USA
`xyn@uga.edu`

⁶ The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China
`{chenjy335,gao57}@mail.sysu.edu.cn`

Abstract. Periodontitis is a prevalent and irreversible chronic inflammatory disease both in developed and developing countries, and affects about 20%-50% of the global population. The tool for automatically diagnosing periodontitis is highly demanded to screen at-risk people for periodontitis and its early detection could prevent the onset of tooth loss, especially in local community and health care settings with limited dental professionals. In the medical field, doctors need to understand and trust the decisions made by computational models and developing interpretable machine learning models is crucial for disease diagnosis. Based on these considerations, we propose an interpretable machine learning method called Deetal-Perio to predict the severity degree of periodontitis in dental panoramic radiographs. In our method, alveolar bone loss (ABL), the clinical hallmark for periodontitis diagnosis, could be interpreted as the key feature. To calculate ABL, we also propose a method for teeth numbering and segmentation. First, Deetal-Perio segments and indexes the individual tooth via Mask R-CNN combined with a novel calibration method. Next, Deetal-Perio segments the contour of the alveolar bone and calculates a ratio for individual tooth to represent ABL. Finally, Deetal-Perio predicts the severity degree of periodontitis given the ratios of all the teeth. The entire architecture could not only outperform state-of-the-art methods and show robustness on two data sets in both periodontitis prediction, and teeth numbering and segmentation tasks, but also be interpretable for doctors to understand the reason why Deetal-Perio works so well.

Keywords: Teeth segmentation and numbering · Interpretable machine learning · Periodontitis diagnosis · Panoramic radiograph

1 Introduction

Periodontitis is a chronic inflammatory disease of periodontium resulting in inflammation within the supporting tissues of the teeth, progressive attachment, and bone loss [11]. Periodontitis is prevalent in both developed and developing countries, and affects about 20%–50% of the global population which makes it a public health concern [12]. Thus, the tool for automatically diagnosing periodontitis is highly demanded to provide the invaluable opportunity to screen at-risk people for periodontitis and its early detection could prevent the onset of tooth loss, especially in local community and health care settings where dentists are not easily accessible [1]. The form of periodontitis is characterized by periodontal ligament loss and resorption of the surrounding alveolar bone caused by severe inflammatory events [14]. Cumulative alveolar bone loss (ABL) results in weakening of the supporting structures of the teeth, and predisposes the patient to tooth mobility and loss [2] (Fig. 1B). Thus ABL is a hallmark of the periodontal disease [17]. To calculate ABL, it is necessary to gather the contours of the individual tooth and the alveolar bone. In this situation, teeth numbering and segmentation are essential and fundamental tasks for periodontitis diagnosis. In addition, dentists usually need to serve numerous patients and read a large number of panoramic radiographs daily. Thus an automatic tool for teeth numbering and segmentation to enhance efficiency and improve the quality of dental care is highly needed [4,18].

Several methods have been proposed to tackle the periodontitis prediction or teeth numbering and segmentation tasks. Joo et al. [8] proposed a classification method for the periodontal disease based on convolutional neural network (CNN) by using periodontal tissue images. This method classified four states of periodontitis and the accuracy on validation data was 0.81. Ozden et al. [13] tested three classification algorithms, artificial neural networks (ANN), support vector machine (SVM), and decision tree (DT) to diagnose periodontal diseases by using 11 measured variables of each patient as raw data. The results showed that DT and SVM were best to classify the periodontal diseases with high accuracy (0.98 of precision). It revealed that SVM and DT appeared to be practical as a decision-making aid for the prediction of periodontal disease.

As for the teeth numbering and segmentation, Wirtz et al. [16] proposed a coupled shape model in conjunction with a neural network by using panoramic radiographs. The network provided a preliminary segmentation of the teeth region which is used to initialize the coupled shape model. Then the 28 individual teeth (excluding wisdom teeth) were segmented and labeled using gradient image features in combination with the model's statistical knowledge. The experimental result showed an average dice of 0.744. Chen et al. [4] used faster regions with CNN features to detect and number teeth in dental periapical films. They proposed three post-processing techniques to improve the numbering performance. Results revealed that mean average precision (mAP) was 0.80 and the

performance of this method was close to the level of junior dentists. Cui et al. [6] used deep CNN to achieve automatic and accurate tooth instance segmentation and identification from cone beam CT (CBCT) images. They extracted the edge map from the CBCT image to enhance image contrast along shape boundaries. Next, the edge map and input images were passed through 3D Mask R-CNN with encoded teeth spatial relationships. Their method produced accurate instance segmentation and identification results automatically.

