Title:
A web-based patient management system in teeth segmentation using deep learning

Authors:
Y. K. Lin, robertlin@nfu.edu.tw, Department of Information Management, National Formosa University, Yunlin, Taiwan
Y. H. Su, syh91006@hotmail.com, Division of Endodontics and Operative Dentistry, Dental Department, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
H. J. Hsu, D80.wu63@msa.hinet.net, Department of Oral Hygiene, College of Dental Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

Keywords:
Web-based, Deep learning, Multi-layer perceptron, Curvature, intelligent healthcare

DOI: 10.14733/cadconfP.2023.189-194

Introduction:
The World Health Organization (WHO) considers malocclusion to be one of the most significant oral health issues after caries and periodontal disease. Its prevalence varies greatly and is estimated to be between 39% and 93% among children and adolescents [1]. It is clear that malocclusion is not only a common issue, but it is also rapidly getting worse in adolescents. Computer-aided design (CAD) is widely adopted by orthodontists for a preoperative planning and postoperative evaluation in the clinical practice of dentistry. Additionally, an ideal treatment depends significantly on a complete and accurate tooth model. However, it is really challenging for the segmentation of the teeth with morphological diversity from optical scanned models due to a large number of manual boundary selections [2-5]. Therefore, the teeth segmentation is a time-consuming task primarily on the user's experience. Moreover, the use of CAD in medical applications requires highly specialized software and hardware. Multiple collaborators, remote collaboration, and interaction between them both synchronously and asynchronously are consequently challenging. Hence, computer-supported collaborative work (CSCW) was proposed as a solution to overcome above issues by Greif and Cashman [6]. With the advancement of Internet technologies, the web-based system plays an important role in CSCW, particularly in the era of COVID-19. It enables users to share knowledge, collaborate on treatment planning, and communicate in a secure environment. Furthermore, users are not frustrated by incompatibilities between different systems or by the inability of applications to support multiple users in different settings. As a consequence, the creation of a cloud system based on deep learning is urgently required in order to improve the accuracy and efficiency of teeth segmentation, reduce the workload of users in practice, and achieve collaborative work.

Some efforts have been devoted to the development of teeth segmentation, which can be roughly categorized into projection-based [7-8] and geometry-based methods [9-11]. In practice, these techniques are typically semi-automatic and reply on user experience. Furthermore, it is sensitive to variations in tooth morphology for their performance. Some deep learning techniques have recently been proposed to enhance the traditional geometry-based methods. For instance, an end-to-end deep network [12-13] is proposed to learn hierarchically multi-scale contextual features and holistic features of the 3D dental surface for labelling. Cui et al. [14] proposes a two-stage algorithm for tooth segmentation, all teeth are detected in the first stage and each tooth is separated in the second stage. A two-stream graph convolutional network (TSGCN) [15] is proposed to learn discriminative geometric features from heterogeneous multi-view inputs for end-to-end tooth segmentation. To train a deep network for the above works, however, curvature information is lacking. Hence, the proposed method integrates the curvature information to enhance the previous works.

Main Idea

An overview of the proposed system is illustrated in Fig. 1. The first step is to collect a studied dataset from desktop scanner or intra-oral scanner from the Department of Dentistry at Kaohsiung Medical University. Data preprocessing, which includes mesh simplification, annotation, and augmentation is the second step. The third step is to calculate an 18-D input vector describing each cell on the dental surface. The 18-D input vector includes the coordinates of cell vertices (9 elements), the normal vector (3 elements), the relative position (3 elements) and the curvatures (3 elements). The 18-D input vector is then to be the input of the deep learning. At the end, the teeth segmentation can be obtained and rendered in the proposed web-based system, as shown in Fig. 1. More details are described in the following sections.

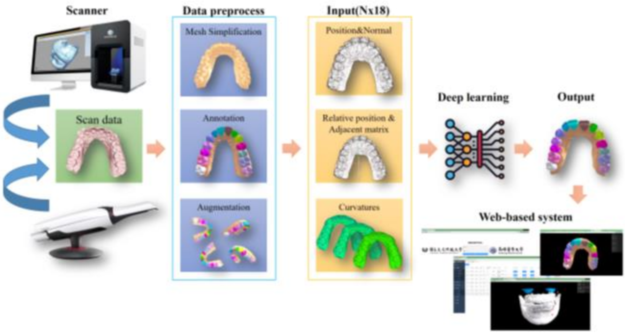


Fig. 1: Overall flowchart of the proposed framework.

