



Root canal treatment planning by automatic tooth and root canal segmentation in dental CBCT with deep multi-task feature learning

Yiwei Wang^{a,1}, Wenjun Xia^{b,1}, Zhennan Yan^{c,1}, Liang Zhao^{d,1}, Xiaohe Bian^a, Chang Liu^d, Zhengnan Qi^a, Shaoting Zhang^{e,f,*}, Zisheng Tang^{a,**}

^a Department of Endodontics, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine; College of Stomatology, Shanghai Jiao Tong University; National Center for Stomatology; National Clinical Research Center for Oral Diseases; Shanghai Key Laboratory of Stomatology; Research Unit of Oral and Maxillofacial Regenerative Medicine, Chinese Academy of Medical Sciences, Shanghai 200011, China

^b Shanghai Xuhui District Dental Center, Shanghai 200031, China

^c SenseBrain Technology, Princeton, NJ 08540, USA

^d SenseTime Research, Shanghai 200233, China

^e Shanghai Artificial Intelligence Laboratory, Shanghai 200232, China

^f Centre for Perceptual and Interactive Intelligence (CPII), Hong Kong Special Administrative Region of China

ARTICLE INFO

Keywords:

CBCT
Segmentation
Tooth
Root canal
Surgical planning

ABSTRACT

Accurate and automatic segmentation of individual tooth and root canal from cone-beam computed tomography (CBCT) images is an essential but challenging step for dental surgical planning. In this paper, we propose a novel framework, which consists of two neural networks, DentalNet and PulpNet, for efficient, precise, and fully automatic tooth instance segmentation and root canal segmentation from CBCT images. We first use the proposed DentalNet to achieve tooth instance segmentation and identification. Then, the region of interest (ROI) of the affected tooth is extracted and fed into the PulpNet to obtain precise segmentation of the pulp chamber and the root canal space. These two networks are trained by multi-task feature learning and evaluated on two clinical datasets respectively and achieve superior performances to several comparing methods. In addition, we incorporate our method into an efficient clinical workflow to improve the surgical planning process. In two clinical case studies, our workflow took only 2 min instead of 6 h to obtain the 3D model of tooth and root canal effectively for the surgical planning, resulting in satisfying outcomes in difficult root canal treatments.

1. Introduction

Root canal (Ordinola-Zapata et al., 2019) treatment is one type of endodontic treatment, that is usually carried out when apical periodontitis has occurred (European Society of Endodontology, 2006). The therapy is used to repair and save a tooth that is badly decayed or infected. This treatment sequence removes the pulp inside the tooth, cleans, disinfects, and shapes the root canals, and places a filling to seal the space. It is a beneficial but dreaded medical procedure. According to statistics by the American Association of Endodontics (Rubino, 2018), more than 15 million root canal treatments were performed from 2005 to 2006. Although root canal treatment is one of the most common procedures, the success rate is not ideal for general practice (only 60%–85%) (Ng et al., 2007). Complex morphology and wide individual variations of root canals is one of the key factors that can influence

treatment outcome. Therefore, a good knowledge of the root anatomy and root canal morphology is essential to obtain a successful outcome that leaves the treated teeth resistant to damage and reinfection.

Cone-beam computed tomography (CBCT) is a commonly used examination in endodontic treatment (Patel et al., 2015; Scarfe et al., 2006, 2009, 2012). It can provide necessary information, such as the anatomical morphology of the affected tooth, the degree and extent of the lesions in the periapical tissue. It can also provide a reference for selecting proper treatment method and equipment. In contrast to periapical radiograph that only provide 2-D information, CBCT can reveal 3-D views of the interested region. The small-field CBCT with a small field of view (several teeth) and high resolution (around 0.1 mm) can provide more precise information of teeth and root canals. From CBCT images, the reconstructed 3D teeth and pulp regions have been

* Corresponding author at: Shanghai Artificial Intelligence Laboratory, Shanghai 200232, China.

** Corresponding author at: Department of Endodontics, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China. E-mail addresses: zhangshaoting@pjlab.org.cn (S. Zhang), tangzisheng163@163.com (Z. Tang).

¹ Yiwei Wang, Wenjun Xia, Zhennan Yan and Liang Zhao are co-first authors.

used in many studies of root canal morphology (Han et al., 2014; Herrero-Hernández et al., 2020; Liu et al., 2014; Neelakantan et al., 2010; Valenti-Obino et al., 2019; Vertucci, 1984), length measurements (Nguyen and Van Pham, 2020; Tchorz et al., 2019), and so on. Such precise knowledge is beneficial in diagnosis, treatment planning and follow up of patients treated for diverse oral conditions, especially for difficult cases with complex and varied root canals.

By using 3D printing technologies, a 3D model of the tooth or root canal from the CBCT image can provide intuitive information and enable preoperative planning and simulations. As an example, Kfir et al. (2013) investigated the use of 3D plastic models printed from CBCT for accurate diagnosis and treatment of a complex case of dens invaginatus. Reymus et al. (2019) utilized CBCT and stereolithographic printer to print 3D tooth replicas for endodontic education purposes. Athirasala et al. (2017) explored a novel 3D bioprinting strategy to construct living dental pulp-like tissue for potentially better root canals.

Despite the potential benefits of 3D models in clinical activities, in practice, the 3D reconstructions are mainly obtained from the manual annotation of CBCT images by experienced doctors. The annotation is usually carried out by going through hundreds of 2D cross-sectional images one by one in some software. Thus, it is time-consuming, several hours per tooth, and often subjective. In order to obtain automatic and objective segmentation of the tooth root canal from CBCT images, several studies tried threshold-based or optimization-based traditional methods on a 2D-image basis (Michetti et al., 2015; Wang et al., 2019; Zichun et al., 2020). Duan et al. (2021) investigated the tooth and pulp segmentation by a 2D U-Net (Ronneberger et al., 2015) and obtained promising results. Despite these pioneer studies, it is still a challenging and open task due to the thin, complex and variable characteristics of root canal, especially in the apical region. Since 2D segmentation methods tend to ignore the spatial correlation between cross sections and lead to discontinuous or unsmooth 3D reconstructions, the accuracy of successive quantitative measurements or surgical planning could be degraded. Therefore, it is of great interest to investigate 3D approach in this use case, which can take advantage of the 3D spatial information for a better segmentation. Besides, the evaluation metrics used in the previous works may not be good enough to assess the most challenging apical region of root canal. Some of them compared the differences of volume sizes (Machado et al., 2019; Sfeir et al., 2017), some of them evaluated the cross-sectional area and the Feret's diameter (Michetti et al., 2017), and some of them used the Dice coefficient of the whole root canal (Hatvani et al., 2018). To better analyze the accuracy around the root tip, it is worth to design new metrics.

In this paper, we propose two novel 3D neural networks for accurate detection and segmentation of tooth and root canal from CBCT images in two stages, respectively. We design different multi-task feature learning strategies in the networks to learn good representations for the segmentation tasks from limited data samples. In the first stage, we formulate the tooth instance segmentation as a clustering task by jointly optimizing spatial embeddings and clustering seed maps. In the second stage, we obtain precise reconstruction of the root canal by integrating an auxiliary regression task for apical foramen into the segmentation network. To properly evaluate the segmentation precision for the thin anatomy of root canal, we present new metrics to assess the distance errors near the apical foramen. Our experimental results indicate the robustness and accuracy of the method. We also conducted two clinical case studies to show that our method can benefit practical root canal treatment by improving the efficiency of personalized root canal treatment planning.

In summary, we review some related work in Section 2, then describe the proposed method in detail in Section 3. Section 4 introduces the experiments and results, which is followed by the discussions and conclusions in Section 5.

2. Related work

2.1. Deep learning for segmentation of tooth and pulp in CBCT

Although many conventional methods using hand-crafted features were designed for tooth segmentation over the past decade (Keyhaninejad et al., 2006; Hiew et al., 2010; Keustermans et al., 2012; Gao and Chae, 2010; Ji et al., 2014; Gan et al., 2015, 2017; Barone et al., 2016; Pei et al., 2016; Evain et al., 2017; Zhou et al., 2018), some recent studies have investigated different deep learning-based approaches for tooth detection, identification and segmentation in CBCT images and achieved promising results (Cui et al., 2019; Lee et al., 2020; Chung et al., 2020; Chen et al., 2020a; Wu et al., 2020; Jang et al., 2021; Yang et al., 2021; Cui et al., 2021). Lately, Cui et al. (2022) presented a fully automatic AI system for tooth and alveolar bone segmentation from CBCT images.

However, there is only a few studies explore the deep learning-based segmentation methods for the root canal in CBCT. Duan et al. (2021) proposed a two-phase deep learning solution for accurate tooth and pulp cavity segmentation. First, the bounding box of single tooth is extracted by using the Region Proposal Network (RPN) (Ren et al., 2015) with Feature Pyramid Network (FPN) (Lin et al., 2017a). Second, U-Net model is adopted for tooth and pulp segmentation of single-rooted tooth and multi-rooted tooth. In another work, Zheng et al. (2021) trained a deep learning model for coarse segmentation of 3D pulp chamber of first molars and utilized a level-set model to refine the coarse segmentation for volume quantification and age estimation.

The root canal space inside the radicular portion of the tooth is critical in the root canal treatment and challenging due to its long thin anatomical shape. In this study, we mainly focus on obtaining precise root canal segmentation for clinical treatment planning. Besides, we develop a fully automatic system that first detects and segments each tooth via instance segmentation and then segments the pulp and root canal in each tooth ROI.

2.2. Multi-task feature learning for problems with small data set

Because deep neural networks are very complex and the number of available training samples is relatively small in most use cases, the learned models tend to have high risk of overfitting. The overfitting risk is even higher for most medical applications due to the challenging annotations for medical data. There are multiple ways of improving model generalization, for example, collecting large-scale dataset (Simpson et al., 2019; Luo et al., 2022a; Da et al., 2022), annotation-efficient learning from unlabeled or noisy-labeled data (Wang et al., 2020, 2021b,a; Luo et al., 2022b; Gu et al., 2022), increasing data samples by data augmentation (Shorten and Khoshgoftaar, 2019), using Dropout layers (Srivastava et al., 2014), applying weight decay (Krogh and Hertz, 1991), and so on. Besides, some studies have showed that learning from multiple related tasks (Misra et al., 2016) through different types of supervision could result in better representations, for example, contrastive self supervision (Chen et al., 2020b), detection and segmentation of a specific target (He et al., 2017), regression and segmentation of object (Tan et al., 2018), and adversarial discrimination (Tan et al., 2019; Nie and Shen, 2020). In another aspect of view, multi-task learning works because tasks act as regularizations to each other.

In our work, we utilize the multi-task learning for both tooth and pulp segmentation. In the tooth segmentation stage, we adopt three branches to learn the spatial embedding, seed map, and identification simultaneously for instance segmentation and identification of teeth. In the pulp segmentation stage, we propose a combination of segmentation task and regression task to improve the delineation of the root canal.

