



Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography

Makoto Murata¹ · Yoshiko Arijii¹ · Yasufumi Ohashi¹ · Taisuke Kawai² · Motoki Fukuda¹ · Takuma Funakoshi¹ · Yoshitaka Kise¹ · Michihito Nozawa¹ · Akitoshi Katsumata³ · Hiroshi Fujita⁴ · Eiichiro Arijii¹

Received: 11 September 2018 / Accepted: 20 November 2018 / Published online: 11 December 2018
© Japanese Society for Oral and Maxillofacial Radiology and Springer Nature Singapore Pte Ltd. 2018

Abstract

Objectives To apply a deep-learning system for diagnosis of maxillary sinusitis on panoramic radiography, and to clarify its diagnostic performance.

Methods Training data for 400 healthy and 400 inflamed maxillary sinuses were enhanced to 6000 samples in each category by data augmentation. Image patches were input into a deep-learning system, the learning process was repeated for 200 epochs, and a learning model was created. Newly-prepared testing image patches from 60 healthy and 60 inflamed sinuses were input into the learning model, and the diagnostic performance was calculated. Receiver-operating characteristic (ROC) curves were drawn, and the area under the curve (AUC) values were obtained. The results were compared with those of two experienced radiologists and two dental residents.

Results The diagnostic performance of the deep-learning system for maxillary sinusitis on panoramic radiographs was high, with accuracy of 87.5%, sensitivity of 86.7%, specificity of 88.3%, and AUC of 0.875. These values showed no significant differences compared with those of the radiologists and were higher than those of the dental residents.

Conclusions The diagnostic performance of the deep-learning system for maxillary sinusitis on panoramic radiographs was sufficiently high. Results from the deep-learning system are expected to provide diagnostic support for inexperienced dentists.

Keywords Deep learning · Maxillary sinusitis · Panoramic radiography · Artificial intelligence · Computed tomography

Introduction

Maxillary sinusitis is often caused by apical and/or periodontal lesions of the maxillary molars [1, 2]. Such pathology with unilateral symptoms and image appearance is regarded as odontogenic maxillary sinusitis, and patients predominantly visit dental offices. Therefore, dentists often

encounter the necessity for diagnosis of maxillary sinus diseases using panoramic images, which are frequently taken at the first visit. The characteristic radiographic findings of maxillary sinusitis are mucosal thickening exceeding 4-mm thickness, fluid collection, and cyst-like appearance of mucus [3, 4]. However, for inexperienced observers such as dental residents, it can be difficult to accurately diagnose lesions in and around the maxillary sinus on panoramic images [5] because anatomical structures like the hard palate and floor of the nasal cavity are complicatedly superimposed on the images and their appearances change with the panoramic image layers [6–8]. Moreover, small changes in the posterior wall cannot always be detected on panoramic radiographs alone [9].

To address this situation, we developed a computer-aided detection (CAD) system for the detection of unilateral sinusitis to support inexperienced observers. With support from the CAD system, the diagnostic performance of inexperienced observers improved to reach equivalence with that of sufficiently experienced radiologists [5]. However, the CAD

✉ Yoshiko Arijii
yoshiko@dpc.agu.ac.jp

¹ Department of Oral and Maxillofacial Radiology, Aichi-Gakuin University School of Dentistry, 2-11 Suemori-dori, Chikusa-ku, Nagoya 464-8651, Japan

² Department of Oral and Maxillofacial Radiology, School of Life Dentistry at Tokyo, Nippon Dental University, Tokyo, Japan

³ Department of Oral Radiology, Asahi University School of Dentistry, Mizuho, Japan

⁴ Department of Electrical, Electronic and Computer Faculty of Engineering, Gifu University, Gifu, Japan

system can only be applied to unilateral sinusitis because it operates by comparing the opacities of the right and left sinuses.

