Automatic Anatomical Annotation of CBCT Scans for Maxillofacial Prosthetics

Bertram Sabrowsky-Hirsch, Stefan Thumfart
RISC Software GmbH, Hagenberg
{first.lastname}@risc-software.at

Thomas Potrusil, Claudia Wittner, Mario Moser
CADS GmbH, Perg
{first.lastname}@cads.at

Josef Scharinger
Johannes Kepler University, Linz
josef.scharinger@jku.at

Abstract

The planning and evaluation of jaw surgery greatly benefit from the availability of accurate 3D models of critical anatomical structures. Automating the creation of these models is difficult, primarily due to the unique challenges introduced by irregularities and traumatic injuries. Existing methods are limited by their focus on specific anatomical structures and have not undergone evaluation in the context of jaw surgery patients. To this end, we have acquired a novel dataset comprising 255 patients who underwent maxillofacial surgery and annotated eight anatomical structures important to surgical planning. We present preliminary results from our annotation method, which leverages a state-of-the-art medical segmentation framework.

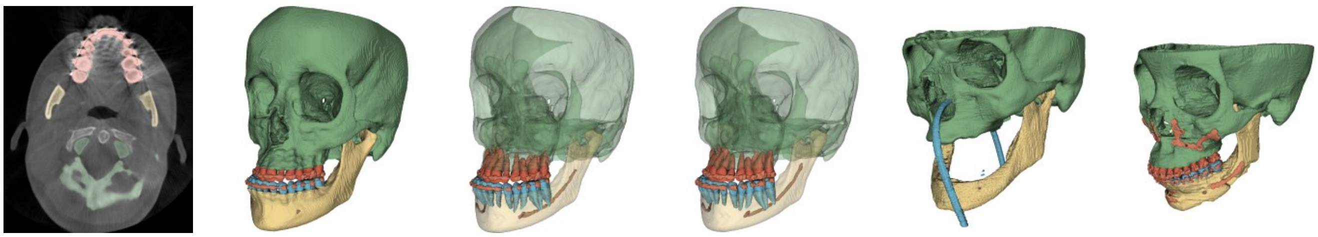
1. Introduction

Maxillofacial surgery aims to correct congenital irregularities of the jaw bones and teeth or restore their aesthetics and proper function after a traumatic injury. Among the corrective interventions, bilateral sagittal split osteotomy (BSSO) is the most commonly performed jaw surgery [1]. The interventions are planned based on the patients' anatomy. Most relevant in this process are the mandibular and maxillary bone and teeth to locate the optimal split positions for aligning the dental arches and designing prostheses. In order to avoid complications, the outline of dental root canals and the inferior alveolar nerve passing through the mandibular canal are crucial anatomical structures to consider. Typically, orthodontic treatment has been performed prior to surgery, which introduces dental implants, braces, and fillings as relevant surrounding structures. The medical experts rely on medical imaging for surgery planning and evaluation, with cone-beam computed tomography (CBCT) as the most commonly used imaging modality for maxillofacial surgery [2]. 3D models of the above-mentioned structures are used in computer-guided treatment planning and the design of prostheses. Medical modelling experts semi-automatically annotate these models based on medical images in a time-consuming process, often exceeding four hours per patient depending on the modelled structures. The annotation process is complex due to strong imaging artifacts caused by metal structures, less severe artifacts inherent to the CBCT acquisition process and the low contrast for the inferior alveolar nerve and dental root canals. The absence of an annotation standard necessitates an ongoing feedback loop between medical experts and modeling specialists. While automatic methods for CBCT annotation exist, e.g. [3], these methods focus on orthodontic treatment planning, therefore lacking a thorough evaluation on cases of severe trauma and disregarding structures critical to maxillofacial surgery. Our goal is to establish a fully automated pipeline that provides the medical experts with immediate means to extract the required anatomical models for maxillofacial surgery planning and evaluation. The first step in achieving this is the acquisition and annotation of a representative dataset, for which we present preliminary results on automatic annotation.

2. Methods and Preliminary Results

We acquired a novel multicenter dataset of 713 clinical scans of maxillofacial surgery patients, of which we identified 255 CBCT scans for our cohort. In terms of demographics, we observed 48% female and 47% male patients (5% categorized as other), from 13 countries (48% US), a median age of 22 years (Interquartile range of 16 years), of which 87% are orthognathic and 13% jaw reconstruction surgery patients.

