

Mandible segmentation for designing patient-specific implants using a deep learning approach

Chou, Markus Chi-Haw
Master Thesis, Biomedical Engineering

FHNW Supervisor: Prof. Dr. Erik Schkommodau
Expert: Stephan Böhringer, mininavident AG
Company Supervisor: Florian Coigny, medartis AG

INTRODUCTION

Patient-specific implants (PSIs) and services offer significant advantages, such as better implant fit and shorter surgery time [1]. As a pre-requisite, the 3D model of the patient's bone structure must be generated from medical imaging data. Since manual bone segmentation is time-consuming and suffers from inter- and intra-rater variability, automatic bone segmentation is paramount. This work investigates mandible segmentation from computed tomography (CT) images for designing PSIs based on a deep learning approach. The CT data were sourced from over 16 hospitals and vary in scan quality, artifacts, body parts scanned, teeth loss, deformed structures due to disease, and traces from previous treatment. The heterogeneity of the CT data distinguishes this work from other literature.

CONCEPT

The 3D U-Net by Çiçek et al. [2], commonly used for medical image segmentation, was modified for mandible segmentation from CT data (s. Fig. 1). Due to hardware limitations, only half the features are extracted. The last layer was changed to one channel, and a sigmoid activation function was added for one binary output. Furthermore, the convolution filter kernels were dilated by five, thereby increasing the field of vision within slices and giving more spatial context.

