Tooth labeling in cone-beam CT using deep convolutional neural network for forensic identification

Yuma Miki[a], Chisako Muramatsu*[a], Tatsuro Hayashi[b], Xiangrong Zhou[a], Takeshi Hara[a], Akitoshi Katsumata[c], Hiroshi Fujita[a]

[a]Department of Intelligent Image Information, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan

[b]Media Co. Ltd., 3-26-6 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[c]Department of Oral Radiology, Asahi University School of Dentistry, 1851 Hozumi, Mizuho, Gifu 501-0296, Japan

ABSTRACT

In large disasters, dental record plays an important role in forensic identification. However, filing dental charts for corpses is not an easy task for general dentists. Moreover, it is laborious and time-consuming work in cases of large scale disasters. We have been investigating a tooth labeling method on dental cone-beam CT images for the purpose of automatic filing of dental charts. In our method, individual tooth in CT images are detected and classified into seven tooth types using deep convolutional neural network. We employed the fully convolutional network using AlexNet architecture for detecting each tooth and applied our previous method using regular AlexNet for classifying the detected teeth into 7 tooth types. From 52 CT volumes obtained by two imaging systems, five images each were randomly selected as test data, and the remaining 42 cases were used as training data. The result showed the tooth detection accuracy of 77.4% with the average false detection of 5.8 per image. The result indicates the potential utility of the proposed method for automatic recording of dental information.

Keywords: Tooth detection, tooth classification, dental cone-beam CT, dental forensic identification, deep convolutional neural network, fully convolutional network

1. INTRODUCTION

Dental information can be employed for forensic identification in large scale earthquakes and tsunami disasters. Forensic dentistry is useful because dental information can be used when body damages are severe, and antemortem x-ray images are more easily accessible than DNA. For dental personal identification, a text record documenting the dentition state, called a dental chart, is usually employed. In general, the dental chart is filed by dentists; however, most dentists are not experienced in recording the dental charts for corpses, and the task can be psychologically challenging. Such work may cause misfiling of records and psychological disorders. In addition, filing dental charts may be a time consuming work when the number of subjects is large.

In order to reduce the risk of mental damage and assist efficient identification, some groups have proposed computerized method to automatically match antemortem and postmortem dental x-ray images [1, 2]. In these methods, antemortem images were searched by matching the tooth contour information. Other groups have proposed computerized methods for detecting and/or numbering teeth on dental x-ray images for automatic filing of dental charts [3-7]. Nassar et al. [3] studied tooth numbering method on bitewing and periapical images. They proposed a two-step method in which the first step comprises tooth separation and classification into 4 tooth types. In the second step, the classification result was used to number teeth. Similarly, Yuniarti et al. [4] first isolated each tooth on bitewing and panoramic images, classified molars and premolars, and numbered these two types of teeth with pattern modification.

Momeni et al. [5] and Hosntalab et al. [6] proposed a three-step method for numbering teeth on CT images. The three steps include segmentation, feature extraction, and classification of teeth into four types. In their method, each tooth was

*chisa@fjt.info.gifu-u.ac.jp; phone +81-58-230-6518; fax +81-58-230-6514

Medical Imaging 2017: Computer-Aided Diagnosis, edited by Samuel G. Armato III, Nicholas A. Petrick, Proc. of SPIE Vol. 10134, 101343E · © 2017 SPIE · CCC code: 1605-7422/17/$18 · doi: 10.1117/12.2254332

precisely segmented, and wavelet-Fourier descriptors were obtained for tooth classification by a supervised classifier. On the other hand, Duy et al. [7] proposed an approach to separate teeth into 16 subregions on CT images without precise segmentation, which were then classified as regions with existing or missing tooth. This method can be considered an initial step for numbering teeth.

We have previously investigated a computerized method to classify tooth types on dental cone-beam CT using convolutional neural network (CNN) [8]. Figure 1 shows an axial slice of a cone-beam dental CT with targeted 7 tooth types. In our previous study, smallest boxes enclosing individual teeth on axial slices were extracted, and these regions of interest (ROIs) were classified into the 7 tooth types by CNN employing AlexNet [9] architecture. The ROIs were obtained manually. In this study, we investigated a method to automatically locate tooth regions using a fully convolutional network (FCN) developed for face recognition, which is called deep dense face detector (DDFD) [10].

