



Deep learning object detection of maxillary cyst-like lesions on panoramic radiographs: preliminary study

Hirofumi Watanabe¹ · Yoshiko Ariji¹ · Motoki Fukuda¹ · Chiaki Kuwada¹ · Yoshitaka Kise¹ · Michihito Nozawa¹ · Yoshihiko Sugita² · Eiichiro Ariji¹

Received: 17 June 2020 / Accepted: 5 September 2020 / Published online: 19 September 2020
© Japanese Society for Oral and Maxillofacial Radiology and Springer Nature Singapore Pte Ltd. 2020

Abstract

Objectives This study aimed to examine the performance of deep learning object detection technology for detecting and identifying maxillary cyst-like lesions on panoramic radiography.

Methods Altogether, 412 patients with maxillary cyst-like lesions (including several benign tumors) were enrolled. All panoramic radiographs were arbitrarily assigned to the training, testing 1, and testing 2 datasets of the study. The deep learning process of the training images and labels was performed for 1000 epochs using the DetectNet neural network. The testing 1 and testing 2 images were applied to the created learning model, and the detection performance was evaluated. For lesions that could be detected, the classification performance (sensitivity) for identifying radicular cysts or other lesions were examined.

Results The recall, precision, and F-1 score for detecting maxillary cysts were 74.6%/77.1%, 89.8%/90.0%, and 81.5%/83.1% for the testing 1/testing 2 datasets, respectively. The recall was higher in the anterior regions and for radicular cysts. The sensitivity was higher for identifying radicular cysts than for other lesions.

Conclusions Using deep learning object detection technology, maxillary cyst-like lesions could be detected in approximately 75–77%.

Keywords Deep learning · Object detection · Maxillary cysts · Radicular cysts · Panoramic radiography

Introduction

Application of artificial intelligence with deep learning system in the field of medical imaging has been increasing [1–5], prompting the appearance of various studies on the computer-assisted detection (CAD) system to diagnose pathology in the dental field [6–12]. Unlike the CAD systems created based on traditional methodology [13], deep learning system, which does not require manual input of imaging characteristics of lesions, has enabled the creation of a learning model by simply importing imaging datasets into the system. Among several functions of deep learning, the object detection, which can automatically detect specific

lesions and conditions, has been applied to panoramic radiographs [9–12]. A recent study verified a high performance for automatic detection of mandibular radiolucent lesions (cysts and benign tumors) on panoramic radiographs [9]. This would be partially attributed to relatively high contrast between the radiolucent lesions and their surrounding mandibular bony structures showing large X-ray attenuation. In the maxilla, however, the cyst-like lesions including cysts and benign tumors would be predictably difficult to detect, because air-containing structures showing radiolucency, such as the nasal cavity and maxillary sinus, would be located adjacent to the lesions and decrease the contrast between them. Moreover, overlapped shadows of the hard palate and inferior nasal concha should be taken into account as obstructive effects [13].

The purpose of this study was to determine the ability of deep learning object detection technology to identify maxillary cyst-like lesions indicating osteolytic benign lesions arising in the maxilla.

✉ Yoshiko Ariji
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 Pathology, Aichi-Gakuin University School of Dentistry, Nagoya, Japan

Materials and methods

This study was approved by the ethics committee of our University (Nos. 496 and 577) and was conducted in compliance with the Declaration of Helsinki.

Subjects

Subjects were patients suspected of having cyst-like lesions (cyst or benign tumors arising in the maxilla). They showed maxillary bone resorption of ≥ 10 mm diameter on panoramic radiographs, and had confirmed histopathological diagnoses (including several odontogenic tumors). Altogether, 412 patients were selected from the imaging database of our dental hospital, from April 2009 to April 2019. Details of the subjects are shown in Table 1. The mean age of the subjects was 44.8 ± 16.6 years, there were 221 men and 191 women.

The categories of lesions were as follows: radicular cysts (299 patients, 323 cysts), nasopalatine duct cysts (44 patients, 44 cysts), dentigerous cysts (36 patients, 37 cysts), odontogenic keratocysts (22 patients, 23 cysts),

ameloblastomas (3 patients, 3 tumors), glandular odontogenic cysts (3 patients, 3 cysts), odontogenic myxomas/myxofibromas (2 patients, 2 tumors), and adenomatoid odontogenic tumors (1 patient, 1 tumor). In all, there were 291 (66.7%) affected sites in the anterior regions, and 145 (33.3%) in the posterior regions.