In recent years, there has been increasing interest in artificial intelligence (AI) with deep-learning procedures. One of the highly established algorithms for deep learning is the convolutional neural network (CNN) [10], which has several processing layers, including learnable operators. Hence, a deep-learning system has the ability to learn and build high-level information from low-level features in an automated manner [11]. Application of deep-learning systems to medical image analysis has been ongoing, with several published studies in areas such as lesion detection [12], classification [13], segmentation [14], image reconstruction [15, 16], and natural language processing [17]. Although a few trials have been conducted using AI systems with deep learning, including classification of the developmental stages of the mandibular third molar [18], this technology has not been applied sufficiently to the dental field.

The purposes of the present study were to apply a deep-learning system to diagnosis of maxillary sinusitis on panoramic images and to clarify its diagnostic performance.

Materials and methods

This study was conducted with approval from the Ethics Committee of our University, and was in accordance with the ethical standards of the Helsinki Declaration. The study was a retrospective study. We posted the research contents on our website and the hospital bulletin board, and gave the subjects the opportunity to refuse to participate in the research.

Subjects

Subjects with or without inflamed lesions in the maxillary sinus whose imaging data were stored in our hospital's image database and who underwent panoramic radiography together with computed tomography (CT) or

cone-beam computed tomography (CBCT) were selected continuously between April 2007 and May 2018. The presence of inflamed lesions was verified by their appearances on CT or CBCT. Sinusitis was defined as mucosal thickening, fluid collection, and/or mucous retention cyst-like appearance extending beyond one-third of the maxillary sinus [3–5]. All subjects with sinusitis visited our hospital because of clinical symptoms, including buccal swelling, pain, and pus discharge around the maxillary teeth. The subjects without sinusitis underwent CT or CBCT for other diseases, such as impacted teeth, jaw deformities, temporomandibular disorders, and cysts limited to the jaw bone. When CT or CBCT images with sufficient inclusion of the entire area of the maxillary sinus showed mucosal thickening of less than 4 mm, the sinus was defined as healthy. Although the panoramic examinations preceded the CT or CBCT examinations by less than 3 weeks, subjects were excluded when no evidence of sinusitis was observed on CT or CBCT images after treatment with antibiotics. Subjects with tumorous lesions, such as fibrous dysplasia and carcinoma, were also excluded. Using CT or CBCT findings as the ground truth, the following diagnostic performances were analyzed on the panoramic radiographs.

For the training data, four separate groups of 100 patients (50 men and 50 women) were retroactively collected from the latest examinations in May 2018, with no difference in mean age among the groups (Table 1, *t* test). The four groups comprised subjects with bilateral healthy sinuses, left inflamed sinuses, right inflamed sinuses, and bilateral inflamed sinuses.

New imaging data that differed from the training datasets were prepared as test datasets, to achieve a situation close to clinical application. These data were obtained for two groups, comprising 30 subjects (15 men and 15 women) with bilateral healthy sinuses and 60 subjects (30 men and 30 women) with unilateral inflamed sinuses, and were continuously collected retroactively from April 2010 to avoid repeated selection of the subjects with images included in the training data. The selected subjects are

Table 1 Subjects

Subject groups	Numbers of subjects			Age (years)	
	Men	Women	Total	Mean ± SD	Range
For training process					
Subjects with bilateral healthy sinuses	50	50	100	50.0 ± 18.1	20–87
Subjects with left inflamed sinus	50	50	100	51.0 ± 18.3	20–87
Subjects with right inflamed sinus	50	50	100	50.6 ± 18.1	20–87
Subjects with bilateral inflamed sinuses	50	50	100	50.6 ± 18.8	20–90
For testing process					
Subjects with bilateral healthy sinuses	15	15	30	48.7 ± 18.4	21–87
Subjects with unilateral inflamed sinuses	30	30	60	48.9 ± 17.4	20–83

summarized in Table 1. There was no difference in mean age between the two groups (*t* test).

