MIDL supplementary material - Joint shape learning and segmentation for medical images using a minimalistic deep network

Balamurali Murugesan[1,2], Kaushik Sarveswaran[2], Sharath M Shankaranarayana[3], Keerthi Ram[2], and Mohanasankar Sivaprakasam[1,2]

[1]**Indian Institute of Technology Madras (IITM), India**
[2]**Healthcare Technology Innovation Centre (HTIC), IITM, India**
[3]**Zasti, India**

ABSTRACT

Recently, state-of-the-art results have been achieved in semantic segmentation using fully convolutional networks (FCNs). Most of these networks employ encoder-decoder style architecture similar to U-Net and are trained with images and the corresponding segmentation maps as a pixel-wise classification task. Such frameworks only exploit class information by using the ground truth segmentation maps. In this paper, we propose a multi-task learning framework with the main aim of exploiting structural and spatial information along with the class information. We modify the decoder part of the FCN to exploit class information and the structural information as well. We intend to do this while also keeping the parameters of the network as low as possible. We obtain the structural information using either of the two ways- i) using the contour map and ii) using the distance map, both of which can be obtained from ground truth segmentation maps with no additional annotation costs. We also explore different ways in which distance maps can be computed and study the effects of different distance maps on the segmentation performance. We also experiment extensively on two different medical image segmentation applications- i.e i) using color fundus images for optic disc and cup segmentation and ii) using endoscopic images for polyp segmentation. Through our experiments, we report results comparable to, and in some cases performing better than the current state-of-the-art architectures and with an order of 2x reduction in the number of parameters.

Keywords: Deep mutitask learning, Fully convolutional network, Optic disc and cup segmentation, Polyp segmentation

1 EXPERIMENTS

1.1 Comparison of segmentation metrics with RIMONE Cup dataset

The proposal made in our manuscript is further validated by performing a comparative study with RIMONE Fumero et al. (2011) cup segmentation dataset. The networks are fine-tuned with models obtained from Origa dataset. Table 1 presents the comparison of Ronneberger et al. (2015), Chen et al. (2016), Tan et al. (2018) with our models (1Enc 1Dec + Conv MC and 1Enc 1Dec + Conv MD). It is clear from the table that our model 1Enc 1Dec + Conv MC and 1Enc 1Dec + Conv MD performs close to that of Chen et al. (2016) and Tan et al. (2018).

1.2 Contour Learning process

In order to understand the learning process of contour by the model. We obtained the predicted contour at each epoch and some of them over the iterations are shown in Figure 1. It can be seen that, during initial epochs, only segmented islands of contour are obtained. Whereas, in the later epochs, whole contour is obtained.

Table 1. Evaluating models of interest in RIMONE dataset (Fumero et al. (2011)).

Method	Dice	Jaccard
1Enc 1Dec M (Ronneberger et al. (2015))	0.8021	0.6894
1Enc 2Dec MC (Chen et al. (2016))	0.8112	0.7004
1Enc 2Dec MD (Tan et al. (2018))	0.8093	0.6972
1Enc 1Dec + Conv MC (Ours)	0.8196	0.7140
1Enc 1Dec + Conv MD (Ours)	0.8048	0.6897

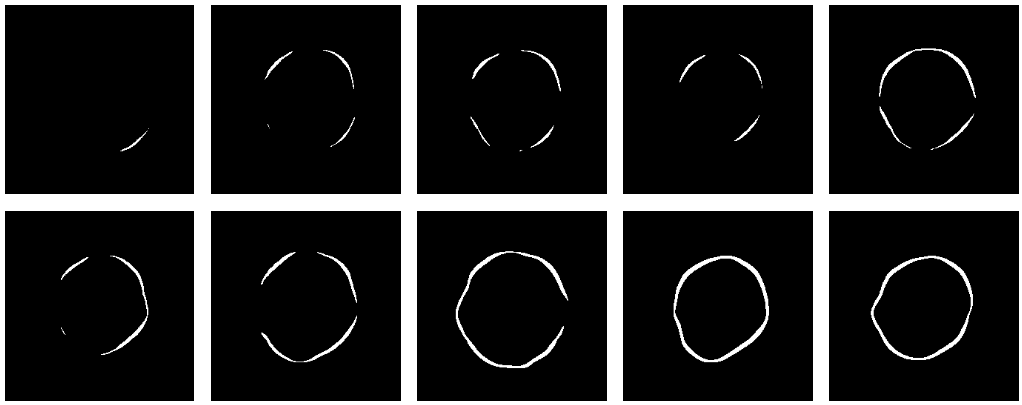

Figure 1. Contour Learning. Illustration of contours predicted by the model 1Enc 1Dec + Conv MC while epochs increase from top left to bottom right.

1.3 Shape and boundary evaluation metrics

1.3.1 Shape Similarity

Along with segmentation evaluation, the network should also produce segmentation maps which are similar to ground truth masks regarding shape Chen et al. (2016). This shape similarity is obtained by Hausdorff distance. Table 2 shows average Hausdorff distance for Cup (Origa and RIMONE dataset) and Polyp (Endovis dataset). It can be seen that 1Enc 1Dec + Conv MC and 1Enc 1Dec + Conv MD performs well compared to 1Enc 1Dec M. This shows the contour and distance map does help in learning shape. Some sample examples are shown in Figure 2.

