
Comparison of deep learning-based techniques for organ segmentation in abdominal CT images

V. Groza

Philips Innovation Labs RUS
Moscow, Russia
vladimir.groza@philips.com

T. Brosch , D. Eschweiler, H. Schulz, S. Renisch and H. Nickisch
Philips GmbH Innovative Technologies
Hamburg, Germany

Abstract

Automatic segmentation of the liver, spleen and both kidneys is an important problem allowing to achieve accurate clinical diagnosis and to improve computer-aided decision support systems. This work presents a computational method for automatic segmentation of liver, spleen, left and right kidney in abdominal CT images using deep convolutional neural networks (CNN) which allows the accurate segmentation of large-scale medical trials. Moreover this work demonstrates the comparison of several CNN based approaches to perform the segmentation of required organs. Validation results on the given dataset show that U-Net based liver, spleen and both kidneys segmentation for transaxial slicing achieves mean Dice similarity scores (DSC) of 94%, 89% and 88% respectively.

1 Introduction

In clinical trials, the interpretation of CT and MRI images obtained in the diagnosis of liver is usually done by manual or semi-automatic segmentation techniques. Such methodologies are subjective, person-dependent and very time-expensive. In order to improve the quality and productivity of radiologists computer-aided techniques have been developed. These systems recently showed a big potential for improving the diagnostic accuracy.

However, an automated robust segmentation of liver, spleen and both kidneys remains still an open problem due to a low-contrast between the organs and lesion, variety of contrast levels, size and abnormalities in tissues.

Deep learning approaches became the new standards in many traditional domains of computer vision like image classification, object detection or segmentation. Deep convolutional neural network methods demonstrated their flexibility and robustness for image processing, what motivates us to apply them for automatic organs segmentation in abdominal CT images. Many recent segmentation methods based on CNN and FCNs were developed for medical image analysis claiming highly competitive results compared to the state-of-the-art [1, 2, 3].

In this work we selected the most popular for biomedical image segmentation U-Net architecture [4] due to its proven success and extended it with modifications in the layers configuration, specific image batch generation and loss function.

Table 1: Comparison of DSCs for single organ (liver,spleen and both kidneys) segmentation with U-Net, F-Net and BRIEF respectively

DSC	Liver	Spleen	Left kidney	Right kidney
F-Net	0.925	0.885	0.890	0.883
BRIEF	0.852	0.751	0.747	0.749
U-Net	0.943	0.887	0.876	0.872

2 Methods

For the segmentation of four organs (liver, spleen and both kidneys) on abdominal CT several deep convolutional neural network architectures were implemented and compared. The dataset for this task comprises 70 CT volumes of abdominal region with corresponding number of ground truth delineations. The original size of images is $512 \times 512 \times 394$ which were resampled to $224 \times 224 \times 394$ in order to proceed several images in one batch. In order to prevent overfitting of the network and obtain reproducible results we implemented a 5-fold cross-validation by random splitting the dataset on training and validation parts with equal size of 14 cases per one fold. The batch normalization layers were excluded from the configuration of 4-levels depth U-Net because the input data were normalized to range $[-1;1]$ during the pre-processing step so additional batch normalization was not necessary.

Initial images were generally normalized by their mean and standard deviation and in order to exclude the artifacts such as outer air/colon gas and kidney stones their contrast values were clipped to manually defined low and high threshold of values 845 and 1295 Hounsfield units (HU). Since a 3D processing of the images was not possible due to memory limitations, images were processed slice-wise as commonly done for medical image processing using deep learning techniques.

Training was performed in mini-batches consisting of 16 images. In medical imaging some masks of organ segmentation may be empty because the target organ is present in a subset of slices only. In order to improve the convergence and accuracy of our approach, we created the batches such that each batch contains at least one image with non-empty mask corresponding to required organ. Since we can not expect that target organ will be present in the input CT volume we have to keep the empty slices in the training process. The network was trained by minimizing the loss function with custom adaptive learning rate scheduler, which decreases the learning rate by the factor of 2 after 5 epochs sticking on plateau. Parameters were optimized by the use of adaptive gradient descent algorithm Adam [5]. Segmentation performance was assessed in terms of Dice similarity scores (DSC) [6], precision and recall rates.