The CT examinations were performed using an Asterion TSX machine (Canon Medical Systems, Otawara, Japan). Axial scans of the maxillary region were taken parallel to the occlusal plane with the following parameters: tube voltage, 120 kV; charge, 50 mAs; slice thickness, 0.5 mm; pitch, 0.3 mm; field of view, 20 cm. The CBCT images were taken with an Alphard Vega scanner (Asahi Roentgen Ind. Co. Ltd., Kyoto, Japan). I-mode (102 × 102 mm) imaging was used with a voxel size of 0.2 mm. The CBCT scans were performed with 360° rotation and the following exposure parameters: tube voltage, 80 kV; tube current, 8 mA; acquisition time, 17 s. Panoramic images were obtained using a Veraview Epocs system (J. Morita Mfg Corp., Kyoto, Japan) with average parameters: tube voltage, 75 kV; tube current, 9 mA; acquisition time, 16 s.

Preparation of training image patches and data augmentation

To create the image patches for the training process, regions of interest (ROIs) were set on the panoramic images as squares of 200 × 200 pixels that included a unilateral maxillary sinus (Fig. 1). In the preliminary analysis of ROI size, the diagnostic performance was higher for setting a respective ROI on one side of the maxillary sinuses than for setting one ROI in both sides of the maxillary sinuses. Thus, the panoramic radiographs were cropped on one side of the maxillary sinuses. The bilateral maxillary sinuses were cropped semi-automatically outside of the ROIs on the panoramic images of all 300 subjects using the macro function of Adobe Photoshop v. 13.0 (Adobe Systems Co. Ltd., USA). As a result, 400 sinus image patches each of healthy and inflamed sinuses were created (Table 2).

To improve the reliability of the learning model, data augmentation was performed to increase the number of training

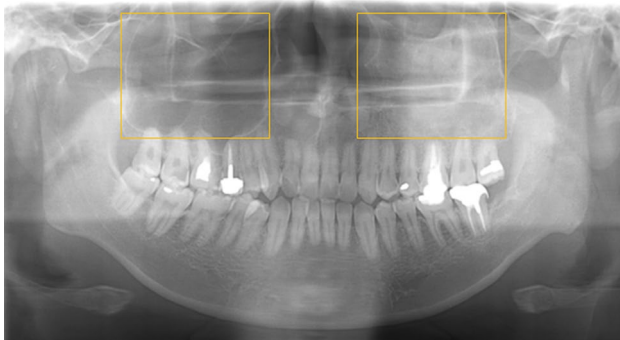


Fig. 1 Segmentation on panoramic radiographs. A region of interest of size 200 × 200 pixels was set, and the maxillary sinuses were cropped out on both sides

Table 2 Image patches

Image patches	Numbers of image patches	
	Cropped patches	After data augmentation
For training process		
Healthy sinuses	400	6000
Inflamed sinuses	400	6000
For testing process		
Healthy sinuses	60	
Inflamed sinuses	60	

image patches using IrfanView software (<https://www.irfanview.com/>). The details for the data augmentation are shown in Table 3. Rotation (horizontal flipping) was not applied because this process was performed automatically in DIGITS. As a result of changing the brightness, contrast, and sharpness, 6000 image patches each were created for healthy and inflamed sinuses.

Preparation of testing image patches

The image patches for testing were prepared in the same manner as those for training (Table 2). Sixty image patches showing healthy sinuses were created from the subjects with bilateral healthy maxillary sinuses, and 60 image patches showing inflamed sinuses were cropped from the affected side of subjects with unilateral inflamed sinuses.

Deep-learning procedure

The deep-learning system was built on a system running Ubuntu OS v. 16.04.2 with 128 GB of memory and 11 GB of GPU memory (NVIDIA GeForce GTX 1080 Ti). As a deep-learning tool, an AlexNet CNN implemented with DIGITS library v. 5.0 (NVIDIA; <https://developer.nvidia.com/digits>) was used on the Caffe framework. AlexNet consists of five convolutional layers and three fully-connected layers

Table 3 Filters used and adapted values for data augmentation

Used filters	Adapted values
None	① None (original)
Automatic tone correction filter	② Automatic tone correction
Brightness filter (−80 to 80)	③ −80, ④ −50, ⑤ 50, ⑥ 80
Contrast filter (−127 to 127)	⑦ −90, ⑧ −50, ⑨ 50, ⑩ 90
Sharpness filter (1–99)	⑪ 50, ⑫ 90
Gamma correction filter (0.01–6.99)	⑬ 1.50
Blur filter (1–99)	⑭ 40
Middle value filter (3–9)	⑮ 9

(Fig. 2). The learning process was performed using the standard classification function in the DIGITS library.