The main limitations of methods mentioned above are as follows: (1) the bias of detecting and numbering teeth in some cases with severe periodontitis due to the disturbance of a large number of missing teeth, (2) the lacking capability of their methods on predicting the number of missing teeth in the shortage data volume of some individual classes, and (3) the lack of interpretability of predicting the severity degree of periodontitis.

In this paper, we try to overcome these limitations through the following contributions. (1) We propose an automatic and interpretable machine learning method called Deetal-Perio to predict the severity degree of periodontitis from dental panoramic radiographs. (2) As a subroutine of Deetal-Perio, we further propose a method for teeth numbering and segmentation which consists of a novel calibration algorithm. (3) Deetal-Perio outperforms state-of-the-art methods and shows the robustness on the two data sets from two hospitals. (4) Deetal-Perio uses ABL as the feature for periodontitis diagnosis and is thus fully interpretable.

2 Materials and Methods

2.1 Data Sets

Suzhou Stomatological Hospital supplied a total of 302 digitized panoramic radiographs (hereinafter referred to as the *Suzhou* dataset). Each radiograph has a high resolution of 1480 pixels by 2776 pixels and was annotated with the consensus taken by three professional dentists following the Fédération Dentaire Internationale (FDI) numbering system to get the contours of teeth and their labels as the ground truth. FDI numbering system divides all teeth into four quadrants where teeth are labeled as 11–18, 21–28, 31–38, 41–48, respectively (Fig. 1A). Among all radiographs, 298 were labeled with the severity degree of periodontitis by dentists, including four categories: 52 of no periodontitis, 189 of mild periodontitis, 43 of moderate periodontitis and 14 of severe periodontitis. These radiographs were randomly divided into two data sets, a training set with 270 images and a test set with 28 images. This random split was repeated for three times and the results were averaged.

To test the robustness and generality of Deetal-Perio, we also collected another dataset from the Sixth Affiliated Hospital, Sun Yat-sen University (hereinafter referred to as the *Zhongshan* dataset). This dataset includes 62 high-resolution panoramic radiographs which are categorized by four classes mentioned above (36 of no periodontitis, 12 of mild periodontitis, 6 of moderate periodontitis and 8 of severe periodontitis).

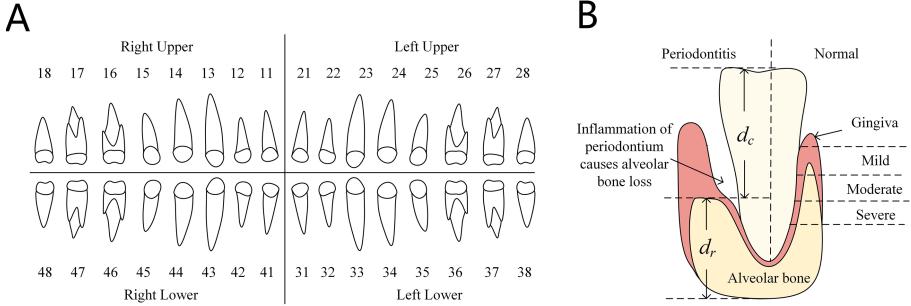


Fig. 1. A. FDI numbering system. It divides all teeth into four quadrants where teeth are labeled as 11–18, 21–28, 31–38, 41–48, respectively. B. The left of tooth shows the appearance of periodontitis and the representation of ABL. The right of tooth shows the appearance of a normal tooth.

2.2 Methods

The architecture of Deetal-Perio is as follows. First, Deetal-Perio segments and numbers individual tooth via Mask R-CNN combined with a novel calibration method. Next, Deetal-Perio segments the contour of the alveolar bone and calculates the ratio for individual tooth which could represent ABL as the key feature to predict periodontitis. Finally, Deetal-Perio uses XGBoost to predict the severity degree of periodontitis by given a vector of ratios from all the numbered teeth. The entire architecture is shown in Fig. 2.

