Tooth and Pulp Chamber Automatic Segmentation with Artificial Intelligence Network and Morphometry Method in Cone-beam CT

INTRODUCTION

Cone-beam computed tomography (CBCT) is widely used in the field of stomatology. It can record abundant information on the maxilla-facial region. With the coordinate information and gray value information, many unexpected details of the disease can be effectively detected. CBCT has been used to assess bone density, bone resorption, and bone augmentation in three dimensions. In dental implants, there are many methods for studying bone repair and reconstruction (Hasan et al., 2015). Furthermore, the periodontal state and shape of the tooth root have also attracted increasing attention (Kapila & Nervina, 2015). The application of 3D segmentation and reconstruction technology has gradually been applied to rebuild the anatomical structures of teeth and surrounding tissues.

In orthodontics and periodontics, there are applicable methods to explore the development of bone resorption and recession. Al-Zahrani et al. (2017) and Zhang et al. (2020) used CBCT data to analyze the progression on of periodontitis disease. For example, the patient’s alveolar bone shrinks owing to disease progression, and may cause multiple teeth to move or loosen. Therefore, CBCT can provide information about bone tissue to help doctors detect signs of periodontitis. If the teeth can be segmented accurately, it can be refined to analyze the changes in periodontal disease, such as periodontal ligament and bone resorption of a certain tooth.

In orthodontic treatment, doctors generally use medical image processing software, such as MIMICS (Materialize, Belgium) and AMIRA (Thermo Fisher Scientific, France) to segment teeth. According to the segmentation results, the doctor can create some corresponding clinical programs. The software can realize automatic threshold segmentation based on gray density values, but cannot realize complete tissue

SUMMARY: This study aims to extract teeth and alveolar bone structures in CBCT images automatically, which is a key step in CBCT image analysis in the field of stomatology. In this study, semantic segmentation was used for automatic segmentation. Five marked classes of CBCT images were input for U-net neural network training. Tooth hard tissue (including enamel, dentin, and cementum), dental pulp cavity, cortical bone, cancellous bone, and other tissues were marked manually in each class. The output data were from different regions of interest. The network configuration and training parameters were optimized and adjusted according to the prediction effect. This method can be used to segment teeth and peripheral bone structures using CBCT. The time of the automatic segmentation process for each CBCT was less than 13 min. The Dice of the evaluation reference image was 98 %. The U-net model combined with the watershed method can effectively segment the teeth, pulp cavity, and cortical bone in CBCT images. It can provide morphological information for clinical treatment.

KEY WORDS: Convolutional neural network; Teeth segmentation; Cone-beam computer tomography; Morphology.
automatic segmentation. Deep learning, a branch of artificial intelligence, has many advantages for analyzing data with speed and precision.

There are three types of tooth segmentation methods: threshold method, active contour model method, and neural network method (Heo et al., 2021). The threshold method is based on different gray values between the tooth and the periodontal ligament. The main problem of segmentation is to obtain the optimal threshold; however, the optimal threshold for different samples is different. Therefore, for different samples, the optimal threshold should be re-explored. Heo & Chae (2004) proposed an optimal threshold scheme for segmentation of teeth. However, the tissue near the tooth root is complicated, and it is difficult to distinguish the alveolar bone from the tooth root with a single threshold range. Akhoondali et al. (2009) proposed a fast and automatic segmentation method using a region-growing method. Because the gray values of the cancellous bone and cortical bone near the root are similar, the teeth cannot be segmented automatically. The active contour method is an interactive segmentation method based on contour reconstruction. The contours of the tooth are marked by the shortest diagonal method or level set algorithm, and the 3D tooth model is reconstructed along the contour of the tooth (Barone et al., 2016; Gan et al., 2018; Fan et al., 2019; Amorim et al., 2020). The research on teeth segmentation combining morphology features revealed that the actual measured Root Mean Square (RMS) value is 0.39 mm, less than 0.4 mm (Yang & Wang, 2020).

Lee et al. (2020) used a histogram-based method as a preprocessing step to compute the average gray density level of the bone and tooth regions. Simultaneously, they developed a posterior probability function (PPF) with CNN models to improve the segmentation performance. Experimental results showed that the proposed method is better than the existing methods (Lee et al.). In addition, Lee et al. considered the impact of metal artifacts during segmentation by adjusting the posture distribution of the teeth, marking all the photos of two CBCT samples and five photos of other CBCTs, and then proceeding to the teeth one by one. The results showed that the convolutional neural networks used can reduce the inter-overlapping area between boxes (Chung et al., 2020).