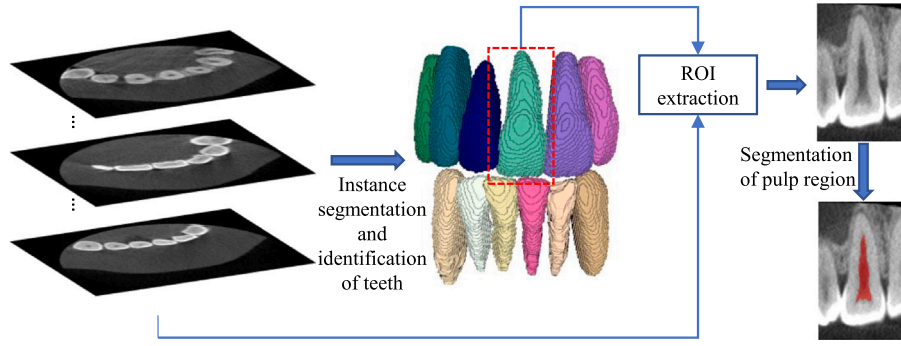


Fig. 1. The framework of the proposed method, which consists of tooth instance segmentation and pulp region segmentation for each individual tooth.

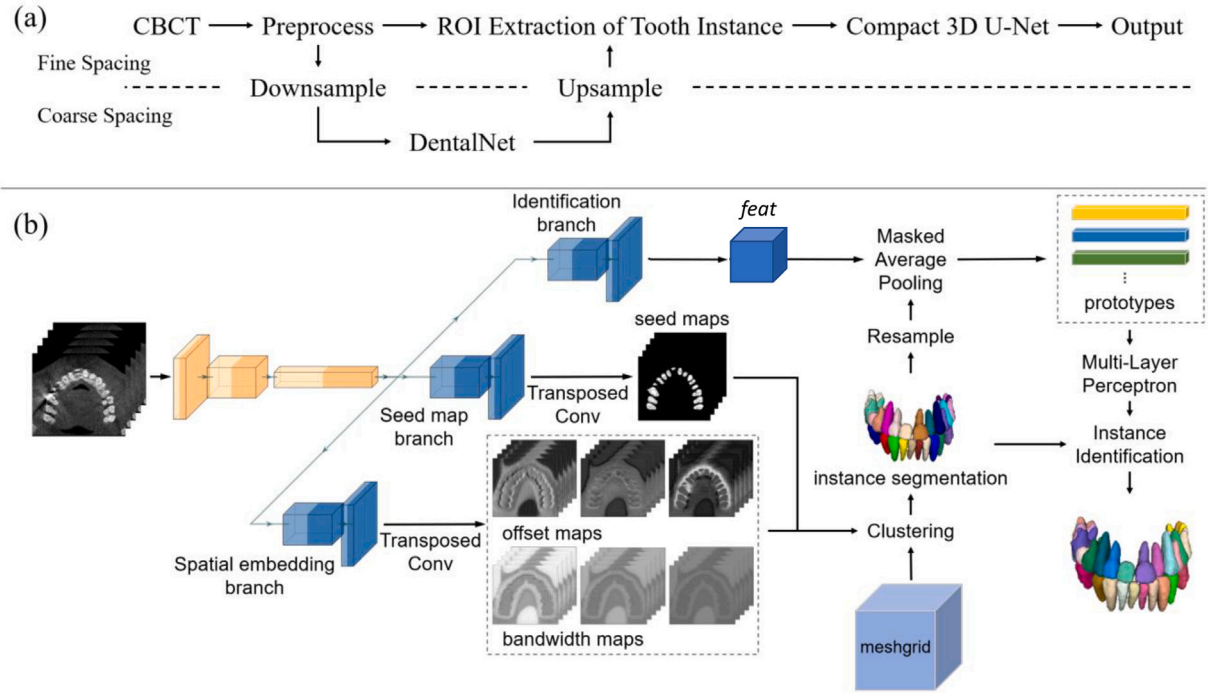


Fig. 2. The proposed framework for tooth instance segmentation and identification. (a) The overall pipeline of the proposed framework. (b) The architecture of DentalNet. The preprocessed CBCT is first put into the encoder and then decoded by three decoders. The outputs of the spatial embedding branch and seed map branch are utilized to generate instance segmentation. Prototypes of each instance are obtained by masked average pooling on the output feature map of the identification branch. The label of tooth instances is identified with the corresponding prototype.

3. Methods

Our framework consists of two major stages, instance segmentation and identification of teeth and segmentation of pulp region, as shown in Fig. 1.

3.1. Teeth segmentation

As illustrated in Fig. 2(a), we proposed a coarse-to-fine pipeline for tooth instance segmentation and identification on CBCT. We first employ DentalNet to obtain the instance segmentation and identification of teeth in the coarse spacing. With the prediction of DentalNet, ROIs of each tooth could be extracted from high-resolution CBCT. We could obtain fine-grained segmentation of each tooth by employing a compact 3D U-Net to these ROIs.

DentalNet is an Encoder-Decoder architecture illustrated in 2(b). In this work, we use 3D ERFNet (Romera et al., 2017) as the backbone. The decoder is split into three branches for instance segmentation and identification. Two branches are utilized for instance segmentation by jointly optimizing spatial embeddings and clustering bandwidth (Neven

et al., 2019). For the convenience of depiction, these two branches are named as spatial embedding branch and seed map branch respectively. A third branch is designed for tooth identification, which identifies the label of each tooth instance.

Inspired by Neven et al. (2019), the tooth instance segmentation is formulated as a clustering task. Considering an instance S_k of tooth as a set of voxels v_i . The spatial embedding branch predicts the offset vector o_i for each voxel v_i , so that $e_i = v_i + o_i$ represents the estimate of the cluster centroid that v_i belongs to. Besides, this branch also predicts the 3D bandwidth $\sigma_i = (\sigma_{ix}, \sigma_{iy}, \sigma_{iz})$. The centroid \hat{C}_k of instance S_k is estimated by $\hat{C}_k = \frac{1}{|S_k|} \sum_{v_i \in S_k} e_i$, while the instance bandwidth is calculated as $\hat{\sigma}_k = \frac{1}{|S_k|} \sum_{v_i \in S_k} \sigma_i$. The centroid and the bandwidth together can define an ellipsoid-shape instance embedding, which can represent different teeth with varied sizes. For each voxel v_i , we can compute a pseudo probability of belonging to a instance S_k by using its embedding e_i : $\phi_k(e_i) = \exp(-\frac{\|e_i - \hat{C}_k\|^2}{2\hat{\sigma}_k^2})$. Thus, we can employ binary segmentation loss to optimize the spatial embedding and clustering estimation for each instance. Let gt_k denotes the ground truth mask of S_k , we used Lovász-hinge loss (Berman et al., 2018; Yu and Blaschko,

2015) to supervise the training of this branch:

$$Loss_{emb} = \frac{1}{K} \sum_k^K Loss_{Lovász}(\phi_k, gt_k) \quad (1)$$

We added a smoothness loss (Neven et al., 2019) during the training stage to ensure $\sigma_k \approx \hat{\sigma}_k = \frac{1}{|S_k|} \sum_{v_i \in S_k} \sigma_i$. For input data contains K instances, the smoothness loss is computed as Eq. (2).

$$Loss_{smooth} = \frac{1}{K} \sum_k^K \frac{1}{|S_k|} \sum_{v_i \in S_k} \|\sigma_i - \hat{\sigma}_k\|^2 \quad (2)$$

The seed map branch predicts a probability s_i for each v_i . The s_i indicates how likely the voxel being the centroid of any instance. Suppose there are N voxels in input data and bg represents the background. The loss of the seed map branch is formulated as:

$$Loss_{seed} = \frac{1}{N} \sum_i^N \mathbb{1}_{\{v_i \in S_k\}} \|s_i - \phi_k(e_i)\|^2 + \mathbb{1}_{\{v_i \in bg\}} \|s_i - 0\|^2 \quad (3)$$

The identification branch is included in the network to classify each instance as one of the tooth classes. The output of the identification branch is a feature map denoted as *feat*. We apply a masked average pooling (MAP) operation on the *feat* $\in R^{C \times D \times H \times W}$ by using the binary mask of each instance $B_k \in R^{1 \times D \times H \times W}$ from the instance segmentation to compute the corresponding prototype $p_k \in R^C$ of the instance. Specifically, the prototype is obtained by $p_k = AveragePooling(B_k \odot feat)$, where \odot is the elementwise multiplication. Considering the scan volume ($D \times H \times W$) could be very large, we remove the last transposed convolution layer of the identification branch to save GPU memory usage. Thus, the spatial dimensions of the *feat* are half of the original sizes and the instance segmentation result is downsampled by a factor of 2 before the MAP operation. These prototypes are classified into 34 classes (background + supernumerary teeth + 32 normal teeth) by a multi-layer perceptron (MLP). The MLP and identification branch are jointly optimized by cross-entropy loss and focal loss (Lin et al., 2017b), denoted as $Loss_{id} = Loss_{ce} + Loss_{focal}$. During the training stage, the ground truth of the input is used as S . While in inference, S is the result of spatial embedding clustering.

To improve the intra-class consistency of prototypes, a homogenization loss $Loss_{homo}$ (Eq. (4)) is applied to the identification branch. Symbol definitions in Eq. (4) are explained as follows: \bar{v}_i is a voxel on the downsampled ground truth, g_i is the ground truth label at \bar{v}_i , f_i is the corresponding feature embedding on *feat* at \bar{v}_i , μ_j is the mean feature embedding with label j , and N_c is the number of labels in the ground truth.

$$Loss_{homo} = \frac{1}{N} \sum_{j=0}^N \frac{1}{|\{f_i | g_i = j\}|} \sum_{f_i \in \{f_i | g_i = j\}} (f_i - \mu_j) \quad (4)$$

$$Loss_{total} = Loss_{emb} + \alpha \cdot Loss_{seed} + \beta \cdot Loss_{smooth} + \gamma \cdot Loss_{id} + \eta \cdot Loss_{homo} \quad (5)$$

The tooth instance segmentation is summarized in Algorithm 1. Then, we can crop each tooth with a fixed size ROI ($80 \times 80 \times 128$) from the high-resolution CBCT. Only the tooth at the patch center is treated as foreground. A compact 3D U-Net is trained with these patches for binary segmentation. This network can refine the prediction of the DentalNet if the fine-grained segmentation is required.

