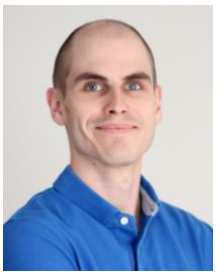
Deep Learning-based Approach for Variable Candidate Segmentation

Jan Riedo

Supervisors: Prof. Dr. Mauricio Reyes, MSc Alain Jungo
Institution: Insel Data Science Center, University Hospital Bern (Inselspital), University of Bern, Switzerland
Examiners: Prof. Dr. Mauricio Reyes, Dr. Raphael Meier



Introduction

Highly specialzied medical diagnostics and therapies require accurate semantic segmentations of certain body parts. In radiation therapy for brain tumor treatment, organs at risk (OAR) need to be marked to prevent them from irradiation during the procedure. The gold standard in therapy planning and treatment are time-consuming manual single-rater segmentations. To enhance the reliability and clinical acceptance of automated deep neural network segmentations, rater behavior need to be taken into account.
Therefore, we propose several networks which aim at generating variable candidate segmentations, imitating the inter-rater variability and going beyond. Further, we show that our segmentation samples can be used for additional steps in the medical imaging pipeline by the example of automated quality control.

Materials and Methods

We employ two datasets for training and testing our proposed approaches. For a proof of concept we use a syntheticly generated dataset with five artificial rater segmentations provided for each image. Clinical examination is conducted on contrast-enhanced T1-weighted MR images of an OAR dataset (3 raters) with focus on the brainstem.
Our proposed approach (see Fig. 1) to introduce variability in the segmentation output consists of a state-of-the-art U-Net [1] with an incorporated variational autoencoder [2]. This architecture provides an interface to a standard Gaussian distribution, from which a continuous sampling of candidate segmentations is possible.
The variable candidate segmentations are employed to train a regression network to estimate the quality of produced segmentations. The perceived quality is evaluated by a clinical expert.

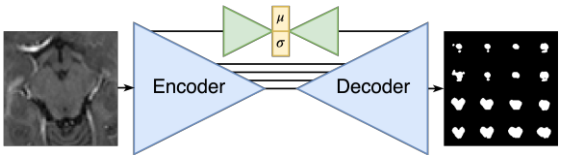


Fig. 1 Proposed VU-Net model consisting of a U-Net with incorporated variational autoencoder for variable candidate segmentations shown in tiled mode over the latent space.

The samples are assessed on a precision weighted variability metric and the maximum Dice similarity coefficient (DSC) between the candidates and the expert segmentations. The Pearson correlation coefficient is used to evaluate the quality estimates.

Results

The experiments on synthetic data confirmed the feasibility of employing a U-Net for variable candidate segmentation. Applied on OAR data (Fig. 2), the variability of segmentations was increased by a factor of ten compared to the experts (0.077 vs. 0.743), while keeping the max DSC at inter-rater level (0.944 vs. 0.945).
The regression network was able to lift the correlation coefficient from 0.790 when trained on deterministic data to 0.983 with variable segmentations. Clinical examination verified the proposed segmentations and DSC estimates.

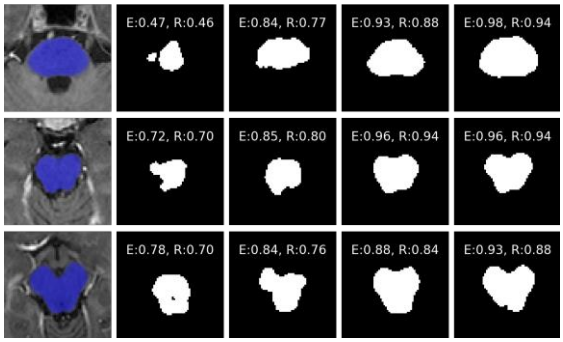


Fig. 2 Deterministic segmentations of the brainstem by the basic U-Net (first column) and sample segmentations obtained from the VU-Net. E: estimated max DSC obtained from the quality control pipeline. R: Real max DSC between sample and any of the raters.

Discussion

We demonstrated the feasibility of employing U-Net based architectures for variable but sensible segmentations, based on label and feature map variability. The networks are successful on synthetic data, as well as on OAR patches and whole slices. The incorporation of variability on whole slices was more challenging, with the results being inferior.
Technical and clinical validation of the proposed quality controlled sampling was successful and showed the need for a training set with segmentations of diverse quality. The success of quality controlled sampling is one step towards artificially generated valid rater segmentations.

References

[1] O. Ronneberger, P.Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” MICCAI, Springer, 2015.

[2] D. P. Kingma and M. Welling, Auto-encoding variational bayes, ICLR 2013.