1.3.2 Segmentation around boundaries

According to our manuscript, the network with distance map (1Enc 1Dec + Conv MD) as a auxilliary task should give a smooth boundary compared to other networks. Smooth boundaries indicate a better segmentation around the boundary. We evaluated the segmentation accuracy around boundary with the method adopted in Krähenbühl and Koltun (2011). Specifically, we count the relative number of misclassified pixels within a narrow band ("trimap") surrounding actual object boundaries, obtained from the accurate ground truth images. As can be seen in Figure 2, the network with distance map (1Enc 1Dec + Conv MD) has less error for trimaps of different widths.

Table 2. Hausdorff Distance measure comparison

Architecture	Cup(ORIGA)	Cup(RIMONE)	Polyp(Endovis)
1Enc 1Dec M (Ronneberger et al. (2015))	14.832	15.379	24.133
1Enc 1Dec + ConvMC (Ours)	14.775	14.367	22.737
1Enc 1Dec + ConvMD (Ours)	14.814	13.619	22.686

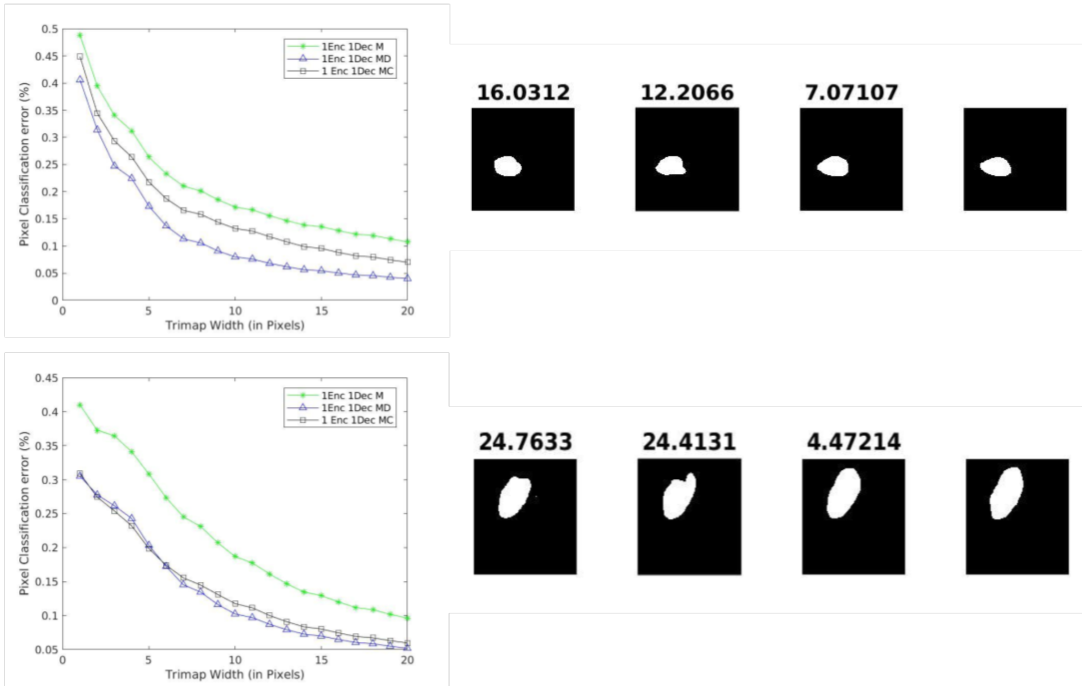

Figure 2. Left: Trimap Right: Respective hausdorff distances (from left to right - 1Enc 1Dec M (Ronneberger et al. (2015))(Green), 1Enc 1Dec + Conv MC (Ours)(Black), 1Enc 1Dec + Conv MD (Ours)(Blue)

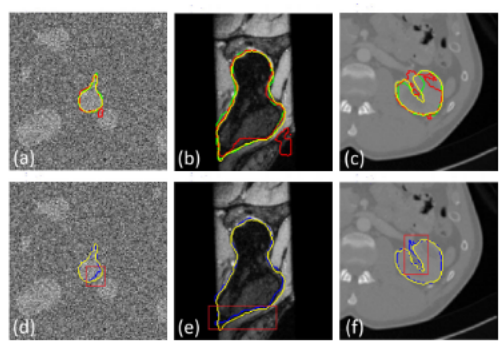

Fig. 3. 2D visual comparisons for simulated data, femur bone and kidney. Green, red, blue and yellow lines are for the GT, U-net, DCAN and proposed method, respectively.

Table 1. Quantitative comparisons.

Method	Simulation		Femur Bone		Kidney	
	DSC	RE	DSC	RE	DSC	RE
U-net	0.90	0.19	0.90	0.19	0.70	0.62
DCAN	0.92	0.15	0.91	0.18	0.83	0.31
Ours	**0.96**	**0.09**	**0.93**	**0.14**	**0.90**	**0.19**

Figure 3. Comparison of Chen et al. (2016) and Tan et al. (2018) as taken from Tan et al. (2018)

REFERENCES

Chen, H., Qi, X., Yu, L., and Heng, P. (2016). DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2487–2496.

Fumero, F., Alayón, S., Sánchez, J. L., Sigut, J. F., and González-Hernández, M. (2011). Rim-one: An open retinal image database for optic nerve evaluation. *2011 24th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 1–6.

Krähenbühl, P. and Koltun, V. (2011). Efficient inference in fully connected crfs with gaussian edge potentials. In Shawe-Taylor, J., Zemel, R. S., Bartlett, P. L., Pereira, F., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 24*, pages 109–117. Curran Associates, Inc.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241. Springer International Publishing.

Tan, C., Zhao, L., Yan, Z., Li, K., Metaxas, D., and Zhan, Y. (2018). Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1221–1224.