Computers in Biology and Medicine 80 (2017) 24–29

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Classification of teeth in cone-beam CT using deep convolutional neural network

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, Gifu 501-1194, Japan*
[b] *Media Co., Ltd., 3-26-6 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[c] *Department of Oral Radiology, School of Dentistry, Asahi University, 1851 Hozumi, Mizuho, Gifu 501-0296, Japan*

ARTICLE INFO

Keywords:
Deep convolutional neural networks
Tooth classification
Dental cone-beam CT
Forensic identification
Dental chart

ABSTRACT

Dental records play an important role in forensic identification. To this end, postmortem dental findings and teeth conditions are recorded in a dental chart and compared with those of antemortem records. However, most dentists are inexperienced at recording the dental chart for corpses, and it is a physically and mentally laborious task, especially in large scale disasters. Our goal is to automate the dental filing process by using dental x-ray images. In this study, we investigated the application of a deep convolutional neural network (DCNN) for classifying tooth types on dental cone-beam computed tomography (CT) images. Regions of interest (ROIs) including single teeth were extracted from CT slices. Fifty two CT volumes were randomly divided into 42 training and 10 test cases, and the ROIs obtained from the training cases were used for training the DCNN. For examining the sampling effect, random sampling was performed 3 times, and training and testing were repeated. We used the AlexNet network architecture provided in the Caffe framework, which consists of 5 convolution layers, 3 pooling layers, and 2 full connection layers. For reducing the overtraining effect, we augmented the data by image rotation and intensity transformation. The test ROIs were classified into 7 tooth types by the trained network. The average classification accuracy using the augmented training data by image rotation and intensity transformation was 88.8%. Compared with the result without data augmentation, data augmentation resulted in an approximately 5% improvement in classification accuracy. This indicates that the further improvement can be expected by expanding the CT dataset. Unlike the conventional methods, the proposed method is advantageous in obtaining high classification accuracy without the need for precise tooth segmentation. The proposed tooth classification method can be useful in automatic filing of dental charts for forensic identification.

1. Introduction

Dental records play an important role in forensic identification after large-scale disasters [1–3]. Forensic dentistry is an important field because dental information can be used for identifying a person even when his/her body has been severely damaged; moreover, antemortem (AM) x-ray images are easier to collect than DNA samples. For dental identification, postmortem (PM) dental findings and teeth conditions are recorded in a dental chart. However, most dentists are inexperienced at recording the dental chart for corpses, and it can cause a psychiatric burden. Such a psychiatric stress may also lead to incorrect data recording and psychiatric disorders.

For overcoming these drawbacks, studies have proposed automatically obtaining dental information from the dental x-ray images, creating panoramic-like images from CT data for better image comparison, and/or matching the AM and PM images [4–10]. Jain et al. [4] investigated a computerized method for matching the AM and PM dental images. Each tooth was first isolated from its neighbors and the tooth contour was extracted on the basis of intensity. The corresponding image was then searched for by matching the extracted contours with rigid transformation. Among 38 AM/PM image pairs, 25 were correctly matched while the genuine AM image was selected as the second-best match in 5 of the remaining 13 cases. Zhou et al. [5] proposed a 3-step method to retrieve matched images. Images were first classified into bitewing, periapical, or panoramic images, and the teeth on bitewing images were segmented using a top-hat filter and an active contour method. The corresponding image was searched for by matching the boundary shape.

* Corresponding author.
E-mail address: chisa@fjt.info.gifu-u.ac.jp (C. Muramatsu).

http://dx.doi.org/10.1016/j.compbiomed.2016.11.003
Received 7 September 2016; Received in revised form 7 November 2016; Accepted 9 November 2016
0010-4825/

Lin et al. [8] proposed a method to classify teeth on bitewing images. After tooth segmentation, the length and width ratio and crown size were used as features to classify each tooth as a molar or premolar by using a support vector machine. They achieved an overall classification accuracy of 95% when using 47 images containing 369 teeth. Hosntalab et al. [10] proposed a multi-stage technique for the classification and numbering of teeth on multi-slice CT images. The 3 step process included the segmentation of tooth regions using a variation level set, feature extraction of the wavelet-Fourier descriptor, and classification of teeth into 4 groups using a supervised classifier. Using this technique, they achieved high classification accuracies above 94% for 804 teeth from 30 CT cases [10].

In this study, as a component of automated dental chart filing system, we investigated an automated method for classifying tooth types on dental cone-beam CT images using a deep convolutional neural network (DCNN). Since the success of Krizhevsky et al. [11] in the ImageNet 2012 competition, DCNNs have shown their outstanding ability in object recognition and natural image classification. The application of DCNNs to medical images has been increasingly investigated by many groups that have achieved certain degrees of success [12–17]. However, a successful application procedure has not yet been established, and to our knowledge DCNNs have only been applied to dental image processing in one study. Wang et al. [18] reported the comparison of dental radiography analysis algorithms for the grand challenges held in IEEE international Symposium on Biomedical Imaging 2015, in which Ronneberger et al. employed u-shaped deep convolution neural network for segmentation of bitewing radiographs for caries detection. In this preliminary study, we investigated the utility of a DCNN in classifying teeth into 7 types by using rectangular regions of interest (ROIs), each of which enclosed a tooth from an axial slice.