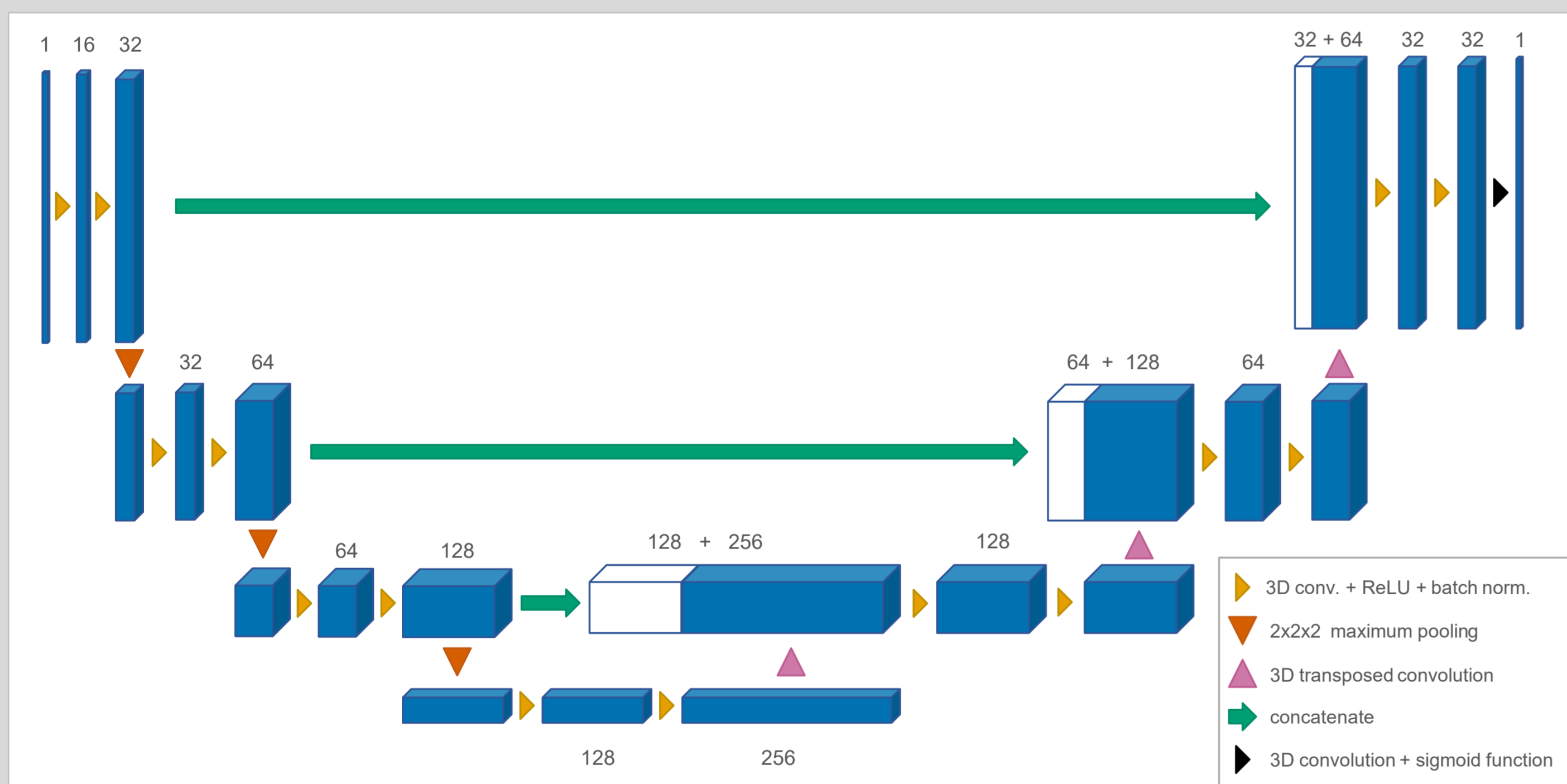


Fig. 1: Modified 3D U-Net architecture, adapted from [2]. Numbers indicate number of features extracted.

Considering that scans vary in the number of slices and cannot be segmented whole, the CT scans were split into stacks of eight slices, taken chronologically and without overlap. The network was trained for 25 epochs only on slices with mandibles with increased weight on higher slices that contained condylar and coronoid processes. The network was then trained on mandible slices for another ten epochs without weight to even out the training. Lastly, the network trained on all slices for five more epochs to learn to ignore other bones. The model was trained on an NVIDIA RTX 3060 with a batch size of one, Adam optimizer, and binary cross-entropy loss.

For evaluating the segmentations, the Dice score is used. Because the surface is the most important for designing PSIs, two new surface accuracy metrics were introduced. The Edge-score specifies how well the surfaces of the manual segmentations done by experts and the automatic segmentations from the network align. The Edge-within-margin score indicates if the surface of the prediction deviates at most half a millimeter from the manual segmentation. Higher deviations mean the PSIs will not fit. The ten test scans were evaluated after post-processing the segmentation and removing small, unattached objects. The 3D visualization was generated with Marching Cubes, Taubin filtering, and mesh decimation.

RESULTS

The training progress, measured on a validation dataset, is depicted in Fig. 2 for the training on mandible slices. The first value of the Edge-within-margin score is considered an outlier. The test segmentation results are listed in Table 1 and exemplary segmentations presented in Fig. 3.

Table 1: Test accuracy scores after 40 epochs of training. Metrics over 0.90 are marked bold.

Case ID	T-0001	T-0002	T-0003	T-0004	T-0005	T-0006	T-0007	T-0008	T-0009	T-0010
Dice score	0.8619	0.9265	0.7396	0.9384	0.7622	0.7622	0.8357	0.8165	0.8083	0.9417
Edge score	0.6919	0.8652	0.6672	0.8341	0.8549	0.5964	0.8814	0.7794	0.6480	0.8687
Edge-within-margin score	0.8535	0.9287	0.7779	0.8875	0.9528	0.7167	0.9718	0.9496	0.7828	0.9682

