Deep learning systems for detecting and classifying the presence of impacted supernumerary teeth in the maxillary incisor region on panoramic radiographs

Chiaki Kuwada, DDS,[a] Yoshiko Ariji, DDS, PhD,[b] Motoki Fukuda, DDS,[c] Yoshitaka Kise, DDS, PhD,[c] Hiroshi Fujita, PhD,[d] Akitoshi Katsumata, DDS PhD,[e] and Eiichiro Ariji, DDS, PhD[f]

Objective. This investigation aimed to verify and compare the performance of 3 deep learning systems for classifying maxillary impacted supernumerary teeth (ISTs) in patients with fully erupted incisors.

Study Design. In total, the study included 550 panoramic radiographs obtained from 275 patients with at least 1 IST and 275 patients without ISTs in the maxillary incisor region. Three learning models were created by using AlexNet, VGG-16, and DetectNet. Four hundred images were randomly selected as training data, and 100 images were assigned as validating and testing data. The remaining 50 images were used as new testing data. The sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve were calculated. Detection performance was evaluated by using recall, precision, and F-measure.

Results. DetectNet generally produced the highest values of diagnostic efficacy. VGG-16 yielded significantly lower values compared with DetectNet and AlexNet. Assessment of the detection performance of DetectNet showed that recall, precision, and F-measure for detection in the incisor region were all 1.0, indicating perfect detection.

Conclusions. DetectNet and AlexNet appear to have potential use in classifying the presence of ISTs in the maxillary incisor region on panoramic radiographs. Additionally, DetectNet would be suitable for automatic detection of this abnormality. (Oral Surg Oral Med Oral Pathol Oral Radiol 2020;130:464−469)

Panoramic radiography, which enables the visualization of wide areas of the maxillofacial region, is used frequently in dental clinics. Various conditions and lesions are incidentally detected in panoramic radiographs; however, these may be difficult to interpret because of the complexity of the relationships between the anatomic structures and the panoramic image layer, and critical diseases may sometimes be overlooked.[1] Computer-assisted detection or diagnosis systems have been developed to support inexperienced observers,[2] such as new graduates, as well as busy general practitioners. Recently, systems based on artificial intelligence (AI), which use deep learning (DL) algorithms with convolutional neural networks (CNNs), have rapidly replaced computer-assisted detection or diagnosis systems created by traditional methodology that uses characteristic features abstracted by human observers.[3] DL systems are based on neural networks that mimic the mechanism of human neurons. Using neural networks with multiple layers, characteristic features included in the data can be learned automatically step by step. Among various DL functions, the object detection and classification functions are frequently applied to panoramic radiography. The object detection function enables detection of the positions of specified objects in an image. The classification function can automatically divide various objects into certain fixed categories. These functions have been reported to perform well in the detection and classification of diseases and conditions such as caries,[4] maxillary sinusitis,[5] unusual root morphology,[6] suspected osteoporosis,[7] vertical root fracture,[8] and radiolucent lesions.[9]

Images in the maxillary incisor region are occasionally difficult to interpret because of the overlapping of the cervical spine and the narrow panoramic image layer.[10] Impacted supernumerary teeth (ISTs) are frequently observed in the maxillary incisor region, where they are known as *mesiodens*. Maxillary ISTs cause many complications in the mixed dentition, such as delayed eruption and root resorption of the permanent incisors.[11-16] ISTs are often observed incidentally,

Statement of Clinical Relevance

Deep learning systems could help successfully detect and classify impacted supernumerary teeth in the maxilla in patients with fully erupted maxillary permanent incisors on panoramic radiographs. The results of this study indicated the potential of deep learning systems to provide diagnostic support in the interpretation of panoramic radiographs.

