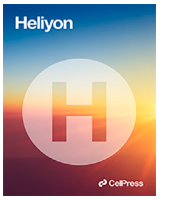
Heliyon 9 (2023) e17651

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Whole mandibular canal segmentation using transformed dental CBCT volume in Frenet frame

Huanmiao Zhao [a,b], Junhua Chen [c], Zhaoqiang Yun [a,b], Qianjin Feng [a,b], Liming Zhong [a,b,*], Wei Yang [a,b,*]

[a] School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China
[b] Guangdong Provincial Key Laboratory of Medical Image Processing, Guangzhou, 510515, China
[c] Stomatology Hospital of Guangzhou Medical University, Guangzhou, 510140, China

ARTICLE INFO

Keywords:
Mandibular canal
Frenet frame
CBCT
Segmentation
Deep learning

ABSTRACT

Accurate segmentation of the mandibular canal is essential in dental implant and maxillofacial surgery, which can help prevent nerve or vascular damage inside the mandibular canal. Achieving this is challenging because of the low contrast in CBCT scans and the small scales of mandibular canal areas. Several innovative methods have been proposed for mandibular canal segmentation with positive performance. However, most of these methods segment the mandibular canal based on sliding patches, which may adversely affect the morphological integrity of the tubular structure. In this study, we propose whole mandibular canal segmentation using transformed dental CBCT volume in the Frenet frame. Considering the connectivity of the mandibular canal, we propose to transform the CBCT volume to obtain a sub-volume containing the whole mandibular canal based on the Frenet frame to ensure complete 3D structural information. Moreover, to further improve the performance of mandibular canal segmentation, we use clDice to guarantee the integrity of the mandibular canal structure and segment the mandibular canal. Experimental results on our CBCT dataset show that integrating the proposed transformed volume in the Frenet frame into other state-of-the-art methods achieves a 0.5% ~ 12.1% improvement in Dice performance. Our proposed method can achieve impressive results with a Dice value of 0.865 (±0.035), and a clDice value of 0.971 (±0.020), suggesting that our method can segment the mandibular canal with superior performance.

1. Introduction

The mandibular canal, which is located in the mandible, is continuous with the mental foramen (opening in front of the mandible) and mandibular foramen (on the inside of the mandibular ramus) (Fig. 1). The mandibular canal consists of the inferior alveolar nerve (IAN), inferior alveolar artery (IAA), and inferior alveolar vein (IAV), which is a major supplier of sensation and blood to the lower jaw [1]. In clinical practice, cone beam computed tomography (CBCT) is the most commonly used for 3D diagnostics and operation planning in dental and maxillofacial radiology due to its lower doses of radiation and cost [2]. The accurate location of the mandibular canal is of great importance to assist in implant placement and other maxillofacial surgeries. Injuries to the mandibular canal during oral surgery can lead to adverse consequences such as facial nerve dysfunction and rupture of blood vessels. The

* Corresponding authors at: School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China.
E-mail addresses: limingzhongmindy@gmail.com (L. Zhong), weiyanggm@gmail.com (W. Yang).

https://doi.org/10.1016/j.heliyon.2023.e17651
Received 15 May 2023; Received in revised form 29 May 2023; Accepted 24 June 2023

Available online 28 June 2023

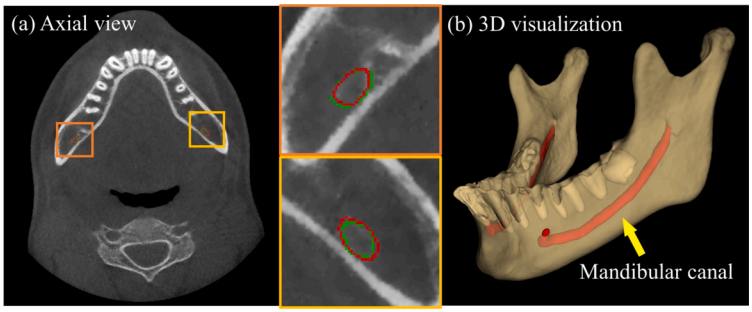


Fig. 1. Mandibular canal segmentation in CBCT images. (a) represents the CBCT image in the axial view. The green curve indicates the ground truth of the mandibular canal and the red curve indicates the mandibular canal mask segmented using our method. (b) A 3D visualization of the mandibular canal (indicated in red) in the mandible (indicated in yellow) via the proposed method.

radiologist's manual localization of the mandibular canal is undoubtedly tedious and time-consuming. A computer-aided diagnostic system can present a 3D view of the target object that is more useful to the clinician in understanding its anatomy compared with a 2D anatomical view. However, automatic segmentation of the mandibular canal is still challenging because of low contrast or even invisible radiographic appearance of the mandibular canal on CBCT images. Existing computer-aided segmentation methods achieve modest performance due to the complexity of the mandibular canal and surrounding tissue anatomy, especially in the molar and submandibular glandular fossa (SDF) areas [3–5]. Therefore, it would be necessary to develop an accurate mandibular canal segmentation method to effectively assist dental clinicians in implant placement and maxillofacial surgery.

Several methods are proposed to automatically segment the mandibular canal in CBCT images based on the prior knowledge of shape models [6–10]. These methods typically use statistical shape models (SSM) to segment the mandible and then segment the mandibular canal using fast matching or tracking algorithms. For example, Kainmueller et al. [6] proposed to segment the mandibular canal in CBCT images by combining a statistical shape model to segment the mandibular surface and the mandibular canal with a Dijkstra-based algorithm optimization. Moris et al. [7] proposed to detect the mandibular foramen and mental foramen first by searching for high-contrast gaps in the mandible and then tracing the path of the mandibular canal between the two points using a matching method. In [8,9], Abdolali et al. achieved automatic segmentation of the mandibular canal based on a statistical shape model in combination with shape context registration and the Lie group, respectively. A similar solution as [6] was also proposed by Abdolali et al. [10]. This framework was proposed for CBCT image enhancement using low-rank matrix decomposition, and a combination of statistical shape model and fast matching method was employed for mandibular bone and canal segmentation. The results of these classical methods were promising. However, the aforementioned methods depend on prior knowledge and a robust mandibular segmentation model. Due to the low contrast of the CBCT images, using predefined thresholds based on the grayscale values for image preprocessing in these methods will limit the accuracy of segmentation in the mandibular canal.