The prepared training image patches were applied to the deep-learning system, and the learning process was performed for 200 epochs. The testing image patches were applied to the created learning model, and its diagnostic performance was calculated. The accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were obtained. A receiver-operating characteristic (ROC) curve was drawn and the area under the curve (AUC) was determined.

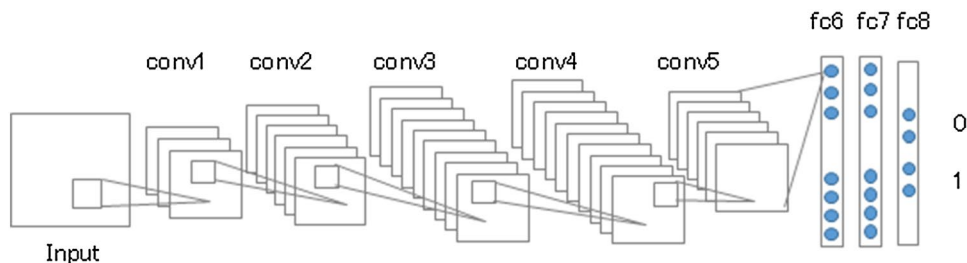
Comparison with diagnostic performance of human observers

Two radiologists with more than 20 years of experience and two dental residents with 1 year of experience individually evaluated the panoramic images for the presence or absence of inflamed lesions in the maxillary sinuses. After practice on several panoramic radiographs, actual observations were performed. A total of 120 sinuses selected for testing data were evaluated on a personal monitor (RadiForce G20; Eizo Nanao Corp., Ishikawa, Japan), with a size of 20.1 inches and resolution of 1600×1200 pixels. The observers evaluated the probability for the presence of inflamed lesions on a four-point rating scale: 1, absent; 2, probably absent; 3, probably present; 4, present. The accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and AUC of the evaluation were calculated. For the evaluation of performance, scores 1 and 2 were regarded as healthy sinuses and scores 3 and 4 were regarded as inflamed sinuses. The diagnostic performance of the deep-learning system was compared with that of the human observers.

Statistical analysis

Differences in mean age between the groups were compared by a *t* test. Comparisons of AUC values were performed by the Chi-squared test. Interobserver Cohen kappa coefficients were obtained. Diagnostic performance was compared between the experienced radiologists and dental residents by the Mann–Whitney *U* test. Values of $p < 0.05$ were considered statistically significant.

Fig. 2 Convolution neural network. AlexNet consists of five convolutional (conv) layers and three fully connected (fc) layers



Results

The time required for the deep-learning process was as follows: 26 s to set up each dataset, and 28 min to input 12,000 training datasets into the deep-learning system using AlexNet, learn for 200 epochs, and create the learning model. The time required to input 120 test datasets into the learning model and judge the presence or absence of maxillary sinusitis was 9 s.

The accuracy and learning losses for the 200 epochs of the learning process are shown in Fig. 3. The accuracy, sensitivity, and specificity of the deep-learning system were 87.5%, 86.7%, and 88.3%, respectively (Table 4). The corresponding values of the radiologists were 89.6%, 90.0%, and 89.2%, respectively, while those of the dental residents were 76.7%, 78.3%, and 75.0%, respectively (Table 4). The diagnostic performance of the deep-learning system was similar to those of the radiologists.

The AUCs were 0.875, 0.896, and 0.767 for the deep-learning system, radiologists, and dental residents, respectively (Fig. 4). The AUC for the dental residents was significantly lower than the AUCs for the deep-learning system and radiologists (Chi-squared test: deep-learning system vs. residents, $p = 0.0018$; radiologists vs. residents, $p = 0.0001$).