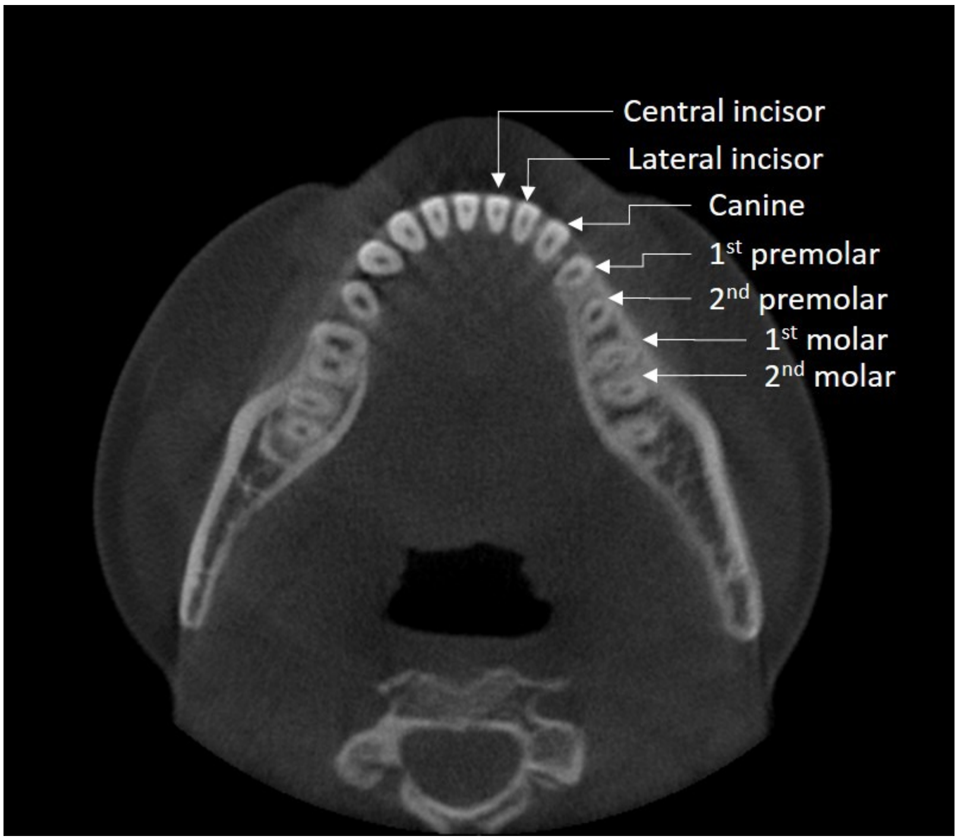

Figure 1. Dental cone-beam CT slice with seven tooth types. 3rd molar was not included for analysis because of the small number of samples.

2. METHODS

2.1 Overview

Farfade et al. proposed a face recognition method using FCN which is based on AlexNet but its inner product layers were converted to convolutional layers [3]. A heat-map responding to faces, or objects of interest, is output by the network, and bounding boxes are detected based on the heat-map. We applied this technique to the tooth detection task in this study.

First, the smallest boxes enclosing single teeth are manually extracted from axial slices as foreground training samples. For background samples, box regions were randomly obtained from extra dental regions. Using these samples, regular AlexNet was trained to classify between two groups.

After the training, inner product layers were converted into convolutional layers by reshaping the weight parameters to design FCN. By transforming the network, the heat-map of tooth probability was obtained which was then used to locate bounding boxes.

2.2 Sample extraction

From axial slices of dental cone-beam CTs, a square bounding box was located on each tooth as shown in Fig. 2. These boxes became foreground samples. Boxes were randomly located on regions excluding the tooth areas as background samples. The original AlexNet and DDFD take input images of 256 x 256 pixels, which were randomly cropped to 227 x 227 pixels to reduce overfitting before entered to the first layer. In this study, the first layer was adjusted to take images of 128 x 128 pixels, which were similarly cropped to 113 x 113 pixels considering the size of teeth. Therefore, each training sample was resized to 128 x 128 pixels. After data augmentation by shifting boxes pixel by pixel for foreground samples, the total number of training samples was 122,222 boxes.