Preparation of images

Panoramic radiographs were taken using a Veraview epocs system (J. Morita Mfg Corp., Kyoto, Japan) with the standard parameters, including a tube voltage of 75 kV, tube current of 9 mA, and acquisition time of 16 s. Panoramic radiographs were downloaded in Bitmap form (.BMP) from the imaging database. All images were standardized to squares of 900×900 pixels and 24 bits per pixel based on the image width.

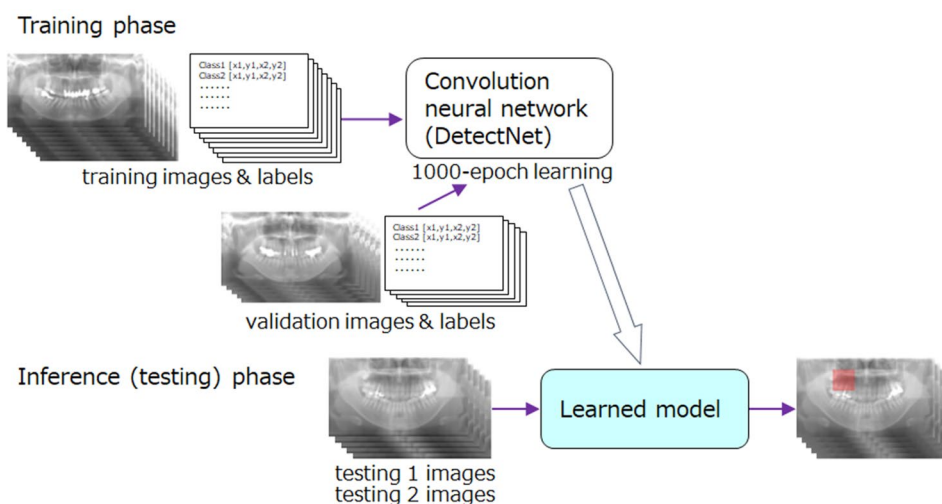
Assignment of training and test datasets

All images were arbitrarily assigned to the training, testing 1, and testing 2 datasets (Fig. 1). The testing 1 dataset was used for testing and validation. Validation was used to

Table 1 Summary of subjects and assignment to datasets

Lesions	No of patients	Age (mean \pm sd)	Gender (M/F)	No of lesions	Regions		Assignment to datasets		
					Anterior	Posterior	Training dataset	Testing 1 dataset	Testing 2 dataset
Radicular cysts	299	45.6 ± 15.1	148/151	323	214	109	248	50	25
Nasopalatine duct cysts	44	48.0 ± 14.8	30/14	44	44	0	32	8	4
Dentigerous cysts	36	35.6 ± 22.7	25/11	37	21	16	27	7	3
Odontogenic keraocysts	22	46.3 ± 21.6	12/10	23	8	15	17	4	2
Others	9	41.7 ± 20.3	4/5	9	4	5	6	2	1
Total	410	44.9 ± 16.6	219/191	436	291	145	330	71	35

Fig. 1 Description of each dataset and deep learning workflow. All images were randomly assigned to the training, testing 1, and testing 2 datasets. Using training images and labels, the learning process of 1000 epochs were conducted with the convolutional neural network "Detectnet". Validation was used to adjust the parameters to determine the final model in the learning process. The testing 1 and testing 2 images were applied to the created learning model, and therefore, the lesional areas were depicted as rectangles



adjust parameters to determine the final model in the learning process. The testing 2 dataset was used only for testing as new external data. Each lesion was set to be entered in each dataset in equal proportions. The numbers of lesions in each dataset are shown in Table 1.

Annotation

On training and validation (testing 1) images, annotation (labeling) procedure was performed. The labels containing the type of lesions (class name) and the coordinates of lesions were created in text format. The class names of the lesions were determined as jaw1 for radicular cysts and jaw2 for other lesions. The coordinates of the upper left and lower right corners of the square regions of interest surrounding the lesions were recorded per image. (Fig. 2).

Construction of deep learning

The deep learning system was implemented on Ubuntu OS and 11MB GPU (Nvidia GeForce GTX). The deep learning process was performed using the DetectNet neural network implemented with the DIGITS library on the Caffe framework.