Recently, convolutional neural networks (CNNs), especially U-type networks, have achieved superior performance in medical image segmentation [11–16]. For 2D medical images, the segmentation of tubular structures is well developed, especially for retinal blood vessel segmentation [17,18]. Implementing CNNs for automatically segmenting tubular structures in 3D medical images is still challenging in medical image segmentation for structures such as liver vessels [19] and tracheas [20]. Current segmentation of tubular structures on 3D medical images [19–21] often requires an a priori shape of the tubular structure. Shit et al. [22] proposed a topology-preserving similarity metric (clDice) for tubular structural segmentation using its differentiable form to obtain more accurate connectivity information for segmentation. Some studies in recent years have used CNNs for mandibular canal segmentation [23–25]. Jaskari et al. [23] trained a full convolutional network for automatic mandibular canal segmentation and obtained higher accuracy than previous classical segmentation methods. In [24], Kwak et al. used threshold processing to obtain the mandibular region and further trained 2D and 3D segmentation networks to obtain the mandibular canal. Note that the roughly annotated mandibular canal used in [23,24] greatly hinders developing an accurate mandibular canal segmentation model in clinical practice. In [25], the voxels of the input to the network are restricted to the mandibular range to improve the segmentation accuracy. Segmenting the areas with low contrast between the mandibular canal and the surrounding tissue on CBCT images is challenging. Considering the large spatial dimensions of the CBCT image and the computational efficiency, these 3D CNN-based methods mainly sample local patches from the whole image. However, patch-based segmentation methods cannot sufficiently learn the global information of the mandibular canal because of the low proportion of voxels in the mandibular canal on CBCT images. Furthermore, due to the variation of the mandibular canal structure, the ability to distinguish the mandibular canal near the mandibular foramen and mental foramen is limited by these CNN-based methods. Moreover, none of these methods can guarantee the integrity of the mandibular canal structure.

To tackle the above issues, we propose whole mandibular canal segmentation using transformed dental CBCT volume in the Frenet frame. We first extract the mandibular centerline via automatic segmentation of the mandible and localization of the mandibular foramen and mental foramen. The sub-volumes containing the mandibular canal information are then obtained using a double reflection method based on the Frenet frame. The transformed sub-volumes are fed into the 3D segmentation network, and the clDice is used to constrain the topology of the mandibular canal. Lastly, the prediction masks are inversely transformed back into the original CBCT images to obtain final segmentations.

Overall, the main contributions of this paper are summarized as follows:

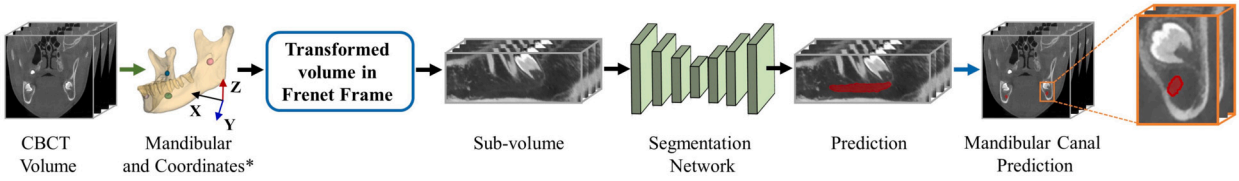


Fig. 2. Workflow of the proposed segmentation of the mandibular canal on CBCT. * represents the coordinates of the mandibular foramen and mental foramen. The green arrow indicates the process of obtaining the mask of the mandibular mask and the coordinates of mandibular foramen and mental foramen. The blue arrow indicates the sub-volume mask conversion to the original image, and the orange box shows the prediction details.

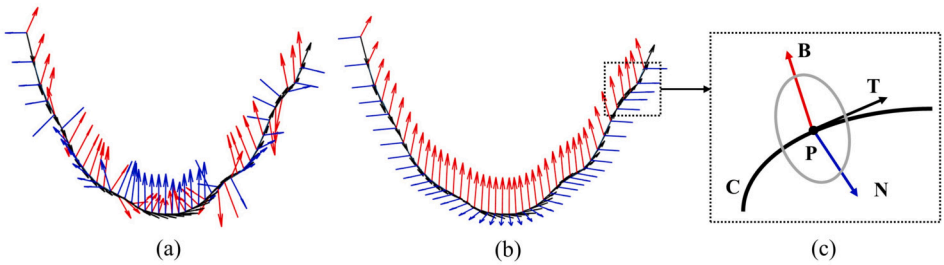


Fig. 3. (a) Diagram of the mandibular centerline under the Frenet frame; (b) Combined with the double reflection method; (c) Illustration of the point P. The gray areas indicate the surface perpendicular to the tangent vector at point P.

(1) We propose to transform the CBCT volume based on the Frenet frame for 3D whole mandibular canal segmentation, which guarantees the global information of the mandibular canal.

(2) Our proposed mandibular canal segmentation framework can achieve the automatic segmentation of mandible and mandibular canal and the automatic detection of the mandibular foramen and mental foramen.

(3) Experimental results demonstrate that by integrating the proposed transformed volume in the Frenet frame into other state-of-the-art methods, the performance of these methods improves significantly on mandibular canal segmentation.

2. Methods

Our proposed mandibular canal segmentation method is shown in Fig. 2. We first segment the mandible and locate the mandibular foramen and mental foramen. Then, using the mandibular mask and the coordinates of the mandibular foramen and mental foramen in the mandible, the sub-volume containing the mandibular canal is obtained by the proposed transformed volume in Frenet frame. Finally, the transformed sub-volume is fed into the 3D segmentation network and the prediction mask is inverse-transformed back into the original CBCT image to obtain the final segmentation. The details of the method are described in the following subsections.

2.1. Transformed volume in Frenet frame

2.1.1. Frenet frame

In Euclidean $\mathbb{R}^3$-space, a point P on a curve C in the Frenet frame consists of N, B, and T, three associated orthogonal unit vector fields [26]. For curve C, T is the unit tangent vector, N is the principal normal and B is the binormal (Fig. 3(c)).

$$T = \frac{P'}{|P|}, N = \frac{P' \times P''}{|P' \times P''|}, B = T \times N \tag{1}$$

In the 3D Frenet frame, there are two curvatures, κ and τ. The constant, κ is called the curvature, τ is called the torsion. The Frenet equations are:

$$\frac{\mathrm{d}T}{\mathrm{d}C} = \kappa N, \frac{\mathrm{d}N}{\mathrm{d}C} = -\kappa T + \tau B, \frac{\mathrm{d}B}{\mathrm{d}C} = -\tau N \tag{2}$$

The Frenet frame is very convenient for calculating any point on the curve. However, the Frenet frame is undefined at curve inflection points or smooth variations, such as where $\kappa = 0$. The curvature vector can be on either side of the inflection point, causing a sharp distortion in the progression of the Frenet frame.

2.1.2. Transformed CBCT volumes

Suppose we extract the central line, a 3D curve $C : \{c_i | i = 1, ..., n\}$, of the mandible along the mandibular canal in the area where it exists between the mandibular foramen and mental foramen. If we compute the curve C directly and analytically in the Frenet frame, the curvature vector will twist uncontrollably and violently because of the presence of near-linear sections of C (Fig. 3(a)). Inspired by [27,28], we propose using the double reflection method to obtain the frame to transform mandibular canal voxels. We take the endpoint of the curve C as the initial point c_0. The right-handed orthogonal frame at c_0 is used as the initial frame $u_0(n_0, b_0, t_0)$. The

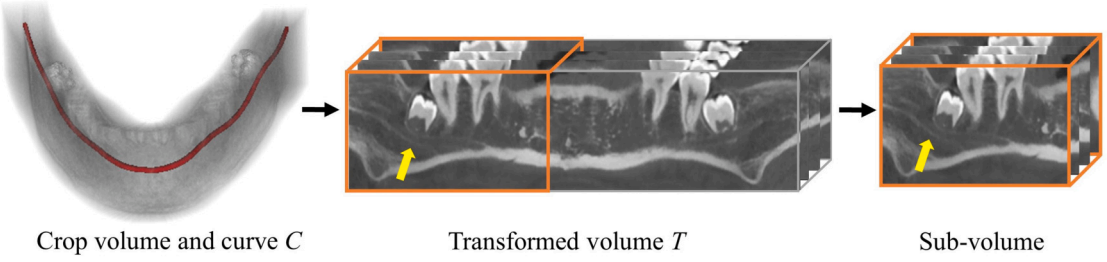


Fig. 4. Illustration of the transformed volume.