Annotations were performed manually by a third-party professional medical annotation provider. The following eight structures were annotated individually: the mandibular bone (MDB); the remaining visceral skeleton and neurocranium (VSN); the maxillary teeth including dental root canals (MXT); the mandibular teeth including dental root canals (MDT); the mandibular canal (MC); dental fillings; dental braces; implants including existing prostheses. The latter three structures have been combined into a single label for metallic structures (MS) in our experimental setup. Where structures were deemed too ambiguous by the annotators, they were omitted, primarily affecting the



(a) (b) (c) (d) (e) (f)

Figure 1: Visualizations of annotated scans. A female orthognathic surgery patient is shown in a sagittal slice (a), the annotated 3D model (b) with transparent bone to show the dental root canals and the inferior alveolar nerve (c), and the respective prediction by the trained model (d). Models of a male jaw reconstruction surgery patient (e) and a female post-operative orthognathic surgery patient (f).

mandibular canal in low-contrast scans. Figure 1 shows a number of exemplary annotations.

For automatic annotation, we used the *nnU-Net* framework [4], which has been recognized as a strong baseline for medical image segmentation through its state-of-the-art performance in multiple medical segmentation challenges.

Table 1 presents the preliminary results for the experimental dataset, using the *3d_fullres* model set to a uniform spacing of 0.4mm and a patch-size of 112x160x128. The models were trained and evaluated in a 5-fold cross-validation, 1000 epochs each, on an NVidia 4090 RTX (24 GB). Folds have been stratified by patient sex and indication. We evaluated the following commonly used medical image segmentation metrics: Dice Similarity Score (DSC), the average Hausdorff distance (AHD) and its 95^{th} percentile (HD95). Mean (M), median (MD) and interquartile range (IQR) are shown, respectively.

Table 1: Preliminary evaluation results.

	DSC			**AHD (mm)**			**HD95 (mm)**		
	M	**MD**	**IQR**	**M**	**MD**	**IQR**	**M**	**MD**	**IQR**
MDB	0.96	0.96	0.02	0.58	0.52	0.20	0.03	0.02	0.01
VSN	0.92	0.93	0.02	1.44	0.89	0.44	0.15	0.06	0.04
MXT	0.91	0.92	0.03	0.83	0.60	0.35	0.07	0.03	0.02
MDT	0.91	0.92	0.04	0.77	0.57	0.38	0.07	0.04	0.02
MC	0.63	0.65	0.17	3.53	2.00	2.83	0.63	0.34	0.47
MS	0.70	0.73	0.15	6.51	1.50	4.34	1.41	0.28	0.58

3. Discussion and Conclusion

The preliminary results for our dataset are promising, especially for bone (MDB and VSN) and dental (MXT, MDT) structures, while metallic structures (MS) are of reasonable quality. However, the annotation of the mandibular canal (MC) still has potential for improvement. The latter two structures (MS and MC) are notably influenced by outliers, also indicated by a high IQR. This is due to sparse ground-truth data in areas where the annotators considered the structures too ambiguous and therefore omitted them.

Key takeaways and insights gained from our project were the advantages of efficiently assessing large datasets of images and annotations through visualizations such as max-intensity projection images and 3D renderings; the importance of maintaining a close and responsive feedback cycle with annotators and the potential of integrating early-stage trained models in the annotation process as an additional tool for identifying mistakes in the ground-truth annotations.

Our next steps are to gather feedback from maxillofacial surgeons on the quality of annotations and outline requirements for clinical practicability and use. Furthermore, we plan to evaluate our model on publicly available CBCT dataset such as provided by Cui *et al.* [3], and vice-versa evaluate available models on our dataset in order to provide a thorough evaluation and identify potential improvements over the baseline *nnU-Net* framework. One option could be introducing the method proposed by Usman *et al.* [5] to improve the segmentation of the mandibular canal.

Ultimately, our goal is to provide a functional pipeline and tool to our medical partners in the field of maxillofacial prosthetics to improve their clinical routines.

References

[1] Monson, Laura A. “Bilateral Sagittal Split Osteotomy.” Seminars in Plastic Surgery 27, no. 3 (August 2013): 145–48. https://doi.org/10.1055/s-0033-1357111.

[2] Andriola, Fernando de Oliveira, et al. “Computed Tomography Imaging Superimposition Protocols to Assess Outcomes in Orthognathic Surgery: A Systematic Review with Comprehensive Recommendations.” Dentomaxillofacial Radiology 51, no. 3 (March 2022): 20210340. https://doi.org/10.1259/dmfr.20210340.

[3] Cui, Zhiming, et al. “A Fully Automatic AI System for Tooth and Alveolar Bone Segmentation from Cone-Beam CT Images.” Nature Communications 13, no. 1 (April 19, 2022): 2096. https://doi.org/10.1038/s41467-022-29637-2.

[4] Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods, 18(2), 203-211. https://doi.org/10.1038/s41592-020-01008-z

[5] Usman, Muhammad, et al. “Dual-Stage Deeply Supervised Attention-Based Convolutional Neural Networks for Mandibular Canal Segmentation in CBCT Scans.” Sensors 22, no. 24 (January 2022): 9877. https://doi.org/10.3390/s22249877.