BABEȘ BOLYAI UNIVERSITY, CLUJ NAPOCA, ROMÂNIA

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

**Medical Imaging for Dentistry Using Artificial Intelligence**

Jurjiu-Tandrau Darius-Stefan

2023-2024

**Abstract**

Medical imaging in general and dental imaging in particular are fields where Artificial Intelligence can be applied in order to provide quicker and more accurate results. Providing patients with better care and giving practitioners an opportunity to direct their focus where it is needed most are just two of the objectives such an approach strives towards.

This project uses deep learning techniques for segmenting images and providing dentists with a baseline diagnostic which they can perfect with their expertise, as well as automating screenings. It aims to provide mandible segmentations, as well as automatic teeth measurements.

**Contents**

**1 Introduction**

1.1 What? Why? How?

1.2 Paper structure and original contribution(s)

**2 Scientifc Problem**

2.1 Problem definition

**3 State of the art/Related work**

**4 Investigated approach**

**5 Application**

**6 Conclusion and future work**

**1. Introduction**

**1.1**

The problem being tackled here is reliable and fast segmentation of dental images. Progress in this area can lead to an increased quality of life for countless people by providing automated screenings, as well as assistance to practitioners.

The approach is based on deep learning, which has shown significant improvements compared to other methods in the past years. Particularly, this project aims to tweak some existing approaches for the given problem: (Kamnitsas, et al., 2017), (Qin, et al., 2020), (Weber, et al., 2021).

**1.2**

The research presented in this paper advances the theory, design, and implementation of several particular models. The main contribution of this report is to present an intelligent algorithm for solving the problem of dental imaging. The second contribution of this report consists of building an application where medical practictioners as well as patients can have seamless access to these imaging techniques. The application is also meant to encourage the usage of the developed models by integrating them in an intuitive and simple interface meant to digitalize the treatment process.

The present work is structured into five chapters as follows:

The first chapter is introductory, presenting general information and motivation.

The second chapter formulates the scientific problem.

The third chapter presents the current state of the art and work related to this topic.

The fourth chapter presents the particular approaches being investigated in this work.

The fifth chapter presents the application that integrates the developed intelligent algorithms.

**2. Scientific Problem**

**2.1**

**Problem definition**

The problem this project is aiming to solve is that of accurate and efficient segmentation of dental images. The aim is provide doctors with aid in separating the mandible from the skull, as well as spotting dental assymetries through precise pixel-wise measurements.

The success of deep learning methods in the past years is the motivation for mainly focusing on such approaches in the present work.

As stated above, a reliable algorithm solving this problem can provide significant improvements to people’s lives’, by improving the reliability, as well as accessibility of diagnostics.

**3. State of the art / Related Work**

Previous work in this area can be broadly considered to be anything regarding image segmentation. DeepMedic (Kamnitsas, et al., 2017) uses Convolutional Neural Networks and Conditional Random Fields to segment brain images. U2Net (Qin, et al., 2020) employs Residual U-blocks (RSU) for general image segmentation. Deeplab (Weber, et al., 2021) is a well-known library comprising several projects employing different deep learning techniques. None of these target dental imaging in particular, but they can be adapted to this task. The Segment Anything model (Kirillov, et al., 2023) is a very popular baseline model. The YOLO models are also known for their efficiency

Regarding work on the specific problem tackled here, U-net has been used for mandible segmentation (Abdi, Kasaei, & Mehdizadeh, 2015), as well as segmenting individual teeth (HELLİ & HAMAMCI), combined with non-deeplearning techniques to achieve pixel-wise measurements. The main dataset used in these endeavours is (Abdi A., 2017), with some modification for the tooth segmentation.

**4. Investigated approach**

First of all, it is important to make a distinction between the two problems tackled here. First, there is mandible segmentation, which is as straight-forward as it sounds: given an X-ray, we want to segment the pixels corresponding to the patient’s mandible. Second, there is pixelwise tooth measurement, which aims to measure the width and height of a all teeth given an X-ray.

This project aims to investigate different approaches in order to improve on previous developments in this area. The first experiment involved using the U2-NET architecture for both mandible segmentation and tooth measurements and comparing this to U-NET. Regarding mandible segmentation, an F1 score of 95% was achieved for U2-NET, which is a marginal improvement over the 94% of U-NET. This result has been deemed satisfactory and most future efforts have been focused on the tooth measurement problem.