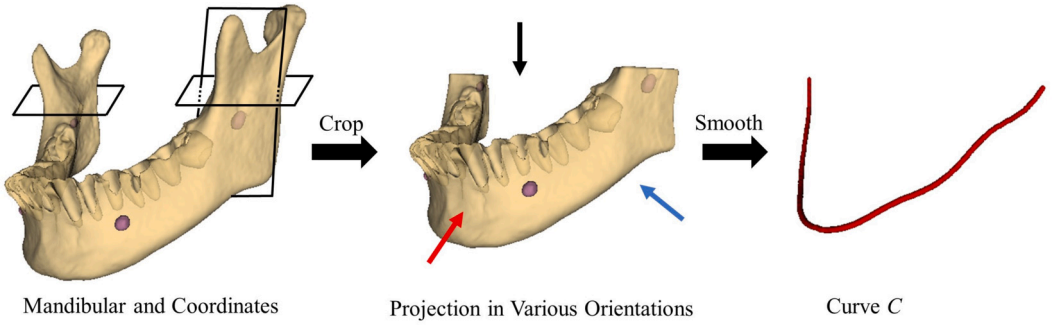


Fig. 5. Illustration of mandibular centerline extraction.

point c_i on the curve C has the unit tangent vector t_i, the primary normal n_i and the subnormal b_i, i.e. the frame $u_i(n_i, b_i, t_i)$. We obtain the left-handed orthogonal frame $u_i^L(n_i^L, b_i^L, t_i^L)$, using the reflection vector $v_1 = c_{i+1} - c_i$, with the following equations:

$$t_i^L = t_i - [2/(v_1 \cdot v_l)] * (v_1 \cdot t_i) * v_1 \tag{3}$$

$$n_i^L = n_i - [2/(v_1 \cdot v_l)] * (v_1 \cdot n_i) * v_1 \tag{4}$$

Next, we derive the frame $u_{i+1}(n_{i+1}, b_{i+1}, t_{i+1})$ of the point c_{i+1} by another reflection vector $v_2 = t_{i+1} - t_i^L$:

$$n_{i+1} = n_i^L - [2/(v_2 \cdot v_2)] * (v_2 \cdot n_i^L) * v_2 \tag{5}$$

$$b_{i+1} = t_{i+1} \times n_{i+1} \tag{6}$$

We obtain the frame $U : \{u_i(n_i, b_i, t_i) | i = 1, ..., n\}$ of the curve C via the double projection method by correcting the unwanted rotation of the Frenet frame, (Fig. 3(b)). We perform the volume transformation based on frame U along the curve C to obtain the transformed volume $T = C + U' * V$. V is artificially set in the region of the centerline C by counting the sizes of the mandibular canal volumes contained in all cases. Here, we take advantage of the symmetry of the anatomical position of the mandibular canal to obtain the final sub-volume containing only one mandibular canal (Fig. 4).

2.2. Mandibular centerline extraction

To obtain the mandibular centerline, we implement a 3D CNN for the automatic segmentation of the mandible and localization of the mandibular foramen and mental foramen. We cropped the acquired mandibular mask at the locations of the mandibular foramen and mental foramen. Then we projected it in each of the three views to obtain the corresponding centerline C (Fig. 5). We utilize nnUNet as the backbone network to segment the mandibular bone and a joint loss composed of Dice loss and Cross-Entropy loss to train the mandibular segmentation network. The mandibular foramen and mental foramen detection are based on regressing heatmap images. Heatmap regression is a popular method for landmark detection due to its robustness and accuracy [29]. The network learns to regress heatmaps by minimizing the differences between the predicted heatmaps $p_i(x)$ and the corresponding target heatmaps $g_i(x)$:

$$\min_x \sum_i^N \sum_x \|p_i(x) - g_i(x)\|_2^2 \tag{7}$$

We set the mean squared error (MSE) as the regression loss to train the network. In the inference, we get the predicted coordinates by the location of the highest heatmap value obtained.

2.3. The mandibular canal segmentation network

The sub-volumes containing the whole mandibular canal information obtained according to the above method are used as input to the 3D segmentation network. We utilize the currently popular medical image processing network UNet as the backbone network. The UNet structure comprises an encoding path, a decoding path and skip connections between them. The encoding path of the mandibular segmentation network consists of five down-sampling operations to extract high-level image features. Correspondingly, the decoding path has five up-sampling operations to restore the space size of the segmented output. Skip connections between the encoding and the decoding paths concatenate the hidden layers along the channel dimension to propagate spatial information and refine the outcomes. Here, we set the kernel size of each convolutional layer to $3 \times 3 \times 3$ and every convolutional operation is followed by a batch normalization and a ReLU nonlinear. Each up-sampling operation is implemented by a transposed convolution with $2 \times 2 \times 2$ kernel size and $2 \times 2 \times 2$ stride. To further maintain the connectivity of the mandibular canal, we use clDice loss [22] to train the mandibular canal segmentation network. The total loss function is formulated as bellow:

$$L_{total} = (1 - \lambda)L_{dice} + \lambda L_{cldice} \tag{8}$$

where λ is the scale factor,

$$L_{cldice} = -2 \times \frac{T_{prec}(S_P, V_L) \times T_{sens}(S_L, V_P)}{T_{prec}(S_P, V_L) + T_{sens}(S_L, V_P)} \tag{9}$$

and where V_P is the prediction of the mandibular canal segmentation network and V_L is the real mask. S_P and S_L are obtained from V_P and V_L by the soft-skeleton method. $T_{prec}(S_P, V_L)$ and $T_{sens}(S_L, V_P)$ are defined bellow (ϵ is set to 1):

$$T_{prec}(S_P, V_L) = \frac{|S_P \cap V_L| + \epsilon}{|S_P| + \epsilon} \tag{10}$$

$$T_{sens}(S_L, V_P) = \frac{|S_L \cap V_P| + \epsilon}{|S_L| + \epsilon} \tag{11}$$

3. Data and experiments

3.1. Data

Our CBCT dataset used in this paper were acquired at the Stomatological Hospital of Southern Medical University, Guangzhou, China. A total of 46 images were collected by a NewTom VGi (QR s.r.l., Verona, Italy) scanner with the following imaging parameters: 110 kVp and 3-8 mA (pulse mode). All the CBCT images were acquired under natural bite or closed bite conditions. Each CBCT image varies in size, with the number of slices ranging from 376-533 and the axial width and length from 376-512 and 470-555, respectively. The dataset was randomly divided into 29 images as the training set, 7 images as the validation set, and another 10 images as the independent test set. We resampled each volume to an isotropic resolution of $1 \times 1 \times 1$ mm^3 from $0.3 \times 0.3 \times 0.3$ mm^3. We calculated the Hounsfield unit (HU) values for all voxels in the foreground of the training data, clipped the images to the 0.5^{th} and 99.5^{th} percentiles of these values, and performed z-score normalization based on the mean and variance of these values.