Because the Hounsfield unit (HU) values of different CBCT images are inconsistent, different segmentation correction methods are used during segmentation (Pauwels et al., 2015). In this study, we used the U-net model to segment CBCT data for dental hard tissue, pulp cavity, cortical bone, and cancellous bone simultaneously. To relieve the overfitting problem as training samples, generative adversarial networks can be applied to medical images (Chen et al., 2021).

MATERIAL AND METHOD

Data acquisition and preprocessing. In this study, we used data from the Peking University School of Stomatology, which were scanned with a CBCT machine (DCT Pro; Vatech, Co., Ltd, South Korea). Voxel size was 250×250×250 mm. The data were saved in the Digital Imaging and Communications in Medicine (DICOM) 3.0 format. CBCT manufacturers and software providers present gray scales as the HU (Razi et al., 2014). With normalization, the HU value of CBCT data was set to 0-1. Tensorflow (Abadi et al., 2016) and Keras (Chollet, 2018) libraries were integrated in the dragonfly software (version 4.3, Objects Research Systems, Montreal, QC, Canada) (Reznikov et al., 2020). A computer with an Intel® Xeon W-2145 E5 CPU, an NVIDIA Quadro P4000 (8GB) graphics card, and 64GB RAM was used.

Normalization. To reduce the amount of calculated data, preprocessing should be performed. First, the data were normalized. In CBCT, the images are presented with HU (Pauwels et al.). HU is a dimensionless unit universally used in computed tomography (CT) scanning to express CT numbers obtained from a linear transformation of the measured attenuation coefficients (Hounsfield, 1980). The results are based on arbitrarily assigned densities of air and pure water. It is in a scale running from -1000 HU for air and over 3000 for metals (Glide-Hurst et al., 2013). In this research, normalization was performed by a linear function:

\[ y = mx + b \]

where x refers to the HU obtained from the CBCT machine, with a minimum value (-1000) and a maximum value (3000-20000), which was dependent on different machine and scan parameters and radiation dose; y is the set HU value in the range of 0-1. In these data, the average HU value was 0.16. The value of the standard deviation was 0.15. The probability distribution map of a sample’s CBCT HU is shown in Figure 1, and after normalization, the probability distribution map of HU is shown in Figure 2.

Data annotation. The goal of semantic image segmentation is to label each pixel of an image with a corresponding class of representation. The output itself is a high-resolution image (typically of the same size as the input image) in which each pixel is classified into a particular class. Every image was marked with five different colors that were used to represent five classes (teeth, pulp chamber, cortical bone, trabecular bone, and background). The images were normalized, as shown by the left image in Figure 3, and the manually marked result is presented by the right image in Figure 3. The image was divided into five classes, where the teeth were green,
pulp chamber was red, cortical bone was pink, trabecular bone was blue, and background was black. The images were annotated in the transverse plane, as shown in Figure 4. The operator can annotate different regions of interest (ROIs) using the local Otsu, Atlas, or threshold segmentation method. Five images after annotation are shown in Figure 5 in the transverse direction of the CBCT.

**Data augmentation.** Data augmentation is a method for artificially generating more training samples to increase the diversity of the training data. This can be achieved by applying affine transformations (e.g., rotation and scaling), flipping vertically and horizontally to the original labeled samples. The brightness of the image was randomly changed by providing a brightness factor because of the different HU values of different CBCT data. The brightness factor was chosen randomly in the range [-0.2, 2]. The change in brightness allows a model to generalize across images trained
on different illumination levels. The data were augmented 10 times using different factors. An example of data augmentation is shown in Figure 6.

Semantic image segmentation with U-net. The goal of this step is to label each pixel of an image with a corresponding class. The 2D U-net model consists of an encoder and a decoder part (Ronneberger et al., 2015). The training parameters in the U-net are as follows: the number of layers in the model = 4, Patch size = 64, Batch size = 32, and Loss function = Dice loss. The training is conducted with 100 epochs, however, it stops if there is no improvement in ‘value loss’ for 10 consecutive epochs. The overall network structure is illustrated in Figure 7. The set was partitioned into learning and validation subsets (80 % and 20 %, respectively). After the training, we previewed the results of applying a deep learning model to a selected dataset. Thereafter, we applied the model to the whole slices of CBCT and obtained a multi-ROI.