3.2. Pulp segmentation

After tooth instance segmentation, we extract the ROI of each tooth to segment its pulp region, which consists of the pulp chamber inside the dental crown and the root canal space inside the radicular portion of the tooth. Due to varied shapes and low contrast boundaries of pulps in various teeth, and metal artifacts in CBCT images, pulp region segmentation is a challenging task, especially with a limited number of training samples. To ease the problem, we first align the tooth

Algorithm 1 tooth instance segmentation

Input: The CBCT image I with N voxels; minimum tooth volume V_{min} .
Output: Instance segmentation results $S = \{S_1, S_2, S_3, \dots, S_K\}$ and identification $L = \{l_1, l_2, \dots, l_K\}$

- 1: Feed I into tooth segmentation framework, and obtain the prediction of offset map $o = \{o_1, o_2, o_3, \dots, o_N\}$, bandwidth map $\sigma = \{\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_N\}$, seed map $s = \{s_1, s_2, s_3, \dots, s_N\}$, and feature map of the identification *feat*
- 2: calculate spatial embedding map $e_i = v_i + o_i$, $i \in \{1, 2, 3, \dots, N\}$
- 3: filter seed map: $s = \{s_i | s_i \geq 0.5\}$
- 4: **while** $|s| \geq V_{min}$ **do**
- 5: select j with the largest s_j from s
- 6: estimate cluster parameters: $C_k = e_j$ and $\sigma_k = \sigma_j$
- 7: calculate the pseudo probability map: $\phi_k(e_i) = \exp(-\frac{\|e_i - C_k\|^2}{2\sigma_k^2})$
- 8: clustering instance: $S_k = \{v_i | \phi_k(e_i) > 0.5\}$
- 9: update sets: $s = s \setminus S_k$ and $S = S \cup \{S_k\}$
- 10: **end while**
- 11: **for** $S_k \in S$ **do**
- 12: compute masked average pooling $p_k = \frac{\sum_{v_i \in S_k} feat(v_i)}{|S_k|}$
- 13: $f = MLP(p_k)$
- 14: classification of the k th instance: $l_k = \argmax_l f_l$
- 15: **end for**

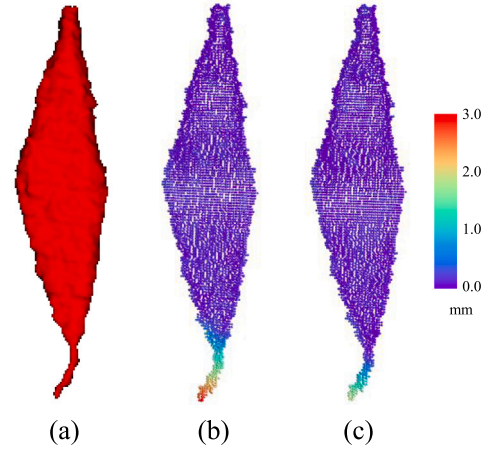


Fig. 3. The 3D error maps of comparing methods on a tooth pulp: (a) ground truth segmentation, (b) surface error map of baseline prediction, (c) surface error map of the proposed method.

ROI into a common spatial space by employing principal component analysis (PCA). The ROI alignment can alleviate the variance of tooth orientation as shown on the left of Fig. 4.

From preliminary studies, we found that the most difficult part of pulp segmentation is around the apical foramen, the small opening at the apex of the root of a tooth. This region is thin and long anatomy, which is crucial for a successful root canal procedure. As shown in Fig. 3(b), a baseline method can achieve very good segmentation except for the severe surface errors around the apical foramen.

Motivated by this observation, we designed a network demonstrated in Fig. 4. The network consists of an encoder and two decoder components. The first decoder outputs the prediction of the pulp mask. The second decoder is responsible for the detection of the landmark of apical foramen as an auxiliary task to help the segmentation branch to better capture the thin structure. The regression task is implemented by estimating an inverse distance and an offset vector to the landmark

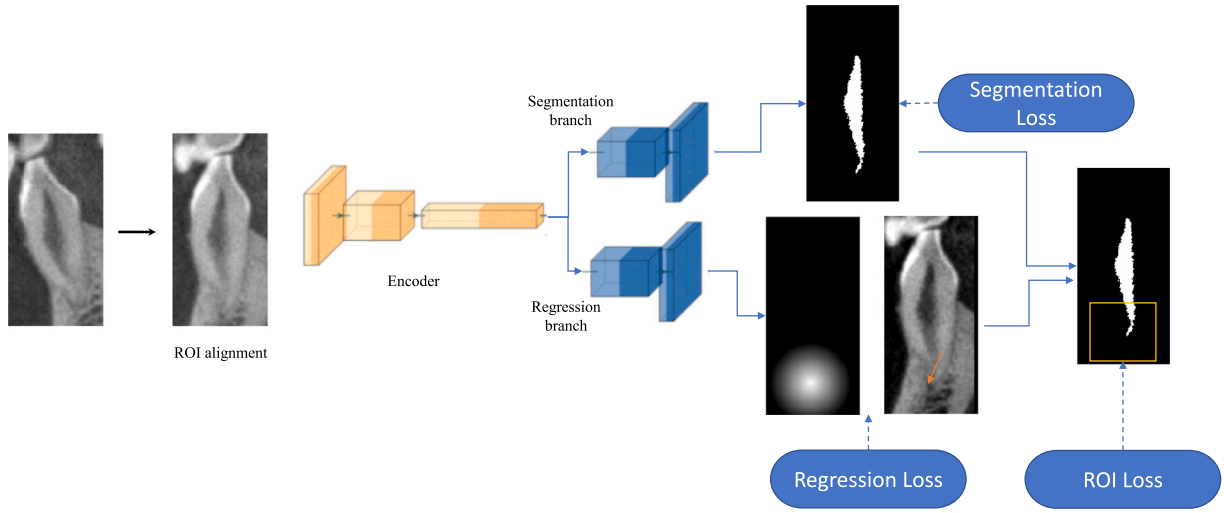


Fig. 4. The proposed segmentation network for pulp region segmentation.

from each voxel. Besides, we can further extract an ROI of the apex of a tooth root based on the regression prediction and emphasize the importance by introducing an additional penalty in the ROI.

We supervise the segmentation task by a combination of Dice loss and Focal loss (Lin et al., 2017b):

$$L_{seg}(i) = DICE(P_i^s, G_i^s) + FOCAL(P_i^s, G_i^s) \quad (6)$$

where P_i^s is the prediction of segmentation branch for voxel i , and G_i^s is the manually labeled mask, which is treated as the ground truth (GT).

The regression loss is defined by two components in Eq. (7). The first one is used to learn a heatmap of the labeled landmark. The ground-truth heatmap is formulated as the inverse distance from voxel v_i to the landmark v_{LM} : $G_i^h = \max(0, 1 - \|v_i - v_{LM}\|_2 / MaxD)$, where $MaxD$ is a pre-defined threshold controlling the responsive region around the landmark. We adopt cross-entropy (CE) loss for the heatmap prediction P_i^h . The second component is designed to learn offset vector at each voxel i : $G_i^o = (v_{LM} - v_i) / MaxD$. We use the shrinkage loss (Lu et al., 2018) to regularize the offset prediction P_i^o inside the heatmap region of the landmark with $G_i^h > 0$. The a and c are two hyper-parameters for the shrinkage loss.

$$L_{reg}(i) = CE(P_i^h, G_i^h) + \begin{cases} Shrink(|P_i^o - G_i^o|), & \text{if } G_i^h > 0 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

$$Shrink(l) = \frac{l^2}{1 + \exp(a(c - l))} \quad (8)$$

To further emphasize the regularization of the most difficult region around apical foramen, we extract a patch centered by the prediction of the landmark, denoted as $ROI(\hat{v}_{LM})$, during the feed-forward process and calculate an ROI loss as Eq. (9) in this small region to boost the performance. We use $2 \times MaxD$ as the patch size. During the training, we adopt a dynamic patch-cropping strategy. In the beginning, the ground-truth position of the landmark is used as the ROI center to stabilize training since the prediction is totally random. When the distance error of the landmark prediction is smaller than the predefined threshold $MaxD$, we use landmark prediction as the center to increase the diversity of the training samples.

$$L_{ROI} = \sum_{i \in ROI(\hat{v}_{LM})} L_{seg}(i) + L_{reg}(i) \quad (9)$$

Since the target landmark of our regression task is only a single point, the heatmap is designed to enlarge the effect of the single point and used as confidence scores to ensemble the offset predictions. The prediction of the landmark is computed as the weighted mean of the

predicted coordinates, where the weights come from the predicted heat map:

$$\hat{v}_{LM} = \frac{\sum_i \omega_i (P_i^o + v_i)}{\sum_i \omega_i}, \quad (10)$$

$$\omega_i = \begin{cases} P_i^h, & \text{if } P_i^h \geq 0.5 \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

The total loss is defined as Eq. (12), which combines the segmentation and regression tasks.

$$L_{total} = \sum_i (L_{seg}(i) + L_{reg}(i)) + \lambda L_{ROI} \quad (12)$$

4. Experiment and results

4.1. Datasets

In this study, we collected CBCT data retrospectively between 2017 and 2020 from the Endodontics Department of the Ninth People's Hospital, which is affiliated with Shanghai Jiaotong University School of Medicine with informed consent from patients. All DICOM images were anonymized before algorithm development and data analysis. The manual annotations were obtained by 3 endodontists.

For tooth instance segmentation, we collected a dataset with 201 anonymous clinical CBCT volumes for experiments. The field of view (FOV) varies from 9×15 cm to 15×15 cm with isotropic spatial resolution from 0.2 mm to 0.3 mm. Dataset is split into three folds for cross-validation. Ground truth of teeth on CBCT was delineated and annotated by an experienced dentist. The ground truth includes all the teeth and prostheses on CBCT. All the teeth were labeled following the FDI notation system (Dentistry, 2020). And prostheses were also labeled with the corresponding tooth position. For supernumerary teeth, labels from 60 were assigned. In this paper, tooth labels are visualized with the colormap shown in Fig. 5(a), and supernumerary teeth are mapped to red.