The publicly available dataset [30] included 347 CBCT images. The spatial resolution of each CBCT scan varies in size from $148 \times 265 \times 312$ to $178 \times 423 \times 463$, with a pixel resolution of $0.3 \times 0.3 \times 0.3$ mm^3. Only 91 cases in this dataset are densely annotated. We performed the same preprocessing on this part of the dataset as on ours. In addition, as in [30], only 15 cases were split as the test set.

3.2. Experiment details

To ensure the integrity of the single mandibular canal in the sub-volume, we expanded the marginal area of the cropped CBCT volume by 32 pixels based on the coordinates of the mandibular foramen and mental foramen. Transforming the CBCT volume containing only the mandibular canal according to the mandibular 3D centerline, we set the size of each slice orthogonal to the tangent vector [64, 96]. The number of slices in the transformed sub-volumes ranged from 249 to 321. Therefore, 58 and 14 volumes with only the single mandibular canal information were obtained as the training and validation sets for the mandibular canal segmentation model, respectively. All sub-volumes were resized to $64 \times 96 \times 280$ during the training phase. In the mandibular foramen and mental foramen landmark localization model, we generated Gaussian heatmaps based on manually labeled coordinates, with σ taken as 3. To improve the performance of the segmentation model, we augmented the training dataset using the data augmentation method in nnUNet. We used test time augmentation for the test set data by mirroring along all axes to further improve the predictions. All models were implemented using PyTorch on NVIDIA RTX 2080 Ti (12G). SGD was adopted as the optimizer with momentum ($\mu = 0.99$) during training. We set the initial learning rate to 0.01, decreasing with training. The batch size was set as 2, and the epoch was 1000 for all models.

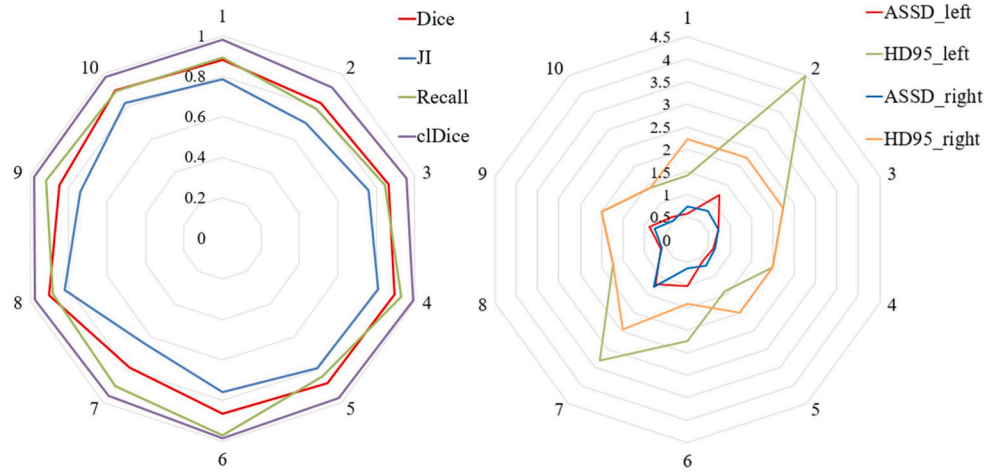


Fig. 6. Quantitative evaluation of the segmentation results for independent test sets. ASSD and HD95 show the results of the left and right mandibular canals respectively.

3.3. Experiment metrics

We adopted four common metrics to evaluate the performance of the segmentation model: Dice similarity coefficient (Dice), Jaccard index (JI), Recall and intactness rate (IR). In particular, we defined a new metric to assess the performance of the model by counting the number of broken or incomplete mandibular canals, namely intactness rate (IR).

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{12}$$

$$JI = \frac{TP}{TP + FP + FN} \tag{13}$$

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

$$IR = \frac{W}{C} \tag{15}$$

TP, FP, and FN denote the number of true positives, false positives and false negatives, respectively. W denotes the number of whole mandibular canals obtained by inference, and C denotes the number of all mandibular canals in the test set.

In addition, for the mandibular canal segmentation model, we also used two metrics based on distance evaluation: the average symmetric surface distance (ASSD), and the 95^{th} percentil of the robust Hausdorff distance (HD_{95}).

$$ASSD = \frac{1}{|S(A) + S(B)|}\left(\sum_{a \in S(A)} d(a, S(B)) + \sum_{b \in S(B)} d(b, S(A))\right) \tag{16}$$

$$HD_{95} = \max\{\max_{P_{95}} d(a, S(B)), \max_{P_{95}} d(b, S(A)) | a \in S(A), b \in S(B)\} \tag{17}$$

$d(,)$ is the shortest Euclidean distance; $S(,)$ is the surface pixel set.

For the performance of the mandibular foramen and mental foramen landmark localization model, we evaluated the mean (Mean) and standard deviation (Std) of the distances between the ground truth and predicted locations as the localization accuracy.

4. Results

This section presents the results of the experiments, including three sections on the overall qualitative and quantitative assessment of our proposed method, ablation studies, and a comparison of the results with other methods. We report results on the independent test set, and all experimental results indicate a positive effect of our proposed method on mandibular canal segmentation.

4.1. Overall segmentation performance of our method

For the mandibular canal segmentation results of our proposed method, the mean Dice was 0.865 (±0.035), the mean JI was 0.764 (±0.053), the mean Recall was 0.886 (±0.053), the mean clDice was 0.971 (±0.020) and the high IR. The standard deviation results for all evaluation metrics indicate that our model performs stably. Additionally, our method showed promising results on distance evaluation metrics. To show this in more detail, the segmentation results for the 10 images are shown in Fig. 6. In the segmentation results of the independent test set, the highest Dice reached 0.903. For subjects with unclear images, the lowest Dice was 0.786. These results show that our method performs well on more complex segmented data. We show a specific example of the mandible

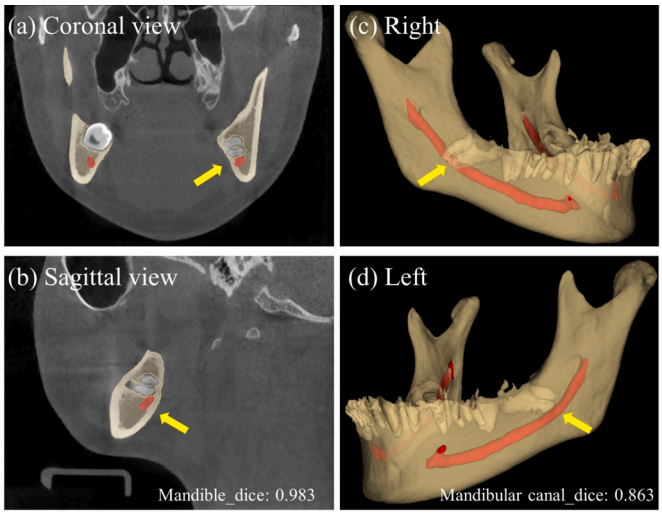


Fig. 7. An example of mandible and mandibular canal segmentation results.