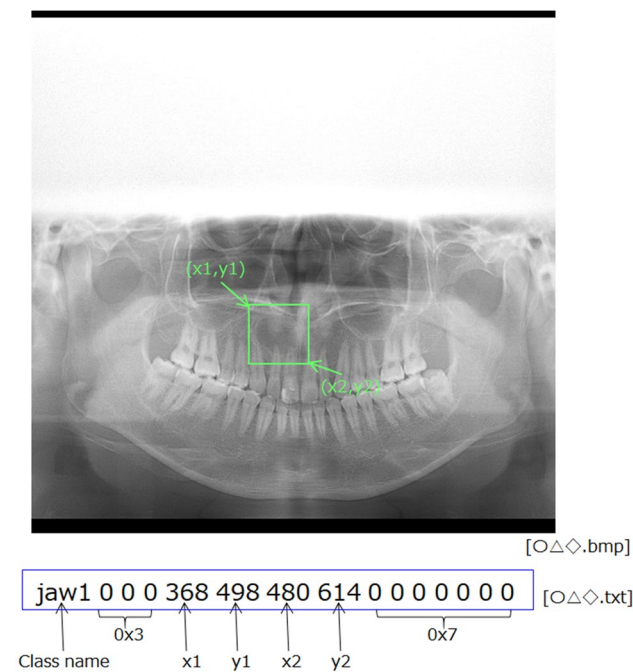


Fig. 2 Example of panoramic radiographs used (900×900 pixels) and the annotation method. A rectangle region of interest of an arbitrary size was set for the lesion, and the x and y coordinates of the upper left corner and the lower right corner were recorded. A label including the lesion class name (jaw1) and the x and y coordinates of the upper left and lower right corners was created in text format

DetectNet is a convolutional neural network developed for object detection. This subnetwork has the same structure as GoogLeNet, except for the data input, final pooling, and output layers [14]. The ADAM (Adaptive Moment Estimation) solver with 0.0001 as the base learning rate was used for computing adaptive learning rates for each parameter.

Detection performance

The training images and labels were imported into Digits, and the learning process was conducted for 1000 epochs. We confirmed that the loss was sufficiently small and the detection performance was sufficiently large and stable in the learning curve of 1000 epochs. A validation using the testing 1 dataset was performed for adjusting parameters and determining the final model (Fig. 1).

After applying, the testing 1 and testing 2 images were applied to the created learning model, radicular cysts were depicted as red rectangles and the other lesions as blue rectangles on each testing panoramic radiograph. Thereafter, the detection performance was evaluated. It was determined that a lesion was detected when the following intersection over union (IoU) was ≥ 0.6 .

$$IoU = S(P \cap G) / S(P \cup G)$$

where P is the predicted bounding box in which the learning model was predicted to have a lesion; G is the ground-truth bounding box, which actually has a lesion; $S(P \cap G)$ is the overlapping area of two bounding boxes; and $S(P \cup G)$ is the total combined area of two bounding boxes.

The affected areas were confirmed histopathologically, and the ground-truth bounding boxes were determined by two experienced radiologists. The lesion-free areas were defined as being clinically asymptomatic and having undergone observation for more than half a year. A true positive (TP) means that the model correctly detected the lesional area as a lesion. A false positive (FP) means that the model incorrectly detected the lesion-free area as a lesion. A false negative (FN) means that the model could not detect the lesional area. For evaluation of detection performance, the following indexes were calculated [15, 16].

$$\text{Recall (sensitivity)} = TP / (TP + FN)$$

$$\text{Precision (positive predictive value)} = TP / (TP + FP)$$

$$F1 \text{ score} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$$

Recall and precision are in a trade-off relation and their harmonic mean (F1-score) is used when the both values should be taken into account.

Classification performance

For lesions that could be detected, the sensitivity to correctly identify the lesion class as jaw 1 (radicular cysts) or jaw 2 (other lesions) was calculated.

Results

It took 19 s to import the training dataset to Digits, 10 h to perform the deep learning process for 1000 epochs, and 12 s to apply each testing datasets to the learning model.