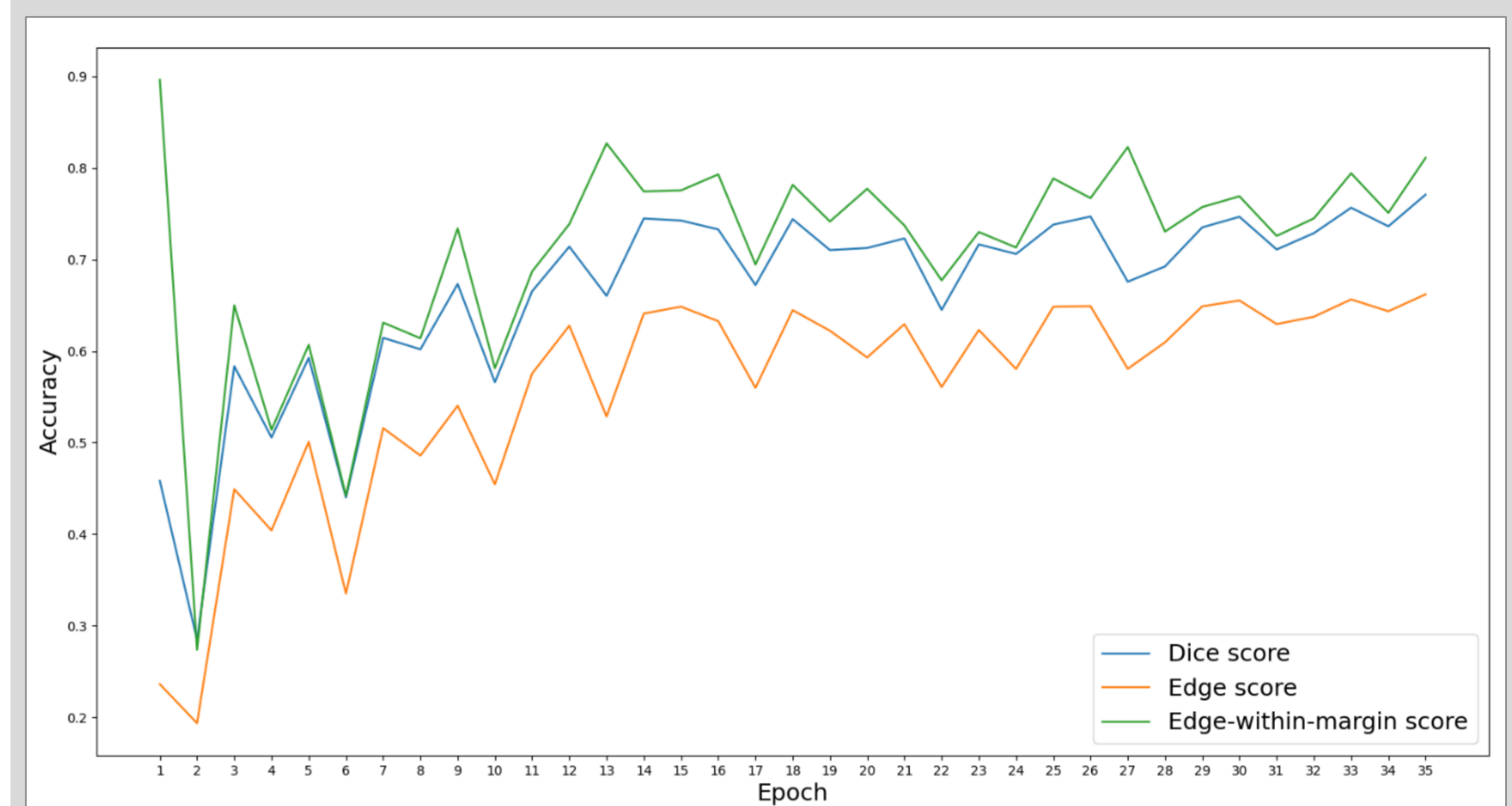


Fig. 2: Validation accuracy scores over 35 epochs of training on only mandible slices

