# Segmentation of costa in 2d MRI scans utilizing the U-Net convolutional network architecture

Helge Fredriksen*[1] and Arthur Schuchter†[1]

[1]UiT - The Arctic University of Tromsø

## 1 Introduction

Segmentation of costa (rib bones) in magnetic resonance images is not a widely researched area. However, it may have significant clinical value, since it can act as a stable reference for multiple analysis and quantification tasks [9]. In the research presented here we utilize the U-Net deep learning convolutional network architecture to perform the segmentation. The main concept of the U-Net network pipeline is the heavy reliance on data augmentation combined with up sampling to allow connection between convolution and deconvolution layers. [7].

## 2 Results

We performed a series of tests on MRI scans, where we trained an U-Net CNN architecture utilizing a subset of 30 512x512 monochrome MRI images from a single 3d axial scan of the chest region consisting of 344 images. Masks for the training images were created by labelling pixels either as background or as costa cross-sections. Each image and mask in the training set contained from 4 to 10 of labelled costa cross-sections, varying in size, location, and orientation. Augmentation was supported by utilizing the preprocessor module in Keras. The model was trained using 300 epochs with 30 steps per epoch, and with a constant learning rate of $3 \times 10^{-5}$.

An initial test was performed on another subset of 30 images from the same MRI scan to see if the algorithm would be able to detect the structures in an equal environment than initially trained upon. Results, which were produced as output masks with a probability measure between 0 and 1 in each pixel, were quite promising. In Figure 1, we see one of the test results, where the left image is the original and the right is a color mapped overlay of the produced probability mask. For validation purposes we utilized the trained model on 4 other samples from an axial MRI scan slightly skewed and with another cross-section size compared to the scan originally trained on. Results from this test showed that the model performed equally well (see Figure 2). We utilized a dice coefficient [1] to measure overlap between ground truth binary images and predictions on these images, setting 1 on pixels that had a probability above 0.5, and 0 on those below:

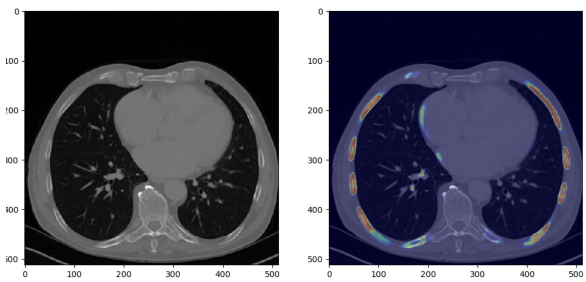

Figure 1: Segmentation of costa in an image taken from the same scan as the original training images

$$D = \frac{2\sum_i^N p_i \cdot g_i}{\sum_i^N p_i + \sum_i^N g_i} \tag{1}$$

where the sum is over all pixels in the masks and $p_i$ and $g_i$ are binary values on predicted and ground truth masks respectively. We found the value of this coefficient to be 0.79, 0.72, 0.80 and 0.64 on the subset of validation images we chose on a random basis.

*Corresponding Author: helge.fredriksen@uit.no

†Corresponding Author: arthur.schuchter@uit.no

# 3 Discussion

Our ambition is to transfer the successful implementation of U-Net to include other tasks that may aid radiologists in detecting various lesions and give input to volumetric calculations. To achieve this, we need to extend the segmentation into a 3d representation. This may be achieved by exploring models such as the V-net architecture [6], exploiting the full volume of the MRI scan as data for a training a neural net model. Although this may seem appealing, the approach can result in high memory footprints and may have challenging setups for annotation i 3D. Thus, having access to a trained 2d model seemingly working efficiently on individual cross-sections, representing a 3d model should be a possible. Interpolating such 3d continuous models from 2d results would require axial distances between scans to be below the scale of significant variation in the xy-placement of detected costa features between neighboring scans. As such, we foresee a robust 3d representation would emerge as a result of connecting the 2d segmented images in a scan. This is similar to a voxel-like presentation of MRI scans as described in [2]. A voxel is defined as the three-dimensional counterpart of a 2d pixel, where the stacking of neighboring 2d images along the z-axis represent the third dimension.