[a]Part-time Lecturer, Department of Oral and Maxillofacial Radiology, Aichi-Gakuin University School of Dentistry, Nagoya, Japan.
[b]Associate Professor, Department of Oral and Maxillofacial Radiology, Aichi-Gakuin University School of Dentistry, Nagoya, Japan.
[c]Assistant Professor, Department of Oral and Maxillofacial Radiology, Aichi-Gakuin University School of Dentistry, Nagoya, Japan.
[d]Professor, Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University, Gifu, Japan.
[e]Professor, Department of Oral Radiology, Asahi University School of Dentistry, Mizuho, Japan.
[f]Professor, Department of Oral and Maxillofacial Radiology, Aichi-Gakuin University School of Dentistry, Nagoya, Japan.
Received for publication Jan 14, 2020; returned for revision Mar 12, 2020; accepted for publication Apr 25, 2020.

2212-4403/$-see front matter
https://doi.org/10.1016/j.oooo.2020.04.813

even after the age of complete eruption of permanent incisors. In such instances, an IST may disturb orthodontic tooth movement[17] or become a focus of infection and cyst formation.[11,12,18,19]

Despite the importance of identifying maxillary ISTs on panoramic radiographs, no studies have addressed the use of DL systems for this task. The aim of the present study was to verify and compare the performances of 3 DL systems for detecting and classifying anterior maxillary ISTs arising in patients with fully erupted incisors.

MATERIALS AND METHODS

This study was approved by the ethics review board of our university (No. 496) and was performed in accordance with the tenets of the Declaration of Helsinki.

Patients

Panoramic radiographs of 275 patients (111 females and 164 males; average age 39.2 years) with at least 1 IST in the maxillary incisor region were selected from images stored in our hospital image database retroactively from August 2019 (Table I). Of the 275 patients, 235 had 1 IST, 39 had 2 ISTs, and 1 had 3 ISTs. As controls, 275 panoramic radiographs of age- and gender-matched patients without ISTs were selected from the same database. The inclusion criteria were the presence of more than 3 erupted permanent incisors and complete or almost complete root formation. Patients with permanent incisors showing root resorption by the ISTs or with restored crowns were included. Patients with impacted permanent incisors, congenital anomalies and syndromes, or cysts or tumors in this region were excluded. Patients who were in the mixed dentition stage and those who had undergone surgical intervention in this region were also excluded. All images were reviewed by 2 radiologists (Y.A. and E.A.), each with greater than 30 years of experience in interpreting panoramic radiographs. They came to a consensus on the presence or absence of at least 1 IST in the maxillary incisor region.

The panoramic radiographs were exposed by using a Veraview Epocs unit (J. Morita Mfg Corp., Kyoto, Japan) or an AUTO III NTR unit (Asahi Roentgen Industry, Kyoto, Japan). In total, 250 radiographs with ISTs and 250 without ISTs were randomly selected to create the learning models. Of these 500 radiographs, 400 were randomly assigned as training data, and 100 were used as testing data (Testing Data 1). These were also used as validating data in the learning process. The remaining 50 radiographs were used as newly assigned testing data (Testing Data 2) to verify the performance of the created learning models.

Table I. Numbers of radiographs in the data set

	Data set			
	Training Data	*Testing Data 1**	*Testing Data 2*	*Total*
With IST	200	50	25	275
Without IST	200	50	25	275
Total	400	100	50	550

*These data were also used as validating data in the learning process.
IST, impacted supernumerary tooth.

DL architecture

The learning models were created by using 3 CNNs. Two classification models (Models A and B) were created by using AlexNet and VGG-16 CNNs, which had only the object classification function in the system used in the present study. They were used for classifying manually cropped images into 2 groups: those with ISTs and those without ISTs. The third model (Model C) was developed by using DetectNet, which had both object detection and classification functions. The model detected the incisor regions on the radiographs and then classified them into 2 groups: those with ISTs and those without ISTs.

