JPIS
Journal of Periodontal & Implant Science

Research Article

Check for updates

Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm

Jae-Hong Lee [1,*], **Do-hyung Kim** [1], **Seong-Nyum Jeong** [1], **Seong-Ho Choi** [2]

[1]Department of Periodontology, Daejeon Dental Hospital, Institute of Wonkwang Dental Research, Wonkwang University College of Dentistry, Daejeon, Korea
[2]Department of Periodontology, Research Institute for Periodontal Regeneration, Yonsei University College of Dentistry, Seoul, Korea

OPEN ACCESS

*Correspondence:
Jae-Hong Lee
Department of Periodontology, Daejeon Dental Hospital, Wonkwang University College of Dentistry, 77 Dunsan-ro, Seo-gu, Daejeon 35233, Korea.
E-mail: ljaehong@gmail.com
Tel: +82-42-366-1114
Fax: +82-42-366-1115

ORCID iDs
Jae-Hong Lee
https://orcid.org/0000-0002-2375-0141
Do-hyung Kim
https://orcid.org/0000-0001-7846-6175
Seong-Nyum Jeong
https://orcid.org/0000-0003-4890-989X
Seong-Ho Choi
https://orcid.org/0000-0001-6704-6124

Author Contributions
Conceptualization: Jae-Hong Lee, Seong-Ho Choi; Data curation: Jae-Hong Lee, Do-hyung Kim, Seong-Nyum Jeong, Seong-Ho Choi; Formal analysis: Jae-Hong Lee, Do-hyung

ABSTRACT

Purpose: The aim of the current study was to develop a computer-assisted detection system based on a deep convolutional neural network (CNN) algorithm and to evaluate the potential usefulness and accuracy of this system for the diagnosis and prediction of periodontally compromised teeth (PCT).

Methods: Combining pretrained deep CNN architecture and a self-trained network, periapical radiographic images were used to determine the optimal CNN algorithm and weights. The diagnostic and predictive accuracy, sensitivity, specificity, positive predictive value, negative predictive value, receiver operating characteristic (ROC) curve, area under the ROC curve, confusion matrix, and 95% confidence intervals (CIs) were calculated using our deep CNN algorithm, based on a Keras framework in Python.

Results: The periapical radiographic dataset was split into training (n=1,044), validation (n=348), and test (n=348) datasets. With the deep learning algorithm, the diagnostic accuracy for PCT was 81.0% for premolars and 76.7% for molars. Using 64 premolars and 64 molars that were clinically diagnosed as severe PCT, the accuracy of predicting extraction was 82.8% (95% CI, 70.1%–91.2%) for premolars and 73.4% (95% CI, 59.9%–84.0%) for molars.

Conclusions: We demonstrated that the deep CNN algorithm was useful for assessing the diagnosis and predictability of PCT. Therefore, with further optimization of the PCT dataset and improvements in the algorithm, a computer-aided detection system can be expected to become an effective and efficient method of diagnosing and predicting PCT.

Keywords: Artificial intelligence; Machine learning; Periodontal diseases; Supervised machine learning

INTRODUCTION

Periodontal disease (PD), in its acute and chronic forms, constitutes a widespread intraoral pathology and the sixth most common type of inflammatory disease [1]. The continuous progression of PD results in the destruction of all periodontal supporting tissues, including the

alveolar bone, gingiva, and periodontal ligaments around the tooth, and PD has been reported to be the most widespread cause of tooth loss in adults [2,3]. Many epidemiological and experimental studies have shown that systemic chronic inflammation caused by periodontal pathogens is a risk factor or risk indicator for comorbid diseases, such as cardiovascular disease, diabetes mellitus, obesity, osteoporosis, erectile dysfunction, and cancer [4-8].

Various non-surgical and surgical methods have been devised and improved for the treatment of periodontally compromised teeth (PCT) and supporting structures, and numerous studies have also been conducted on regenerative periodontal tissues [9,10]. Despite advances in treatment modalities, there has not yet been a significant improvement in the methodology for diagnosing and predicting PCT. Although panoramic/periapical radiographs and periodontal probes, which are widely used as objective diagnostic tools for diagnosing and predicting PCT, are used in conventional periodontal examinations, clinical diagnostic and prognostic judgment depends heavily on empirical evidence [11].