For pulp segmentation, we collected another dataset of 14 small-field CBCT volumes from 14 patients of high spatial resolution 0.1 mm and small FOV (50×50 mm) involving any maxillary and mandibular anterior teeth and premolar regions. In total, 67 teeth were extracted from CBCT images. We excluded the teeth with metal fillings, crowns, and radioactive filling materials in the root canal, which caused severe image artifacts, and teeth with a partial-visible root canal or double roots. The endodontists used the ITK-SNAP software to manually label the voxels belonging to the pulp and root canal and recorded

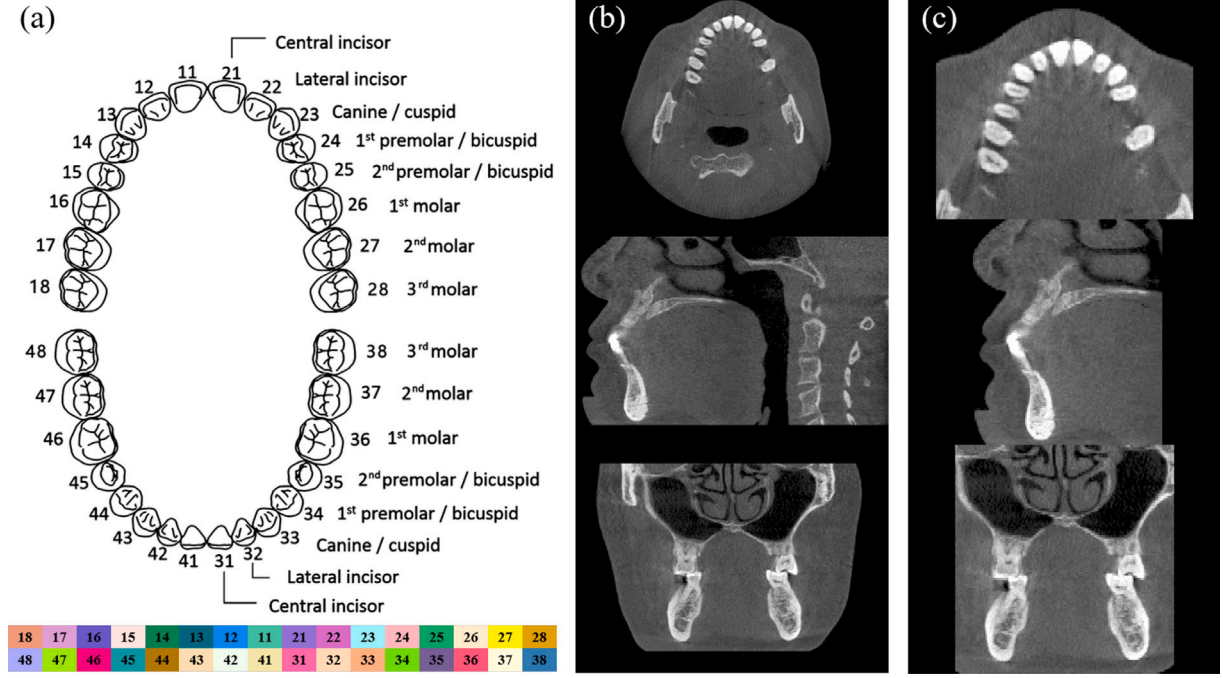


Fig. 5. (a) FDI Tooth annotation system and visualization colormap. (b) An example of fine-grained CBCT. (c) The cropped coarse spatial resolution image of (b).

the coordinates of the apical foramen for each tooth. The labeling results were then reviewed and revised by a team composed of three endodontists, two of whom have more than five years of experience, to ensure the accuracy of the manual annotations. As the result, 56 teeth were labeled. The data set is randomly divided into a training set of 5 subjects with 21 teeth, a validation set of 2 subjects with 8 teeth, and a test set of 7 subjects with 27 teeth.

4.2. Implementation details and metrics

In the data preprocessing, to alleviate the impact of diverse image quality, all the wide-field CBCT volumes for teeth segmentation are resampled to a uniform fine-grained spacing ($0.3 \times 0.3 \times 0.3$ mm). Intensities of each CBCT are then normalized to $[0, 1]$. The bulky usage of GPU memory often confines the processing of fine-grained images with 3D networks. To reduce the GPU memory usage of DentalNet, the normalized CBCT is resampled to a coarse spatial resolution ($0.6 \times 0.6 \times 0.6$ mm) and cropped with a fixed size ROI ($128 \times 128 \times 144$), which includes all the teeth.

The proposed framework is implemented in Pytorch and trained on a server with an Nvidia Tesla V100 GPU. In our experiment of tooth instance segmentation, $\alpha = 10$, while $\beta = \gamma = \eta = 1$ in Eq. (5). Adam (Kingma and Ba, 2014) is employed to optimize networks with an initial learning rate of 0.0001. During the training stage, we follow a cosine annealing schedule to adjust the learning rate.

To assess the performance of tooth segmentation, symmetric best dice (SBD) (Schar et al., 2016) and average symmetric surface distance (ASSD) were calculated. As for identification performance, we measured identification accuracy (FA) (Cui et al., 2019) and average instance dice (AID) between the ground truth and prediction. FA and AID are defined by Eq. (13), where G is the set including all the teeth in ground truth, P is the set of all the teeth in prediction, and L is the set of teeth that are correctly identified. For set X , $|X|$ indicates the number of its elements.

$$FA = \frac{|L|}{|P \cup G|} \quad \text{and} \quad AID = \frac{1}{|P \cup G|} \sum_{i \in L} Dice_i \quad (13)$$

In our experiment of pulp segmentation, $a = 10$, $c = 0.2$ in Eq. (8), and $\lambda = 0.5$ in Eq. (12). We set $MaxD = 50$ empirically in the landmark

Table 1

Comparison of quantitative results of tooth segmentation, in terms of the SBD, ASSD (mm), FA, and AID. Bold text indicates the best result for each metric. The arrow \uparrow means a larger value is better, while \downarrow means a smaller value is better.

Methods	SBD \uparrow	ASSD \downarrow	FA \uparrow	AID \uparrow
<i>baseline_T</i>	0.9331	0.1841	0.9674	0.9209
insDentalNet	0.9532	0.1271	–	–
baseDentalNet	0.9579	0.1213	0.9779	0.9403
DentalNet	0.9584	0.1180	0.9797	0.9425

detection. We use four metrics to validate the performance of the pulp segmentation, including the Dice, Hausdorff distance (HD), 99 percent HD (HD99), and minimum distance from the ground truth landmark to the predicted segmentation surface (D_{LM2S}). The euclidean distance between the landmark prediction and its ground truth, noted as D_{LM} , is used to assess the performance of the landmark detection subtask.

4.3. Performance of proposed method

4.3.1. Tooth segmentation

DentalNet framework achieved excellent performance on tooth segmentation and identification. Voxels near intercuspation surfaces and fuzzy boundaries could be separated into correct instances by this framework. In our experiments, the DentalNet framework also shows exceptional generalization to deal with some complex components, such as wisdom teeth, supernumerary teeth, and prostheses. Fig. 6 shows examples of DentalNet predictions, which include complex components. In these examples, all the teeth are delineated accurately and identified with correct labels. The DentalNet framework is an efficient method, of which the GPU usage is less than 800 MB during inference.

For quantitative evaluation, we reimplemented an existing method (Ezhov et al., 2019) for comparison. In our reimplementation (denoted as *baseline_T*), we used the same 3D ERFNet backbone as in our method and added one more class in its semantic segmentation to identify supernumerary teeth. This baseline approach outputs both instance segmentation and their identities. It is similar to using the identification branch only in our method without the separate instance segmentation generated from the seed map and spatial embedding

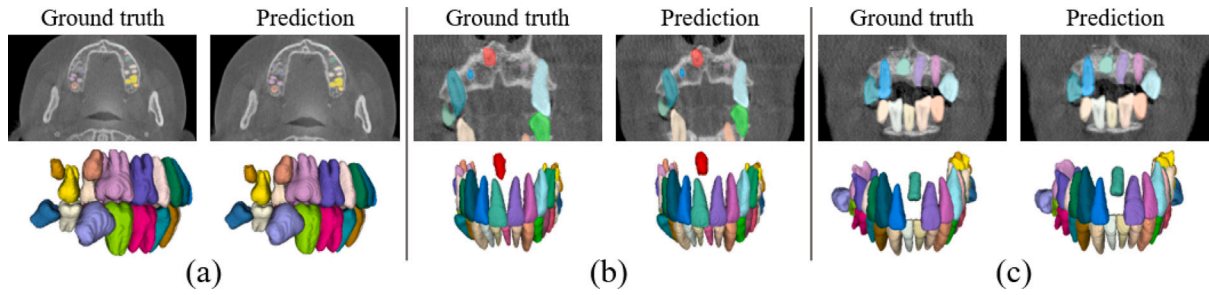


Fig. 6. Visualization of DentalNet prediction for some complex components: (a) wisdom teeth, (b) a supernumerary tooth, (c) a prosthesis.

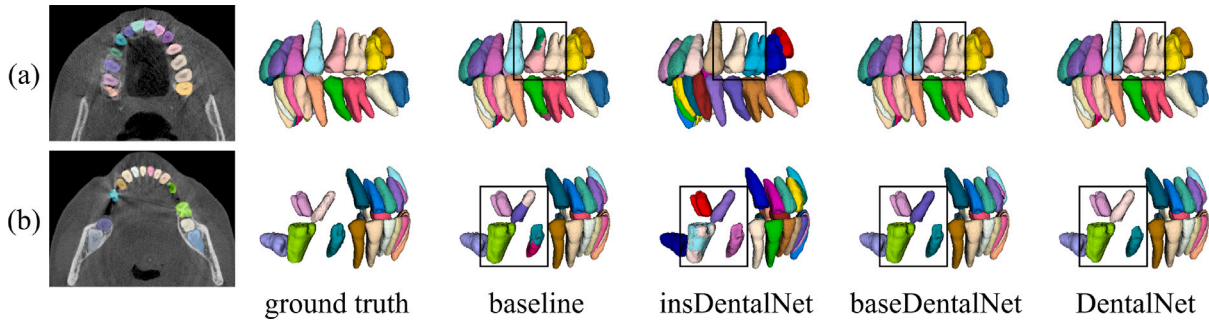


Fig. 7. Predictions of different methods. (a) DentalNet frameworks outperform the baseline method on tooth instance segmentation. (b) Improvement induced by the identification branch and the homogenization loss.

Table 2

Comparison of quantitative results of pulp segmentation (mean \pm std), in terms of the Dice, HD (mm), HD99 (mm), D_{LM2S} (mm), and D_{LM} (mm). Bold text indicates the best result for each metric. The arrow \uparrow means a larger value is better, while \downarrow means a smaller value is better.

Methods	Dice \uparrow	HD \downarrow	HD99 \downarrow	D_{LM2S} \downarrow	D_{LM} \downarrow
VNet	0.863 \pm 0.102	1.33 \pm 1.567	0.539 \pm 0.756	0.87 \pm 1.674	–
baseline _p	0.882 \pm 0.031	1.348 \pm 0.952	0.492 \pm 0.384	1.21 \pm 1.014	–
base+reg1	0.866 \pm 0.039	1.273 \pm 0.787	0.501 \pm 0.313	1.12 \pm 0.896	1.721 \pm 1.222
base+reg2	0.876 \pm 0.036	1.241 \pm 0.887	0.483 \pm 0.368	1.074 \pm 0.933	1.179 \pm 0.752
base+reg3	0.879 \pm 0.027	1.222 \pm 0.808	0.452 \pm 0.306	1.099 \pm 0.898	1.017 \pm 0.597
PulpNet*	0.878 \pm 0.027	0.982 \pm 0.518	0.372 \pm 0.194	0.779 \pm 0.609	0.980 \pm 0.495
PulpNet	0.875 \pm 0.031	0.956 \pm 0.511	0.364 \pm 0.188	0.671 \pm 0.589	0.949 \pm 0.536

branches. Thus, the baseline result can imply the contribution of the identification branch to a certain extent. Besides the baseline approach, we conducted two ablation studies to validate the contribution of different components in our framework. Since the seed map branch and spatial embedding branch are designed together to produce the instance segmentation, we trained the DentalNet without identification branch to test the instance segmentation only, noted as insDentalNet. We also trained a DentalNet without the homogenization loss, noted as baseDentalNet. The same compact U-Net was used to refine the coarse segmentation in these experiments.