2.3 Pre-training

A CNN was trained to classify boxes into foreground and background samples**.** The regular AlexNet was employed for this task. AlexNet comprises 5 convolutional layers, 3 pooling layers, and 3 fully connected layers. The detail of the architecture can be found elsewhere [9].

2.4 Box detection

FCN-AlexNet was constructed by converting the trained fully connected layers to convolutional layers as shown in Fig. 3. By this process, images with any size can be input to the network, and their resulted heat-maps can be output. In this study, an original axial slice of a dental CT was input to the FCN-AlexNet, which output the heat-map responding to the tooth regions. In providing the input images, slice images were magnified by 2 so that the tooth regions on the slices were approximately in 113 x 113 pixel range, which was the size of the cropped training samples. Each pixel value on the output heat-map corresponds to the probability of being a part of tooth regions. Using this result, those having above 95% confidence were considered as tooth regions, and bounding boxes were placed. For combining the overlapped boxes, non-maximum suppression module (NMS) with two strategies called NMS-max and NMS-ave [10] were employed.

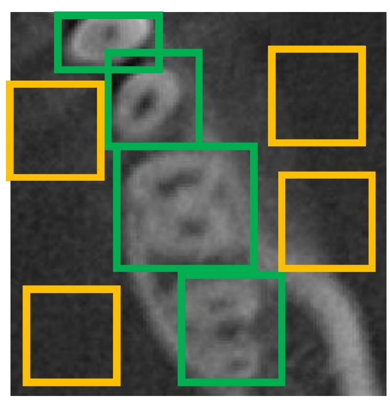

Figure 2. Training samples cropped from a CT slice. Green and orange boxes correspond to foreground and background samples, respectively.

3. IMAGE DATASET

Dental cone-beam CT volumes used in this study were obtained with two CT units: 33 cases imaged with Veraviewepocs3D (J.Morita Mfg, Corp., Kyoto, Japan) and 19 cases obtained with Alphard VEGA (Asahi Roentgen Ind. Co., Ltd., Kyoto, Japan). Spatial resolution ranged from 0.1 to 0.39 mm, and the field of view ranged from 51 to 200 mm.

Because dental cone-beam CT images do not employ Hounsfield units, pixel values are not standardized. Therefore, we selected a model case and manually set the window level and window width so that the dental region is clearly visualized. Window settings of all other cases were adjusted so that the intensity levels appear similar to that of the model image. For standardizing the size of teeth, pixel size was unified to 0.125 mm by linear interpolation/extrapolation.