The DL system for classification was created on Ubuntu OS version 16.04.2 with an 11 GB graphics processor unit (NVIDIA GeForce GTX 1080 Ti; NVIDIA, Santa Clara, CA). AlexNet and VGG-16 CNN on the DIGITS library version 5.0 (NVIDIA, Santa Clara, CA; https://developer.ndivia.com/digits) were used with the Café framework. AlexNet comprised 5 convolutional layers and 3 fully connected layers. The VGG-16 CNN system consisted of 13 convolutional and 5 pooling layers and was fine-tuned by using a pretrained CNN. The system for DetectNet provided on the DIGITS library version 5.0 was also implemented on Ubuntu OS with GPS of GeForce GIX. The adaptive moment estimation (Adam) solver was used with 0.0001 as the base learning rate. DetectNet consisted of 5 main parts, including a data ingest and augmentation part, a fully convolutional network, a loss function measurement part, a bounding box clustering part, and a mean average precision calculation part.[20]

Development and evaluation of Models A and B

Image patches were manually cropped from the downloaded panoramic images with an arbitrarily sized rectangular region of interest (ROI) set by a radiologist (C.K.) (Figure 1). The ROIs included the bilateral contact points between the canines and the lateral incisors in the mesiodistal direction, and the inferior line of the nasal cavity and the incisal edges of these teeth in the superoinferior direction. When the canines and/or the lateral incisors were missing, the contact points were assumed by the radiologist for the ROI setting.

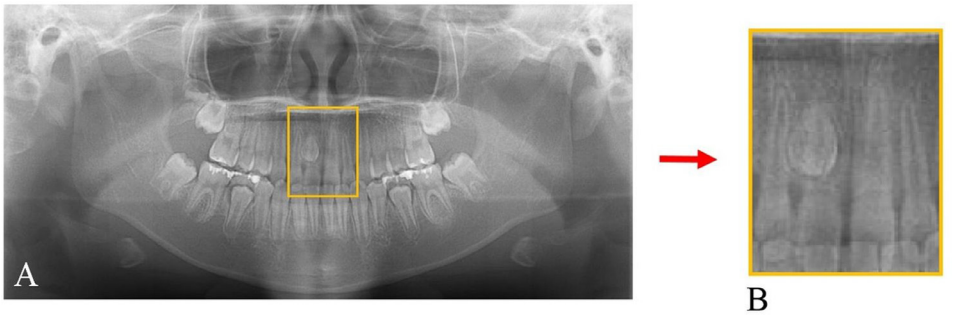

Fig. 1. A region of interest (ROI) is outlined by the yellow rectangle in the maxillary incisor region **(A)**. The ROI includes the bilateral contact points between the canine and the lateral incisor in the mesiodistal direction, and the inferior line of the nasal cavity and the incisal edges in the superoinferior direction. Cropped image patch **(B)**.

Consequently, 400 patches (200 with IST and 200 without IST) were prepared for training, and 100 patches (50 with IST and 50 without IST) were prepared for validating. The training data were augmented to 9000 patches by using IrfanView software (Irfan Škiljan, Austria; https://www.irfanview.com/). The learning process was performed in 200 epochs both for the AlexNet and VGG-16 CNN systems. After these processes, two learning models (Models A and B) were developed and tested by using Testing Data 1 and 2. Diagnostic performance was analyzed for sensitivity, specificity, and accuracy. Based on sensitivity and specificity, the area under the receiver operating characteristic curve (AUC) was determined for Testing Data 1 and 2.

Development and evaluation of Model C

DetectNet had both object detection and classification functions. We used the same panoramic radiographs as those used for creating Models A and B. The panoramic radiographs were downloaded in JPEG format with 900 × 900 pixels. Initially, the incisor regions were labeled by setting arbitrarily sized rectangular ROIs to contain the maxillary incisors, using the same areas as in Models A and B (Figure 2). The coordinates of the upper left (x1, y1) and lower right (x2, y2) corners of the ROIs were determined by using ImageJ software (National Institutes of Health, Bethesda, MD) and were converted to text form (see Figure 2). Next, the regions were assigned to 2 classes according to the presence or absence of ISTs, indicated as class 0 for the regions with ISTs and 1 for the regions without ISTs. These 400 images and labeled data were imported into the training process with 100 testing (validating) data, and 1000 epochs of training were performed. This model was also tested by Testing Data 2. A resultant image of the testing process was provided for each panoramic radiograph with a rectangular box (Figure 3) when the model detected the maxillary incisor region. Simultaneously, the box was colored red when the model assigned it as class 0 (with ISTs). When assigned to class 1, the box was colored blue. The detection performance was evaluated by using the statistical values of recall, precision, and F-measure as follows:

