Detection of unilateral and bilateral cleft alveolus on panoramic radiographs using a deep learning system

Chiaki Kuwada (✉ chiaki@dpc.agu.ac.jp)

Aichi-Gakuin University School of Dentistry

Yoshiko Arijji

Osaka Dental University

Yoshitaka Kise

Aichi-Gakuin University School of Dentistry

Motoki Fukuda

Aichi-Gakuin University School of Dentistry

Jun Ota

Aichi Gakuin University School of Dentistry, Dental Hospital

Hisanobu Ohara

Aichi Gakuin University School of Dentistry, Dental Hospital

Norinaga kojima

Aichi Gakuin University School of Dentistry, Dental Hospital

Eiichiro Arijji

Aichi-Gakuin University School of Dentistry

Research Article

Keywords: deep learning, cleft alveolus, bilateral CA, unilateral CA

Posted Date: October 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-948208/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Dentomaxillofacial Radiology on January 25th, 2022. See the published version at <https://doi.org/10.1259/dmfr.20210436>.

Abstract

The purpose of this study was to create an effective deep learning (DL) model for detecting cleft alveolus (CA), including bilateral CA (BCA), on panoramic radiographs by comparing the detection performance of a DL model based on unilateral CA (UCA) and normal data with a model developed by combining UCA, BCA and normal data. We created two models using DetectNet. Model A was created using only UCA and normal images, and Model B was created using UCA, BCA and normal images. The performance of Models A and B was evaluated with the same testing data, and compared with two human observers. The total detection sensitivities were 0.55, 0.85, and 0.86 for Model A, Model B, and human observers, respectively. The ratios of detected and undetected CAs were significantly different among these three evaluators ($p < 0.001$). Regarding the UCA group, no significant differences were found in the ratio of those between Models A and B ($p = 0.248$). However, in the BCA group, the ratios were significantly different between the models ($p < 0.001$). The DL model created with the data including the BCA (Model B) achieved high detection performance for the testing data of both the UCA and BCA.

Introduction

Deep learning (DL) techniques with convolutional neural networks (CNN) have often been used for the automatic detection and classification of various oral and maxillofacial diseases on panoramic radiographs, such as radiolucent lesions in the mandible [1], root fractures [2], maxillary sinus lesions [3], and impacted supernumerary teeth [4].

Cleft lip and palate is one of the most common congenital anomalies in the maxillofacial region in the Japanese population [5] and is frequently associated with unilateral or bilateral cleft alveolus (CA). CA patients are usually treated with bone graft techniques at the age of 8 to 10 years, when the maxilla has grown sufficiently for this surgery [6]. Therefore, patients are followed up regularly from birth through physical and imaging examinations. Panoramic radiography plays an essential role in evaluating the status of the CA because of its low level of radiation exposure to patients and low cost compared with computed tomography (CT) or cone-beam CT for dental use (CBCT) [7].

Although CA status can be easily recognized during a physical examination, oral and maxillofacial radiologists, who have to interpret many panoramic radiographs routinely, cannot always perform such examinations and are forced to diagnose the presence of clefts by the panoramic appearance alone. In such cases, a computer-aided diagnosis/detection system created using DL with CNN would help radiologists, especially those who are inexperienced, to avoid overlooking clefts. To date, the only DL model reported has been for automatic detection of unilateral CA (UCA) on panoramic radiographs [8]. Although this DL model appears to perform well in detecting UCA on panoramic radiographs, the suitability of its application to bilateral CA (BCA) has not been verified. If the panoramic appearance is identical between the UCA and BCA when a cleft is examined, the DL model for detecting the UCA, which was created using only the UCA data, would be effective also for detecting the BCA. If this is not the case, another model should be created based on data including BCA radiographs.

The purpose of this study was to create an effective DL model for detecting both UCA and BCA. For this purpose, we compared the detection performances between a DL model based solely on UCA and normal data and a DL model developed by combining UCA, BCA, and normal data.

Materials And Methods

Informed consent was obtained from all patients for being included in the study. This study was approved by the ethics committee of our university (no. 496) and was performed in accordance with the Declaration of Helsinki.

Patients

Panoramic radiographs of 383 patients (169 female and 214 male) with UCA and 108 patients (45 female and 63 male) with BCA were retrospectively selected from our hospital image database between August 2004 and July 2020. Among the patients with UCA and BCA, 209 (54.5%) and 90 (83.3%) patients, respectively, had a cleft palate. The mean age of CA patients was 8.5 years. All patients were examined repeatedly with panoramic radiography. The radiographs taken immediately before bony transplant surgery were selected. All patients were verified as having unilateral or bilateral CA by medical records and CT examinations. As a normal group without CA, 210 patients who matched the mean age and sex distribution of the CA patients were selected from the same database during the same period. These patients were examined for other purposes, such as pre-examination for orthodontic treatment.