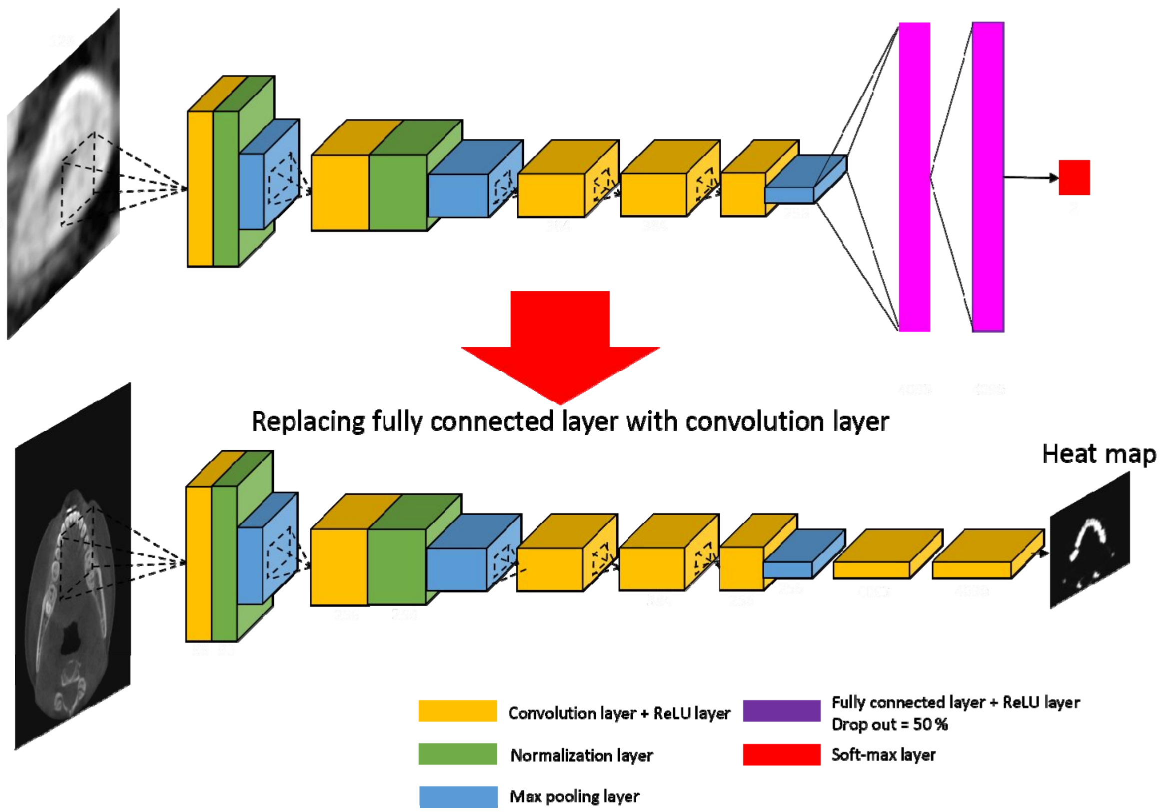

Figure 3. Network conversion from AlexNet with fully connected layers to FCN-AlexNet.

4. RESULTS

Out of 52 cases, 42 cases were employed for training and remaining 10 cases were used for testing by random sampling. As a preliminary investigation, from the 10 cases, six slices near the center of teeth that were not strongly affected by metal artifact were selected for testing the proposed network. For evaluation, a box including more than 2/3 of a tooth region and not including more than 2/3 of the neighboring teeth was considered as positive detection, and others were considered false positive detections.

The result is shown in Table 1. The average detection rate was 77.4% with 5.8 false positives per image. When true positive detection boxes were applied to the tooth-type classification network trained in our previous study [8], the average classification accuracy was 77.1%. The classification accuracy was lower than in the previous study (91.0%). This result may be due to the impreciseness of the detected bounding boxes and a small number of test samples.

Figure 4 shows an example of detection result. In some cases, a tooth was separately detected by two boxes, and neighboring teeth were detected together in one box. The possible reason may be that small teeth such as incisors and large teeth such as molars were trained together in one network. In addition, interdental spaces are very narrow in frontal teeth, and it could be difficult to detect adjacent teeth separately. Such misdetection could be improved by including a preprocessing method for separating nearby teeth.