We report the performance of different methods on the tooth dataset by three-fold cross-validation in Table 1. According to Table 1, insDentalNet outperforms the baseline method on SBD and ASSD. It implies that the DentalNet framework is effective in tooth instance segmentation. As listed in the second and third rows of Table 1, there is a slight improvement in segmentation performance while the identification branch is added to DentalNet. The design of the identification branch not only enables DentalNet to identify tooth instances but also benefits the segmentation performance. The benefit of the homogenization loss is demonstrated by the comparison between baseDentalNet and DentalNet. Through these comparisons, we can conclude that the instance segmentation branch, identification branch, and the homogenization loss are all necessary to achieve the best performance. Fig. 7 visualized the prediction of different methods in two cases.

We conducted another experiment to evaluate the accuracy of the DentalNet on small-FOV CBCT images as in the pulp segmentation

dataset. We randomly cropped two $256 \times 256 \times 256$ patches with 0.2 mm isotropic spacing from each large-FOV tooth scan and tested our DentalNet model. The performance downgraded to SBD = 0.744, ASSD = 0.283, FA = 0.783, AID = 0.710. However, we observed that the inferior performance was mainly due to the incomplete teeth around the boundaries of the small FOV. This is not a concern in the clinical process because the teeth of interest were intentionally acquired in the middle of the scan. After filtering out the incomplete teeth, the average metrics became SBD = 0.901, ASSD = 0.210, FA = 0.958, and AID = 0.887. Despite a slight performance drop, the DentalNet can still extract satisfying tooth ROI as the input of the PulpNet.

4.3.2. Pulp segmentation

In pulp segmentation, we use the same backbone network as in the tooth segmentation and compare it with a well-known network in medical image segmentation, VNet (Milletari et al., 2016). We also conducted ablation studies to analyze the effectiveness of the components in the PulpNet network (Fig. 4). The baseline_p is the result of the binary segmentation of pulp by the 3D ERFNet (segmentation branch only). The base+reg1 approach adds the regression branch to the baseline, but only predicts the heatmap. The landmark is predicted by the median coordinates of the highest responses in the heatmap. The base+reg2 approach is similar to base+reg1, but it predicts offset only in the regression branch. The base+reg3 predicts both the heatmap and the offset in the regression branch, but is trained without the ROI loss. The row of PulpNet* represents the results of another variant of

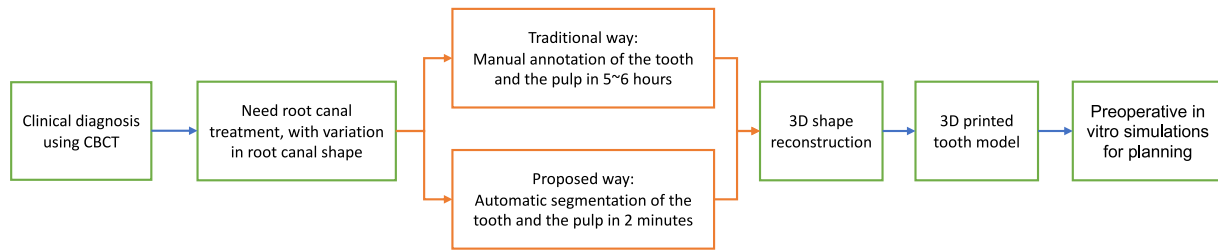


Fig. 8. Diagram of clinical workflow of root canal treatment planning. The orange blocks are alternative ways to obtain the segmentation of the tooth and the pulp.

our proposed method, in which the ROI loss in Eq. (9) is computed in a fixed patch centered by the ground truth of the landmark during the training process. Table 2 summarizes the performances of different methods, including metrics of Dice, HD, HD99, and D_{LM2S} for the segmentation subtask, and D_{LM} for the regression subtask. We use more metrics on the segmentation task because the regression subtask is introduced as an auxiliary to help the optimization of segmentation.

From Table 2, we have several observations. First, only using the segmentation branch ($baseline_p$) can achieve better performances compared with the VNet. The VNet's results have a larger standard deviation indicating its relatively unstable generalization in this use case. The results in the second row show that the segmentation branch of our method is already a good starting point in terms of the region-based metric (Dice). However, the distance-based metrics are not satisfactory, especially the D_{LM2S} . Second, introducing the regression branch (the third, fourth, and fifth rows) can help improve the segmentation results in terms of the distance-based metrics, e.g. HD and D_{LM2S} . Third, the supervision by the offset regression (the row of base+reg2) can help achieve better results with smaller distance errors than the supervision by heatmap only (the row of base+reg1). The combined supervision of offset and heatmap can further improve the distance errors for both the regression task (the metric D_{LM}) and the segmentation task (HD, HD99, and D_{LM2S}). Furthermore, the dynamic patch cropping strategy in computing the ROI loss of Eq. (9) can achieve more improvement than the fixed-patch strategy (PulpNet*) due to the increased diversity of the training samples. Last but not least, the proposed method PulpNet with all the necessary components produces the best overall results. Note that, although the $baseline_p$ approach yields the highest value in terms of Dice, the difference in the Dice score between $baseline_p$ and PulpNet is not significant ($p = 0.18$ by a paired t-test). In contrast, the proposed method is significantly better than the $baseline_p$ regarding the distance-based metrics, i.e. HD ($p = 0.007$) and D_{LM2S} ($p = 8 \times 10^{-5}$). The results show that the distance-based metrics are more sensitive to subtle shape changes than the region-based metric for the small pulp region. We can see from the distribution of errors in Fig. 3(b) that the long and thin shape of the root canal (the bottom of the pulp) is the most challenging part. Because of the critical role of an accurate shape of the root canal in the treatment planning process, we focus on improving the segmentation accuracy in this small region. However, the commonly used scale-value metric, e.g. average surface distance or HD, cannot directly reflect the accuracy in this particular area. Therefore, we propose the metric D_{LM2S} to evaluate how close the surface of root canal segmentation is to the apical foramen. We believe an ideal root canal shape can reach the apical foramen, and thus the distance D_{LM2S} should be close to 0. For example, Fig. 3(b) has a D_{LM2S} value of 2.57, while a better segmentation result in Fig. 3(c) has a D_{LM2S} value of 1.52. Table 2 shows that the proposed method PulpNet can achieve much better segmentation and regression results than all the other comparison methods in terms of the distance-based metrics. These quantitative results indicate the effectiveness of all the components in our method in obtaining an accurate shape of the root canal. A visualization of the surface error map of an example is shown in Fig. 3, which demonstrates the advantage of our method in terms of accuracy in the region close to the apical foramen.

4.4. Real case study

Since in small-field CBCT images, the density difference between the pulp cavity and dentin, cementum, and alveolar bone is relatively small, especially in the small apical part, it is often difficult to distinguish them. Accurate segmentation may be obtained by clinically experienced physicians by manually annotating the 2D images slice by slice. It is time-consuming and labor-intensive because each tooth has about 200–300 slices, and the annotation usually takes 5–6 h per root canal. The automatic segmentation method established in this study based on 3D deep learning can obtain the reconstruction of the tooth shape and internal pulp cavity in only a few minutes, which significantly improves the segmentation efficiency (Fig. 8). At the same time, compared with the 2D segmentation, the intelligent 3D segmentation method proposed in this study can obtain more continuous and smooth 3D reconstructed images, and can restore the root canal's shape more accurately in the apical region.

For clinically difficult root canal cases, we used the DentalNet and PulpNet, integrated in the SenseCare platform (Duan et al., 2020), to perform rapid automatic segmentation and 3D reconstruction of CBCT images of the affected tooth and its pulp cavity, respectively. Based on the 3D reconstruction, a high-precision 3D printer was used to obtain a 3D printed tooth model. The model not only has the same external shape as the tooth but also restores the inside shape of the pulp cavity and root canal as much as possible. Using this 3D printed tooth model, we can perform surgical simulation in vitro to pre-assess the possible risks and surgical outcomes during surgery, formulate a more scientific and personalized treatment plan and improve its success rate. We demonstrate the application of this automatic segmentation technique in actual clinical practice through the following two clinical cases.

4.4.1. Case 1: severely curved root canal

A 19-year-old young male patient presented with a 4-year protruding left upper anterior tooth and required full crown restoration after root canal treatment to improve aesthetics. The small-field CBCT (Fig. 9(a)) was taken before the operation, and it was found that the root of the left upper central incisor was curved toward the labial side, and the curvature exceeded 90 degrees. The root canal treatment was extremely difficult, mainly in the following two aspects: (1) how to choose the appropriate access opening, (2) how to choose appropriate root canal preparation instruments, and effectively prepare the root canal wall while maintaining the original root canal shape, and avoid instrument separation.

In this regard, we conduct preoperative preparation and practical surgery according to the following steps. (1) We adopted the proposed DentalNet and PulpNet to intelligently segment teeth and their root canal from CBCT images. The left upper central incisor and its pulp cavity were automatically segmented as shown in Fig. 9(b). The overall segmentation process was completed within 2 min. (2) According to the results of the 3D reconstruction of the segmented images, an in vitro 3D printed model of the upper left central incisor was obtained by a high-precision 3D printer, shown in Fig. 9(c). (3) The clinician designed eight surgical plans by using different access opening positions and