Convolutional neural networks (CNNs), which are the latest, core model of artificial neural networks and deep learning in computer vision, have developed rapidly since roughly 2010 [12]. Since medical data are digitally stored and accumulated quantitatively and qualitatively, deep CNNs with computer-aided detection (CAD) systems have clear opportunities to be applied in the medical field, and this very fast-growing new area of research has yielded impressive results in terms of diagnosis and prediction in radiological and pathological research [13,14]. Therefore, most recently reported artificial intelligence performance has been based on deep learning and was developed mainly for medical image classification [15,16].

Although radiographic image analysis is conventionally and widely used to diagnose and predict PD, it still tends to be used as an auxiliary means of clinical diagnosis and prediction, and studies on the diagnosis of PCT using deep CNNs with CAD are limited [17]. Therefore, the purpose of the current study was to evaluate the potential usefulness and accuracy of deep CNN algorithms for diagnosing and predicting PCT.

MATERIALS AND METHODS

Dataset collection
This study was approved by the Institutional Review Board of Daejeon Dental Hospital, Wonkwang University (approval No. W1723/001-001), and was carried out at the Department of Periodontology, Daejeon Dental Hospital, Wonkwang University. We collected periapical radiographic datasets between January 2015 and December 2016, and all images were de-identified. Periapical radiographs of patients with PD and those aged 12 years or younger, as well as images with severe noise or haziness or showing teeth that were partially present or severely distorted, were excluded. Teeth with more than 4 roots, those that had undergone root canal treatment, those that had undergone apical surgery with root resection, those with moderate to severe caries, those with a full restorative crown, and teeth with a shape that deviated from normal anatomical structures were also excluded.

Diagnosis and prediction of PCT
All periapical radiographic datasets and electronic dental records were evaluated by 3 calibrated board-certified periodontists, who collected, deciphered, and categorized them to determine the severity of PCT. All periapical radiographs for which the diagnosis of the 3 examiners did

not agree were excluded. Teeth with a clinical attachment level (CAL) of less than 3 mm in a clinical examination using a World Health Organization-standardized community periodontal index probe were classified as healthy teeth. Teeth with bleeding on probing during the clinical examination and a CAL of less than 6 mm or a bone loss of less than 4 mm on radiography were classified as moderate PCT, and teeth with a CAL of greater than 6 mm and a bone loss of more than 4 mm were classified as severe PCT [18,19]. Among severe PCT, teeth that were extracted immediately after clinical and radiological examinations or during the follow-up period of 3 months were defined as hopeless teeth. Because the differential diagnosis between healthy teeth and incipient PCT was made using only periapical radiographs, this study did not diagnose or distinguish between healthy teeth and incipient PCT.

Preprocessing and image augmentation

All the selected periapical radiographic images were cropped and resized to 224×224 pixels (from the original 1,440×1,920 pixels), and converted into PNG format. In addition, all the maxillary teeth were vertically flipped to the form of mandibular teeth. A pretrained VGG-19 network was used for preprocessing, and the dataset was augmented using the Keras framework based on the ImageDataGenerator function [20]. This was randomly performed with a rotation range of 15°, a width and height shift range of 0.1, a shear range of 0.5, and 100 images were generated for each tooth to obtain a total of 104,400 training dataset images.

Architecture of the deep CNN algorithm

CNN is a type of machine learning that is used in various fields, especially in image and sound recognition. Deep CNNs imitate the connectivity patterns of neurons in the animal visual cortex. CNNs consist of 1 or more convolutional layer, a pooling layer, and a fully connected layer. Every convolutional layer responds to stimuli only in a restricted region of the visual field known as the receptive field. This structure is distinguished from conventional image classification algorithms and other deep learning algorithms, since CNN can learn the type of filter that is hand-crafted in conventional algorithms.