Detection performance

The recall, precision, and F-1 score for detecting maxillary cyst-like lesions were 74.6%/77.1%, 89.8%/90.0%, and 81.5%/83.1% for the testing 1/testing 2 datasets, respectively (Table 2). Compared by sites, the detection performances in the anterior regions was higher than that in the posterior regions for both testing datasets. Compared by lesions, the recall was higher for radicular cysts than for other lesions.

Representative examples are shown: well-defined radicular cysts in the posterior regions could be correctly detected (Fig. 3a), whereas odontogenic keratocysts extending for almost the whole maxillary sinus could not be detected (Fig. 3b). The false-positive results were found mostly in the anterior region where bone was thin (Fig. 4a), and in the posterior regions where the maxillary sinus expanded to around the alveolar ridge (Fig. 4b).

Classification performance

The sensitivities for identifying radicular cysts were 94.9 and 100.0%, respectively, for the testing 1 and testing 2 datasets, whereas those of the other lesions were 64.3 and 71.4%, respectively. The sensitivity was higher for radicular cysts than for other lesions.

Representative examples are shown: Fig. 5a shows a nasopalatine duct cyst that was successfully detected and

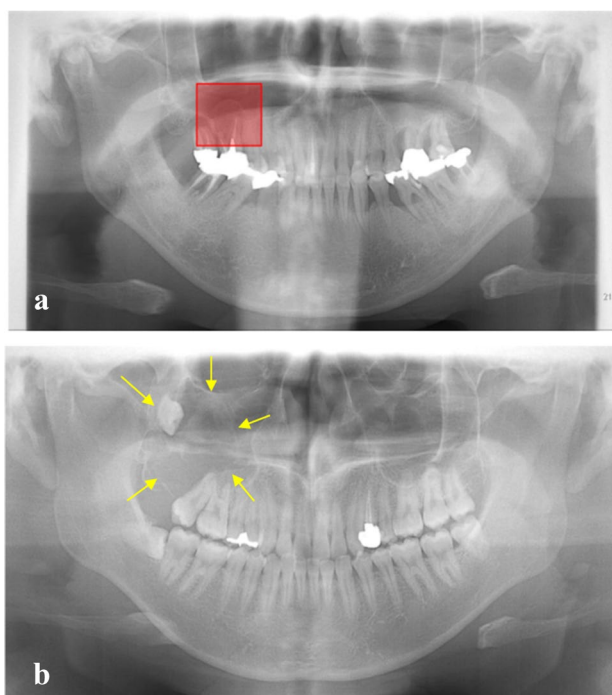


Fig. 3 Examples with success (a) and failure (b) of detecting lesions in the molar regions. **a** Well-defined radicular cyst was detected. **b** Odontogenic keratocyst that extends the almost throughout the maxillary sinus could not be detected

classified. Figure 5b shows an ameloblastoma in the molar region, where the lesion was detected but was incorrectly classified as a radicular cyst.

Discussion

Several neural networks for object detection have been reported [4, 9, 10, 17–20]. DetectNet, which is one of them, has a subnetwork similar to that of GoogLeNet, except for the data input, final pooling, and output layers [14]. It has the great advantage of using a pre-trained GoogLeNet model, thereby reducing the training time and improving the

Table 2 Detection performance results

	Recall (sensitivity) for testing 1/testing 2 datasets	Precision (positive predictive value) for testing 1/testing 2 datasets	F-1 scores for testing 1/testing 2 datasets
Sites			
Anterior region	80.0%/100.0%	94.7%/91.7%	86.7%/95.7%
Posterior region	65.4%/46.7%	85.0%/87.5%	73.9%/60.9%
Lesions			
Radicular cysts	78.0%/80.0%	88.6%/87.0%	83.0%/83.3%
Other lesions	66.7%/70.0%	93.3%/100.0%	77.8%/82.4%
Total	74.6%/77.1%	89.8%/90.0%	81.5%/83.1%