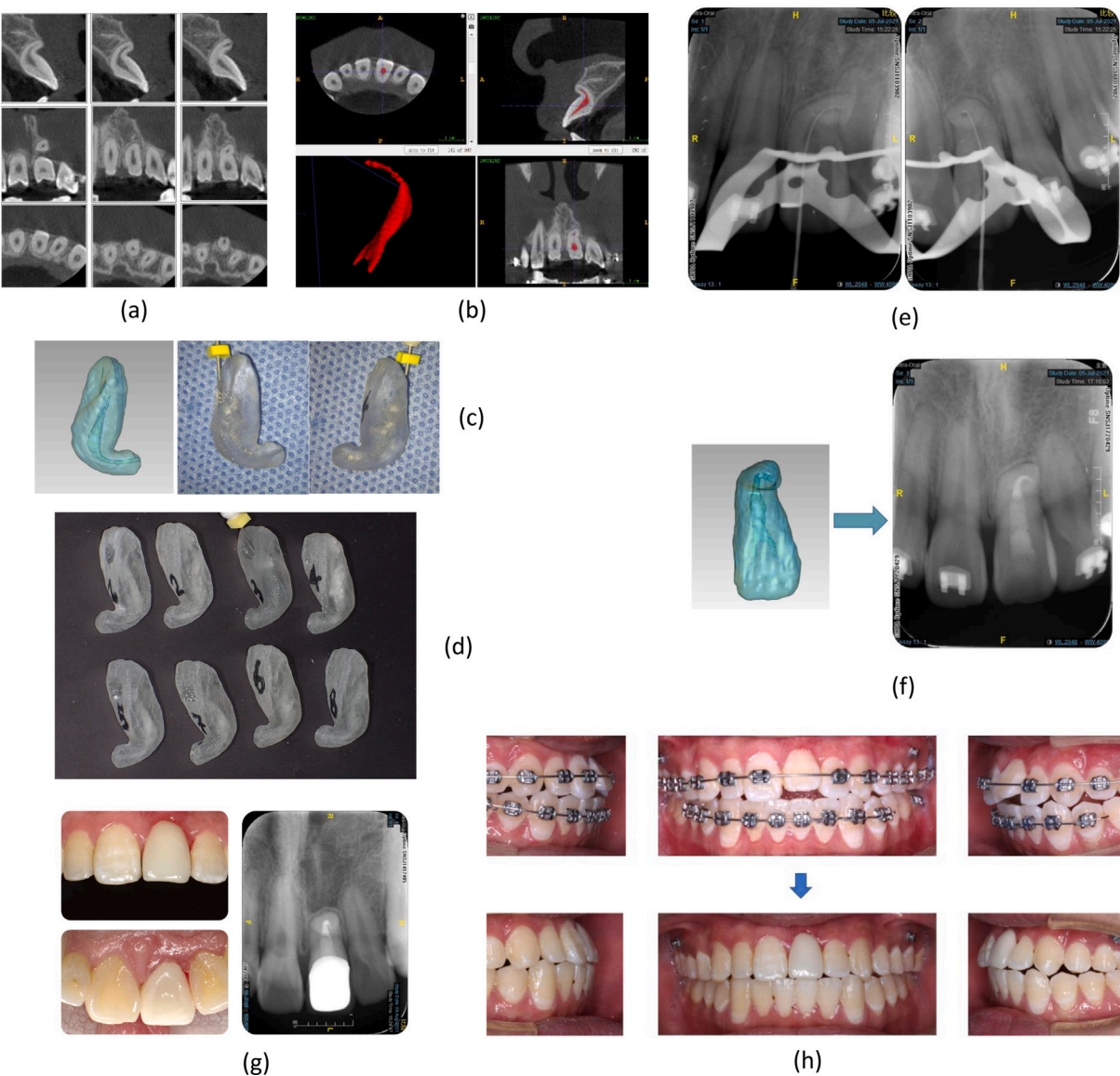


Fig. 9. Study of case 1. (a) The CBCT images of the left upper central incisor in sagittal (1st row), coronal (2nd row), and axial views (3rd row), respectively. (b) The segmentation of the left upper central incisor's root canal from CBCT images. (c) The 3D reconstruction and printed model of the left upper central incisor. (d) Eight 3D printed root canal tooth models for in vitro simulation. (e) X-ray images during root canal therapy. (f) X-ray images immediately after the treatment. (g) Intraoral photos and X-ray image one month post-operatively (the tooth has been restored with a full crown). (h) Comparison of intraoral photos before and after the surgery.

Table 3
Root canal surgical planning based on the in vitro 3D printed tooth models.

Plan ID	Access opening position	Preparation instrument	Results
1	labial side	X-Gold	It is difficult for the instrument to reach the root apex.
2	lingual side		
3	labial side	X-Blue	It is a little difficult. X2 can reach the apex, but it is easy to get stuck.
4	lingual side		
5	labial side	X-Gray	Instrument separation occurs in the middle and lower part of the root canal.
6	lingual side		
7	labial side	Hyflex CM	Instrument separation occurs in the apical part of the root canal.
8	lingual side		
			Barely completion of the full-length root canal preparation until 04/30

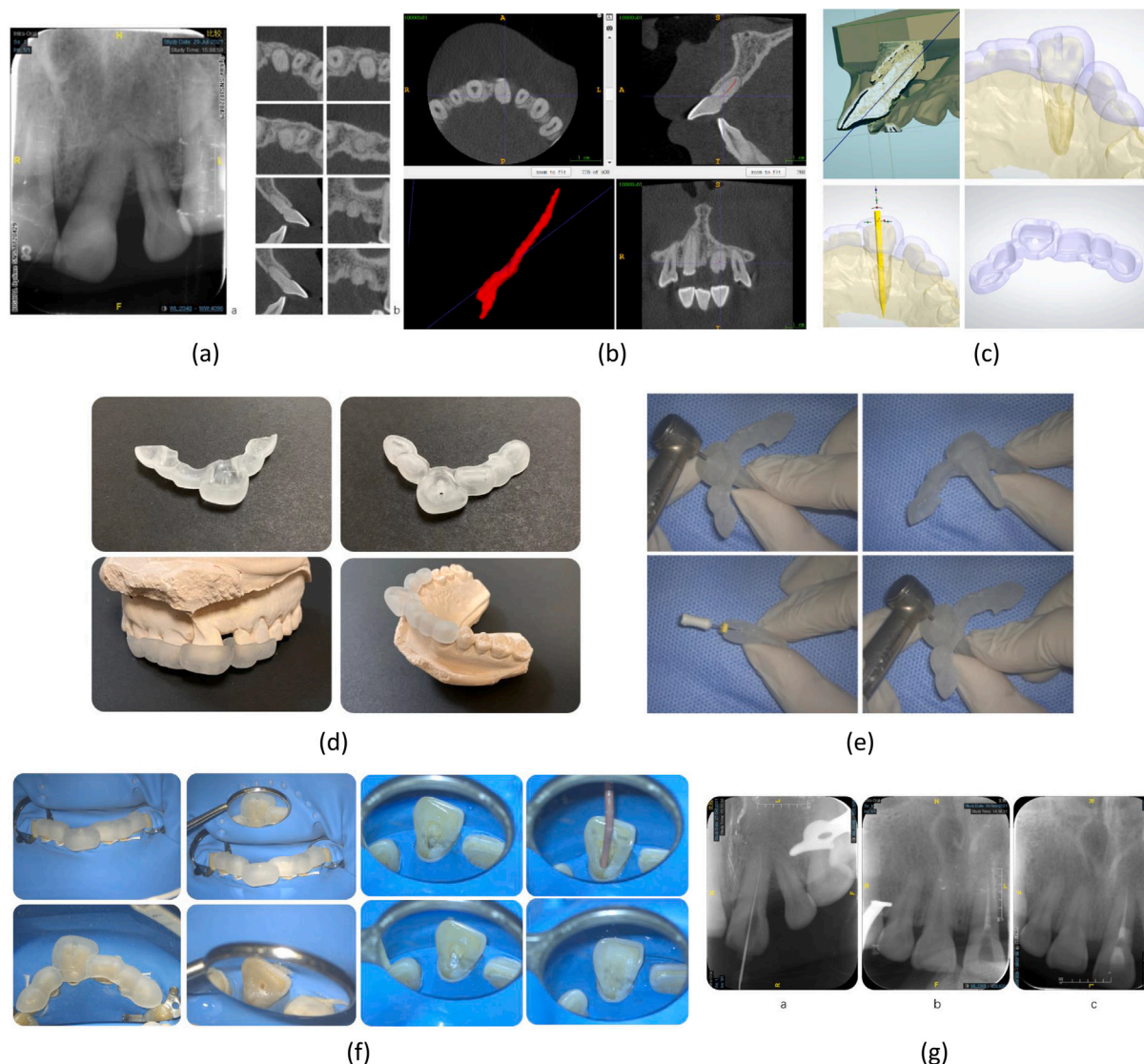


Fig. 10. Study of case 2. (a) Preoperative images of X-ray and CBCT. (b) The segmentation of the left upper central incisor's root canal from CBCT images. (c) The design of access opening guide. (d) 3D printed access opening guide. (e) Verification of the access opening guide by simulating on the 3D printed root canal tooth model. (f) The clinical process of access-opening-guided root canal treatment. (g) Intraoperative localization X-ray, immediate postoperative X-ray, and X-ray image one month post-operatively.

different preparation instruments² in Table 3, and performed in vitro experiments on eight 3D printed root canal tooth models respectively (Fig. 9(d)). The results showed that in plan 6, adopting the lingual access opening and the X-Gray nickel–titanium file with better flexibility could establish a smooth root canal access, and complete the full-length preparation of the severely curved root canal by pre-bending. (4) After determination of the best surgical plan (lingual access opening and X-Gray nickel–titanium file), the clinician can strengthen the proficiency by performing preparation exercises on the 3D printed model to optimize the best surgical plan. (5) We adopted the best surgical plan in the actual clinical application, successfully completed the root canal treatment, and restored the patient's anterior tooth with a post crown (Fig. 9(e–h)).

In this case, DentalNet and PulpNet technologies have achieved rapid and accurate automatic segmentation of the affected tooth and its pulp cavity. The obtained 3D printed root canal tooth model is highly

consistent with the actual tooth in terms of external root shape and internal root canal shape. The patient-specific surgical plan determined by the in vitro simulations makes the clinically difficult root canal treatment successful.

4.4.2. Case 2: severely calcified root canal

A 36-year-old male patient with left upper anterior teeth protruding for more than 20 years required full-crown restoration after root canal treatment to improve the appearance. The patient had a history of supernumerary tooth extraction and dental trauma in the left upper anterior tooth area more than 10 years ago. The preoperative X-ray showed that the left upper central incisor had no obvious root canal in Fig. 10(a) left. Then, from a small-field CBCT (Fig. 10(a) right), it was found that the calcification in the pulp cavity and root canal was severe, and only some blurred root canal images could be seen. For this severely calcified root canal, we utilized 3D printed model to design the access opening, locate the root canal orifice, and establish a smooth root canal access for the success of the access-opening-guided root canal surgery.

The whole workflow consists of five steps. (1) We obtained the automatic segmentation of the left upper central incisor and its pulp

² The instruments, X-Gold, X-Blue, and X-Gray, are products of Shenzhen Superline Technology Co.Ltd., <http://en.supline.com/>. The instrument Hyflex CM is product of COLTENE, <https://www.coltene.com/>.

cavity in the CBCT images, which was completed within 2 min by using DentalNet and PulpNet models (Fig. 10(b)). (2) We imported the 3D reconstruction of the left upper central incisor and the patient's intraoral 3D model into the 3Shape Appliance Designer 2020 software, and designed the personalized access opening guide (Fig. 10(c)). (3) We printed the root canal tooth model and access opening guide of the upper left central incisor by using a high-precision 3D printer (Fig. 10(d)). (4) The clinician placed the 3D printed guide template on the 3D printed root canal tooth model, and simulated the access opening treatment. Through the assist of the guide template, the position of the root canal orifice can be found successfully, and a smooth root canal path can be established (Fig. 10(e)). (5) After verifying the accuracy of the access opening guide by in vitro simulation, we applied it in clinical practice. With the help of the precise positioning of the guide, the root canal orifice of the left upper central incisor was successfully accessed after access opening, and the calcified root canal was prepared and obturated (Fig. 10(f-g)).