A. Data Preprocessing

The studied dataset consists of 85 scanned models. Each raw model contains approximately over 100,000 mesh cells, which were down-sampled to 8,000 mesh cells while preserving the original topology. As the input of the proposed method, the network input is an N x 18 matrix (e.g., N=8,000). On the down-sampled surface, the ground-truth annotations of the entire studied dataset were defined according to the experienced dentists' advice. In practice, the entire studied dataset is divided into two parts: 70 of the datasets are selected for training, and the remaining for validation. Furthermore, we also augment the training dataset with random translation between -20 and 20, random rescaling between 0.7 and 1.3, and random rotation along the x and y axis between $-\frac{\pi}{4}$ and $\frac{\pi}{4}$, and along z axis between $-\pi$ and π.

B. Deep learning

In this paper, the proposed network is an extension of MeshSegNet. Firstly, the coordinates of three adjacent cell vertices, the normal, curvatures and relative positions of cells with respective to the whole surface are adopted as the primary network input. Secondly, multi-scale graph-constrained learning modules are used to hierarchically model multi-scale local geometric contexts on surfaces. Thirdly, a dense fusion strategy is done for dense skip connections combing local-to-global features. Finally, crown segmentation is predicted an end-to-end prediction.

Proceedings of CAD'23, Mexico City, Mexico, July 10-12, 2023, 189-194

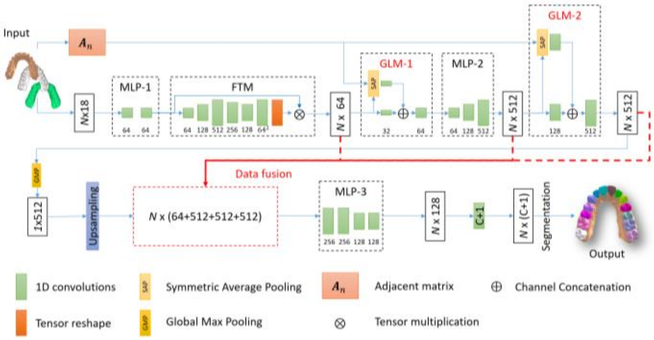


Fig. 2: Proposed deep learning network.

The input raw surface is a $N \times 18$ matrix $\mathbf{F}^0$, where N denotes the number of mesh cells. After the deep learning procedure, $\mathbf{P}$ is obtained as an $N \times (C+1)$ matrix with each row denoting the probabilities of the respective cell belonging to different classes. The multi-layer perceptrons (MLPs) is adopted to extract increasingly higher-level geometric features for each mesh cell. In this paper, MLP-1 includes two 1D Convs with 64 channels. For MLP-2, three 1D Convs with 64, 128, and 512 channels are set respectively. MLP-3 has four 1D Convs with 256, 256, 128, and 128 channels, respectively. Besides, a feature-transformer module (FTM) is used to predict a transformation matrix and align the initial features for different cells onto a canonical feature space. Hence, it could improve the learning performance in subsequent layers [16]. Inspired by [13], two graph-constrained learning modules (GLMs) are employed to explicitly model the local geometric context on the input surface data. To construct GLM-1 and GLM-2, each cell is regarded as the centroid to define a 3D ball with a specific radius. The cells within the respective ball are selected as the neighbors for each centroid. A $N \times N$ adjacent matrix is then constructed for the input mesh data. The radius in this paper is set 0.15 empirically. In comparison to MeshSegNet, the proposed network eliminates the additional 3D ball to reduce the calculation time. Based on A_N, GLM-1 first applies the *symmetric average pooling* (SAP), a graph-based fusion operation, to propagate the local contextual information onto each centroid cell. Assume $\tilde{A}_N = A_N + \mathrm{I}$ is the adjacency with self-loops, and $\widetilde{D}_S^{-\frac{1}{2}} \tilde{A}_N \widetilde{D}_S^{-\frac{1}{2}}$ is the corresponding symmetric-normalized adjacency, where $\widetilde{D}_S$ denotes the diagonal degree matrix. The SAP operation in terms of $\hat{F}^1$ given A_N is defined as

$$\begin{cases} \tilde{F}^n = (\widetilde{D}_S^{-\frac{1}{2}} \tilde{A}_N \widetilde{D}_S^{-\frac{1}{2}}) \hat{F}^n \\ \qquad n = 1, 2 \end{cases} \tag{1}$$

where $\tilde{F}^1$ is the updated cell-wise feature matrix encoding local geometric contexts. After SAP, $\tilde{F}^1$ and $\hat{F}^1$ are further squeezed by the 1D Convs with 32 channels. Another 1D Conv with 64 channels is then used for fusion with the resulting feature matrices. On the other hands, a larger receptive field is adopted to learn multi-scale contextual features in GLM-2. Hence, the MLP-2 output is a N x 512 matrix $\hat{F}^2$ and a $\tilde{F}^2$ is obtained from the second SAP operation. Both $\hat{F}^2$ and $\tilde{F}^2$ are then squeezed by the 1D Convs with 128 channels, which are finally concatenated across channels and fused by another 1D Conv with 512 channels.