The interobserver kappa coefficient of the two radiologists was 0.733, indicating substantial agreement, while that of the two dental residents was 0.508, indicating moderate agreement.

Discussion

The deep-learning system showed sufficiently high diagnostic performance close to that of the two radiologists, with accuracy, sensitivity, and specificity of 87.5%, 86.7%, and 88.3%, respectively, and AUC of 0.875. Meanwhile, the dental residents had significantly lower values than those obtained for both the deep-learning system and the radiologists. These findings indicate that the newly-developed AI system has the potential to provide diagnostic support for inflamed maxillary sinus diseases on panoramic images, especially for inexperienced observers.

Fig. 3 The results of learning using image patches of each ROI. The blue line shows the loss on the training data set. This loss represents the fit between a prediction and the ground-truth label, and decreases over time as training improves. The orange line shows the accuracy during the training process. The accuracy increases over time, reaching a final accuracy of about 80% at the final epoch

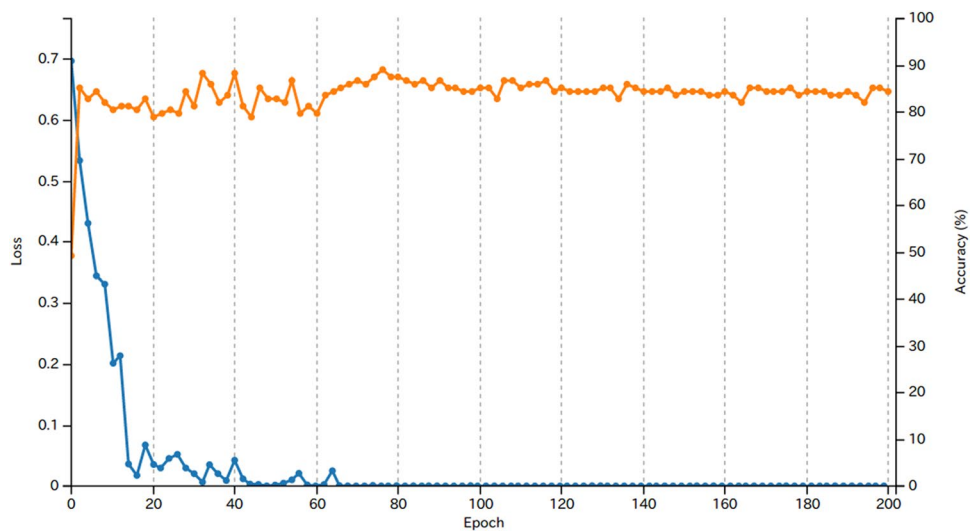


Table 4 Diagnostic performances of deep-learning system, radiologists, and residents

	Deep-learning system	Radiologists	Residents
Accuracy	87.5	89.6	76.7
Sensitivity	86.7	90.0	78.3
Specificity	88.3	89.2	75.0
Positive predictive value	88.1	89.3	75.8
Negative predictive value	86.9	89.9	77.6

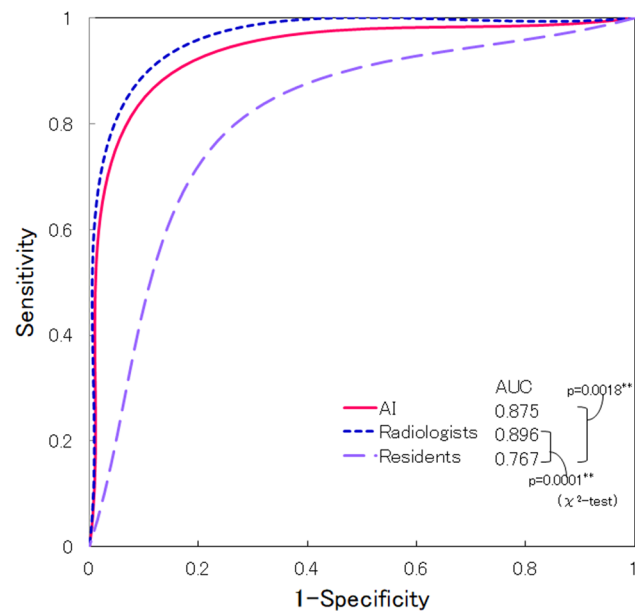


Fig. 4 Receiver-operating characteristic curves for the deep-learning artificial intelligence (AI) system and human observers. The area under the curve (AUC) values were 0.875, 0.896, and 0.767 for the AI system, radiologists, and dental residents, respectively