- Recall = number of correctly detected regions / number of all regions

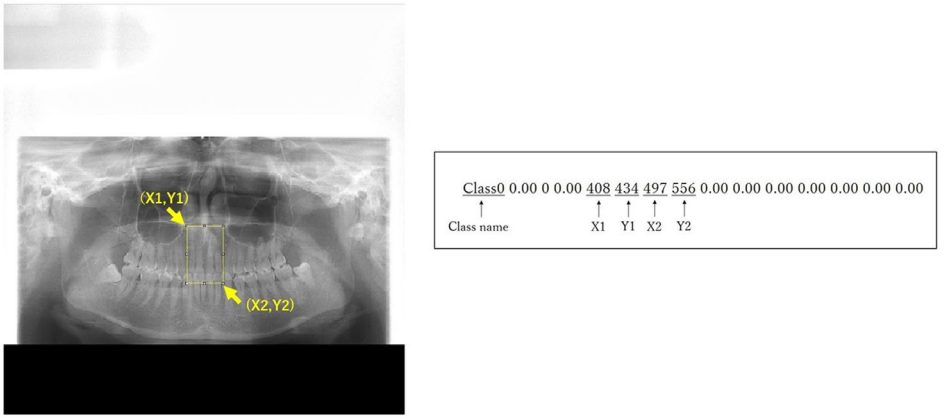

Fig. 2. The label is created in the maxillary incisor region setting a region of interest (ROI) by using the same method shown in Figure 1. The coordinates of the upper left (x1, y1) and lower right (x2, y2) corners of the ROIs were determined and were converted to text form.

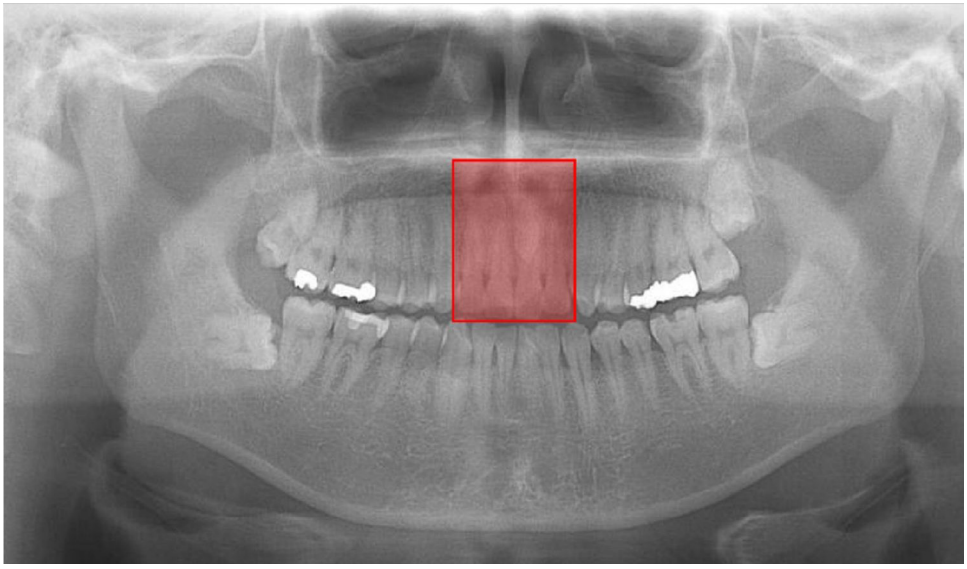

Fig. 3. An example of a successfully detected and classified panoramic radiograph. The maxillary incisor region is correctly detected. The detected region is shown as a red rectangle, indicating that the model correctly assigns it as class 0 with an impacted supernumerary tooth.