This study used 16 convolutional layers and 3 fully connected dense layers; this network is illustrated in Figure 1. Each convolutional layer was designed with a kernel size of 3×3 pixels, the same padding, and a rectified linear unit activation function. The maximum pooling layers were designed with strides of 2×2 pixels. After extracting the feature quantities of images using convolutional layers, we used the maximum pooling layers to reduce the
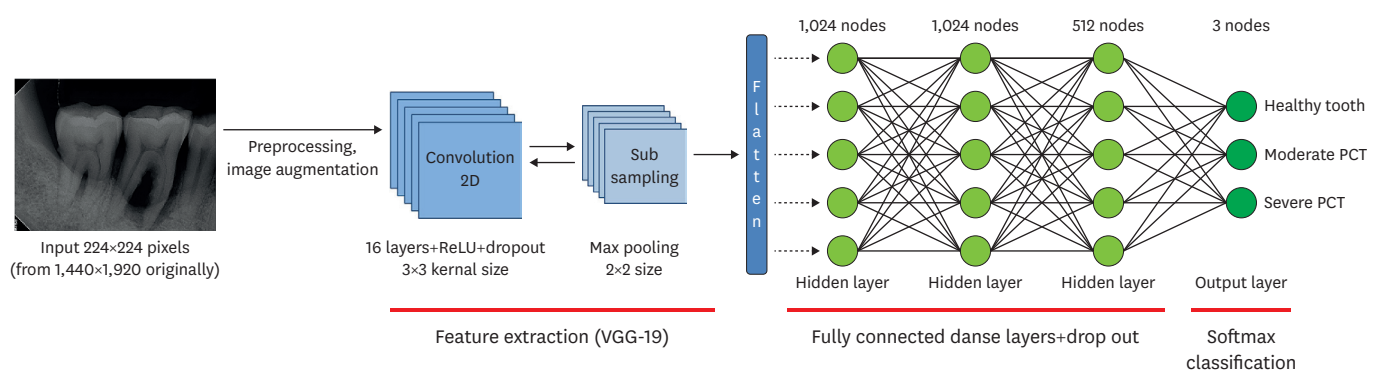
Figure 1. Overall architecture of the deep CNN model. The dataset for the PCT images (224×224 pixels) is labeled as the input. Each of the convolutional layers is followed by a ReLU activation function, dropout, maximum pooling layers, and 3 fully connected layers with 1,024, 1,024, and 512 nodes, respectively. The final output layer performs 3 classifications using the Softmax function.
CNN: convolutional neural network, PCT: periodontally compromised tooth, ReLU: rectified linear unit.

position sensitivity problem and to allow for more generic recognition capability. Next, 3 fully connected deep hidden layers with 1,024, 1,024, and 512 nodes, respectively, were connected to remove spatial information and to statistically determine the key classification of PCT [21]. Dropout, which is a typical method of regularization (rescaling the deep CNN weights to a more effective range), was set to 0.5, and the final output layer was classified in terms of PCT using the Softmax classifier [22]. A total of 500 epochs were used and training network weights were learned using the Adam algorithm (learning rate=0.0001), a stochastic gradient descent method [23]. To produce better diagnosis and prediction of PCT, after 20 epochs of this training phase, fine-tuning was performed in order to optimize the weights and to improve the results by adjusting the hyperparameters of layers [24,25].

Statistical analysis

A randomization sequence was generated using the RAND function in the Excel spreadsheet (Microsoft Corporation, Redmond, WA, USA), and used to divide the periapical radiographic image dataset into a training dataset (n=1,044; 60%), a validation dataset (n=348; 20%), and a test dataset (n=348; 20%). The training and validation datasets were directly used to analyze the PCT and supporting structures, and then to create optimal weights for a deep CNN algorithm model. The test dataset was used to calculate the chi-square, diagnostic and predictive accuracy, sensitivity, specificity, positive predictive value, negative predictive value, receiver operating characteristic (ROC) curve, area under the ROC curve (AUC), and confusion matrix using our deep CNN algorithm, based on a Keras framework in Python (Python 3.6.1, Python Software Foundation, Wilmington, DE, USA). The 95% confidence intervals (CIs) were calculated. P values of less than 0.05 were considered to indicate statistical significance.