When image patches are input into a multi-layered CNN, a deep-learning system repeats the learning process and automatically builds a learning model [11]. The resultant models have been applied to various tasks, including lesion detection [12], classification [13], segmentation [14], and natural language processing [17]. In the present study, a classification task was performed for diagnosis of maxillary sinusitis on panoramic radiography using the AlexNet CNN system, which consists of five convolution layers and three fully-connected layers [19]. Although more advanced technologies, such as GoogleNet [12], may be applicable, AlexNet showed sufficient effectiveness on the 12,000 image patches used in the present study, and could be operated in a short time without any problems. Data augmentation is often used in deep-learning studies when clinical datasets of limited size are available [19–22]. By changing the image filters, including brightness, contrast, and noise, it is possible to increase the size of the training dataset by several times. Although we did not perform confirmation methods to determine the number of training image patches required to achieve a sufficiently qualified learning model, the present dataset appeared appropriate because of the high diagnostic performance obtained.

In our previous study, a CAD system was developed for diagnosis of unilateral maxillary sinusitis on panoramic images [5]. Inflamed lesions in the unilateral maxillary sinuses can be detected by comparing the opacities of the left and right sinuses using a subtraction technique. With the assistance of this CAD system, the diagnostic performance of inexperienced observers using panoramic images was improved to resemble that of experienced observers. However, experienced observers showed high performance even without the CAD system, and their performance was not improved by use of the CAD system. Although the present results cannot be directly compared with those of the CAD

system, experienced radiologists in both studies showed similarly high performances equivalent to that of the AI system in the present study. The CAD system was solely used to assist observers in rendering interpretations, whereas the AI system can classify sinus diseases by itself and appears to be a possible tool for automated diagnosis because of its high diagnostic performance.

During diagnosis of maxillary sinus diseases on panoramic images, radiologists often pay attention to sinus appearance changes, such as morphology and density, and simultaneously compare them with those of the contralateral sinus. According to these considerations and our previous study on the CAD system, it may be preferable that the image patches are cropped from the ROIs to include the bilateral sinuses in cases of unilateral sinusitis. In clinics, however, there are patients with bilateral sinus diseases or bilateral healthy sinuses. Moreover, the diagnostic process using AI may differ from that by human radiologists. In addition, image reversion was used during the data augmentation process, indicating that the left and right sides of the image patches were treated in the same manner. Therefore, the learning model in the present study was created for single sinuses and the results showed sufficiently high diagnostic performance. However, additional verification with ROIs including both sinuses may be necessary for more advanced results.

The present study can be assigned as part of a body of research aiming to develop a completely automated system for diagnosis on panoramic images. Many issues need to be resolved to accomplish this purpose. Although the image patches were created semi-automatically using commercially available software in the present study, development of a fully automated image segmentation procedure is a significant task required to reach the final goal. Further research should be promoted with the aim of clinical application, by increasing training datasets to include images obtained from different panoramic units in other facilities and creating more accurate deep-learning models.

Compliance with ethical standards

Conflict of interest Makoto Murata, Yoshiko Arijii, Yasufumi Ohashi, Taisuke Kawai, Motoki Fukuda, Takuma Funakoshi, Yoshitaka Kise, Michihito Nozawa, Akitoshi Katsumata, Hiroshi Fujita, and Eiichiro Arijii declare that they have no conflict of interest.

Human rights statement All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and later versions.

Animal rights statement This article does not contain any studies with animal subjects performed by any of the authors.

Informed consent Informed consent was obtained from all patients for being included in the study.