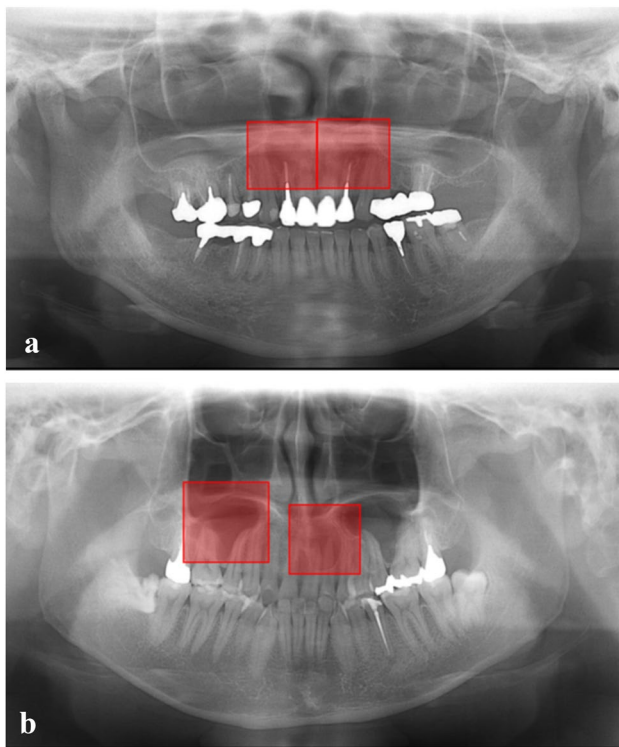


Fig. 4 Examples with false-positive results. **a** Radicular cyst in the right lateral incisor was detected correctly. Although there was no lesion in the left lateral incisor, deep learning indicated the region as a lesion (false positive). The false-positive identification was made in the anterior region where there was thin bone. **b** Radicular cyst in the left incisors was detected correctly. Although there was no lesion in the right molars, deep learning indicated the region as a lesion (false positive). The false-positive identification was made in the posterior region where the maxillary sinus expanded to around the alveolar ridge

accuracy of the final model. DetectNet enables object detection and classification by extracting features and predicting object classes and bounding boxes for each grid square.

Our previous study achieved a high detection sensitivity (recall) of 88% for identifying mandibular radiolucent lesions on panoramic radiographs using deep learning object detection technology [9]. Based on the results of the present study, the detection recall for maxillary lesions was 75 and 77%. As predicted, this result was lower than that for mandibular radiolucent lesions [9], which may be due to low contrast between the lesions and surrounding structures, and the lesions being overlapped by the nasal cavity, maxillary sinus, and other anatomical structures [13].

A comparison between anterior and posterior lesions revealed that the detection performances (recall, precision, and F1-score) were higher in the anterior regions, possibly because there were more anterior lesions and they might have been more sufficiently learned. In addition, large lesions of the posterior regions (e.g., odontogenic

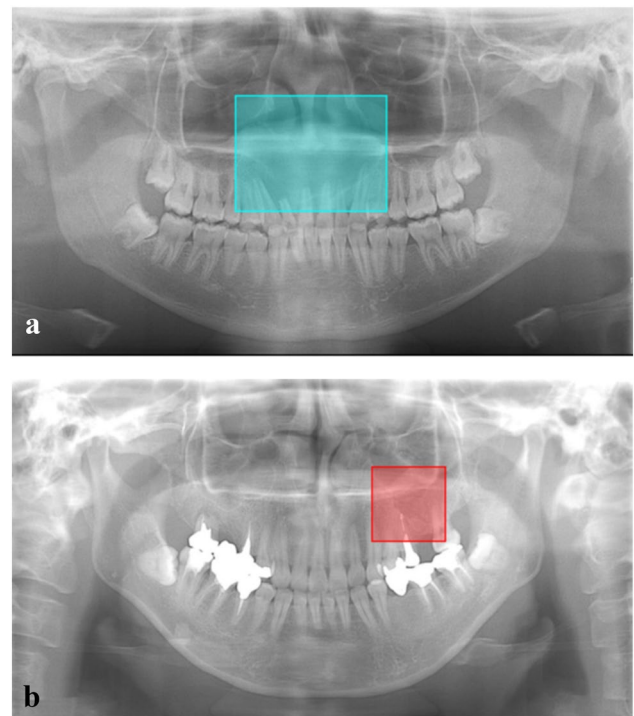


Fig. 5 Examples of success (**a**) and failure (**b**) of classifying lesions. **a** Nasopalatine duct cyst was successfully detected and classified represented in a blue box. **b** Ameloblastoma in the molar region was detected, but falsely classified as a radicular cyst represented in a red box

keratocysts, dentigerous cysts, and other odontogenic tumors) were mostly found in the maxillary sinus (Fig. 3b) and so might have been misdiagnosed as maxillary sinus lesions or sinus itself, and therefore not detected.