- Precision = number of correctly detected regions / (number of correctly detected regions + number of falsely-detected regions)
- F-measure = 2 (recall × precision) / (recall + precision)

For the correctly detected incisor regions, the classification performance for the positive evaluation of an IST was assessed for sensitivity, specificity, accuracy, and AUC for Testing Data 1 and 2.

Statistical analysis

The differences in the AUCs were statistically assessed by using the χ^2 test, with significance established at $P < .05$.

RESULTS

The sensitivity, specificity, accuracy, and AUC of Models A through C are shown in Table II. For the detection performance of Model C, the recall, precision, and F-measure for detecting the incisor region were all 1.0, indicating perfect detection for both Testing Data 1 and 2. The highest accuracy was provided by Model C, followed by model A, for both Testing Data 1 and 2. Model B yielded relatively low values in all performance measures (Figure 4). Significant differences were found in the AUCs between Models A and C ($P = 0.010$) and between Models B and C ($P < 0.001$) for Testing Data 1, and between Models A and B ($P < 0.001$) and between Models B and C ($P < 0.001$) for Testing Data 2.

DISCUSSION

There has been a significant increase in the number of studies on DL systems with various functions, including in the field of oral and maxillofacial radiology. Among the various DL functions, the classification function is frequently applied to panoramic radiography and is reported to perform well in the evaluation of various diseases and conditions.[5,6] Hiraiwa et al.[6] evaluated 2 DL systems with

Table II. Diagnostic performance of deep learning systems for classifying the incisal regions with IST versus those without IST

Created model (CNN used)	*Testing data used*	*Sensitivity*	*Specificity*	*Accuracy*	*AUC*
Model A (AlexNet)	Testing Data 1	0.74	0.86	0.80	0.80
	Testing Data 2	0.84	0.96	0.90	0.90
Model B (VGG-16)	Testing Data 1	0.74	0.70	0.72	0.73
	Testing Data 2	0.44	0.60	0.52	0.52
Model C (DetectNet)	Testing Data 1	0.90	0.96	0.93	0.93
	Testing Data 2	0.92	1.00	0.96	0.96

AUC, area under the receiver operating characteristic curve; *CNN,* convolutional neural network; *IST,* impacted supernumerary teeth.

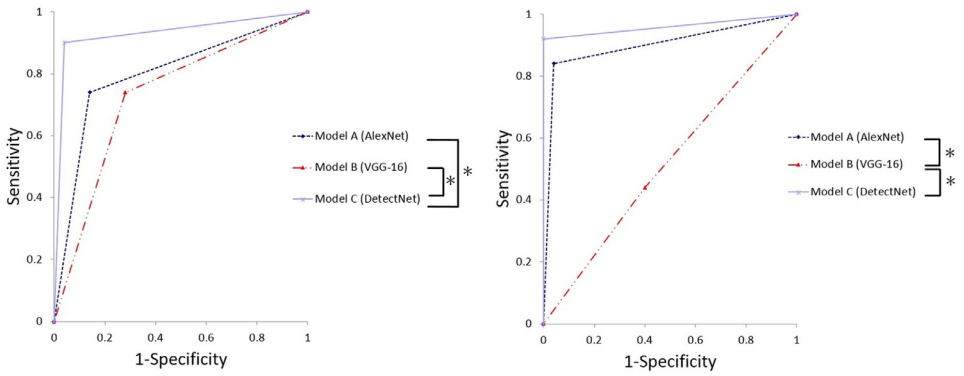

Fig. 4. Receiver operating characteristic curves of models for Testing Data 1 **(A)** and 2 **(B)**.* statistically significant difference with $P < .05$.