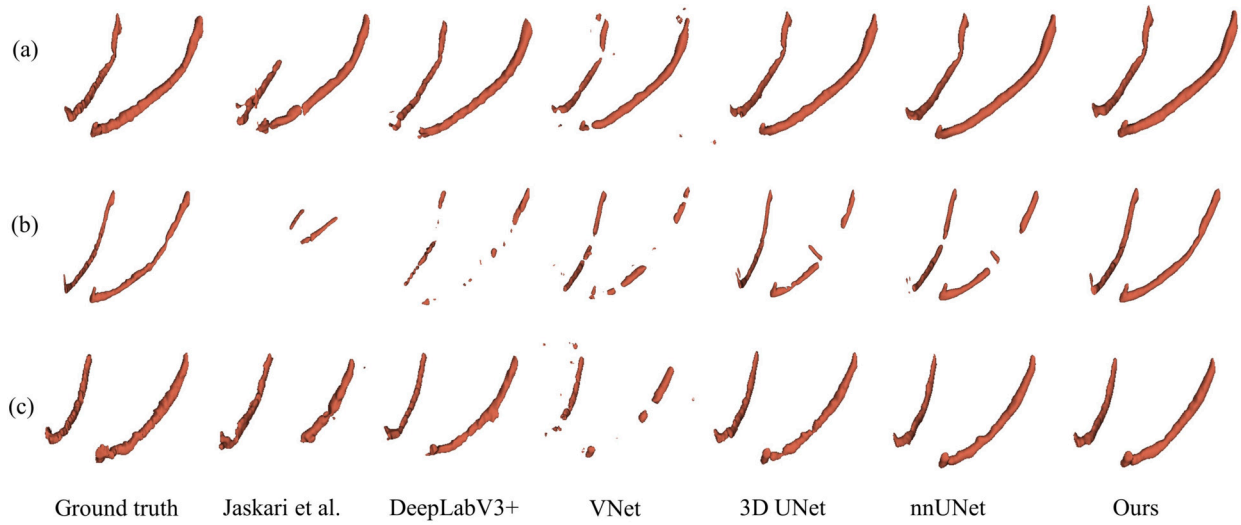


Fig. 8. Examples of 3D visualization with different segmentation methods.

Table 1
Segmentation performance of different methods (mean ± standard deviation). * is a metric for tubular structure, as proposed in [22]. - shows that the result is calculated as NAN.

Method	Dice	JI	Recall	clDice*	ASSD (mm)	HD_{95} (mm)	IR (%)
Jaskari et al.	0.679±0.237	0.550±0.203	0.719±0.203	-	16.242±40.208	31.651±55.171	50
DeepLabV3+	0.730±0.161	0.594±0.157	0.730±0.157	0.882±0.166	1.833±1.501	5.187±5.721	90
VNet	0.769±0.095	0.634±0.118	0.742±0.118	0.756±0.097	6.079±4.013	31.405±49.953	45
3D UNet	0.813±0.038	0.686±0.053	0.864±0.059	0.955±0.055	1.133±0.636	3.780±4.780	90
nnUNet	0.860±0.036	0.757±0.055	0.875±0.055	0.965±0.044	0.836±0.457	2.964±3.430	90
Ours	**0.865±0.035**	**0.764±0.053**	**0.886±0.053**	**0.971±0.020**	**0.772±0.226**	**2.098±0.707**	**100**

and mandibular canal segmentation in Fig. 7(a-d). Besides, our method includes mandibular segmentation, mandibular foramen and mental foramen landmark localization. For accurate extraction of the mandibular centerline, the mandibular segmentation model achieved good performance with Dice, JI, and Recall at 0.986 (±0.002), 0.973 (±0.005), and 0.989 (±0.005) respectively. The results show that the mandibular segmentation model is highly reliable. The landmark localization model of the mandibular foramen and mental foramen achieves an average accuracy of 3.131 (±2.065) mm. This result satisfies our subsequent experiments. Lastly, according to our proposed method of transforming mandibular canal volumes, all the transformed sub-volumes were confirmed to contain whole mandibular canals.

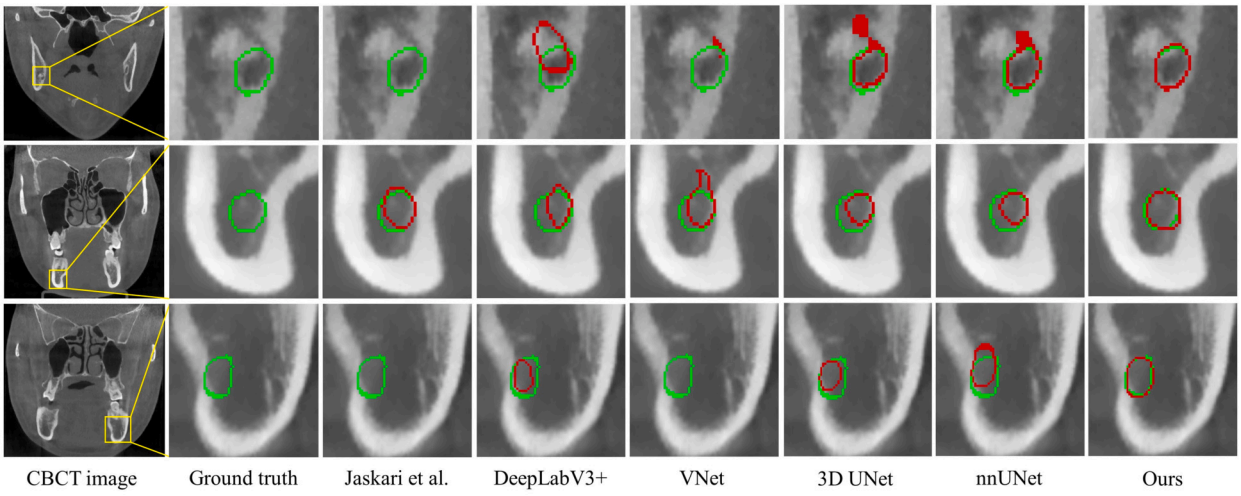


Fig. 9. Illustration of the segmentation results in the coronal view. Green represents the ground truth and red indicates the results predicted by different methods. Only green indicates that the mandibular canal is not segmented here.

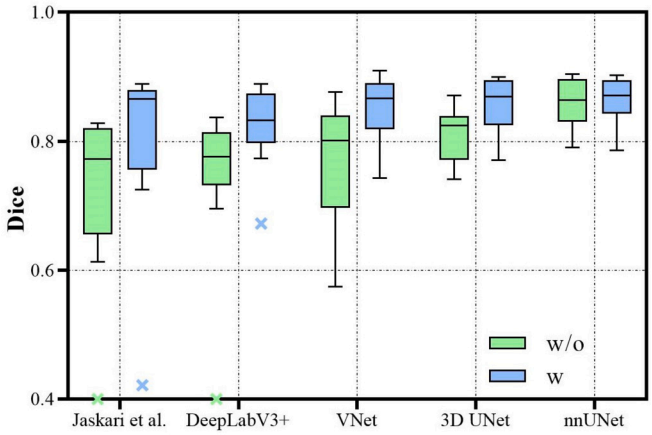


Fig. 10. Comparison of Dice values for different segmentation networks. (w) indicates the use of the transformed volume method, (w/o) indicates no use.