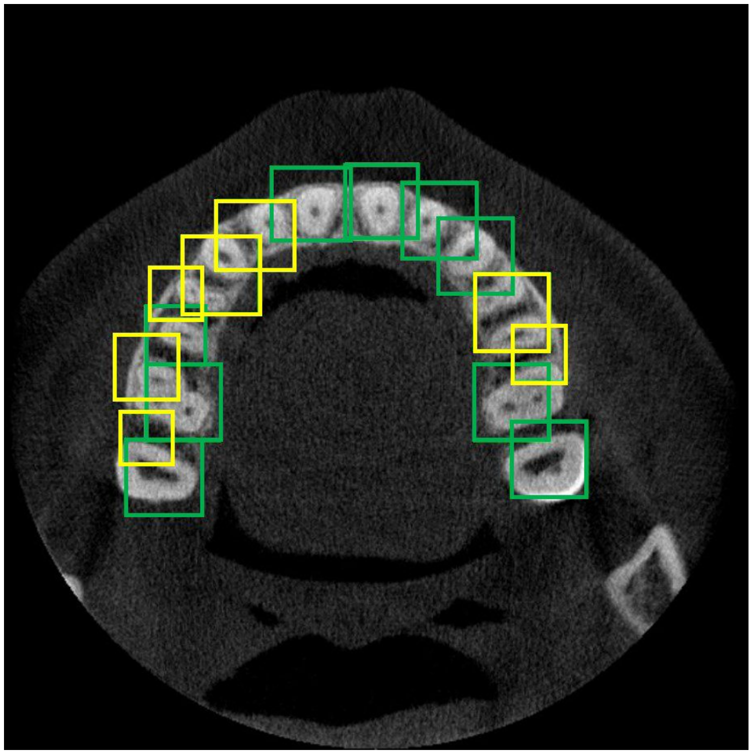

Figure 4. The result of tooth detection. Green and yellow boxes correspond to true and false detections, respectively.

Table 1. Tooth detection result for each test slice. The average detection rate was 77.4% with 5.8 false positive detection boxes per image. (TP: true positives, FP: false positives)

	Number of Teeth	Number of Detected box	TP	FP	Detection rate[%]
Case 1	16	16	14	2	87.5
Case 2	16	18	10	8	62.5
Case 3	14	18	10	8	71.4
Case 4	14	16	9	7	64.3
Case 5	12	17	12	5	100
Case 6	14	16	11	5	78.6

5. CONCLUSION

We investigated an application of CNN with a full convolution architecture for detection of tooth regions on dental cone-beam CT slices. As the preliminary study with a small number of test samples, the average detection rate was 77.4% with 5.8 false positives per image. Using the detected boxes as input data for tooth classification CNN, the classification accuracy was 77.1%. These results indicate a potential usefulness of the proposed method for tooth localization and classification on dental cone-beam CT images. Optimization of network architecture and inclusion of preprocessing and post-processing techniques will be investigated in the future work.

ACKNOWLEDGEMENTS

This study was partly supported by a Grant-in-Aid for Scientific Research (B) (No. 26293402) by Japan Society for the Promotion of Science and a Grant-in-Aid for Scientific Research on Innovative Areas (Multidisciplinary Computational Anatomy, No. 26108005) by Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES

[1] Jain, A. K. and Chen, H., "Matching of dental X-ray images for human identification," Pat. Recogn. 37, 1295-1305 (2015).

[2] Zhou, J. and Abdel-Mttaleb, M., "A content based system for human identification based on bitewing dental X-ray images," Pat. Recogn. 38, 2132-2142 (2005).

[3] Nassar, D. E., Abaza, A., Li, X., and Ammar, H., "Automatic construction of dental charts for postmortem identification," IEEE Transactions on Information Forensics and Security 3, 234-246 (2008).

[4] Yuniarti, A., Nugroho, A. S., Amaliah, B., and Arifin, A. Z., "Classification and numbering of dental radiographs for an automated human identification system," TELKOMNIKA 10, 137-146 (2012).

[5] Momeni, M. and Zoroofi, R. A., "Automated dental recognition by wavelet descriptors in CT multi-slices data," Int J CARS 3, 533-542 (2008).

[6] Hosntalab, M., Zoroofi, R. A., Tehrani-Fard, A. A., and Shirani, G., "Classification and numbering of teeth in multi-slice CT images using wavelet-Fourier descriptor," Int J CARS 5, 237-249 (2010).

[7] Duy, N. T., Lamecker, H., Kainmueller, D., and Zachow, S., "Automatic detection and classification of teeth in CT data," MICCAI 7510, 609-616 (2012).

[8] Miki, Y., Muramatsu, C., Hayashi, T., Zhou, X., Hara, T., Katsumata, A., and Fujita, H., "Classification of teeth in cone-beam CT using deep convolutional neural network," Computer in Biology and Medicine 80, 24-29 (2017).

[9] Krizhevsky, A., Sutskever, I., and Hinton, G. E., "ImageNet classification with deep convolutional neural network," In: Advances in Neural Information processing Systems 25, 1106-1114 (2012).

[10] Farfade, S. S., Saberian, M., and Li, L. J., "Multi-view face detection using deep convolutional neural networks," In: Proc. ACM on Int Conf Multimedia Retrieval 2015, 643-650 (2015).