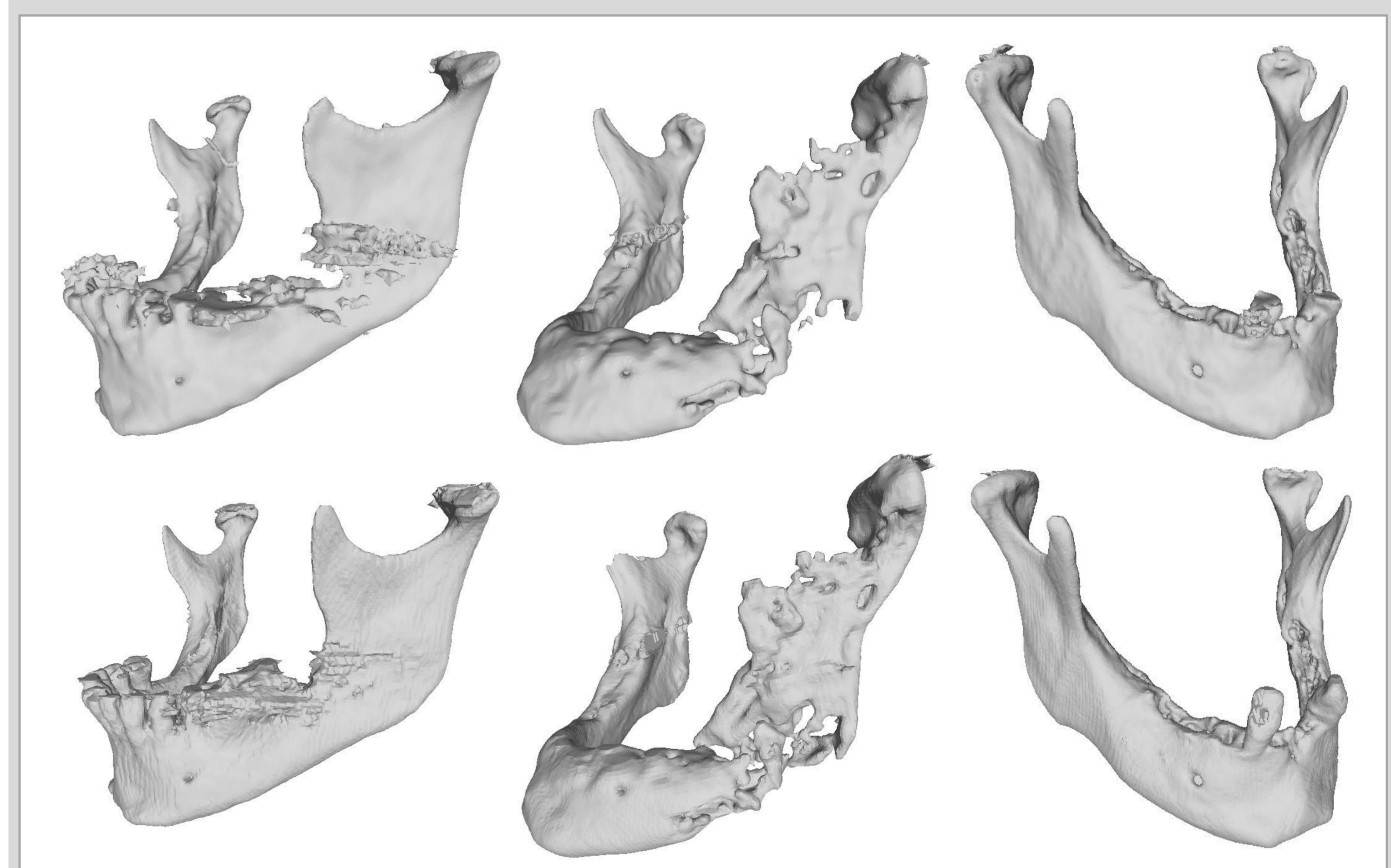


Fig. 3: Comparison of test cases T-0002, T-0005 and T-0010 (from left to right). Top: automatic segmentation from the modified 3D U-Net. Bottom: manual segmentation by experts.

When artifacts are present, the attempted removal of them leads to mis-segmentation of the bone, as shown in case T-0002. Furthermore, condylar and coronoid processes may be segmented inaccurately. However, with minimal artifacts and no implants present, e.g., case T-0010, the automatic segmentation is close to the manual segmentation, therefore allowing designing PSIs.

CONCLUSION

The modified 3D U-Net produces suitable segmentations for designing PSIs. With more training data in the future, especially scans with implants and artifacts, continued training is expected to yield even better segmentation results.

REFERENCES

- [1] P. Dérand, L.-E. Rännar, and J.-M. Hirsch, 'Imaging, Virtual Planning, Design, and Production of Patient-Specific Implants and Clinical Validation in Craniomaxillofacial Surgery', *Craniomaxillofacial Trauma Reconstr.*, vol. 5, no. 3, pp. 137–143, Sep. 2012, doi: 10.1055/s-0032-1313357.
- [2] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, '3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation', in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, vol. 9901, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds. Cham: Springer International Publishing, 2016, pp. 424–432. doi: 10.1007/978-3-319-46723-8_49.