The first experiment was based on simple segmentation models such as those listed before. After segmenting individual teeth, some erosion was done in order to better delineate the teeth. Then, in order to get measurements, the connex components were separated.

X-ray of a person's chest

Description automatically generatedA yellow and purple image of a mask

Description automatically generated

A x-ray of teeth and gums

Description automatically generatedFor consistency’s sake the F1 score for this experiment is around 95% for both U-NET and U2-NET architectures. However, the F1 score or other segmentation metrics would not be very telling for this particular approach, as any segmented connection between two teeth would lead to the algorithm measuring them together. A more appropriate metric would be measuring the ratio between the predicted number of teeth and the actual number of teeth. On the validation dataset, this method managed to detect 514 out of 796 teeth, divided between 28 images. The ratio is, therefore, 64% This ratio might seem like an oversimplifcation of the evaluation process, but upon viewing the model results, it becomes clear that the measurements are nearly perfect everywhere and that a tooth rarely gets segmented where there is none. Therefore, the main challenge is getting the model to segment teeth that are very near to each other separately. The following image illustrates this point.

It should be noted that the U-NET architecture seems to do better than U2-NET on segmenting individual teeth, even though U2-NET has a slightly better F-score, by 1%. Once again, this is due to the fact that individual teeth measurements require precise boundaries between the teeth.

Since the main problem was the segmentation of the boundaries between the teeth, the next investigated approach was to try giving the model a prompt for how many teeth should be measured. The SAM model was ideal for this, since it can accept bounding boxes as prompts. In order for this to work, a model for providing it with bounding boxes had to be trained. Also, a new dataset had to be created based on the already existing segmentation. Therefore, using the same erosion and connex component separation method as before, bounding boxes were extracted from the ground truth masks. A YOLOv8 model was trained on this new dataset

\*For now, some remarks during training.

I had great hopes for using the SAM model combined with prompts from YOLO, but so far it hasn’t turned out great. Using soft thresholding with the sigmoid function makes the decoder forget to segment teeth that have overlapping segmentation masks. Not using thresholding makes it unable to recognize boundaries between instances that have overlap. This is due to the fact that instead of returning a single segmentation mask, the model return a set of masks I have to then add along the detection axis to get the full mask. There is overlap in the central teeth, where the boundaries are very fine. Havng watched the predictions during training, this is the best intuitive explanation I could find for why this approach is performing poorly.

Oriented bounding boxes, on the other hand, have been surprisingly effective. The regular bounding boxes didn’t have orientation and therefore had some overlap between teeth that were heavily angled. Current status: comparing the angled bounding boxes with the unet model. I also have to look into GNNs.

**5. Application**

\*To be Developed

**6. Conclusion and future work**

# References

Abdi, A. H., Kasaei, S., & Mehdizadeh, M. (2015). Automatic segmentation of mandible in panoramic x-ray. *Journal of Medical Imaging, 2*, 044003–044003.

HELLİ, S., & HAMAMCI, A. (n.d.). Tooth Instance Segmentation on Panoramic Dental Radiographs Using U-Nets and Morphological Processing. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 10*, 39–50.

Hyttinen, J., Fält, P., Jäsberg, H., Kullaa, A., & Hauta-Kasari, M. (2020). Oral and Dental Spectral Image Database—ODSI-DB. *Applied Sciences, 10*. doi:10.3390/app10207246

Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., . . . Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical image analysis, 36*, 61–78.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., . . . others. (2023). Segment anything. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, (pp. 4015–4026).

Lu, Y., Chen, Y., Zhao, D., Liu, B., Lai, Z., & Chen, J. (2021). CNN-G: Convolutional Neural Network Combined With Graph for Image Segmentation With Theoretical Analysis. *IEEE Transactions on Cognitive and Developmental Systems, 13*, 631-644. doi:10.1109/TCDS.2020.2998497

Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O., & Jagersand, M. (2020). U2-Net: Going Deeper with Nested U-Structure for Salient Object Detection., *106*, p. 107404.

Weber, M., Wang, H., Qiao, S., Xie, J., Collins, M. D., Zhu, Y., . . . Chen, L.-C. (2021). DeepLab2: A TensorFlow Library for Deep Labeling. *arXiv: 2106.09748*.