We removed image noise from the ROIs, smoothed the data, and created a 3D model reconstruction from the ROIs. Small disconnected noises that were less than 0.8 nm³ were excluded. Thereafter, the final ROI was converted into a ‘mesh’ data.

Individual tooth segmentation -Watershed method to segment each tooth. Marker-controlled watershed transform (WMT) was adopted to segment individual tooth (Chen et al., 2020). A different class of models, known as instance segmentation models, which distinguishes between separate objects of the same class, was adopted to segment individual teeth. We set the pulp as the marker and used the distance map of the teeth (the union of Class 1 and Class 5) as a landscape. Thereafter, watershed segmentation was performed to obtain the individual tooth.

RESULTS

This study used deep learning methods to automatically segment and extract teeth and surrounding tissues in CBCT images. The input data for neural network training were CBCT images and five marked classes, and the output data were multi-ROI. A 2D U-net neural network model was adopted. The parameters were optimized and adjusted according to the prediction effect. Neural network training, image segmentation, and reconstruction were completed using Dragonfly software. This method can segment teeth and peripheral regions in CBCT. After training, the CBCT image sequence was segmented automatically. The Dice score (DSC) of the valuation data (20 % pixels in the marked images) was 98.59 %.
TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

\[ DSC = \frac{2TP}{2TP + FP + FN} \]

By optimizing the parameters step by step and modifying the annotation data, the optimal segmentation result can be obtained. The network trained for 100 epochs took 4 h and 15 min, while for 21 epochs took 55 min and 3 s. Using the normalized CBCT data as the test set, all segmentation with the trained model took 12 min and 20 s. The segmented results are presented in Figure 8. The U-net model of Dragonfly software can effectively segment the teeth, pulp cavity, cancellous bone, and cortical bone in the CBCT image. Furthermore, it can provide morphological information for clinical treatment (Figs. 9 to 11).
DISCUSSION

The quality of tooth segmentation determines the accuracy of treatment in orthodontic and periodontitis treatment planning. In this study, we proposed an automated tooth and bone segmentation method for CBCT.

Lee et al. segmented teeth by marking all the photos of two CBCT samples and five photos of other CBCTs to train a convolutional neural network (Chung et al.). In this study, we realized semi-automatic or automatic labeling of multi-dimensional data based on five photos of CBCT samples with U-net and data augmentation methods. To evaluate the performance of the model, we used data from the Peking University School of Stomatology to show that the proposed model can segment the whole CBCT data to several ROIs in other data.

Normalization is important for image preprocessing. If the image is not preprocessed, the network will diverge during the training process, resulting in the loss function value reaching infinity. Normalization is performed by a linear change, and the resulting value is between 0 and 1. If the normalized range is different, the result will be affected. However, with data augmentation in the training process, the trained U-net model can be used in a new CBCT.

Gan et al. proposed a method to extract the connected region of the tooth and alveolar bone from CT images using a global convex level set model. However, the method can only be used in subjects whose teeth are in an open bite position. In this study, we proposed a U-net combined with a data augmentation method to achieve automatic segmentation of different tissues in a closed bite.

Lahoud et al. (2021) used the AI algorithm to segment teeth in CBCT; however, it has some limitations, such as the network can only be used in premolars, but not in molars. Furthermore, it segments one or more teeth; thus, in different tooth segmentations, corresponding errors will occur. In addition to tooth segmentation, we segmented the pulp chamber (Fig. 12) and bones, which will provide more morphometric information for oral surgery, orthodontics, or guided endodontics.

CONCLUSION

In this study, training the network by manually labeling five pictures significantly improves the efficiency of image segmentation. In the future, by expanding the amount of data, a fully automatic segmentation and identification of multiple tissues in the oral cavity can be performed. It can provide morphological guidance for orthodontics, implants, and periodontal diseases. The accuracy of the automatic segmentation can be improved, which will greatly improve the automated diagnostics in dentistry.

ACKNOWLEDGEMENTS. This work is supported by PKU-Baidu Fund(2020BD037).

REFERENCES