Based on the automatic segmentation technology proposed in this study, we can not only obtain an accurate root canal tooth model that is highly consistent with the affected tooth's shape in a short time for in vitro root canal surgical simulation, but also combine it with other software to design personalized access opening guide that can assist in the precise localization of calcified root canals during surgery.

5. Discussion and conclusion

In clinical practice of root canal treatment, when the actual condition of the affected tooth cannot be determined by periapical X-ray, CBCT examination can be supplemented to observe the anatomical shape and lesion range of the affected tooth to clarify the treatment plan. Since the cost and radiation dose of CBCT are higher than the X-ray, we aim to develop an efficient and practical way to fully utilize the information from CBCT in this study. In summary, we adopt small-field CBCT to obtain high-resolution images of the affected tooth. We propose a novel and fully-automatic deep learning-based method to accurately extract the tooth and its root canal from CBCT data. The reconstructed 3D model can then be obtained to benefit root canal treatment planning by using 3D printing technologies. We validated the accuracy of the proposed segmentation method quantitatively and qualitatively. Then, we demonstrated in two real-case studies that the proposed segmentation method can be applied in practical root canal treatment. The clinical studies reflect the broad application prospects of the proposed technology in difficult root canal treatments.

The key contributions of the proposed method lie in three aspects. Firstly, the accurate and automatic segmentation that can be done in 2 min greatly improves the efficiency of the doctor by avoiding the tedious efforts of manual annotations. Therefore, the doctor can have more time to deal with other practices. Secondly, the accurate 3D reconstruction of the tooth model can be 3D printed and used in preoperative surgical planning. The clinician can use the 3D printed models for personalized planning and simulations, which can in turn help to achieve successful outcomes. Last but not least, this efficient segmentation technology can enable large data analysis of anatomical morphology of tooth and root canal in a population. Such data analysis may help other practices in health care, for example, improvement of instruments or treatments.

Due to limited data samples, we focus on single-root teeth and do not cover variant root canals in all teeth in this study. In future work, some further improvements can be considered, such as standardization of the scanning process and data set expansion to include more data with consistent manual annotations.

CRediT authorship contribution statement

Yiwei Wang: Conceptualization, Resources, Validation, Writing – original draft. **Wenjun Xia:** Resources, Validation, Writing – original draft. **Zhennan Yan:** Methodology, Validation, Writing – original draft. **Liang Zhao:** Methodology, Validation, Writing – original draft. **Xiaohe Bian:** Resources, Writing – original draft. **Chang Liu:** Methodology, Validation, Writing – original draft. **Zhennan Qi:** Resources, Writing – review & editing. **Shaoting Zhang:** Project administration, Supervision, Writing – review & editing. **Zisheng Tang:** Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Chang Liu, Liang Zhao has patent Tooth image processing method and device, electronic equipment and storage medium pending to SHANGHAI SHANGTANG INTELLIGENT TECH CO LTD.

Data availability

The data that has been used is confidential.

Acknowledgments

This work was supported by funds from the National Natural Science Foundation of China (82071105, 81870749, 21934007), the Cross-disciplinary Research Fund of Shanghai Ninth People's Hospital, Shanghai Jiao Tong university School of Medicine (JYJC202006), the CAMS Innovation Fund for Medical Sciences (CIFMS) (Project No. 2019-I2M-5-037), the Clinical Research Program of Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine (JYLJ201815), the Centre for Perceptual and Interactive Intelligence (CPIL) Ltd under the Innovation and Technology Commission (ITC)'s InnoHK.

References

- Athirasala, A., Lins, F., Tahayeri, A., Hinds, M., Smith, A.J., Sedgley, C., Ferracane, J., Bertassoni, L.E., 2017. A novel strategy to engineer pre-vascularized full-length dental pulp-like tissue constructs. *Sci. Rep.* 7 (1), 1–11.
- Barone, S., Paoli, A., Rationale, A.V., 2016. CT segmentation of dental shapes by anatomy-driven reformation imaging and B-spline modelling. *Int. J. Numer. Methods Biomed. Eng.* 32 (6), e02747.
- Berman, M., Triki, A.R., Blaschko, M.B., 2018. The Lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 4413–4421.
- Chen, Y., Du, H., Yun, Z., Yang, S., Dai, Z., Zhong, L., Feng, Q., Yang, W., 2020a. Automatic segmentation of individual tooth in dental CBCT images from tooth surface map by a multi-task FCN. *IEEE Access* 8, 97296–97309.
- Chen, T., Kornblith, S., Norouzi, M., Hinton, G., 2020b. A simple framework for contrastive learning of visual representations. In: *International Conference on Machine Learning*. PMLR, pp. 1597–1607.
- Chung, M., Lee, M., Hong, J., Park, S., Lee, J., Lee, J., Yang, L.-H., Lee, J., Shin, Y.-G., 2020. Pose-aware instance segmentation framework from cone beam CT images for tooth segmentation. *Comput. Biol. Med.* 120, 103720.
- Cui, Z., Fang, Y., Mei, L., Zhang, B., Yu, B., Liu, J., Jiang, C., Sun, Y., Ma, L., Huang, J., Liu, Y., Zhao, Y., Lian, C., Ding, Z., Zhu, M., Shen, D., 2022. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. *Nature Commun.* 13 (1), 1–11.
- Cui, Z., Li, C., Wang, W., 2019. ToothNet: Automatic tooth instance segmentation and identification from cone beam CT images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 6368–6377.
- Cui, Z., Zhang, B., Lian, C., Li, C., Yang, L., Wang, W., Zhu, M., Shen, D., 2021. Hierarchical morphology-guided tooth instance segmentation from CBCT images. In: *International Conference on Information Processing in Medical Imaging*. Springer, pp. 150–162.
- Da, Q., Huang, X., Li, Z., Zuo, Y., Zhang, C., Liu, J., Chen, W., Li, J., Xu, D., Hu, Z., et al., 2022. DigestPath: A benchmark dataset with challenge review for the pathological detection and segmentation of digestive-system. *Med. Image Anal.* 102485.