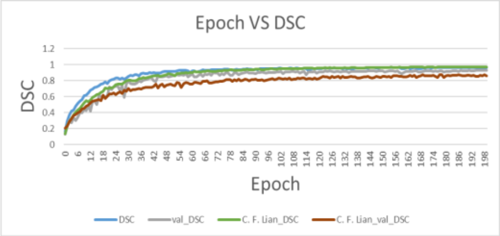
A global max pooling (GMP) is applied to the GLM-2 output for generating the translation-invariant holistic features. These features encode the semantic information of the whole input surface. Furthermore, the local-to-global features from FTM, GLM-1, GLM-2 and upsampling GMP are concatenated densely. Then, a MLP-3 is used to generate a N x 128 feature matrix. Finally, a $N \times (C+1)$ probability matrix $\mathbf{P}$ is predicted by the MLP-3 output and a 1D Conv layer with softmax activation. The batch normalization (BN) and ReLU activation are adopted in all 1D Convs of these MLPs and GLMs.

C. Web-based system

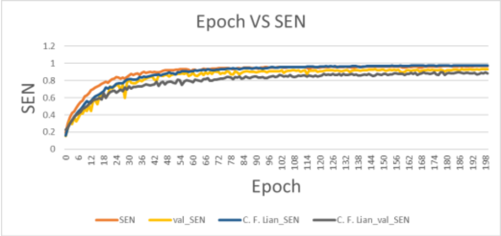
In this paper, a web-based system is developed for the telemedicine, collaborative treatment, and follow-up in different regions or countries. Django serves as the backend, Vue.js and three.js serve as the frontend, and MySQL serves as the database. All models could be uploaded and visualized in the system.

D. Implementation

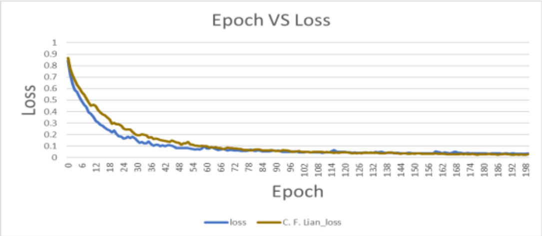
The proposed deep network was implemented using Python and PyTorch. It is trained by minimizing the generalized Dice loss with the AMSGrad variant of the Adam optimizer (mini-batch size:10; number of epochs: 200). All experiments are performed on an Intel i7 processor with 16 GB of RAM and an 8 GB GeForce RTX 3060Ti GPU. In this paper, we compare the proposed technique with MeshSegNet[13] using the same experimental setup, loss function and optimizer. A 3-fold cross-validation on the studied dataset is carried out. Based on the ground-truth annotations, the outcomes of the crown segmentation are evaluated by Dice similarity coefficient (DSC), sensitivity (SEN) and training loss. As depicted by the convergence curve in Fig. 3. As can be seen, the proposed network's training and validation times in DSC coverage and SEN coverage are faster than MeshSegNet. In addition, , the proposed network could reach 0.1 loss in 36th epoch and the original MeshSegNet reach 0.1 loss in 60th epoch. According to the observation, the loss decreasing in the proposed network is 1.67 times as fast as the one in the original method. We will add these in our final version.



(a)



(b)

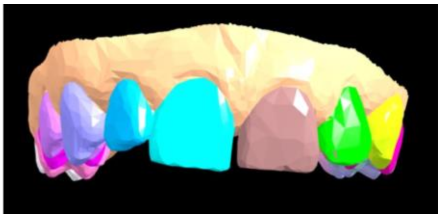


(c)

Fig. 3: The training and validation loss of the proposed method: (a) DSC relative to number of training epochs, (b) SEN relative to number of training epochs, (c) Total loss relative to number of training epochs.

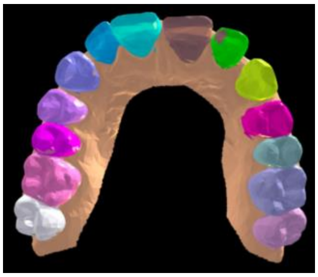
Finally, the crown segmentation from the optical scanned models can be obtained by the training model and visualized on the proposed web-based system, as shown in Fig. 4.