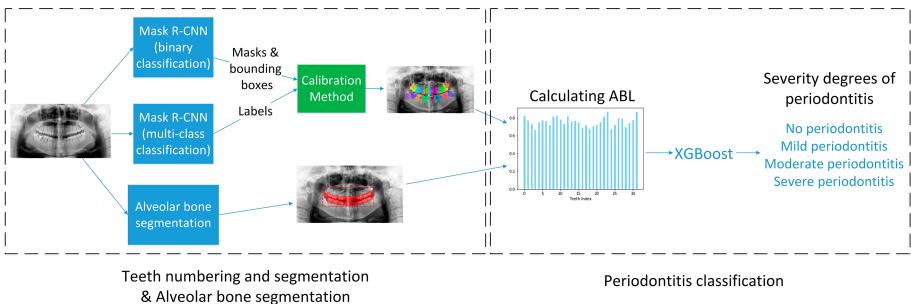


Fig. 2. The workflow of Deetal-Perio.

Teeth Segmentation and Numbering. Inspired by the state-of-the-art architecture in object classification and segmentation called Mask R-CNN [7], we tried to segment the teeth with binary classification via Mask R-CNN. That was, we wanted to differentiate teeth from the background image. The result revealed that Mask R-CNN could detect almost all of the teeth given a radiograph. Then, we tried to segment and number the teeth with multi-class classification via Mask R-CNN and this time, we wanted to identify the labels of these

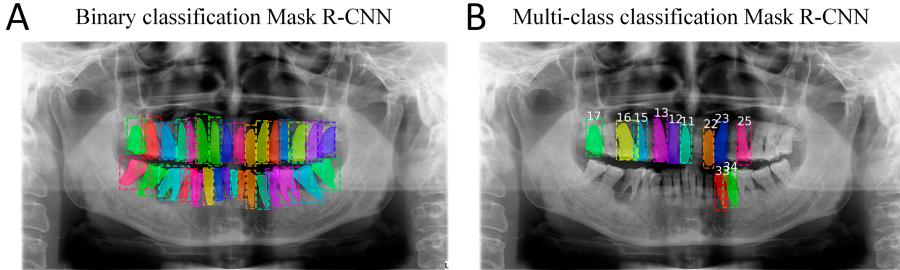


Fig. 3. A. The result of binary classification Mask R-CNN. B. The result of multi-class classification Mask R-CNN.

teeth. The result showed that only a minority of teeth could be detected and numbered due to the limitation of data from individual classes, but compared with the ground truth, most detected teeth were numbered correctly. Figure 3 shows these two results, which provide complementary information to each other. Thus, we proposed to combine the results from the binary and multi-class Mask R-CNN models together. We extracted the bounding boxes and masks from the binary classification results, and the labels of numbered teeth from the multi-class classification results. We further proposed a calibration method to integrate the results from the two classifiers, refine the labels of numbered teeth, and infer the labels of unnumbered teeth.

The calibration method is designed as follows: first, $B = \{B_1, B_2, \dots, B_m\}$ and $M = \{M_1, M_2, \dots, M_n\}$ represent m and n of center points of teeth's bounding boxes from the results of binary and multi-class classification Mask R-CNN, respectively. Then, we found the closest tooth to each tooth of M in B by calculating the Euclidean distance and assigned the labels of teeth from M to B . Next, we iterated each tooth in B to judge whether neighboring teeth are labeled and calibrate its own label until all the teeth in B satisfied two conditions: each tooth had been labeled and all labeled results followed the FDI numbering system. The details of algorithm are given in Algorithm 1. Finally, all teeth are labeled in B which are considered to be the results of teeth segmentation and numbering step.

