Improving Weakly Supervised Lesion Segmentation Using Multi-Task Learning

Tianshu Chu[1]*, Xinmeng Li[1]*, Huy V. Vo[2], Ronald M. Summers[3], Elena Sizikova[1] * - contributed equally

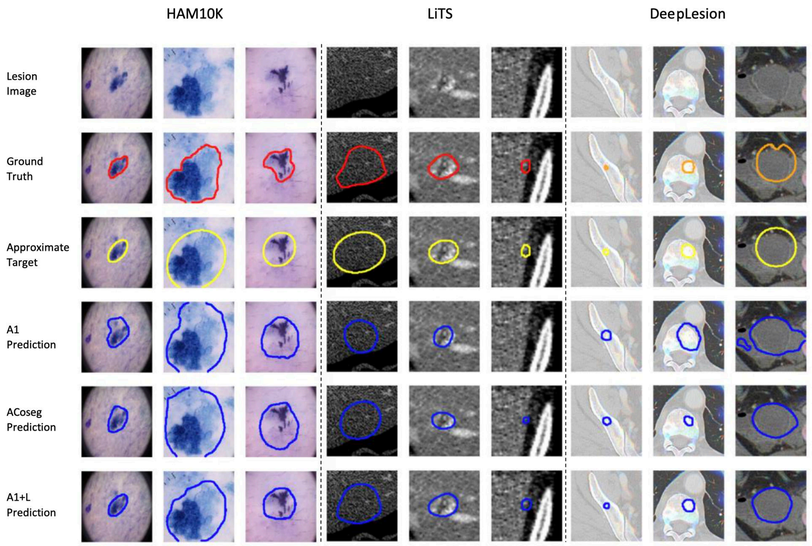
[1]Center for Data Science, New York University; [2]Ecole Normale Superieure; [3]Diagnostic Radiology, NIH Clinical Center

Introduction

Given that we do not have full supervision, how can we leverage the existing partial data to learn a good model? Such problems are common in medical imaging, where acquiring full supervision requires specialized training and specialized knowledge. RECIST is a standard way to roughly annotate the lesion extent in CT, but no dense masks are typically available.

In this work, we propose to replace the segmentation alone and co-segmentation (Agarwal et al., 2020a) process by a join classification/ segmentation network.

Visualization

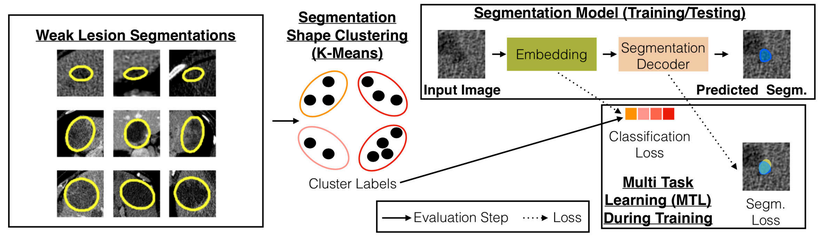


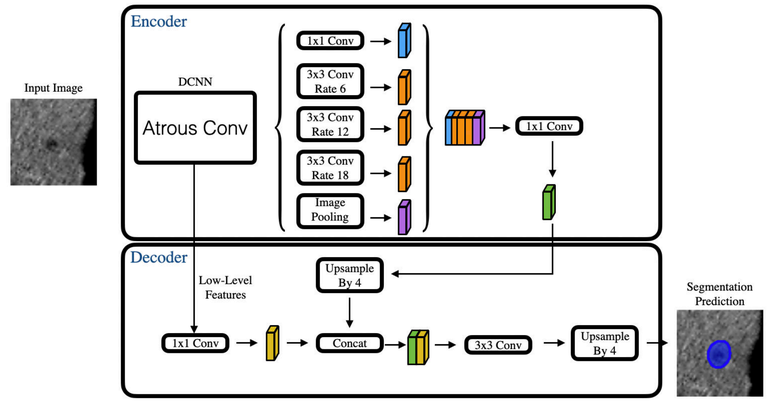
The prediction masks of our method are closer to the relative rotation, shape and size of ground truth than A1 and ACoseg results.