- Dentistry, 2020. Dentistry — Designation system for teeth and areas of the oral cavity. <https://www.iso.org/standard/41835.html>. (Last Accessed 11 July 2020).
- Duan, W., Chen, Y., Zhang, Q., Lin, X., Yang, X., 2021. Refined tooth and pulp segmentation using U-Net in CBCT image. *Dentomaxillofacial Radiol.* 49, 20200251.
- Duan, Q., Wang, G., Wang, R., Fu, C., Li, X., Gong, M., Liu, X., Xia, Q., Huang, X., Hu, Z., et al., 2020. SenseCare: A research platform for medical image informatics and interactive 3D visualization. *arXiv preprint arXiv:2004.07031*.
- European Society of Endodontology, 2006. Quality guidelines for endodontic treatment: Consensus report of the European society of endodontology. *Int. Endontic. J.* 39 (12), 921–930.
- Evain, T., Ripoche, X., Atif, J., Bloch, I., 2017. Semi-automatic teeth segmentation in cone-beam computed tomography by graph-cut with statistical shape priors. In: 2017 IEEE 14th International Symposium on Biomedical Imaging. ISBI 2017, IEEE, pp. 1197–1200.
- Ezhov, M., Zakirov, A., Gusarev, M., 2019. Coarse-to-fine volumetric segmentation of teeth in cone-beam CT. In: 2019 IEEE 16th International Symposium on Biomedical Imaging. ISBI 2019, IEEE, pp. 52–56.
- Gan, Y., Xia, Z., Xiong, J., Li, G., Zhao, Q., 2017. Tooth and alveolar bone segmentation from dental computed tomography images. *IEEE J. Biomed. Health Inf.* 22 (1), 196–204.
- Gan, Y., Xia, Z., Xiong, J., Zhao, Q., Hu, Y., Zhang, J., 2015. Toward accurate tooth segmentation from computed tomography images using a hybrid level set model. *Med. Phys.* 42 (1), 14–27.
- Gao, H., Chae, O., 2010. Individual tooth segmentation from CT images using level set method with shape and intensity prior. *Pattern Recognit.* 43 (7), 2406–2417.
- Gu, R., Zhang, J., Wang, G., Lei, W., Song, T., Zhang, X., Li, K., Zhang, S., 2022. Contrastive semi-supervised learning for domain adaptive segmentation across similar anatomical structures. *IEEE Trans. Med. Imaging* 42 (1), 245–256.
- Han, T., Ma, Y., Yang, L., Chen, X., Zhang, X., Wang, Y., 2014. A study of the root canal morphology of mandibular anterior teeth using cone-beam computed tomography in a Chinese subpopulation. *J. Endodontics* 40 (9), 1309–1314.
- Hatvani, J., Horváth, A., Michetti, J., Basarab, A., Kouamé, D., Gyöngy, M., 2018. Deep learning-based super-resolution applied to dental computed tomography. *IEEE Trans. Radiat. Plasma Med. Sci.* 3 (2), 120–128.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. In: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2961–2969.
- Herrero-Hernández, S., López-Valverde, N., Bravo, M., Valencia de Pablo, Ó., Peix-Sánchez, M., Flores-Fraile, J., Ramírez, J.M., Macedo de Sousa, B., López-Valverde, A., 2020. Root canal morphology of the permanent mandibular incisors by cone beam computed tomography: A systematic review. *Appl. Sci.* 10 (14), 4914.
- Hiew, L., Ong, S., Foong, K.W., Weng, C., 2010. Tooth segmentation from cone-beam CT using graph cut. In: *Proceedings of the Second APSIPA Annual Summit and Conference*. ASC, Singapore, pp. 272–275.
- Jang, T.J., Kim, K.C., Cho, H.C., Seo, J.K., 2021. A fully automated method for 3D individual tooth identification and segmentation in dental CBCT. *IEEE Trans. Pattern Anal. Mach. Intell.* 1. <http://dx.doi.org/10.1109/TPAMI.2021.3086072>.
- Ji, D.X., Ong, S.H., Foong, K.W.C., 2014. A level-set based approach for anterior teeth segmentation in cone beam computed tomography images. *Comput. Biol. Med.* 50, 116–128.
- Keustermans, J., Vandermeulen, D., Suetens, P., 2012. Integrating statistical shape models into a graph cut framework for tooth segmentation. In: *International Workshop on Machine Learning in Medical Imaging*. Springer, pp. 242–249.
- Keyhaninejad, S., Zoroofi, R.A., Setarehdan, S.K., Shirani, G., 2006. Automated segmentation of teeth in multi-slice CT images. In: 2006 IET International Conference on Visual Information Engineering. pp. 339–344. <http://dx.doi.org/10.1049/cp:20060553>.
- Kfir, A., Telishevsky-Strauss, Y., Leitner, A., Metzger, Z., 2013. The diagnosis and conservative treatment of a complex type 3 dens invaginatus using cone beam computed tomography (CBCT) and 3D plastic models. *Int. Endontic. J.* 46 (3), 275–288.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Krogh, A., Hertz, J., 1991. A simple weight decay can improve generalization. *Adv. Neural Inf. Process. Syst.* 4.
- Lee, S., Woo, S., Yu, J., Seo, J., Lee, J., Lee, C., 2020. Automated CNN-based tooth segmentation in cone-beam CT for dental implant planning. *IEEE Access* 8, 50507–50518.
- Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017a. Feature pyramid networks for object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 2117–2125.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017b. Focal loss for dense object detection. In: *Proceedings of the IEEE International Conference on Computer Vision*. pp. 2980–2988.
- Liu, J., Luo, J., Dou, L., Yang, D., 2014. CBCT study of root and canal morphology of permanent mandibular incisors in a Chinese population. *Acta Odontol. Scand.* 72 (1), 26–30.
- Lu, X., Ma, C., Ni, B., Yang, X., Reid, I., Yang, M.-H., 2018. Deep regression tracking with shrinkage loss. In: *Proceedings of the European Conference on Computer Vision*. ECCV, pp. 353–369.
- Luo, X., Liao, W., Xiao, J., Chen, J., Song, T., Zhang, X., Li, K., Metaxas, D.N., Wang, G., Zhang, S., 2022a. WORD: A large scale dataset, benchmark and clinical applicable study for abdominal organ segmentation from CT image. *Med. Image Anal.* 82, 102642.
- Luo, X., Wang, G., Liao, W., Chen, J., Song, T., Chen, Y., Zhang, S., Metaxas, D.N., Zhang, S., 2022b. Semi-supervised medical image segmentation via uncertainty rectified pyramid consistency. *Med. Image Anal.* 80, 102517.
- Machado, J.F., Pires, P.M., dos Santos, T.M.P., de Almeida Neves, A., Lopes, R.T., Visconti, M.A.P.G., 2019. Root canal segmentation in cone-beam computed tomography: Comparison with a micro-CT gold standard. *Braz. J. Oral Sci.* 18, e191627.
- Michetti, J., Basarab, A., Diemer, F., Kouame, D., 2017. Comparison of an adaptive local thresholding method on CBCT and μ CT endodontic images. *Phys. Med. Biol.* 63 (1), 015020.
- Michetti, J., Georgelin-Gurgel, M., Mallet, J.-P., Diemer, F., Boulanaou, K., 2015. Influence of CBCT parameters on the output of an automatic edge-detection-based endodontic segmentation. *Dentomaxillofacial Radiol.* 44 (8), 20140413.
- Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision. 3DV, IEEE, pp. 565–571.
- Misra, I., Shrivastava, A., Gupta, A., Hebert, M., 2016. Cross-stitch networks for multi-task learning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3994–4003.
- Neelakantan, P., Subbarao, C., Ahuja, R., Subbarao, C.V., Gutmann, J.L., 2010. Cone-beam computed tomography study of root and canal morphology of maxillary first and second molars in an Indian population. *J. Endod.* 36 (10), 1622–1627.
- Neven, D., Brabandere, B.D., Proesmans, M., Gool, L.V., 2019. Instance segmentation by jointly optimizing spatial embeddings and clustering bandwidth. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 8837–8845.
- Ng, Y.-L., Mann, V., Rahbaran, S., Lewsey, J., Gulabivala, K., 2007. Outcome of primary root canal treatment: Systematic review of the literature—part 1. Effects of study characteristics on probability of success. *Int. Endontic. J.* 40 (12), 921–939.
- Nguyen, P.N., Van Pham, K., 2020. Endodontic length measurements using different modalities: An in vitro study. *J. Int. Soc. Prevent. Commun. Dentistry* 10 (6), 752.
- Nie, D., Shen, D., 2020. Adversarial confidence learning for medical image segmentation and synthesis. *Int. J. Comput. Vis.* 128 (10), 2494–2513.
- Ordinola-Zapata, R., Versiani, M.A., Bramante, C.M., 2019. Root canal components. In: *The Root Canal Anatomy in Permanent Dentition*. Springer International Publishing, Cham, pp. 31–46. http://dx.doi.org/10.1007/978-3-319-73444-6_3.
- Patel, S., Durack, C., Abella, F., Shemesh, H., Roig, M., Lemberg, K., 2015. Cone beam computed tomography in endodontics—A review. *Int. Endontic. J.* 48 (1), 3–15.
- Pei, Y., Ai, X., Zha, H., Xu, T., Ma, G., 2016. 3D exemplar-based random walks for tooth segmentation from cone-beam computed tomography images. *Med. Phys.* 43 (9), 5040–5050.
- Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* 28, 91–99.
- Reymus, M., Fotiadou, C., Kessler, A., Heck, K., Hickel, R., Diegritz, C., 2019. 3D printed replicas for endodontic education. *Int. Endontic. J.* 52 (1), 123–130.
- Romera, E., Alvarez, J.M., Bergasa, L.M., Arroyo, R., 2017. Erfnet: Efficient residual factorized convnet for real-time semantic segmentation. *IEEE Trans. Intell. Transp. Syst.* 19 (1), 263–272.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 234–241.
- Rubino, D., 2018. Working to dispel the root canal myth. URL: <https://www.aae.org/patients/2018/03/16/working-dispel-root-canal-myth/>.
- Scarfe, W.C., Farman, A.G., Sukovic, P., 2006. Clinical applications of cone-beam computed tomography in dental practice. *J. Can. Dent. Assoc.* 72 (1), 75.
- Scarfe, W.C., Levin, M.D., Gane, D., Farman, A.G., 2009. Use of cone beam computed tomography in endodontics. *Int. J. Dent.* 2009.
- Scarfe, W.C., Li, Z., Aboelmaaty, W., Scott, S., Farman, A., 2012. Maxillofacial cone beam computed tomography: Essence, elements and steps to interpretation. *Aust. Dent. J.* 57, 46–60.
- Scharr, H., Minervini, M., French, A.P., Klukas, C., Kramer, D.M., Liu, X., Luengo, I., Pape, J.-M., Polder, G., Vukadinovic, D., Yin, X., Tsafaris, S.A., 2016. Leaf segmentation in plant phenotyping: A collation study. *Mach. Vis. Appl.* 27 (4), 585–606.
- Sfeir, R., Michetti, J., Chebaro, B., Diemer, F., Basarab, A., Kouamé, D., 2017. Dental root canal segmentation from super-resolved 3D cone beam computed tomography data. In: 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference. NSS/MIC, IEEE, pp. 1–2.
- Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. *J. Big Data* 6 (1), 1–48.
- Simpson, A.L., Antonelli, M., Bakas, S., Bilello, M., Farahani, K., Van Ginneken, B., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., et al., 2019. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv preprint arXiv:1902.09063*.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15 (56), 1929–1958, URL: <http://jmlr.org/papers/v15/srivastava14a.html>.

- Tan, C., Yan, Z., Zhang, S., Li, K., Metaxas, D.N., 2019. Collaborative multi-agent learning for MR knee articular cartilage segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 282–290.
- Tan, C., Zhao, L., Yan, Z., Li, K., Metaxas, D., Zhan, Y., 2018. Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging. ISBI 2018, IEEE, pp. 1221–1224.
- Tchorz, J., Wrbas, K., Von See, C., Vach, K., Patzelt, S., 2019. Accuracy of software-based three-dimensional root canal length measurements using cone-beam computed tomography. *Eur. Endod. J.* 4 (1), 28.
- Valenti-Obino, F., Di Nardo, D., Quero, L., Miccoli, G., Gambarini, G., Testarelli, L., Galli, M., 2019. Symmetry of root and root canal morphology of mandibular incisors: A cone-beam computed tomography study in vivo. *J. Clin. Exp. Dent.* 11 (6), e527.
- Vertucci, F.J., 1984. Root canal anatomy of the human permanent teeth. *Oral. Surg. Oral. Med. Oral. Pathol.* 58 (5), 589–599.
- Wang, L., Guo, D., Wang, G., Zhang, S., 2020. Annotation-efficient learning for medical image segmentation based on noisy pseudo labels and adversarial learning. *IEEE Trans. Med. Imaging* 40 (10), 2795–2807.
- Wang, L., Li, J.-p., Ge, Z.-p., Li, G., 2019. CBCT image based segmentation method for tooth pulp cavity region extraction. *Dentomaxillofacial. Radiol.* 48 (2), 20180236.
- Wang, W., Xia, Q., Hu, Z., Yan, Z., Li, Z., Wu, Y., Huang, N., Gao, Y., Metaxas, D., Zhang, S., 2021a. Few-shot learning by a cascaded framework with shape-constrained pseudo label assessment for whole heart segmentation. *IEEE Trans. Med. Imaging* 40 (10), 2629–2641.
- Wang, G., Zhai, S., Lasio, G., Zhang, B., Yi, B., Chen, S., Macvittie, T.J., Metaxas, D., Zhou, J., Zhang, S., 2021b. Semi-supervised segmentation of radiation-induced pulmonary fibrosis from lung CT scans with multi-scale guided dense attention. *IEEE Trans. Med. Imaging* 41 (3), 531–542.
- Wu, X., Chen, H., Huang, Y., Guo, H., Qiu, T., Wang, L., 2020. Center-sensitive and boundary-aware tooth instance segmentation and classification from cone-beam CT. In: 2020 IEEE 17th International Symposium on Biomedical Imaging. ISBI, IEEE, pp. 939–942.
- Yang, Y., Xie, R., Jia, W., Chen, Z., Yang, Y., Xie, L., Jiang, B., 2021. Accurate and automatic tooth image segmentation model with deep convolutional neural networks and level set method. *Neurocomputing* 419, 108–125.
- Yu, J., Blaschko, M., 2015. Learning submodular losses with the Lovász hinge. In: International Conference on Machine Learning. PMLR, pp. 1623–1631.
- Zheng, Q., Ge, Z., Du, H., Li, G., 2021. Age estimation based on 3D pulp chamber segmentation of first molars from cone-beam-computed tomography by integrated deep learning and level set. *Int. J. Legal Med.* 135 (1), 365–373.
- Zhou, X., Gan, Y., Xiong, J., Zhang, D., Zhao, Q., Xia, Z., 2018. A method for tooth model reconstruction based on integration of multimodal images. *J. Healthc. Eng.* 2018.
- Zichun, Y., Qunfei, Z., Zisheng, T., Wenjun, X., 2020. CBCT image segmentation of tooth-root canal based on improved level set algorithm. In: Proceedings of the 2020 International Conference on Computers, Information Processing and Advanced Education. pp. 42–51.