The panoramic radiographs were taken using a Veraview Epocs unit (J. Morita Mfg. Corp., Kyoto, Japan), with a tube voltage of 75 kV, tube current of 8 mA, and exposure time of 16.2 s, or an AUTO III NTR unit (Asahi Roentgen Industry, Kyoto, Japan), with a tube voltage of 75 kV, tube current of 12 mA, and exposure time of 12 s.

DL architecture

The DL process was performed on Ubuntu OS version 16.04.2 with an 11 GB graphics processor unit (NVIDIA GeForce GTX 1080 Ti; NVIDIA, Santa Clara, CA, USA), and Digits version 5.0 training system (NVIDIA) with a customized DetectNet (<https://devblogs.nvidia.com/detectnet-deep-neural-network-object-detection-digits/>) with object detection and classification functions was used. The adaptive moment estimation (Adam) solver was used with 0.0001 as the base learning rate.

Development of learning models

For the learning process, panoramic image data with annotated labels were required as the training and validating data. The panoramic images were downloaded in JPEG format and cropped to 900 × 900 pixels. To create the labels, the rectangular regions of interest (ROIs) were set and the coordinates of the upper left (x1, y1) and lower right (x2, y2) corners were recorded using ImageJ software (National Institute of Health, Bethesda, MD, USA) (Figure 1a). Thereafter, they were converted to text form (Figure 1b). In the normal group, labels without coordinates were created.

The ROI of the CA area was determined based on the following definitions. The superior distal corner was set at the most distal portion of the nasal cavity lateral wall, and the inferior medial corner was set at the alveolar ridge between the central incisors. Consequently, the BCA group had two ROIs.

We created two models (Models A and B) in the present study. Model A was created using only UCA and normal group images as training and validating data, and Model B was created using UCA, BCA, and normal images as training and validating data (Table 1). In the UCA, BCA, and normal groups, 60, 30, and 30 images were randomly assigned to the test dataset, respectively. The remaining images were used to create the learning model as the training and validating data. The training and validating data were arbitrarily selected with the ratios of 80% and 20%, respectively. This learning process was repeated twice for Models A and B with randomly assigned training and validation data. Consequently, two respective models were created and evaluated for Models A and B using the same testing data to compare their performance with that of two human observers. To create each model, 1000 epochs of the learning process were performed. Applying the testing data to the learning models, a rectangular red box was shown on the testing images when the model detected a CA (Figure 2).

Comparison with the detection performance of human observers

To compare the DL performance with that of human observers, a radiologist and a dental resident evaluated the same testing data that were used for the evaluation of the DL models. They evaluated both sides of the maxillary incisor regions and determined whether CAs were present or absent.

Statistical analysis

The differences in ratios of detected and undetected CAs between the evaluators were tested by the chi-square test. The threshold of significant difference was set as $p < 0.05$.

Results

In the resulting testing images, at most two bounding boxes could be observed. All bounding boxes estimated by the models were located in areas where CAs truly existed or would arise. These areas were

similar to the areas annotated in the training and validating data. No false positive boxes could be found in areas other than these areas.

Testing results and performances of the models and observers are shown in Tables 2 and 3, respectively. The detection sensitivity, which was defined as the number of correctly detected CAs per total number of CAs (240), was 0.55 (133/240 CAs), 0.85 (204/240 CAs), and 0.86 (208/240 CAs) for Model A, Model B, and human observers, respectively. The ratio of detected and undetected CAs was significantly different among the three evaluators ($p < 0.001$). The detection sensitivity of Model B was higher than that of Model A and almost the same as that of the human observers.

When Models A and B were compared while limited to the UCA group, the detection sensitivities were 0.84 (101/120 CAs) and 0.90 (108/120 CAs), respectively. No significant differences were found in the ratio of detected and undetected CAs between the two models ($p = 0.248$). However, for the BCA group, the detection sensitivities of Models A and B were 0.26 (32/120 CAs) and 0.80 (96/120 CAs), respectively. The ratio was significantly different between the models ($p < 0.001$). A typical example is shown in Figure 3.

False positive results were very few for all three evaluators. Each model falsely detected two normal areas as CA in the normal group (Figure 4). However, human observers falsely detected five contralateral normal sides in the UCA group (Figure 5). This indicated that five UCA patients were misdiagnosed as BCA patients.