RESULTS

Baseline characteristics

The baseline characteristics of the study population are presented in Table 1. A total of 651 subjects participated in the present study, consisting of 363 (55.8%) males and 288 (44.2%) females. In terms of age, the number of individuals in their 20s was the smallest (n=22; 3.4%), and the number of those in their 60s was the highest (n=216; 33.2%). The dataset consisted of a total of 1,740 periapical radiographic images of 447 (25.7%) maxillary premolars, 450 (25.9%) maxillary molars, 403 (23.2%) mandibular premolars, and 440 (25.3%) mandibular molars. There were 264 (15.2%) premolars and 394 (16.9%) molars that were diagnosed as healthy teeth, 265 (15.2%) premolars and 297 (17.1%) molars diagnosed as moderate PCT, and 322 (18.5%) premolars and 298 (17.1%) molars that were diagnosed as severe PCT.

Diagnosis of PCT

Figure 2 shows the confusion matrix, with and without normalization, showing the results of the classification of PCT. The diagonal elements are the number of points where the predicted label was the same as the actual label, while the non-diagonal elements were misinterpreted by the classifier. The higher the classification value and the darker the shade of blue, the more accurate was the diagnosis. For premolars, the total diagnostic accuracy was 81.0%, the diagnostic accuracy was the highest for severe PCT (82.8%), and the diagnostic accuracy was the lowest for moderate PCT (77.3%). For molars, the total diagnostic accuracy was 76.7%, the diagnostic accuracy was the highest for severe PCT (81.3%), and the diagnostic accuracy was the lowest for moderate PCT (70.3%).

Table 1. Study population and baseline characteristics of the patients and teeth

Characteristics	Training dataset	Validation dataset	Test dataset	P value
Patients	351 (100)	149 (100)	151 (100)	
Sex				0.590
Male	190 (54.1)	88 (59.1)	85 (56.3)	
Female	161 (45.9)	61 (40.9)	66 (43.7)	
Age group (yr)				0.349
20–29	11 (3.1)	6 (4.0)	5 (3.3)	
30–39	12 (3.4)	12 (8.1)	14 (9.3)	
40–49	44 (12.5)	21 (14.1)	24 (15.9)	
50–59	111 (31.6)	41 (27.5)	39 (25.8)	
60–69	121 (34.5)	46 (30.9)	49 (32.5)	
≥70	52 (14.8)	23 (15.4)	20 (13.2)	
Teeth	1,044 (100)	348 (100)	348 (100)	
Position				
Maxilla				0.914
Premolar	261 (25.0)	87 (25.0)	99 (28.4)	
Molar	264 (25.3)	91 (26.1)	95 (27.3)	
Mandible				0.690
Premolar	253 (24.2)	81 (23.3)	69 (19.8)	
Molar	266 (25.5)	89 (25.6)	85 (24.4)	
Classification of diagnosis				
Healthy teeth				0.319
Premolar	151 (14.5)	53 (15.2)	60 (17.2)	
Molar	182 (17.4)	60 (17.2)	52 (14.9)	
Moderate PCT				0.325
Premolar	169 (16.2)	52 (14.9)	44 (12.6)	
Molar	180 (17.2)	53 (15.2)	64 (18.4)	
Severe PCT				0.615
Premolar	194 (18.6)	64 (18.4)	64 (18.4)	
Molar	168 (16.1)	66 (19.0)	64 (18.4)	

Values are presented as number (%).
PCT: periodontally compromised teeth.

Prediction of hopeless teeth

The accuracy of predicting extraction was evaluated and compared between the CNN and blinded board-certified periodontists using 64 premolars and 64 molars diagnosed as severe PCT in the test dataset. For premolars, the deep CNN had an accuracy of 82.8% (95% CI, 70.1%–91.2%) and an AUC of 82.6% (95% CI, 71.1%–91.1%), while the corresponding values for the periodontists were 79.7% (95% CI, 66.7%–88.5%) and 79.3% (95% CI, 67.4%–88.4%), respectively. Thus, the deep CNN had a higher AUC value, but there was no statistically significant difference in the predictive accuracy between the 2 methods (*P*=0.150). For molars, the deep CNN had an accuracy of 73.4% (95% CI, 59.9%–84.0%) and an AUC of 73.4% (95% CI, 60.9%–83.7%), while the corresponding values for the periodontists were 76.6% (95% CI, 63.2%–86.5%) and 76.4% (95% CI, 64.1%–86.1%), respectively. Thus, the periodontists had a higher AUC value, but as with the premolars, there was no statistically significant difference in predictive accuracy between the 2 methods (*P*=0.151, Table 2).