a classification function and reported significantly better performance than experienced radiologists in specificity and AUC for identifying supernumerary distal roots in the mandibular first molar. For distinguishing inflamed maxillary sinuses from healthy sinuses, AlexNet was reported to provide an accuracy of 0.875 and an AUC of 0.875, representing equal and superior performance to experienced and inexperienced radiologists, respectively.[5] As described in these reports, image cropping was often performed manually, with appropriate areas set by human observers. In some studies, the cropping operations were performed semiautomatically by using commercially available software for the process of making learning models. However, this would have to be accomplished in actual use in clinics by human observers if appropriate procedures had not been developed for automatic cropping.

DetectNet, which has functions for both object detection and classification, might help resolve this problem. As indicated by the results of the present study, it could provide a highly accurate classification of ISTs, together with detection of the maxillary incisor region, by just inputting appropriate panoramic images into the model created. Although DetectNet produced the highest classification performance among the 3 CNNs evaluated, the reason could not be clarified because the CNN did not disclose the features on which it focused. However, this CNN was considered to have the potential for completely automatic evaluation. DetectNet was applied to the detection and classification of radiolucent lesions in the mandible on panoramic images, and high detection sensitivity and a low false-positive rate were reported.[9] Teeth with vertical root fractures were also detected with high sensitivity (recall) when sufficient data could be provided for the training process in the evaluation of mandibular molars.[8] As indicated by these reports, the performance of DL systems generally depends on the amount of data available for training. We also considered another option to create a learning model for detecting and classifying the presence of ISTs. We initially created a learning model with a training process by using only images of the incisor region with ISTs. However, the acquired learning model erroneously detected almost all the regions without ISTs and showed very low specificity when it was applied to the testing data, including the regions without ISTs. Therefore, to replace this model, we created Model C in the present study; this model detected the region regardless of the presence or absence of ISTs and classified them into 2 categories.

Although the performance of the learning models should be verified with a substantial number of data sets, the CNNs built with AlexNet and DetectNet appeared to have potentially impressive performance in actual use. In contrast, the VGG-16 CNN, which was tuned and available as a pretrained network on the DIGITS library version 5.0, exhibited significantly poorer performance. The use of many color images other than radiologic images in pretraining may explain the poor performance of this CNN. Relatively simple CNNs might be more suitable for gray-scale images, such as panoramic radiographs.

The present investigation had some limitations. First, it did not include patients with mixed dentition. In such a dentition, the presence of an IST would be more problematic because the eruption of permanent teeth would frequently be disturbed.[11-16] Detection of ISTs in the mixed dentition would be more difficult because of the presence of unerupted permanent teeth. However, AI technology may lead to the development of an application to solve this challenge. Second, the amount of training and testing data was so small that the final evaluations of AI application could not be determined. The solution would be to collect panoramic radiographs in multi-institutional studies with many different kinds of panoramic machines. Third, the AUCs were not determined on the basis of minutely drawn ROC curves. The ROC analysis could be described in more detail for the learning models created by AlexNet and VGG-16 CNNs because the models could provide the confidence scores of the classification

corresponding to the true-positive fraction and the false-positive fraction for each image patch. However, DetectNet could only determine whether the classes were 0 or 1. Therefore, to compare 3 models, only sensitivity and specificity were used for calculating the AUC.

CONCLUSIONS

DetectNet and AlexNet appear to have a potential use in classifying the presence of ISTs in the maxillary incisor region on panoramic radiographs. Additionally, DetectNet may be suitable for the automatic detection as well as classification of this abnormality.

ACKNOWLEDGMENTS

We thank Helen Jeays, BDSc, AE, from the Edanz Group (https://en-author-services.edanzgroup.com/) for editing a draft of this manuscript.