Fig. 11. 3D visualization of the segmentation results with or without the clDice module.

4.2. Comparison with other methods

This section compares and analyzes our proposed method with other patch-based segmentation methods. Fig. 8(a-c) and Fig. 9 show the qualitative comparison of different methods. The FCN-based method proposed by Jaskari et al. [23], DeepLabV3+ [31], VNet [14], 3D UNet [15], nnUNet [16] and our method evaluated the performance of the independent test set. The mandibular canal tends to show an anatomical representation of a small canal diameter due to the compression of the mandibular canal by the surrounding tissue in the molar region. In this region, the patch-based segmentation methods perform poorly and the predicted mandibular canal masks appear broken. In addition, as these models learn patches outside of the mandibular canal region, they result in incorrect mandibular canal predictions. Our proposed method limits the regions and therefore avoids such errors. Our method also performs superiorly in the part of the mandibular canal near the lower teeth. Table 1 shows the results of the different

Table 2
Quantitative evaluation of ablation studies of the proposed method (mean ± standard deviation). T indicates the proposed transformed volume method. C indicates the clDice module.

	Baseline	Baseline+T	Baseline+T+C
Dice	0.860±0.036	0.863±0.036	**0.865±0.035**
JI	0.757±0.055	0.760±0.054	**0.764±0.053**
Recall	0.875±0.055	0.879±0.054	**0.886±0.053**
clDice	0.965±0.044	0.968±0.030	**0.971±0.020**
ASSD(mm)	0.836±0.457	0.795±0.234	**0.772±0.226**
HD_{95} (mm)	2.964±3.430	2.224±0.839	**2.098±0.707**
IR (%)	90	90	**100**

Table 3
Performances under different convolutional neural networks with transformed volume (mean ± standard deviation).

Method	Dice	JI	Recall	clDice	ASSD (mm)	HD_{95} (mm)	IR (%)
Jaskari et al.	0.800±0.136	0.684±0.157	0.779±0.157	0.900±0.148	1.855±3.540	8.439±21.951	80
DeepLabV3+	0.824±0.060	0.704±0.082	0.837±0.082	0.956±0.039	0.990±0.322	2.736±1.033	95
VNet	0.850±0.051	0.743±0.074	0.868±0.074	0.951±0.053	1.033±0.655	3.575±5.208	85
3D UNet	0.855±0.042	0.749±0.062	0.870±0.062	0.949±0.044	0.822±0.259	2.295±0.903	100
Ours	**0.865±0.035**	**0.764±0.053**	**0.886±0.053**	**0.971±0.020**	**0.772±0.226**	**2.098±0.707**	**100**

Table 4
Comparison of the segmentation results with different λ values (mean ± standard deviation).

λ	Dice	JI	Recall	clDice	ASSD (mm)	HD_{95} (mm)
0.1	0.864±0.037	0.762±0.056	0.883±0.056	0.967±0.028	0.782±0.234	2.155±0.777
0.2	0.863±0.041	0.761±0.061	0.882±0.061	0.966±0.030	0.785±0.247	2.177±0.813
0.3	0.863±0.036	0.761±0.054	0.881±0.054	0.965±0.028	0.809±0.289	2.206±1.016
0.4	0.857±0.753	0.753±0.067	0.879±0.067	0.965±0.037	0.833±0.292	2.338±1.202
0.5	**0.865±0.035**	**0.764±0.053**	**0.886±0.053**	**0.971±0.020**	**0.772±0.226**	**2.098±0.707**

methods for quantitative evaluations. The results of our method outperform other baseline networks. Even surpassing nnUNet in terms of evaluation metrics. The standard deviation values show that our method is more robust and performs well for images that are difficult to segment.

4.3. The effectiveness of the proposed transform volume method

We used the nnUNet framework as a baseline network. Quantitative evaluation metrics showed an improvement in segmentation performance for our proposed transform volume in the Frenet frame method, as shown in Table 2. Table 3 lists the results when integrating the proposed transformed volume in Frenet frame into other methods. The Dice, JI, Recall, clDice, and IR are all improved by the different segmentation methods integrated with the proposed transform volume method. These results show that using the transform sub-volumes effectively improves segmentation performance. Fig. 10 shows the Dice values of with or without transformed volumes using different networks in detail. The comparison shows that our proposed transform volume method positively affects the mandibular canal segmentation on different networks. This revelation shows that the mandibular canal segmentation network can be replaced with any up-to-date network.

4.4. Loss function analysis

The proposed mandibular canal segmentation network is combined with the clDice module. To demonstrate that the clDice module is effective for mandibular canal segmentation, we compared the segmentation metrics with or without it. Table 2 shows the results of the network with or without the clDice module. The model Dice value, JI, Recall, and clDice using the clDice module improved by 0.002, 0.004, 0.007, and 0.003, respectively. Additionally, the ASSD and HD_{95} are all reduced. The standard deviations show that the model stability has also been improved. For segmenting the mandibular canal, better connectivity is achieved with this model (Fig. 11). Table 4 demonstrates the segmentation performance for different values of λ. The results show that the best performance is achieved for $\lambda = 0.5$.

4.5. Comparison on the public dataset

We also compared the segmentation results on the publicly available dataset [30] to demonstrate the effectiveness of our proposed method. Our previous mandibular segmentation model was still effective in segmenting the mandible from the public dataset. Similarly, the locations of the mandibular foramen and mental foramen were detected according to the previous landmark localization

Table 5
Comparison of the segmentation results on the public dataset.

	Jaskari et al. [23]	Cipriano et al. [32]	Usman et al. [33]	Ours
Dice	0.67	0.75	0.77	0.81