DISCUSSION

Recently, several studies have investigated the potential usefulness and accuracy of artificial intelligence approaches in interpreting medical images, such as clinical photographs, X-rays, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography scans. In particular, the deep CNN algorithm has been used most commonly and has yielded promising results [26-28]. Consequently, in this study, supervised deep learning,
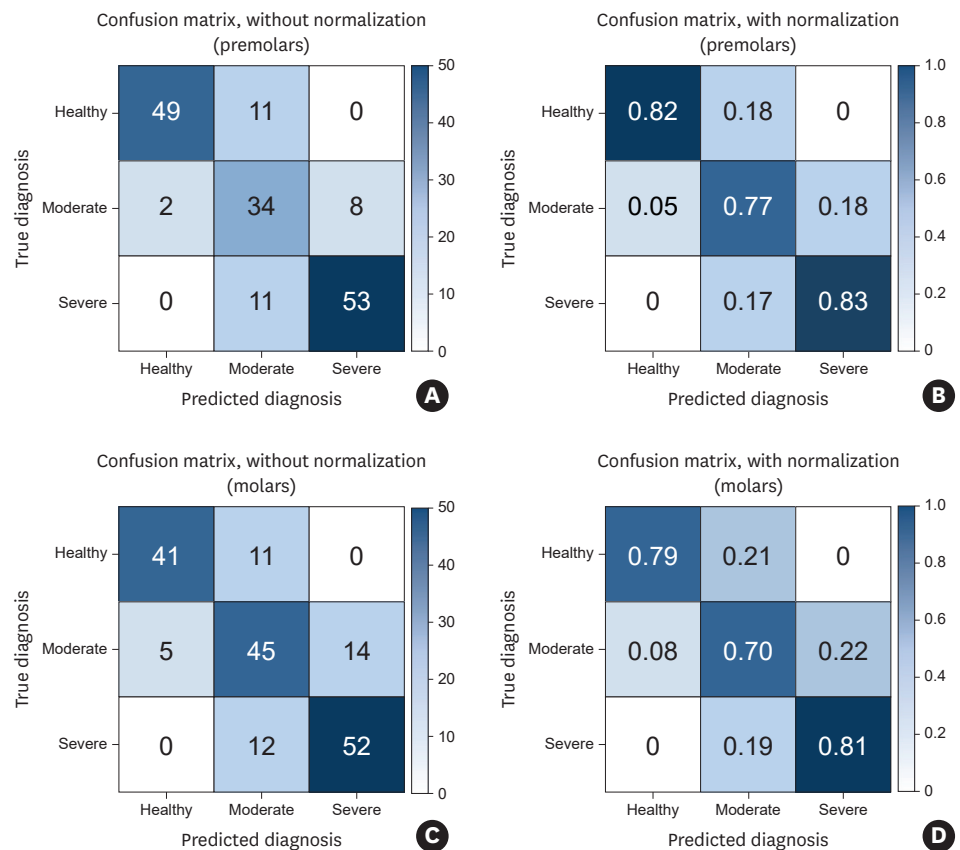
Figure 2. Multiclass classification confusion matrix with and without normalization using a deep CNN classifier. The diagonal elements are the number of points where the predicted label was the same as the actual label, while the non-diagonal elements were misinterpreted by the classifier. The higher the diagonal value and the darker the shade of blue, the more accurate the diagnosis of health and periodontally compromised teeth (A, B) Premolars without/with normalization. (C, D) Molars without/with normalization.
CNN: convolutional neural network.

based on the CNN algorithm, was performed using a prelabeled periapical radiographic dataset. We also confirmed that the results had similar diagnostic and predictive accuracy to those obtained by board-certified periodontists.