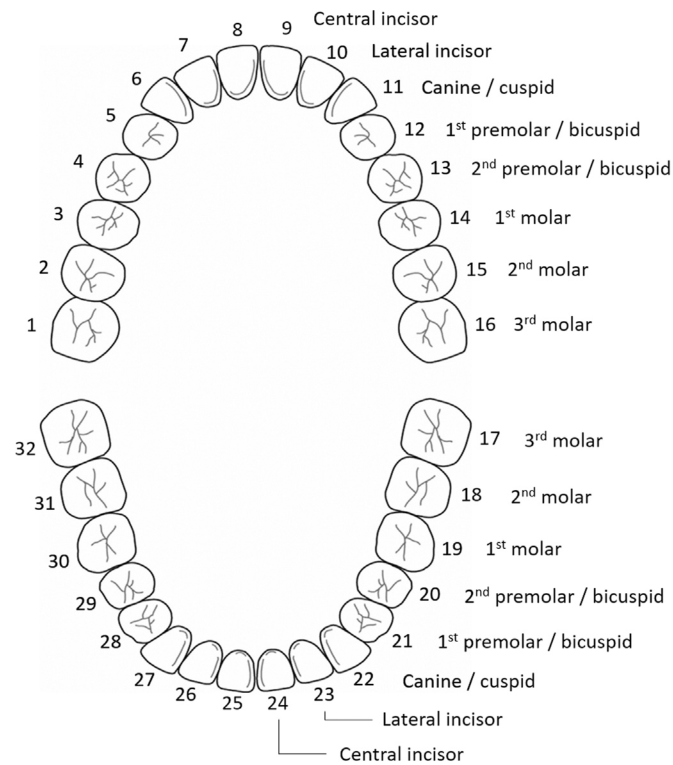

Fig. 1. Dental chart and tooth types.

2. Material and methods

2.1. Image dataset

The images used in this study were obtained using two dental CT units, namely Veraviewepocs 3D (J.Morita Mfg, Corp., Kyoto, Japan) and Alphard VEGA (Asahi Roentgen Ind. Co., Ltd., Kyoto, Japan), which were used to acquire images in 33 and 19 cases, respectively. The images were obtained from Asahi University Hospital, Gifu, Japan. The diameter of the field of view ranged from 51 to 200 mm, and the voxel resolution ranged from 0.1 to 0.39 mm. The institutional review boards of Gifu University and Asahi University approved the study protocol.

In general, medical CT images employ Hounsfield units for representing the gray levels. However, in dental cone beam CT images, gray levels are not standardized. Therefore, in this study, the window level and window width were manually adjusted to have an appearance similar to the model image, in which the dental region was clearly visualized. The average window level was 701 ± 461 and the average window width was 1338 ± 973 for 12-bit images. The gray level was then reduced to 8 bits prior to DCNN training and testing. From the 52 cases, 5 cases each from the two imaging systems were randomly selected and used as the evaluation dataset, and the remaining were used as the training dataset. For examining a sampling effect, the test dataset was sampled 3 times, and the training and testing were repeated.

2.2. ROI extraction

For both the training and test cases, the smallest bounding box enclosing each tooth was manually placed on the CT volume. From the bounding box, all axial slices excluding the upper and lower 20% were used as the training and test ROIs. As this was a preliminary investigation, the ROIs affected by metal artifacts were not included in this study. The number of ROIs obtained from a single tooth ranged from 19 to 171, with an average of 45. The 7 tooth types included the central incisors, lateral incisors, canines, first and second premolars, and first and second molars. The third molars were excluded from this study because of the small number of samples. Fig. 1 shows an example of the dental chart with different tooth types. The total number of ROIs extracted from the CT volumes was 6653, 6766. 7928, 5794, 3346, 2115, and 2657, respectively, for each of the abovementioned 7 tooth types. For the test cases, all ROIs were used for evaluation. The number of test ROIs for the 7 tooth types in each of the 3 samplings is listed in Table 1. For training the DCNN, the number of samples was balanced to the minimum number of ROIs in the 7 tooth types by random sampling. The number of the training ROIs in the 3 samplings was 11354, 12572, and 12985. Fig. 2 shows the extracted sample ROIs.

2.3. Data augmentation

A small number of training cases can often leads to overtraining. Because the number of training cases in our dataset was limited, the training data were augmented by image manipulation [11,19], and the results with and without data augmentation were compared. One method that we investigated in this study was image rotation. Another method was intensity transformation by gamma correction defined as follow:

$$y=I_{max}\bullet\left(\frac{x}{I_{max}}\right)^{1/\gamma},$$

where x and y are the input and output pixel values, respectively, and Imax is the maximum pixel value, which is 255 for the input image. For the image rotation, the ROIs were extracted by rotating the CT volume from −10 to +10° in 5° steps along the x-y plane and replacing the bounding box. This resulted in a sample size 5 times the original sample size. For intensity transformation, γ values of 0.75 and 1.5 were applied, thereby increasing the number of samples by a factor of 3. The sample images are shown in Fig. 3.