Elastic Boundary Projection for 3D Medical Image Segmentation

Tianwei Ni[1], Lingxi Xie[2,3(✉)], Huangjie Zheng[4], Elliot K. Fishman[5], Alan L. Yuille[2]
[1]Peking University [2]Johns Hopkins University [3]Noah's Ark Lab, Huawei Inc.
[4]Shanghai Jiao Tong University [5]Johns Hopkins Medical Institute
{twni2016, 198808xc, alan.l.yuille}@gmail.com zhj865265@sjtu.edu.cn efishman@jhmi.edu

Abstract

Here we show some qualitative comparisons between EBP and two baselines, namely, RSTN (2D) and VNet (3D). The materials were not put in the main article due to the space limit.

Qualitative Results

In this part, we compare the segmentation results by RSTN [2], VNet [1] and EBP. We choose one case from our dataset and one case from the MSD *spleen* dataset, respectively.

To compare the different behaviors between EBP and two baselines, we display some slice-wise segmentation results in Figure 1. We can see that EBP often produces results in a good shape, even when the image is impacted by some unusual conditions in CT scan. In comparison, RSTN and VNet produce segmentation by merging several parts (RSTN: slices, VNet: patches), therefore, in such extreme situations, some parts can be missing and thus segmentation accuracy can be low. On the other hand, the most common issue that harms the accuracy of EBP is the inaccuracy in distinguishing inner pivots from outer pivots. Under regular conditions, EBP is often more sensitive to the boundary of the targets, as it is especially trained to handle these cases – an example comes from the visualization results in the MSD *spleen* dataset, which demonstrates that EBP sometimes produces better results than the ground-truth especially near the boundary areas.

References

[1] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *International Conference on 3D Vision*, pages 565–571. IEEE, 2016. 1

[2] Qihang Yu, Lingxi Xie, Yan Wang, Yuyin Zhou, Elliot K Fishman, and Alan L Yuille. Recurrent saliency transformation network: Incorporating multi-stage visual cues for small organ segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 8280–8289, 2018. 1

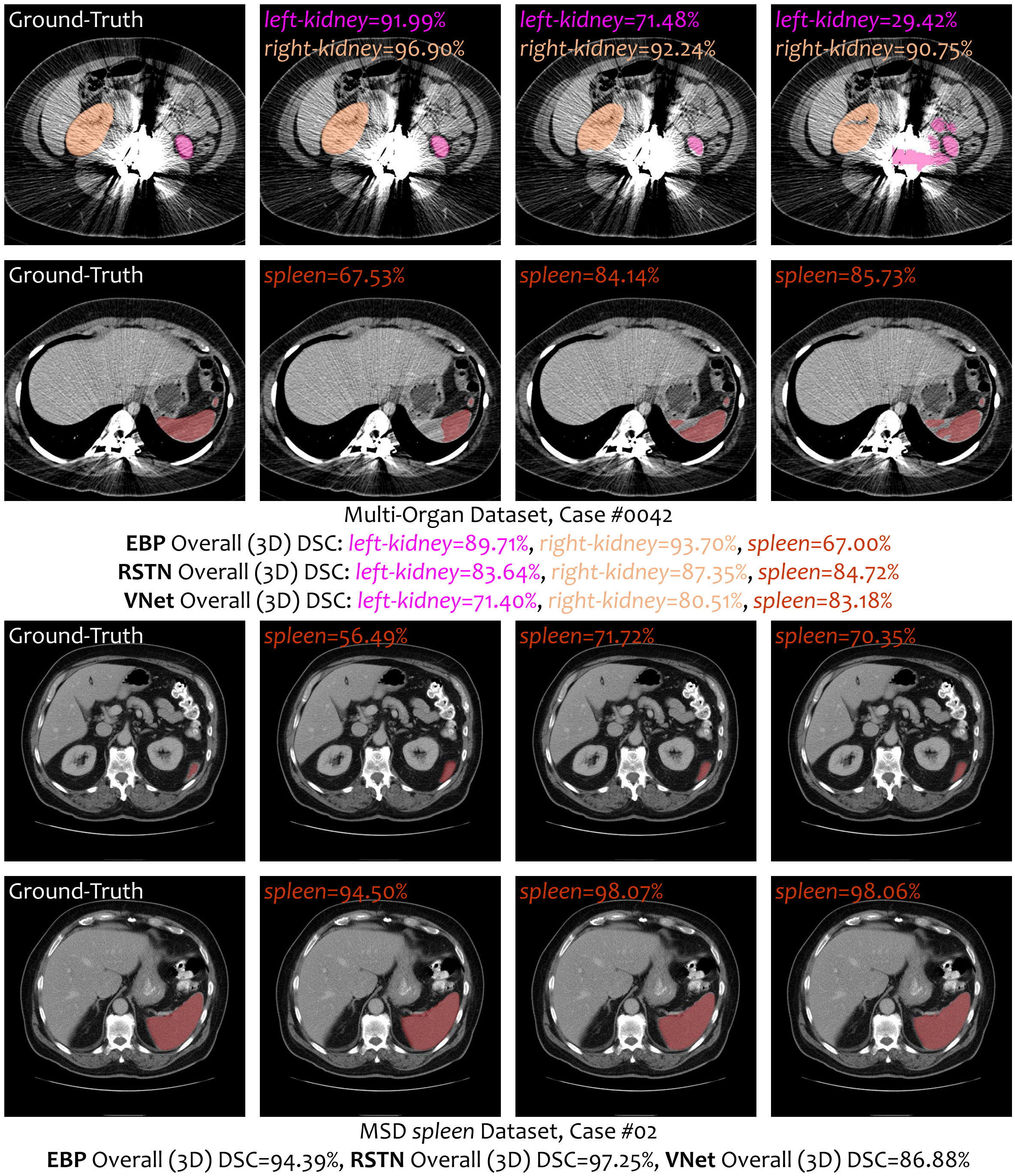

Figure 1. 2D visualization of segmentation results (best viewed in color). In each row, from left to right: ground-truth, EBP, RSTN, VNet. **The top part** shows one special case in our multi-organ segmentation dataset. In this case, the image looks differently compared to most training images, due to some unusual situations during the CT scan. In this case, *kidney* segmentation results of both RSTN and VNet are heavily impacted whereas EBP works reasonably well. EBP produces unsatisfactory results on *spleen* segmentation, mainly because a part of pivots are not recognized as inner pivots. By simply tuning down the threshold by a little bit, EBP reports 86.44% on *spleen* segmentation, which surpasses both RSTN and VNet. **The bottom part** shows a case in the MSD *spleen* dataset, which we can observe how imperfect annotation affects DSC evaluations. In both rows, the ground-truth annotations do not cover the entire *spleen*. RSTN and VNet somehow miss a small margin close to the boundary, while EBP produces obviously better results but gets lower DSC scores. This tells us (i) ground-truth annotations in medical images are often imperfect; (ii) DSC values above the human level (*e.g.*, it can be defined as the average DSC between two individual human labelers, but such numbers are not available in most datasets) do not accurately reflect the absolute quality of segmentation, and in this scenario, a higher DSC does not guarantee better segmentation.