References

1. *Vatsal Agarwal, Youbao Tang, Jing Xiao, and Ronald M Summers. Weakly supervised lesion co-segmentation on ct scans. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pages 203–206. IEEE, 2020a.*
2. *Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (ECCV), pages 801–818, 2018.*

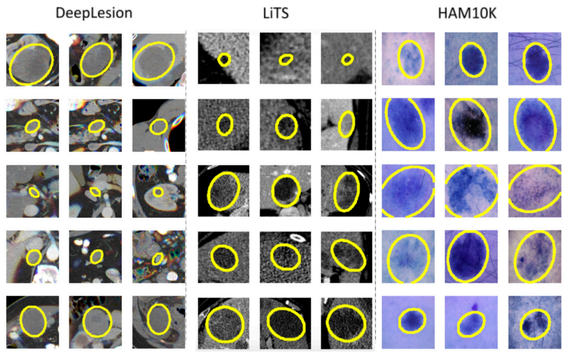
Methods





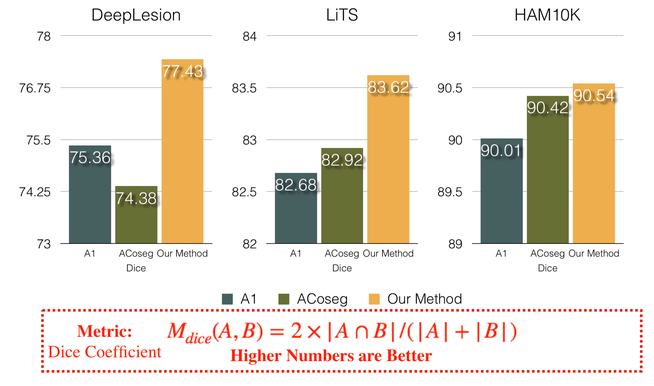
DeepLabV3+ [Chen, ECCV'18]

In this work, we propose to replace the co-segmentation process by a joint classification/segmentation network. We obtain the class labels by clustering lesions' RECIST measurements with K-means. Based on the RECIST, we also generate pseudo-ground truth using the Ellipse algorithm. We perform classification on the encoder output and train our model in a multi-task learning fashion to minimize the joint loss.

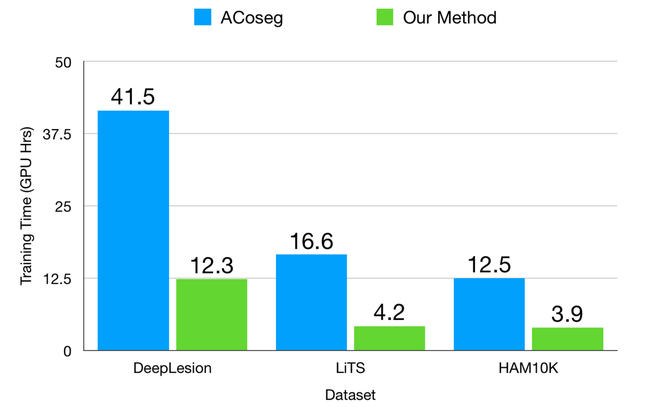


Cluster Visualization

Results

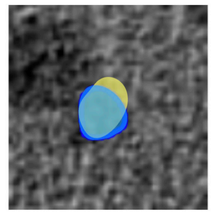


Metric: Dice Coefficient $M_{dice}(A,B) = 2 \times |A \cap B| / (|A| + |B|)$ **Higher Numbers are Better**



We presented an multi-task learning approach that allows more accurate segmentations of computed tomography (CT) and dermatoscopic imaging data while significantly reducing training time.

Conclusion



- New joint detection and segmentation scheme for weakly supervised lesion segmentation
- Systematic analysis on CT and dermatoscopic datasets shows higher accuracy and faster training