Masaryk University, Faculty of Informatics
Centre for Biomedical Image Analysis, Botanická 68a, Brno
Prof. Michal Kozubek Tel: 54949 4023 https://cbia.fi.muni.cz

Brno, November 24, 2021

Review of the doctoral thesis submitted by Oldřich Kodym

The doctoral thesis entitled "Deep Learning for Virtual Patient-Specific Skull Modelling and Reconstruction" deals with one of the challenges in biomedical image analysis that has not been solved in a satisfactory way so far. The ultimate goal is to help design implants to fill in missing parts of the skull resulting from an injury or surgical intervention. Therefore, the research in this area may help in clinical practice. I appreciate the collaboration of the author with TESCAN Medical company that is focused on this topic.

From the image analysis point of view, there are mainly skull segmentation and reconstruction tasks based on a 3D CT image. These two tasks form the main content of the thesis. For each of them, there is a detailed description of the state-of-the-art, suggested methodology, evaluation of results and discussion. In general, the thesis is well written, the text is clear and easy to follow, there are nearly no typos. There are numerous figures and tables illustrating the ideas as well as many references (139) cited in the text including recent works, which shows that the author has followed the recent trends in the field.

Concerning the methodology, the author has managed to come with innovative solutions employing and combining various state-of-the-art components. For example, in the segmentation part, he has successfully combined deep learning with graph cut approach, which is in line with current trends of enhancing machine learning capabilities by integrating them with classical non-learning approaches. He has also dealt with the problem of limited amount of data and annotations, which is typical for medical image analysis. I appreciate the effort invested into creating an own benchmark dataset and releasing it to the community within MICCAI AutoImplant 2021 Challenge. The author of the thesis is actually even a co-organizer of this competition after previously participating himself in 2020 winning the 3rd prize in both tasks. It should be noted that MICCAI is the best conference in the field (with CORE rank A) and MICCAI challenges are the most respected competitions in biomedical image analysis. As a member of the MICCAI Challenge Board, I can confirm that all MICCAI challenges undergo a strict review process and quality check before being accepted.

The evaluation of the performance of the developed methods has been done rigorously considering multiple state-of-the-art metrics. In particular, not only classical Dice score have been used but also surface Dice score introduced recently by Nikolov et al., which is more suitable to capture the boundary precision. The performance in terms of technical measures is also compared to the rating of human experts showing high correlation.

The results reached by the author are promising, he managed to reach excellent performance on simulated data (created from real data of complete skulls by artificially removing some parts of the skull) and very good on real data (due to the lack of real data for training and testing). As mentioned above, he also reached the 3rd place in an international competition in 2020.

Concerning the publication output, the author has published in both conferences and journals. The best publication is the one in Computers in Biology and Medicine (Q1 journal) as the main author in 2020 (with one more being revised for this journal now, again as the main author). A few papers are off topic (focused on text processing) but even without them the publication output is fully adequate.

Questions for the defense

Although the thesis is well written, I still have encountered some minor issues that I would like to make clear:

1) Resolution. In different parts of the thesis, different resolution of processed images is used including resampling. For example, in chapter 7 (creation of SkullBreak benchmark dataset), the resolution of 0.4 x 0.4 x 0.4 mm was used, which is fine. However, the resolution in chapter 5 (segmentation, especially section 5.4), rather high-resolution images (voxel size 0.38 x 0.38 x 0.38 mm to 0.5 x 0.5 x 1.5 mm) were resampled to lower resolution of 1 x 1 x 1 mm. Why wasn't also here the isotropic voxel size of 0.4 mm or 0.5 mm used? This would better correspond to the precision of modeling in terms of surface errors that are typically in the range of 0.5 to 0.7 mm (e.g., section 8.4) as well as the necessity to keep gaps below 0.8 mm (section 8.1).

2) Expected precision/error. Except for the precision/error limit of 0.8 mm stated above, I have not found any values of clinical requirements or state-of-the-art clinical expert performance. I know there is a lack of real data but at least for the limited set of patients, would it be possible to compute inter- and/or intra-annotator variability for human experts to see what the precision/error of a human expert is? In other words, how does computer performance compare to human one?

3) Inner versus outer surface. It is shown in the thesis that manually designed skull patches are thinner than the surrounding skull thickness while computer-generated patches (also in the artificial training data) adhere to the normal skull thickness (e.g., Figure 6.1 on p. 48 or Figure 10.2 on p. 86). In order to avoid evaluation problems, it is stated that only outer surface is taken into account while computing metrics (e.g., section 8.4 on p. 65). Isn't it dangerous to leave the inner surface without any inspection? Also, is it desirable (from clinical point of view) to let the computer generate thicker patches than experts do?

4) MICCAI AutoImplant. The achieved 3rd rank in MICCAI AutoImplant 2020 is surely a success. Nevertheless, I have missed comparison with the top-2 methods in terms of methodology. Also, the thesis was submitted before MICCAI AutoImplant 2021 (that took place on October 1, 2021). Would it be possible to compare with 2021 methods and results during the defense?

Conclusion

To conclude, I believe that Oldřich Kodym has proved to have deep knowledge in the field of image processing and image analysis, especially segmentation and reconstruction. He is able to apply the knowledge to derive new solutions that are better in terms of widely accepted metrics and benchmarks and he even contributes to the creation of new benchmarks.

During his PhD studies he has published multiple papers (including Q1 journal Computers in Biology and Medicine), has been involved in a challenge at MICCAI conference (CORE rank A) and managed to show very promising performance of his methods. In this way he has significantly contributed to the progress in the given field. Therefore, I have no doubt that he deserves the title of "Doctor of Philosophy" in the field of information technology and I **recommend accepting the submitted doctoral thesis** as the basis for such promotion.

……
Michal Kozubek