REFERENCES

1. Yoshida K, Fukuda M, Gotoh K, Ariji E. Depression of the maxillary sinus anterior wall and its influence on panoramic radiography appearance. *Dentomaxillofac Radiol*. 2017;46:20170126.
2. Yilmaz E, Kayikcioglu T, Kayipmaz S. Computer-aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography. *Comput Methods Programs Biomed*. 2017;146:91-100.
3. Ohashi Y, Ariji Y, Katsumata A, et al. Utilization of computer-aided detection system in diagnosing unilateral maxillary sinusitis on panoramic radiographs. *Dentomaxillofac Radiol*. 2016;45:20150419.
4. Tuzoff DV, Tuzova LN, Bornstein MM, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac Radiol*. 2019;48:20180051.
5. Murata M, Ariji Y, Ohashi Y, et al. Deep-learning classification using convolutional neural network for evaluation of maxillary sinusitis on panoramic radiography. *Oral Radiol*. 2019;35:301-307.
6. Hiraiwa T, Ariji Y, Fukuda M, et al. A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. *Dentomaxillofac Radiol*. 2019;48:20180218.
7. Chu P, Bo C, Liang X. Using Octuplet Siamese Network for osteoporosis analysis on dental panoramic radiographs. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:2579-2582.
8. Fukuda M, Inamoto K, Shibata N. Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography [e-pub ahead of print]. doi: 10.1007/s11282-019-00409-x. Accessed May 27, 2020.
9. Ariji Y, Yanashita Y, Kutsuna S, et al. 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;128:424-430.
10. Anthonappa RP, King NM, Rabie AB, Mallineni SK. Reliability of panoramic radiographs for identifying supernumerary teeth in children. *Int J Paediatr Dent*. 2012;22:37-43.
11. Asaumi J, Shibata Y, Yanagi Y, et al. Radiographic examination of mesiodens and their associated complications. *Dentomaxillofac Radiol*. 2004;33:125-127.
12. Colak H, Uzgur R, Tan E, Hamidi MM, Turkal M, Colak T. Investigation of prevalence and characteristics of mesiodens in a non-syndromic 11256 dental outpatients. *Eur Rev Med Pharmacol Sci*. 2013;17:2684-2689.
13. Jung YH, Kim JY, Cho BH. The effects of impacted premaxillary supernumerary teeth on permanent incisors. *Imaging Sci Dent*. 2016;46:251-258.
14. Zhao J, Liu N, Shen ML, Hao XH. Clinical analysis of impacted supernumerary teeth in 115 patients. *Shanghai Kou Qiang Yi Xue*. 2019;28:426-429. [in Chinese].
15. Gregg TA, Kinirons MJ. The effect of the position and orientation of unerupted premaxillary supernumerary teeth on eruption and displacement of permanent incisors. *Int J Paediatr Dent*. 1991;1:3-7.
16. Cosme-Silva L, Costa E, Silva LL, Junqueira MA. Combined surgical removal of a supernumerary tooth and orthodontic traction of an impacted maxillary central incisor. *J Dent Child (Chic)*. 2016;83:167-172.
17. Suga SS, Kruly Pde C, Garrido TM. Radiographic follow-up during orthodontic treatment for early diagnosis of sequential supernumerary teeth. *Case Rep Dent*. 2016;2016:3067106.
18. Shah KM, Karagir A, Adaki S, Pattanshetti C. Dentigerous cyst associated with an impacted anterior maxillary supernumerary tooth. *BMJ Case Rep*. 2013;2013. Bcr2012008329.
19. Lustmann J, Bodner L. Dentigerous cysts associated with supernumerary teeth. *Int J Oral Maxillofac Surg*. 1988;17:100-102.
20. Tao A, Barker J, Sarathy S. Detect Net: deep neural network for object detection in DIGITS. 2016. Available at: https://devblogs.nvidia.com/parallelforall/detectnet-deep-neural-network-object-detection-digits/. Accessed May 27, 2020.

Reprint requests:
Chiaki Kuwada
Aichi-Gakuin University
2-11 Suemori-dori
Chikusa-ku
Nagoya 464-8651
Aichi
Japan.
chiaki@dpc.agu.ac.jp