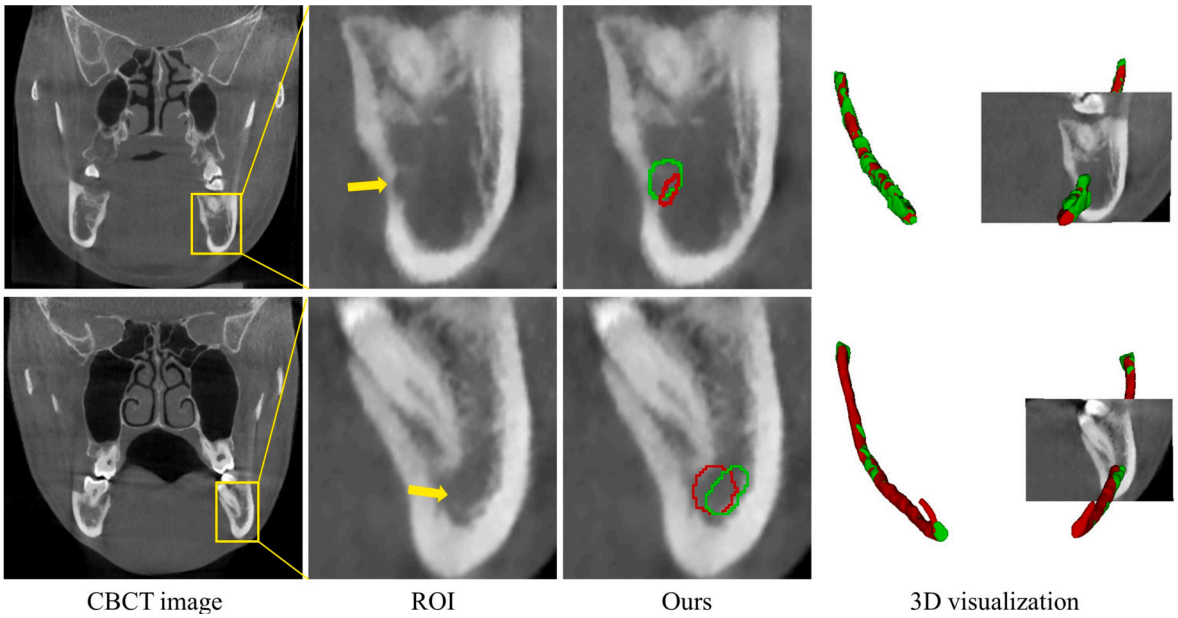


Fig. 12. Examples of poor segmentation results. The green indicates the labeling of experts and the red represents the segmentation results. The last row is the position of the shown slice in 3D.

model. We trained an independent mandibular canal segmentation model according to our proposed framework using publicly data densely annotated cases, as shown in Fig. 2. Compared with the state-of-the-art methods [23,32,33], our proposed method outperforms them in terms of the average Dice score on the test set (as in Table 5). However, since the size of the public dataset is not equal to the size of our dataset, the results in the public dataset did not perform as well as the results in our dataset.

5. Discussion

The automatic segmentation of the mandibular canal in CBCT images is of great importance for oral surgery planning and orthodontics [2–5]. It is still a challenging task for accurate mandibular canal segmentation resulting from the poor contrast between the mandibular canal and the surrounding tissue in CBCT images. Compared with the traditional segmentation techniques [6–8], the CNN-based methods [9,10,23,24] have achieved promising results of mandibular canal segmentation. However, existing deep learning-based methods use sliding windows to extract image patches for mandibular canal segmentation, which inevitably lacks the morphological integrity of the tubular structures. To overcome the problems of the poor contrast in CBCT images and the lacking integrity of the tubular structures, we propose whole mandibular canal segmentation using transformed dental CBCT volume in Frenet frame.

Mandibular segmentation is a critical step in our proposed method. Section 4.1 shows the results of mandibular segmentation with a Dice of 0.986 (±0.002), a JI of 0.973 (±0.005), and a Recall of 0.989 (±0.005). It performed equally well as the previously proposed methods [34,35]. With the superior performance of our mandibular segmentation model, it is sufficient to adopt the segmented mandible for further mandibular canal segmentation. Because the structural details of the mandible segmentation methods are not the primary concern of our study, the comparisons of mandible segmentation methods are not included in our study. Moreover, the high performance of landmark localization of the mandibular foramen and mental foramen meets the requirements for obtaining a sub-volume that contains the whole mandibular canal. Few studies have attempted to develop models for landmark localization of the mandibular foramen and mental foramen.

Extracting the whole mandibular canal, which ensures the complete 3D structural information of the mandibular canal, can improve inference efficiency. Compared with the inference time of the nnUNet [16] (about 2 min), our network inference takes approximately 2 s, and the inverse transformation to the original image size takes about 30 s. Note that our study is based on nnUNet [16] because of the excellent performance of image segmentation. The main contribution of our study is developing the transformed volume in the Frenet frame to extract the whole mandibular canals. Thus, the network structures of the mandibular canal segmentation can be replaced by any state-of-the-art image segmentation methods. Besides, Fig. 10 shows that when integrating

the proposed transformed volume in the Frenet frame into other state-of-the-art methods, these methods achieve 0.5% ~ 12.1% improvement in Dice performance. We use the IR indicators to indicate the positive effect on the maintenance of the mandibular canal topology. As shown in Table 1, all compared methods cannot ensure the integrity of the whole mandibular canal (50%~90%). When integrating the proposed transformed volume in the Frenet frame, significant improvement of IR (80%~100%) can be found in Table 3, which demonstrates that the proposed transformation volume performs well on different segmentation networks.

A limitation of our method is that in cases where the mandibular canal wall is missing or incomplete, there is also some uncertainty in the actual manually labelled mandibular canal by experts, which can also impact the mask obtained from the segmentation model. As shown in Fig. 12, even though our method ensures the integrity of the tubular structure of the mandibular canal, it is less effective in segmentation at locations where the canal wall is missing. To address this limitation, we consider developing image contrast enhancement methods to improve the boundary contrast in CBCT images. Another limitation is the small amount of data used in our method. More image data will be collected in our future study to further improve the accuracy of the model.

6. Conclusions

In this study, we propose a 3D transform volume of the mandibular canal based on the Frenet frame in CBCT images for mandibular canal segmentation. Our proposed mandibular canal segmentation framework can achieve the segmentation of mandible and mandibular canal and the detection of the mandibular foramen and mental foramen. Furthermore, the proposed transformed volume in Frenet frame can extract the whole mandibular canal to ensure the complete 3D structural information. The experimental results of this study can achieve accurate segmentation of the mandibular canal, which can help in planning oral surgery, such as oral orthodontic treatment, and dental implants.

CRediT authorship contribution statement

Huanmiao Zhao: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. **Wei Yang:** Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. **Liming Zhong:** Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper. **Junhua Chen:** Contributed reagents, materials, analysis tools or data; Wrote the paper. **Zhaoqiang Yun; Qianjin Feng:** Contributed reagents, materials, analysis tools or data.

Declaration of competing interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Data availability

We published our code at https://github.com/SMU-MedicalVision/Mandibular-canal-segmentation. The data used is presently confidential. The available public dataset is at https://ditto.ing.unimore.it/maxillo/.

Acknowledgements

This work was partially supported by grants from the National Natural Science Foundation of China (No. 82172020, No. 62101239), and Guangdong Provincial Key Laboratory of Medical Image Processing (No. 2020B1212060039).