The present study found that the deep CNN algorithm had higher diagnostic accuracy for identifying PCT among premolars than among molars. Although we excluded abnormal anatomical variations to minimize the complexity as much as possible, these findings may be due to the higher efficiency of deep learning for premolars, which normally have 1 root, than for mandibular molars with 2 roots or for maxillary molars with 3 roots. In addition, there was a

Table 2. Pairwise comparison between the deep CNN algorithm and periodontists for the prediction of hopeless teeth

Variables	Accuracy (%, 95% CI)	AUC (%, 95% CI)	Difference (%, 95% CI)	*P* value
Premolar			3.3 (−1.2–7.8)	0.150
Deep CNN	82.8 (70.1–91.2)	82.6 (71.1–91.1)		
Periodontist	79.7 (66.7–88.5)	79.3 (67.4–88.4)		
Molar			2.9 (−1.0–6.9)	0.151
Deep CNN	73.4 (59.9–84.0)	73.4 (60.9–83.7)		
Periodontist	76.6 (63.2–86.5)	76.4 (64.1–86.1)		

CNN: convolutional neural network, AUC: area under the receiver operating characteristic curve, CI: confidence interval.

tendency to judge PCT as more severe. The diagnostic accuracy for severe PCT was the highest overall, and the trained deep CNN algorithm was better optimized for the detection of severe PCT. Further studies on the mechanisms underlying deep CNN algorithms are necessary.

Distinguishing between PD and dental caries is crucial, and edge detection needs to be good in cases of destructive PD in order to improve the diagnostic and predictive accuracy for PCT. Unlike traditional shallow learning algorithms, deep CNN algorithms can automatically learn hierarchical feature representations and capture regional patterns from PCT images in their multiple convolutional and hidden layers. Wang [29] reported that deep CNN algorithms could efficiently perform edge detection with only 2 convolutional layers and 1 fully connected hidden layer. Therefore, as deep CNN architecture has a structurally powerful advantage in solving the detection problem, it was chosen for use in the present study [29,30].

The deep CNN algorithm used in the current study was designed based on the VGG-19 network architecture. This architecture, which consists of 16 convolution layers and 2 fully connected layers, is ideal for deep learning and very effective at solving object detection and image classification problems in complex non-medical image data [22,31]. However, since it is necessary to classify black-and-white PCT images of similar size and shape, a potential issue with the VGG-19 network architecture was thought to be that, without correction, learning efficiency might be reduced and the possibility of overfitting might be markedly increased. Therefore, we modified the VGG-19 network architecture by adjusting the number of convolutional and hidden layers and hyperparameters, including the number of epochs, batch size, loss function, optimizer, momentum, and learning rate, to reduce overfitting as much as possible and to facilitate efficient deep learning performance [32].

Fast and accurate diagnosis and prediction is an important element of PD treatment, and optimizing speed and accuracy is an ongoing research problem in CAD. Three principal factors complicate this task. The first limitation of this study is that relatively few images of PCT, including premolars and molars, were used. In order to achieve superior artificial intelligence performance with deep learning, the design of the deep CNN algorithm itself is important, but it is also important to use a high-quality training dataset. In a recent study of the detection of diabetic retinopathy, 54 ophthalmologists repeatedly read 130,000 fundus photographs [27]. In another study of the diagnosis of skin cancer, 18 doctors systematically read 130,000 digital skin images of over 200 skin diseases as a learning dataset [33].

Since unsupervised or semi-supervised deep learning algorithms, including the generative adversarial network (GAN) and reinforcement learning, which do not use any pretrained dataset, have been developing rapidly in recent years, the importance of the training dataset itself is decreasing. Nevertheless, maintaining and securing a high-quality dataset is still important for the deep learning approach. Therefore, to overcome the limitation of the number of images required for deep learning, we collected only high-quality images that were most clearly classified by the 3 experienced periodontists. In addition, learning transfer and preprocessing techniques, including image augmentation and enhancement, were used to avoid overfitting and to normalize the model [34].