References

- Maillet M, Bowles WR, McClanahan SL, John MT, Ahmad M. Cone-beam computed tomography evaluation of maxillary sinusitis. *J Endod*. 2011;37:753–7.
- Obayashi N, Arijii Y, Goto M, Izumi M, Naitoh M, Kurita K, et al. Spread of odontogenic infection originating in the maxillary teeth: computerized tomographic assessment. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2004;98:223–31.
- Maestre-Ferrín L, Galán-Gil S, Carrillo-García C, Peñarrocha-Diogo M. Radiographic findings in the maxillary sinus: comparison of panoramic radiography with computed tomography. *Int J Oral Maxillofac Implants*. 2011;26:341–6.
- Yoshiura K, Ban S, Hijjiya T, Yuasa K, Miwa K, Arijii E, et al. Analysis of maxillary sinusitis using computed tomography. *Dentomaxillofac Radiol*. 1993;22:86–92.
- Ohashi Y, Arijii Y, Katsumata A, Fujita H, Nakayama M, Fukuda M, et al. Utilization of computer-aided detection system in diagnosing unilateral maxillary sinusitis on panoramic radiographs. *Dentomaxillofac Radiol*. 2016;45:20150419.
- Yoshida K, Fukuda M, Gotoh K, Arijii E. Depression of the maxillary sinus anterior wall and its influence on panoramic radiography appearance. *Dentomaxillofac Radiol*. 2017;46:20170126.
- Damante JH, Filho LI, Silva MA. Radiographic image of the hard palate and nasal fossa floor in panoramic radiography. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 1998;85:479–84.
- Suomalainen A, Pakbaznejad Esmaeili E, Robinson S. Dentomaxillofacial imaging with panoramic views and cone beam CT. *Insights Imaging*. 2015;6:1–16.
- Ohba T, Ogawa Y, Shinohara Y, Hiromatsu T, Uchida A, Toyoda Y. Limitations of panoramic radiography in the detection of bone defects in the posterior wall of the maxillary sinus: an experimental study. *Dentomaxillofac Radiol*. 1994;23:149–53.
- Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging*. 2018;9:611–629. <https://doi.org/10.1007/s13244-018-0639-9> (Epub ahead of print).
- Gao XW, Hui R, Tian Z. Classification of CT brain images based on deep learning networks. *Comput Methods Progr Biomed*. 2017;138:49–56.
- Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*. 2017;284:574–82.
- Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. *Radiology*. 2018;286:887–96.
- Christ PF, Elshaer MEA, Ettlinger F, Tatavarty S, Bickel M, Bilic P, et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W, editors. *Proceedings of medical image computing and computer-assisted intervention—MICCAI 2016*. Cham: Springer; 2016. p. 415–23. https://doi.org/10.1007/978-3-319-46723-8_48.
- Kim KH, Choi SH, Park SH. Improving arterial spin labeling by using deep learning. *Radiology*. 2018;287:658–66.
- Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging-based attenuation correction for PET/MR imaging. *Radiology*. 2018;286:676–84.
- Chen MC, Ball RL, Yang L, Moradzadeh N, Chapman BE, Larson DB, et al. Deep learning to classify radiology free-text reports. *Radiology*. 2018;286:845–52.
- De Tobel J, Radesh P, Vandermeulen D, Thevissen PW. An automated technique to stage lower third molar development

- on panoramic radiographs for age estimation: a pilot study. *J Forensic Odontostomatol.* 2017;2:42–54.
19. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst.* 2012;1–9.
 20. Wang H, Zhou Z, Li Y, Chen Z, Lu P, Wang W, et al. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from 18F-FDG PET/CT images. *EJNMMI Res.* 2017;7:11.
 21. Lee H, Tajmir S, Lee J, Zissen M, Yeshiwas BA, Alkasab TK, et al. Fully automated deep learning system for bone age assessment. *J Digit Imaging.* 2017;30:427–41.
 22. Song Q, Zhao L, Luo X, Dou X. Using deep learning for classification of lung nodules on computed tomography images. *J Healthc Eng.* 2017;2017:8314740.