Discussion

The total detection sensitivity was quite low for Model A when compared with Model B and the human observers. This result may be because the BCA group was not added to the learning process in Model A. We formulated the following hypothesis for the present study: "If the panoramic appearance is identical between a UCA and a BCA when a cleft is examined, the DL model for detecting UCA, which was created using only UCA data, would also be effective for detecting BCA". The results of our study disproved our hypothesis because the detection sensitivity of the BCA group was significantly lower in Model A. This suggests that there might be a difference between the UCA and BCA findings on panoramic radiographs. The higher prevalence of associated cleft palate in the BCA group might be a possible reason for the difference in cleft appearance between the UCA and BCA groups. The CAs associated with a cleft palate may show more radiolucency than those without a cleft palate.

In a previous study [8], two DL models created with and without normal data were compared in the learning process for automatic detection of UCA on panoramic radiographs. The results verified that the model with normal data performed better with fewer false positive results. In the present study, therefore, panoramic radiographs without CAs were included as a normal group in the learning process. The low number of false positive results in the present study could be attributed to this procedure.

The false positive evaluations revealed differences between the models and human observers. The models erroneously detected areas in the normal group, while the human observers falsely detected the contralateral side of the UCA area (Figures 4 and 5). The models might learn the CA findings themselves, while the humans might take the right and left asymmetry into account when diagnosing CAs. Therefore, a UCA case with relatively symmetrical features might be misdiagnosed as a BCA by human observers.

The present study had some limitations. First, the number of cases in the BCA group was lower than that of the UCA group. The detection sensitivity of the BCA group could be improved by using more cases in future studies. Second, owing to the differences in the detection sensitivity of the BCA group between Models A and B, there may have been a difference in the panoramic findings between the UCA and BCA groups. However, we did not analyze the differences in the panoramic appearance in great detail. In future research, these differences should be investigated. Third, in the present study, we did not take the presence or absence of cleft palate into account. The presence of a cleft palate might also be a reason for differences in the panoramic appearance.

In conclusion, the DL model created with the data including the BCA group (Model B) achieved high detection performance for the testing data comprising both the UCA and BCA groups.

Declarations

Competing interests :

The authors declare no competing interests.

References

1. Arij, Y. *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*, **128** (4), 424–430 (2019).
2. Fukuda, M. *et al.* Evaluation of an artificial intelligence system for detecting vertical root fracture on panoramic radiography. *Oral Radiol*, **36** (4), 337–343 (2020).
3. Watanabe, H. *et al.* Deep learning object detection of maxillary cyst-like lesions on panoramic radiographs: preliminary study. *Oral Radiol*, **37** (3), 487–493 (2021).
4. Kuwada, C. *et al.* Deep learning systems for detecting and classifying the presence of impacted supernumerary teeth in the maxillary incisor region on panoramic radiographs. *Oral Surg Oral Med Oral Pathol Oral Radiol*, **130** (4), 464–469 (2020).
5. Omiya, T., Ito, M. & Yamazaki, Y. Disclosure of congenital cleft lip and palate to Japanese patients: reported patient experiences and relationship to self-esteem. *BMC Res Notes*, **7**, 924 (2014).
6. Fowler, P. V., Al-Ani, A. H. & Thompson, J. M. D. Bone fill following secondary alveolar bone grafting for children with cleft of the alveolus in New Zealand. *Orthod Craniofac Res*, **22** (3), 153–158 (2019).

7. Jacobs, R. *et al.* Pediatric cleft palate patients show a 3- to 5-fold increase in cumulative radiation exposure from dental radiology compared with an age- and gender-matched population: a retrospective cohort study. *Clin Oral Investig*, **22** (4), 1783–1793 (2018).

8. Kuwada, C. *et al.* Detection and classification of unilateral cleft alveolus with and without cleft palate on panoramic radiographs using a deep learning system. *Sci Rep*, **11** (1), 16044 (2021).

Tables

Tables are available as a download in the supplementary files section.

Figures

Figure 1

(a) A region of interest (ROI) is set at the cleft alveolus (CA) area and its coordinates are recorded. The superior distal corner is set at the most distal portion of the nasal cavity lateral wall, and the inferior medial corner is set at the alveolar ridge between the central incisors. (b)An example of converting the coordinates of the ROIs in a patient with a classification of unilateral CA (UCA). For the bilateral CA (BCA) group, two coordinates are assigned.

Figure 2

A unilateral cleft alveolus (UCA) in the right maxilla is correctly detected and is shown as a red rectangular box

Figure 3

Model A successfully detects the right cleft alveolus (CA) but fails to detect the left side in a bilateral CA (BCA) case (a), while Model B correctly detects both CAs (b)

Figure 4

In a normal case, Models A and B both erroneously detect the right maxillary area as having a cleft alveolus. In this case, the right lateral incisor is absent, resulting in an asymmetrical feature with more radiolucency on the right side

Figure 5

In a case with a unilateral cleft alveolus (UCA) in the left side, two human observers falsely diagnose it as a bilateral cleft alveolus (BCA). Neither Model A nor B can detect the UCA

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table.xlsx](#)