The Representation of ABL. After applying the teeth segmentation and numbering method, we obtained the contours of individual teeth and their labels. We also acquired the contour of the alveolar bone by applying Mask R-CNN. Next, we calculated the ratio $d = d_c/(d_c + d_r)$ for each individual tooth, where d_c denotes the vertical distance from the alveolar bone to the top of the dental crown and d_r denotes the vertical distance from the bottom of the dental root to the alveolar bone (Fig. 1B). Due to the smoothness of the contour of the alveolar bone, we randomly selected two points on this contour to draw a line. Then, we randomly chose 50% points on the contour of the dental crown to calculate the vertical distance from these points to this line respectively and defined the largest

Algorithm 1. The calibration method for teeth numbering

Input: $B = \{B_1, \dots, B_m\}$, $M = \{M_1, \dots, M_n\}$, L_i : the label of tooth i , X_i : the x coordinate of tooth i

Output: calibrated $\{L_{B_1}, \dots, L_{B_m}\}$

- 1: L_{B_i}, \dots, L_{B_m} are set to 0
- 2: **for** $i \leftarrow 1$ to n **do**
- 3: $B_k \leftarrow$ the nearest point to M_i by calculating the Euclidean distance
- 4: $L_{B_k} \leftarrow L_{M_i}$
- 5: **end for**
- 6: **while** exists tooth unlabeled in B or not following FDI numbering system **do**
- 7: **for** each unlabeled tooth B_i **do**
- 8: **if** $L_{B_{i-1}} = 0$ and $L_{B_{i+1}} = 0$ **then**
- 9: continue
- 10: **else if** $L_{B_{i-1}} \neq 0$ and $L_{B_{i+1}} \neq 0$ **then**
- 11: find nearer tooth B_j to B_i between B_{i-1} and B_{i+1} , the other tooth is B_h
- 12: $B_k \leftarrow$ the neighbor tooth of B_j (not B_i)
- 13: **else if** only one of B_{i-1} and B_{i+1} is labeled **then**
- 14: $B_j \leftarrow$ labeled tooth between B_{i-1} and B_{i+1}
- 15: $B_k \leftarrow$ the neighbor tooth of B_j (not B_i)
- 16: **end if**
- 17: $z \leftarrow \text{round}(|X_{B_i} - X_{B_j}| / |X_{B_j} - X_{B_k}|)$
- 18: infer L_{B_i} considering $z - 1$ of missing teeth bewteen B_j and B_i
- 19: **end for**
- 20: **end while**

vertical distance as d_c . We estimated d_r in a similar way. In our case, d_c could be a good estimation to represent ABL showing the level of destruction of the alveolar bone and in order to normalize d_c , we divided it by $d_c + d_r$ which is the estimation of the perpendicular height of the tooth. Thus, d could represent the ABL to reflect the severity degree of periodontitis.

Periodontitis Prediction. After acquiring the ratio d of individual teeth, each radiograph would output a vector of ratios $D = \{d_1, d_2, \dots, d_{32}\}$ where each radio corresponds to a label of tooth. Apparently, some radiographs do not have all the 32 teeth. In such cases, the ratios of teeth which do not exist in the radiograph are set to be the mean value of its neighboring teeth' ratios. We then solved the periodontitis classification task by XGBoost [5] which has gained attention as the algorithm of choice for many winning teams of machine learning competitions [15]. To tackle the class imbalance problem, we used Synthetic Minority Over-sampling Technique (SMOTE) [3] for over-sampling the minority classes. After over-sampling, D was inputted as the feature to classify the four-class severity degree of periodontitis by XGBoost.

3 Result

Figure 4 shows two examples of teeth segmentation and numbering results on the *Suzhou* dataset and *Zhongshan* dataset respectively. We first set out to evaluate

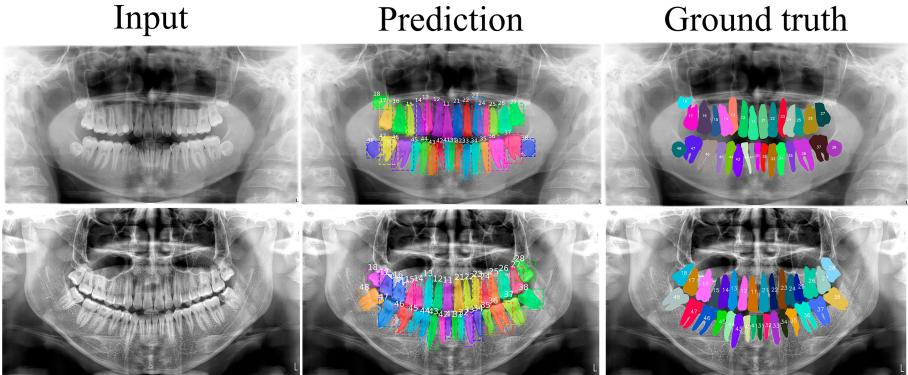


Fig. 4. Two examples from teeth segmentation and numbering results from the *Suzhou* (top) and *Zhongshan* (bottom) datasets respectively.