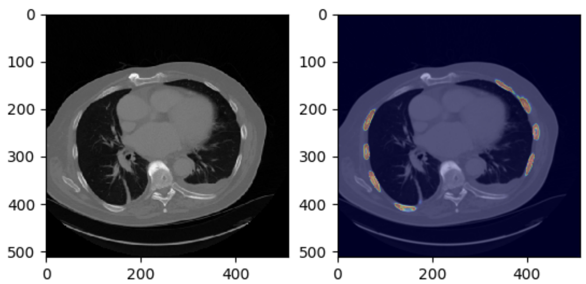


Figure 2: Segmentation of costa in a scan from a different patient than originally trained on

If the voxel-built structure is generated successfully through the utilization of multiple MRI scans, the user will be able to hide certain areas of the scanned object or create a clear cross-section of the body part, making the visualization of the desired feature much easier. This could enable a surgeon to analyze the inner section of the feature and pinpoint locations of tumors for example, leading to a safer and more precise operation in surgery. Moreover, this could aid the estimation of feature volumes, also called volumetry, making radiologists better capable to perform numerical estimations. Such volumentry would be based on simply counting connected neighboring voxels. However, this would require a mapping of the metric size of the voxel based on a calibration procedure [5].

Presentation of segmentation results for visualization in 3d is a challenging issue. However, the popular Integrated Development Environment (IDE) Unity™ can be utilized for such tasks [8]. The concept of voxel generation is based on an application within Unity called the voxel engine. While initially designed for video games, the voxel generator of the voxel engine is capable of creating a 3D digital structure of organs and tissue. Initial tests utilizing this engine for generation of 3d models from 2d segmentations shows promising results (see Figure 3). The results shown are based on K-means [3] and Watershed algorithms [4] for the segmentation, where a variety of filters are utilized prior to the application of the algorithm. These filters actively reduce noise from the image, e.g., salt and pepper noise, providing the segmentation process with better input. Commonly used are the high and low pass filter in addition to mean and median filters. Future research will include U-Net as yet another alternative to K-Means and Watershed algorithms for processing of 2d segmentation prior to the 3d modelling.

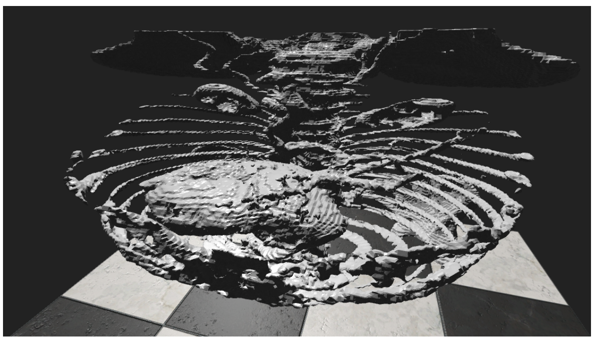

Figure 3: Voxel generated 3d model based on 2d segmented MRI images showing costa and urethra

# References


[1] W. Crum, O. Camara, and D. Hill. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Transactions on Medical Imaging*, 25(11):1451–1461, 2006. doi:10.1109/TMI.2006.880587.

[2] I. Despotović, B. Goossens, and W. Philips. Mri segmentation of the human brain: Challenges, methods, and applications. *Computational and Mathematical Methods in Medicine*, 2015:450341, 2015. doi:10.1155/2015/450341.

[3] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979. doi:10.2307/2346830.

[4] N. Hieu Tat, M. Worring, and R. v. d. Boomgaard. Watersnakes: energy-driven watershed segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):330–342, 2003. doi:10.1109/TPAMI.2003.1182096.

[5] M. Jafar, Y. M. Jafar, C. Dean, and M. E. Miquel. Assessment of geometric distortion in six clinical scanners using a 3d-printed grid phantom. *Journal of Imaging*, 3(3):28, 2017. doi:https://doi.org/10.3390/jimaging3030028.

[6] F. Milletari, N. Navab, and S.-A. Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016. doi:10.1109/3DV.2016.79.

[7] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *18th International Conference on Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015*, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pages 234–241. Springer International Publishing, 2015. doi:10.1007/978-3-319-24571-3.

[8] G. Wheeler, S. Deng, N. Toussaint, K. Pushparajah, J. A. Schnabel, J. M. Simpson, and A. Gomez. Virtual interaction and visualisation of 3d medical imaging data with vtk and unity. *Healthcare technology letters*, 5(5):148–153, 2018. doi:10.1049/htl.2018.5064.

[9] Z. Xu, U. Bagci, C. Jonsson, S. Jain, and D. J. Mollura. Efficient ribcage segmentation from ct scans using shape features. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, 2014:2899–2902, 2014. doi:10.1109/EMBC.2014.6944229.