(a)



(b)

(c)



(d)

Fig. 4: Visualization of representative segmentation results: (a) front view, (b) front view with segmentation, (c) top view, (d) top view with segmentation.

Conclusions

This paper aims to build a deep learning network for automatically segmenting crowns and to visualize the results on a web-based system. The proposed system can efficiently assist dentists in creating a complete and accurate teeth model for preoperative planning and postoperative evaluation. In conclusion, the advantages of the proposed system presented in this paper are as follows:

- Due to a large number of manual boundary selections required by the conventional approach of crown segmentation, it takes a lot of time. In the proposed system, a new structure of deep learning network is designed for automatic crown segmentation. Dentists only need to make minor adjustments for annotation refinement.
- Gaussian, maximum and minimum curvatures are critical characteristics at every point on the surface. Thereby, the features are added to the proposed network to significantly boost performance.
- A web-based system is developed to have multiple collaborators, remote collaboration, and interaction with each other synchronously and asynchronously.

References

[1] Cenzzto N.; Nobili A.; Maspero C.: Prevalence of Dental Malocclusions in Different Geographical Areas: Scoping Review, Dentistry Journal, 9(117), 2021, 1-10. https://doi.org/10.3390/dj910011

[2] 3shape. https://www.3shape.com/en

[3] exoCAD. https://exocad.com/

[4] Maestro3D. https://www.maestro3d.com/

[5] INTEWARE. https://www.inteware.com.tw/zh/%e9%a6%96%e9%a0%81/

[6] Greif I.: How we started CSCW, Nature Electronics, 2(314), 2019. https://doi.org/10.1038/s41928-019-0229-y

[7] Kondo T.; Ong S. H.; Foong K. W. C.: Tooth segmentation of dental study models using range images, IEEE Transactions on Medical Imaging, 23(3), 2004, 350-362. https://doi.org/10.1109/tmi.2004.824235

[8] Wongwaen N.; Sinthanayothin C.: Computerized algorithm for 3D teeth segmentation, International Conference on Electronics and Information Engineering, 2010, 277-280. https://doi.org/10.1109/ICEIE.2010.5559877

[9] Li Z.; Ning X.; Wang Z.: A fast segmentation method for STL teeth model, IEEE/ICME International Conference on Complex Medical Engineering, 2007, 163-166. https://doi.org/10.1109/ICCME.2007.4381713

[10] Yuan T.; Liao W.; Dai N.; Cheng X.; Yu Q.: Single-tooth modeling for 3D dental model, Journal of Biomedical Imaging, 2010(9), 2010, 1-14. https://doi.org/10.1155%2F2010%2F535329

[11] Zou B. J.; Liu S. J.; Liao S. H.; Ding X.; Liang Y.: Interactive tooth partition of dental mesh base on tooth-target harmonic field, Computers in Biology and Medicine, 56(C), 2015, 132-144. https://doi.org/10.1016/j.compbiomed.2014.10.013

[12] Lian C. F.; Wang L.; Wu T. H.; Liu M. X.; Duran F.; Ko C. C.; Shen D. G.: MeshSNet: Deep Multi-scale Mesh Feature Learning for End-to-End Tooth Labeling on 3D Dental Surfaces, Medical Image Computing and Computer Assisted Intervention, 11769, 2019, 837-845. https://doi.org/10.1007/978-3-030-32226-7_93

[13] Lian C. F.; Wang L.; Wu T. H.; Wang F.; Yap P. T.: Deep Multi-Scale Mesh Feature Learning for Automated Labeling of Raw Dental Surfaces From 3D Intraoral Scanners, IEEE Transactions on Medical Imaging, 39(7), 2020, 2440-2450. http://dx.doi.org/10.1109/TMI.2020.2971730

[14] Cui Z.; Li C.; Chen N.; Wei G.; Chen R.; Zhou Y.; Wang W.: TSegNet: An efficient and accurate tooth segmentation network on 3D dental model, Medical Image Analysis, 69, 2021, 101949. https://doi.org/10.1016/j.media.2020.101949

[15] Zhang L.; Zhao Y.; Meng D.; Cui Z.; Gao C.; Gao X.; Lian C.; Shen, D.: TSGCNet: Discriminative Geometric Feature Learning with Two-Stream Graph Convolutional Network for 3D Dental Model Segmentation, Computer Vision and Pattern Recognition Conference, 2021, 6699–6708. https://doi.org/10.48550/arXiv.2012.13697

[16] Qi C. R.; Su H.; Mo K.; Guibas L. J.: Pointnet: Deep learning on point sets for 3d classification and segmentation, Computer Vision and Pattern Recognition Conference, 2017, 77-85. https://doi.org/10.48550/arXiv.1612.00593