References

[1] S.-K. Yu, M.-H. Lee, Y.H. Jeon, Y.Y. Chung, H.-J. Kim, Anatomical configuration of the inferior alveolar neurovascular bundle: a histomorphometric analysis, Surg. Radiol. Anat. 38 (2) (2016) 195–201.
[2] V. Ebenezer, G. Sumathi, et al., Importance of cone beam computed tomography in dental implants: a review, J. Positive Sch. Psychol. 6 (3) (2022) 3798–3800.
[3] C. de Oliveira-Santos, P.H.C. Souza, S. de Azambuja Berti-Couto, L. Stinkens, K. Moyaert, I.R.F. Rubira-Bullen, R. Jacobs, Assessment of variations of the mandibular canal through cone beam computed tomography, Clin. Oral Investig. 16 (2) (2012) 387–393.
[4] J. Iwanaga, M. Katafuchi, Y. Matsushita, T. Kato, K. Horner, R.S. Tubbs, Anatomy of the mandibular canal and surrounding structures: part I: morphology of the superior wall of the mandibular canal, Ann. Anat. 232 (2020) 151580.
[5] J. Iwanaga, K. Shiromoto, T. Kato, T. Tanaka, S. Ibaragi, R.S. Tubbs, Anatomy of the mandibular canal and surrounding structures. Part II: cancellous pattern of the mandible, Ann. Anat. 232 (2020) 151583.
[6] D. Kainmueller, H. Lamecker, H. Seim, M. Zinser, S. Zachow, Automatic extraction of mandibular nerve and bone from cone-beam CT data, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2009, pp. 76–83.
[7] B. Moris, L. Claesen, Y. Sun, C. Politis, Automated tracking of the mandibular canal in CBCT images using matching and multiple hypotheses methods, in: 2012 Fourth International Conference on Communications and Electronics (ICCE), IEEE, 2012, pp. 327–332.
[8] F. Abdolali, R.A. Zoroofi, Mandibular canal segmentation using 3d active appearance models and shape context registration, in: 2014 21st Iranian Conference on Biomedical Engineering (ICBME), IEEE, 2014, pp. 7–11.
[9] F. Abdolali, R.A. Zoroofi, A. Biniaz, Fully automated detection of the mandibular canal in cone beam ct images using Lie group based statistical shape models, in: 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME), IEEE, 2018, pp. 1–6.
[10] F. Abdolali, R.A. Zoroofi, M. Abdolali, F. Yokota, Y. Otake, Y. Sato, Automatic segmentation of mandibular canal in cone beam CT images using conditional statistical shape model and fast marching, Int. J. Comput. Assisted Radiol. Surg. 12 (4) (2017) 581–593.

[11] M.H. Hesamian, W. Jia, X. He, P. Kennedy, Deep learning techniques for medical image segmentation: achievements and challenges, J. Digit. Imag. 32 (4) (2019) 582–596.
[12] T. Lei, R. Wang, Y. Wan, X. Du, H. Meng, A.K. Nandi, Medical image segmentation using deep learning: a survey, 2020.
[13] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2015, pp. 234–241.
[14] F. Milletari, N. Navab, S.-A. Ahmadi V-net, Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 Fourth International Conference on 3D Vision (3DV), IEEE, 2016, pp. 565–571.
[15] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-net: learning dense volumetric segmentation from sparse annotation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19, Springer, 2016, pp. 424–432.
[16] F. Isensee, P.F. Jaeger, S.A. Kohl, J. Petersen, K.H. Maier-Hein, nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, Nat. Methods 18 (2) (2021) 203–211.
[17] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, G. Yang, Hierarchical retinal blood vessel segmentation based on feature and ensemble learning, Neurocomputing 149 (2015) 708–717.
[18] H. Fu, Y. Xu, S. Lin, D.W. Kee Wong, J. Liu, Deepvessel: retinal vessel segmentation via deep learning and conditional random field, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2016, pp. 132–139.
[19] Q. Huang, J. Sun, H. Ding, X. Wang, G. Wang, Robust liver vessel extraction using 3D U-Net with variant dice loss function, Comput. Biol. Med. 101 (2018) 153–162.
[20] C. Wang, Y. Hayashi, M. Oda, H. Itoh, T. Kitasaka, A.F. Frangi, K. Mori, Tubular structure segmentation using spatial fully connected network with radial distance loss for 3D medical images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2019, pp. 348–356.
[21] Y. Wang, X. Wei, F. Liu, J. Chen, Y. Zhou, W. Shen, E.K. Fishman, A.L. Yuille, Deep distance transform for tubular structure segmentation in CT scans, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 3833–3842.
[22] S. Shit, J.C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylka, J.P. Pluim, U. Bauer, B.H. Menze, clDice-a novel topology-preserving loss function for tubular structure segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 16560–16569.
[23] J. Jaskari, J. Sahlsten, J. Jarnstedt, H. Mehtonen, K. Karhu, O. Sundqvist, A. Hietanen, V. Varjonen, V. Mattila, K. Kaski, Deep learning method for mandibular canal segmentation in dental cone beam computed tomography volumes, Sci. Rep. 10 (1) (2020) 1–8.
[24] G.H. Kwak, E.-J. Kwak, J.M. Song, H.R. Park, Y.-H. Jung, B.-H. Cho, P. Hui, J.J. Hwang, Automatic mandibular canal detection using a deep convolutional neural network, Sci. Rep. 10 (1) (2020) 1–8.
[25] P. Lahoud, S. Diels, L. Niclaes, S. Van Aelst, H. Willems, A. Van Gerven, M. Quirynen, R. Jacobs, Development and validation of a novel artificial intelligence driven tool for accurate mandibular canal segmentation on CBCT, J. Dent. 116 (2022) 103891.
[26] R.L. Bishop, There is more than one way to frame a curve, Am. Math. Mon. 82 (3) (1975) 246–251.
[27] J. Bloomenthal, Calculation of reference frames along a space curve, Graph. Gems 1 (1990) 567–571.
[28] W. Wang, B. Juttler, D. Zheng, Y. Liu, Computation of rotation minimizing frames, ACM Trans. Graph. 27 (1) (2008) 1–18.
[29] T. Pfister, J. Charles, A. Zisserman, Flowing convnets for human pose estimation in videos, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1913–1921.
[30] M. Cipriano, S. Allegretti, F. Bolelli, M. Di Bartolomeo, F. Pollastri, A. Pellacani, P. Minafra, A. Anesi, C. Grana, Deep segmentation of the mandibular canal: a new 3D annotated dataset of CBCT volumes, IEEE Access 10 (2022) 11500–11510.
[31] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 801–818.
[32] M. Cipriano, S. Allegretti, F. Bolelli, F. Pollastri, C. Grana, Improving segmentation of the inferior alveolar nerve through deep label propagation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 21137–21146.
[33] M. Usman, A. Rehman, A.M. Saleem, R. Jawaid, S.-S. Byon, S.-H. Kim, B.-D. Lee, M.-S. Heo, Y.-G. Shin, Dual-stage deeply supervised attention-based convolutional neural networks for mandibular canal segmentation in CBCT scans, Sensors 22 (24) (2022) 9877.
[34] P.-J. Verhelst, A. Smolders, T. Beznik, J. Meewis, A. Vandemeulebroucke, E. Shaheen, A. Van Gerven, H. Willems, C. Politis, R. Jacobs, Layered deep learning for automatic mandibular segmentation in cone-beam computed tomography, J. Dent. 114 (2021) 103786.
[35] Y. Fan, R. Beare, H. Matthews, P. Schneider, N. Kilpatrick, J. Clement, P. Claes, A. Penington, C. Adamson, Marker-based watershed transform method for fully automatic mandibular segmentation from CBCT images, Dentomaxillofacial Radiol. 48 (2) (2019) 20180261.