3 Results

For the U-Net type network, DSCs of 0.943, 0.887 and 0.876 were obtained for liver, spleen and both kidneys for transaxial slicing respectively. The neural network was trained from scratch during the 70 epochs for each of 5-fold splitting. We included the results of another approaches BRIEF [7] and F-Net [8] obtained by the evaluation of the methods on the same dataset using the cross-validation with identical split, what allows for a direct comparison.

These network architectures demonstrated similar results – with the use of BRIEF DSCs of 0.852, 0.751, 0.749 were respectively obtained, while for F-Net appropriate scores are 0.925, 0.885, 0.89. No post-processing was performed in order to fairly compare the performance of different approaches. More detailed comparisons in terms of DSCs, precision and recall rates between different approaches for transaxial slicing are presented in the Tables 1, 2 and 3.

In addition, we assessed the impact of using SegNet [9] inspired trick of transferring of maxpooling indices as additional feature information from each encoder layer to decoder. The additional concatenations of this feature maps with standard U-Net concatenated decoder part did not demonstrate statistically significant improvement and had only a minimal effect on the segmentation accuracy, while had a high impact on memory and calculating time consumption.

Table 2: Comparison of precision rates for single organ (liver,spleen and both kidneys) segmentation with U-Net, F-Net and BRIEF respectively

DSC	Liver	Spleen	Left kidney	Right kidney
F-Net	0.932	0.892	0.904	0.895
BRIEF	0.779	0.668	0.690	0.696
U-Net	0.947	0.910	0.903	0.915

Table 3: Comparison of recall rates for single organ (liver,spleen and both kidneys) segmentation with U-Net, F-Net and BRIEF respectively

DSC	Liver	Spleen	Left kidney	Right kidney
F-Net	0.920	0.887	0.892	0.880
BRIEF	0.947	0.880	0.842	0.831
U-Net	0.940	0.868	0.860	0.854

4 Conclusion

This study indicates that segmentation of liver, spleen, left and right kidney using neural networks is feasible with very good accuracy by the use of different techniques and approaches. However depending on specific organ segmentation problem and image type, the choice of the network architecture as well as the hyperparameters might be more sensitive and requires further investigation.

Acknowledgments

The authors of this work would like to acknowledge Irina Fedulova, Joël Valentin Stadelmann and Nina Tuluptceva for many inspiring discussions and their comments on this work.

References

- [1] Brosch, T., Tang, L. Y., Yoo, Y., Li, D. K., Traboulsee, A., and Tam, R., "Deep 3D convolutional encoder networks with shortcuts for multi-scale feature integration applied to multiple sclerosis lesion segmentation", *IEEE transactions on medical imaging* 35(5), 1229–1239 (2016).
- [2] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation, *Medical Image Analysis* 36, 61–78 (2017).
- [3] H. R. Roth, L. Lu, A. Farag, H.-C. Shin, J. Liu, E. B. Turkbey, R. M. Summers, Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation, *MICCAI*, 556–564 (2015).
- [4] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, *MICCAI*, Vol. 9351, 234–241 (2015).
- [5] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).
- [6] Dice, L. R., "Measures of the amount of ecologic association between species", *Ecology* 26(3), 297–302 (1945).
- [7] Heinrich, M. P. and Blendowski, M., "Multi-organ segmentation using vantage point forests and binary context features", *MICCAI 2016, Part II, LNCS 9901*, 598–606, Springer (2016).
- [8] Brosch T., Saalbach A., "Foveal fully convolutional nets for multi-organ segmentation", *Proc. SPIE 10574, Medical Imaging 2018: Image Processing*, 105740U (2018).
- [9] Badrinarayanan V., Handa A. and Cipolla R., "Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling." *arXiv preprint arXiv:1505.07293* (2015).