Another limitation is that it is impossible to make a complete diagnosis and prediction of PD using only 2-dimensional periapical radiographs. For a more accurate diagnosis and prediction of PD, it is necessary to comprehensively review radiographic and clinical data, such as the patient's history, clinical probing depth, CAL, bleeding on probing, mobility,

percussion, and electric pulp test. Therefore, a deep CNN algorithm using periapical radiographic images alone does not provide sufficient evidence, although it may still be used as a reference for the diagnosis and prediction of PCT. Nonetheless, a 3-dimensional deep CNN algorithm using CT and MRI data was implemented in this study, and models related to this algorithm were developed and refined further in this study. A 3-dimensional deep CNN algorithm will be even more helpful for sophisticated and effective diagnosis and prediction.

As a final limitation, a number of previous deep CNN studies have used downscaled low-resolution medical images instead of high-resolution large matrix images due to limitations of computing power, storage space, cost, and training time. The images used in the present study were also cropped and resized to 224×224 pixels, due to practical constraints. Therefore, the use of a low-resolution dataset in this study is considered to have reduced the accuracy of the diagnosis and prediction of PCT [28].

Currently, deep learning algorithms are being developed and improved, with remarkable accuracy in diagnosis and prediction, particularly in the fields of radiology and pathology. In the present study, we evaluated the potential effectiveness of a deep CNN algorithm for diagnosing and predicting PCT, and demonstrated that it was as effective as experienced periodontists for positively diagnosing and predicting PCT. Through the continuous accumulation of high-quality image datasets and by applying improved algorithms, CAD is expected to become an effective and efficient method of diagnosing and predicting PCT.

REFERENCES

1. Tonetti MS, Jepsen S, Jin L, Otomo-Corgel J. Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: a call for global action. J Clin Periodontol 2017;44:456-62.
   PUBMED | **CROSSREF**

2. Lee JH, Lee JS, Choi JK, Kweon HI, Kim YT, Choi SH. National dental policies and socio-demographic factors affecting changes in the incidence of periodontal treatments in Korean: A nationwide population-based retrospective cohort study from 2002-2013. BMC Oral Health 2016;16:118.
   PUBMED | **CROSSREF**

3. Lee JH, Oh JY, Choi JK, Kim YT, Park YS, Jeong SN, et al. Trends in the incidence of tooth extraction due to periodontal disease: results of a 12-year longitudinal cohort study in South Korea. J Periodontal Implant Sci 2017;47:264-72.
   PUBMED | **CROSSREF**

4. Lee JH, Lee JS, Park JY, Choi JK, Kim DW, Kim YT, et al. Association of lifestyle-related comorbidities with periodontitis: a nationwide cohort study in Korea. Medicine (Baltimore) 2015;94:e1567.
   PUBMED | **CROSSREF**

5. Lee JH, Choi JK, Kim SH, Cho KH, Kim YT, Choi SH, et al. Association between periodontal flap surgery for periodontitis and vasculogenic erectile dysfunction in Koreans. J Periodontal Implant Sci 2017;47:96-105.
   PUBMED | **CROSSREF**

6. Lee JH, Oh JY, Youk TM, Jeong SN, Kim YT, Choi SH. Association between periodontal disease and non-communicable diseases: A 12-year longitudinal health-examinee cohort study in South Korea. Medicine (Baltimore) 2017;96:e7398.
   PUBMED | **CROSSREF**

7. Choi JK, Kim YT, Kweon HI, Park EC, Choi SH, Lee JH. Effect of periodontitis on the development of osteoporosis: results from a nationwide population-based cohort study (2003-2013). BMC Womens Health 2017;17:77.
   PUBMED | **CROSSREF**

8. Lee JH, Kweon HH, Choi JK, Kim YT, Choi SH. Association between periodontal disease and prostate cancer: results of a 12-year longitudinal cohort study in South Korea. J Cancer 2017;8:2959-65.
   PUBMED | **CROSSREF**

9. Graziani F, Karapetsa D, Alonso B, Herrera D. Nonsurgical and surgical treatment of periodontitis: how many options for one disease? Periodontol 2000 2017;75:152-88.
PUBMED | CROSSREF