the performance of teeth numbering and segmentation, by using three metrics: (1) Dice (all) denotes the overall dice coefficient of teeth segmentation. (2) Dice (single) denotes the mean value of all dice coefficients from all labeled teeth respectively. (3) The mAP is calculated as the average of the average precisions for all object classes. We compared with the baseline multi-class classification Mask R-CNN, and the state-of-the-art methods, including the methods proposed by Chen et al. [4] and Wirtz et al. [16] which are trained on the *Suzhou* dataset for fair comparison (Table 1). Thanks to the calibration method in Deetal-Perio, we could number teeth much more correctly than other compared methods. Thus, the performance of Deetal-Perio in the segmentation of individual tooth is also better than other methods. To prove the robustness of Deetal-Perio, we tested our method directly on the *Zhongshan* dataset, without re-training the model, and the mAP, dice (all) and dice (single) reached 0.841, 0.852 and 0.748 respectively, which are close to the results from the *Suzhou* dataset. This cross-set validation demonstrates the robustness and generality of our method.

We then evaluated the performance of periodontitis prediction. Table 2 compares our method with several baseline machine learning methods and the method proposed by Joo et al. [8] which is trained on the *Suzhou* dataset for fair comparison. The metrics include F1-score and mean accuracy over the four classes. Our method has a stably good accuracy over all the four classes: 0.932 for no periodontitis, 0.891 for mild periodontitis, 0.952 for moderate periodontitis and 0.992 for severe periodontitis. We also tested it on the *Zhongshan* dataset directly, without re-training, and the F1-score and mean accuracy reached 0.454 and 0.817 respectively. The boosting algorithm of XGBoost made it a strong leaner to enhance the performance compared with the simple decision trees, and the regularization of XGBoost made it robust against the noise and thus outperforming Adaboost. Compared to Joo et al. [8], our method greatly reduced the feature dimension from 1480×2776 to a 1×32 vector instead of simply implementing a CNN for image classification with a large number of disturbing

redundant features. In addition, the intermediate results of teeth segmentation and numbering as well as the geometrically calculated ABL provide dentists completely transparent and interpretable information, so that they not only know that our method works, but also understand how it works.

Table 1. Performance comparison tested on the *Suzhou* dataset between Deetal-Perio and other methods on teeth numbering/segmentation task, by mAP, Dice (all) and Dice (single). The method of Chen, et al. [4] only has tooth number, but not segmentation function.

Methods	mAP	Dice (all)	Dice (single)
Deetal-Perio	0.826	0.868	0.778
Mask R-CNN (multi-class)	0.372	0.496	0.377
Wirtz et al. [16]	0.410	0.763	0.517
Chen et al. [4]	0.195	—	—

Table 2. Performance comparison tested on the *Suzhou* dataset between Deetal-Perio and five methods by F1-score and accuracy on the periodontitis prediction task.

Methods	F1-score	Accuracy
Deetal-Perio	0.878	0.884
CNN	0.659	0.689
Decision Tree	0.792	0.806
SVM	0.655	0.675
Adaboost	0.652	0.657
Joo et al. [8]	0.310	0.641

4 Discussion and Conclusion

In this paper, we proposed a fully automatic and completely interpretable method, Deetal-Perio, for diagnosing the severity degrees of periodontitis from panoramic radiographs using ABL as the key feature. As the intermediate results, our method also accomplished teeth numbering and segmentation. Comprehensive experiments on two datasets show that Deetal-Perio not only dramatically outperforms other compared methods in both teeth segmentation and numbering, and periodontitis prediction, but is also robust and generalizable on independent datasets, which makes Deetal-Perio a suitable method for periodontitis screening and diagnostics.