The false-positive identifications were often seen in the maxillary lateral incisors and canines, or in the molar areas. The former site including the canine fossa often appeared radiolucent because of the thin bone thickness, and therefore, the judgement would have been incorrect (Fig. 4a). In the latter site, the maxillary sinus sometimes extended to the level of the alveolar ridge, and was misdiagnosed as a lesion (Fig. 4b).

When the performances per lesion were compared, the detection recall and classification sensitivities for radicular cysts were slightly higher than those for the other lesions. The difference in the detection performance was because there were more radicular cysts, and so the learning process might have favored their identification. In addition, it was helped by the fact that radicular cysts were frequently found around the tooth root apex.

The lesions of a wide range of radiolucent lesions from apical lesions to tumors were enrolled in this study, because the classification (diagnosis) performances per lesions were compared. Radicular cysts in the anterior regions should be

differentiated from nasopalatine duct cysts, and dentigerous cysts. These lesions might be easily differentiated because of the relations between the lesions and teeth. That is, radicular cysts occur mostly around the root apices, nasopalatine ducts cysts are in the midline region regardless of the teeth's positions, and dentigerous cysts are around the crowns of impacted teeth.

Radicular cysts in the posterior regions must be differentiated from odontogenic keratocysts, dentigerous cysts, and other odontogenic tumors. Except for radicular cysts, the number of cases was small, and thus learning about each of the lesions would not be enough. As already noted, the large cyst-like lesions occupy almost all of the maxillary sinus, such as odontogenic keratocysts, dentigerous cysts, and other odontogenic tumors, might be misdiagnosed as sinus lesions or sinus itself. Conversely, the small ameloblastomas might be misdiagnosed as radicular cysts.

Although the present study could provide a learning model for detecting and classifying maxillary radicular cysts with high performances, there were some limitations. The number of the other lesions should be increased. The large lesions occupying almost all of the maxillary sinuses should be learned in comparison with maxillary sinus lesions. The learning models should include using panoramic radiographs from other institutions that were obtained with different parameters. Furthermore, the maxillary cyst-like lesions should be learned together with mandibular lesions, because it would be convenient if the lesions of both jaws can be detected in a single test using the detection system in clinical practice.

It is unknown whether the results of detection using DetectNet network reflect imaging findings including unilocular vs multilocular, or buccal and lingual expansion. Future networks may be able to provide information on the learning process.

The testing 1 dataset was used for validation and testing. A validation adjusts the parameters during training process and determines the final model. The testing 2 dataset is used for only testing, so it was completely new external data. It will need to increase the number of external data. However, the rates for individual lesions in testing 1 and 2 datasets were almost equal as shown in Table 2. The difference in detection performances between the two datasets may not be from the difference in the number of data, but probably due to the difference in the imaging findings, such as lesion shapes and sizes.

To improve the detection performance, it is necessary to increase the number of training data. Performance will be improved by learning with a large number of training data from multiple facilities to reduce the effects of positioning errors of panoramic radiography. Increasing training data with improved networks may lead to potential clinical applications.

In conclusion, we developed a deep learning model for detecting maxillary cyst-like lesions with a detection recall of 74.6 and 77.1%. The classification performances for radicular cysts were 94.9 and 100.0%, whereas those of the other lesions were lower. Further studies using more data are needed.

Acknowledgements We thank Nancy Schatken, BS, MT(ASCP), from Edanz Group (<https://en-author-services.edanzgroup.com/>) for editing a draft of this manuscript.

Compliance with ethical standards

Conflict of interest Hirofumi Watanabe, Yoshiko Ariji, Motoki Fukuda, Chiaki Kuwada, Yoshitaka Kise, Michihito Nozawa, Yoshihiko Sugita, and Eiichiro Ariji declare that they have no conflicts of interest.

Human rights statements and informed consent All procedures were performed 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.