10. Martins SH, Novaes AB Jr, Taba M Jr, Palioto DB, Messora MR, Reino DM, et al. Effect of surgical periodontal treatment associated to antimicrobial photodynamic therapy on chronic periodontitis: A randomized controlled clinical trial. J Clin Periodontol 2017;44:717-28.
PUBMED | CROSSREF

11. Ainamo J, Barmes D, Beagrie G, Cutress T, Martin J, Sardo-Infirri J. Development of the World Health Organization (WHO) community periodontal index of treatment needs (CPITN). Int Dent J 1982;32:281-91.
PUBMED

12. Sklan JE, Plassard AJ, Fabbri D, Landman BA. Toward content based image retrieval with deep convolutional neural networks. Proc SPIE Int Soc Opt Eng 2015;9417.
PUBMED

13. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv e-print 2017;arXiv:1711.05225.

14. Rajpurkar P, Hannun AY, Haghpanahi M, Bourn C, Ng AY. Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv e-print 2017:arXiv:1707.01836.

15. Garcia-Hernandez JJ, Gomez-Flores W, Rubio-Loyola J. Analysis of the impact of digital watermarking on computer-aided diagnosis in medical imaging. Comput Biol Med 2016;68:37-48.
PUBMED | CROSSREF

16. Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60-88.
PUBMED | CROSSREF

17. Kim TS, Obst C, Zehaczek S, Geenen C. Detection of bone loss with different X-ray techniques in periodontal patients. J Periodontol 2008;79:1141-9.
PUBMED | CROSSREF

18. Armitage GC. Periodontal diagnoses and classification of periodontal diseases. Periodontol 2000 2004;34:9-21.
PUBMED | CROSSREF

19. Page RC, Eke PI. Case definitions for use in population-based surveillance of periodontitis. J Periodontol 2007;78 (7 Suppl):1387-99.
PUBMED | CROSSREF

20. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 2016;35:1285-98.
PUBMED | CROSSREF

21. Ohsugi H, Tabuchi H, Enno H, Ishitobi N. Accuracy of deep learning, a machine-learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment. Sci Rep 2017;7:9425.
PUBMED | CROSSREF

22. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv e-print 2014:arXiv:1409.556.

23. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th International Conference on International Conference on Machine Learning; 2010 Jun 21–24; Haifa. Madison (WI): Omnipress; 2010. p.807-14.

24. Chollet F. Keras [Internet]. San Francisco (CA): GitHub, Inc.; 2017 [cited 2018 Mar 19]. Available from: https://github.com/keras-team/keras.

25. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. arXiv e-print 2016:arXiv:1603.04467.

26. Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA Intern Med 2015;175:1828-37.
PUBMED | CROSSREF

27. 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;316:2402-10.
PUBMED | CROSSREF

28. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 2017;284:574-82.
PUBMED | CROSSREF

29. Wang R. Edge detection using convolutional neural network. In: Cheng L, Liu Q, Ronzhin A, editors. Advances in neural networks – ISNN 2016. 13th International Symposium on Neural Networks, ISNN 2016; 2016 Jul 6–8; Saint Petersburg. Cham: Springer International Publishing; 2016. p.12-20.

30. Ouyang W, Wang X. Joint deep learning for pedestrian detection. 2013 IEEE International Conference on Computer Vision (ICCV); 2013 Dec 1–8; Sydney. Piscataway (NJ): IEEE; 2013. p.2056-63.

31. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 26–Jul 1; Las Vegas Valley (NV). Piscataway (NJ): IEEE; 2016. p.2818-26.

32. Keskar NS, Mudigere D, Nocedal J, Smelyanskiy M, Tang PT. On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. arXiv e-print 2017;arXiv:1609.0483.

33. 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;542:115-8.
   PUBMED | **CROSSREF**

34. Peng X, Sun B, Ali K, Saenko K. Learning deep object detectors from 3D models. 2015 IEEE International Conference on Computer Vision (ICCV); 2015 Dec 7–13; Santiago. Piscataway (NJ): IEEE; 2015. p.1278-86