Despite the success of Deetal-Perio, the performance of teeth numbering relies on the numbering results from the multi-class classification Mask R-CNN model in Deetal-Perio. This can cause issues when there are radiographs with severe periodontitis which have only few or abnormal shapes of teeth. To overcome this limitation, more data, especially the one with these special cases, need to be obtained to further improve the performance, for which few-shot learning might be a helpful way to deal with such special situations [9, 10].

Acknowledgement. We thank He Zhang, Yi Zhang, and Yongwei Tan at Suzhou Stomatological Hospital for providing the data. The research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. FCC/1/1976-04, FCC/1/1976-06, FCC/1/1976-17, FCC/1/1976-18, FCC/1/1976-23, FCC/1/1976-25, FCC/1/1976-26, and REI/1/0018-01-01.

References

1. Balaei, A.T., de Chazal, P., Eberhard, J., Domnisch, H., Spahr, A., Ruiz, K.: Automatic detection of periodontitis using intra-oral images. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3906–3909 (2017). <https://doi.org/10.1109/EMBC.2017.8037710>
2. Bhatt, A.A., et al.: Contributors. In: Cappelli, D.P., Mobley, C.C. (eds.) Prevention in Clinical Oral Health Care, pp. v–vi. Mosby, Saint Louis (2008). <https://doi.org/10.1016/B978-0-323-03695-5.50001-X>, <http://www.sciencedirect.com/science/article/pii/B97803230369550001X>
3. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: Smote: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002)
4. Chen, H., et al.: A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci. Rep.* **9**(1), 3840 (2019). <https://doi.org/10.1038/s41598-019-40414-y>
5. Chen, T., Guestrin, C.: Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM (2016)
6. Cui, Z., Li, C., Wang, W.: ToothNet: automatic tooth instance segmentation and identification from cone beam CT images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6368–6377 (2019)
7. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969 (2017)
8. Joo, J., Jeong, S., Jin, H., Lee, U., Yoon, J.Y., Kim, S.C.: Periodontal disease detection using convolutional neural networks. In: 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 360–362 (2019). <https://doi.org/10.1109/ICAIIC.2019.8669021>
9. Li, H., et al.: Modern deep learning in bioinformatics. *J. Mol. Cell Biol.*, June 2020. <https://doi.org/10.1093/jmcb/mjaa030>, mjaa030
10. Li, Y., Huang, C., Ding, L., Li, Z., Pan, Y., Gao, X.: Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. *Methods* **166**, 4–21 (2019)
11. Lindhe, J., et al.: Consensus report: chronic periodontitis. *Ann. Periodontol.* **4**(1), 38 (1999)
12. Nazir, M.A.: Prevalence of periodontal disease, its association with systemic diseases and prevention. *Int. J. Health Sci.* **11**(2), 72–80 (2017). <https://www.ncbi.nlm.nih.gov/pubmed/28539867>, www.ncbi.nlm.nih.gov/pmc/articles/PMC5426403/
13. Ozden, F.O., Ozgonenel, O., Ozden, B., Aydogdu, A.: Diagnosis of periodontal diseases using different classification algorithms: a preliminary study. *Niger. J. Clin. Practice* **18**(3), 416–421 (2015)

14. de Pablo, P., Chapple, I.L.C., Buckley, C.D., Dietrich, T.: Periodontitis in systemic rheumatic diseases. *Nature Rev. Rheumatol.* **5**(4), 218–224 (2009). <https://doi.org/10.1038/nrrheum.2009.28>
15. Volkovs, M., Yu, G.W., Poutanen, T.: Content-based neighbor models for cold start in recommender systems. *Proc. Recommender Syst. Challenge* **2017**, 1–6 (2017)
16. Wirtz, A., Mirashi, S.G., Wesarg, S.: Automatic teeth segmentation in panoramic X-Ray images using a coupled shape model in combination with a neural network. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11073, pp. 712–719. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00937-3_81
17. Yang, M., Nam, G.E., Salamat, A., Baldwin, M., Deng, M., Liu, Z.J.: Alveolar bone loss and mineralization in the pig with experimental periodontal disease. *Heliyon* **4**(3), e00589 (2018)
18. Zhou, L., et al.: A rapid, accurate and machine-agnostic segmentation and quantification method for CT-based covid-19 diagnosis. *IEEE Trans. Med. Imaging* **99**, 1 (2020)