References

1. Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, Summers RM, Giger ML. Deep learning in medical imaging and radiation therapy. *Med Phys*. 2019;46:e1–e36.
2. Cao Z, Yang G, Chen Q, Chen X, Lv F. Breast tumor classification through learning from noisy labeled ultrasound images. *Med Phys*. 2020;47:1048–57.
3. Wang X, Liang X, Jiang Z, Nguchu BA, Zhou Y, Wang Y, Wang H, Li Y, Zhu Y, Wu F, Gao JH, Qiu B. Decoding and mapping task states of the human brain via deep learning. *Hum Brain Mapp*. 2020;41:1505–19.
4. Cao Z, Duan L, Yang G, Yue T, Chen Q. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Med Imaging*. 2019;19:51.
5. Chen S, Zhong X, Hu S, Dorn S, Kachelrieß M, Lell M, Maier A. Automatic multi-organ segmentation in dual energy CT (DECT) with dedicated 3D fully convolutional DECT networks. *Med Phys*. 2020;47:552–62.
6. Ariji Y, Fukuda M, Kise Y, Nozawa M, Yanashita Y, Fujita H, et al. Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using deep learning system of artificial intelligence. *Oral Sur Oral Med Oral Pathol Oral Radiol*. 2019;127:458–63.
7. Kise Y, Ikeda H, Fujii T, Fukuda M, Ariji Y, Fujita H, et al. Preliminary study on the application of deep learning system to diagnosis of Sjogren's syndrome. *Dentomaxillofac Radiol*. 2019;48:20190019.
8. Ariji Y, Sugita Y, Nagao T, Nakayama A, Fukuda M, Kise Y, et al. CT evaluation of extranodal extension of cervical lymph node metastases in patients with oral squamous cell carcinoma using deep learning classification. *Oral Radiol*. 2020;36:148–55.
9. Ariji Y, Yanashita Y, Kutsuna S, Muramatsu C, Fukuda M, Kise Y, Nozawa M, Kuwada C, Fujita H, Katsumata A, Ariji E. Automatic detection and classification of radiolucent lesions in the

- mandible on panoramic radiographs using a deep learning object detection technique. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2019;128:424–30.
10. Fukuda M, Inamoto K, Shibata N, Arijji Y, Yanashita Y, Kutsuna S, Nakata K, Katsumata A, Fujita H, Arijji E. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. *Oral Radiol.* 2020;36:337–343. <https://doi.org/10.1007/s11282-019-00409-x>.
 11. Murata M, Arijji Y, Ohashi Y, Kawai T, Fukuda M, Funakoshi T, Kise Y, Nozawa M, Katsumata A, Fujita H, Arijji E. Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. *Oral Radiol.* 2019;35:301–7.
 12. Hiraiwa T, Arijji Y, Fukuda M, Kise Y, Nakata K, Katsumata A, Fujita H, Arijji E. A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. *Dentomaxillofac Radiol.* 2019;48:20180218.
 13. Ohashi Y, Arijji Y, Katsumata A, Fujita H, Nakayama M, Fukuda M, Nozawa M, Arijji E. Utilization of computer-aided detection system in diagnosing unilateral maxillary sinusitis on panoramic radiographs. *Dentomaxillofac Radiol.* 2016;45:20150419.
 14. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going Deeper with Convolutions. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015; 1–9. <https://arxiv.org/abs/1409.4842>. Accessed 20 Dec 2015.
 15. Lee JH, Han SS, Kim YH, Lee C, Kim I. Application of a fully deep convolutional neural network to the automation of tooth segmentation on panoramic radiographs. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2020;129:635–42.
 16. Ghazi P, Hernandez AM, Abbey C, Yang K, Boone JM. Shading artifact correction in breast CT using an interleaved deep learning segmentation and maximum-likelihood polynomial fitting approach. *Med Phys.* 2019;46:3414–30.
 17. Chen H, Zhang K, Lyu P, Li H, Zhang L, Wu J, Lee CH. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci Rep.* 2019;9:3840.
 18. Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, Sveshnikov MM, Bednenko GB. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac Radiol.* 2019;48:20180051.
 19. Kats L, Vered M, Zlotogorski-Hurvitz A, Harpaz I. Atherosclerotic carotid plaque on panoramic radiographs: neural network detection. *Int J Comput Dent.* 2019;22:163–9.
 20. Yu J, Sharpe SM, Schumann AW, Boyd NS. Detection of broad-leaf weeds growing in turfgrass with convolutional neural networks. *Pest Manag Sci.* 2019;75:2211–8.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.