One-Shot Uncertainty Estimation for Deep Network Based Image Segmentation

Y. MIN [1] and D. RUAN [1]

1 UCLA School of Medicine, Los Angeles, CA

David Geffen School of Medicine UCLA

JULY 12–16 2020 VIRTUAL JOINT AAPM | COMP MEETING EASTERN TIME [GMT-4]

PURPOSE

The advancement of highly conformal and targeted radiotherapy delivery mechanisms imposes an urgent need for automatic segmentation with high accuracy and efficiency, where deep neural networks (DNN) had yielded great promise. In addition, it is important to understand the confidence associated with such contour set. While Monte Carlo Dropout (MCDO) approaches with Bayesian network setting [1] can yield uncertainty estimation, we hypothesize that treating the contours in a pseudo-probabilistic setting could yield comparable information with much better computational effectiveness.

METHOD

A 2D Dense-Unet [2] like architecture was utilized as segmentation model for this uncertainty estimation experiment. Similar to MCDO on Bayesian SegNet, the benchmark multi-sample based approach defines

$$p_c = \frac{\sum_{t=1}^{T} p_{t,c}}{T}, \qquad uncertainty_c = Variance(p_{t,c})$$

where T stands for the number of stochastic network samples and $\mathbb{C}$ is a set of k classes and $c \in \mathbb{C}, t \in T$. We set T=30 for the benchmark.

Although the soft membership assigned from intermediate layers of a segmentation DNN is arguably a probability measure [3], we utilize it as "free" by product, and endow it with statistical interpretations to define uncertainty. Specifically, trace of estimated multivariate covariance (TCV) and multiple variant entropy (VE) were investigated as uncertainty indicators.

Estimated multivariate covariance matrix is defined as

$$\begin{bmatrix} p_1(1-p_1) & \cdots & p_1(1-p_k) \\ \vdots & \ddots & \vdots \\ p_k(1-p_1) & \cdots & p_k(1-p_k) \end{bmatrix}$$

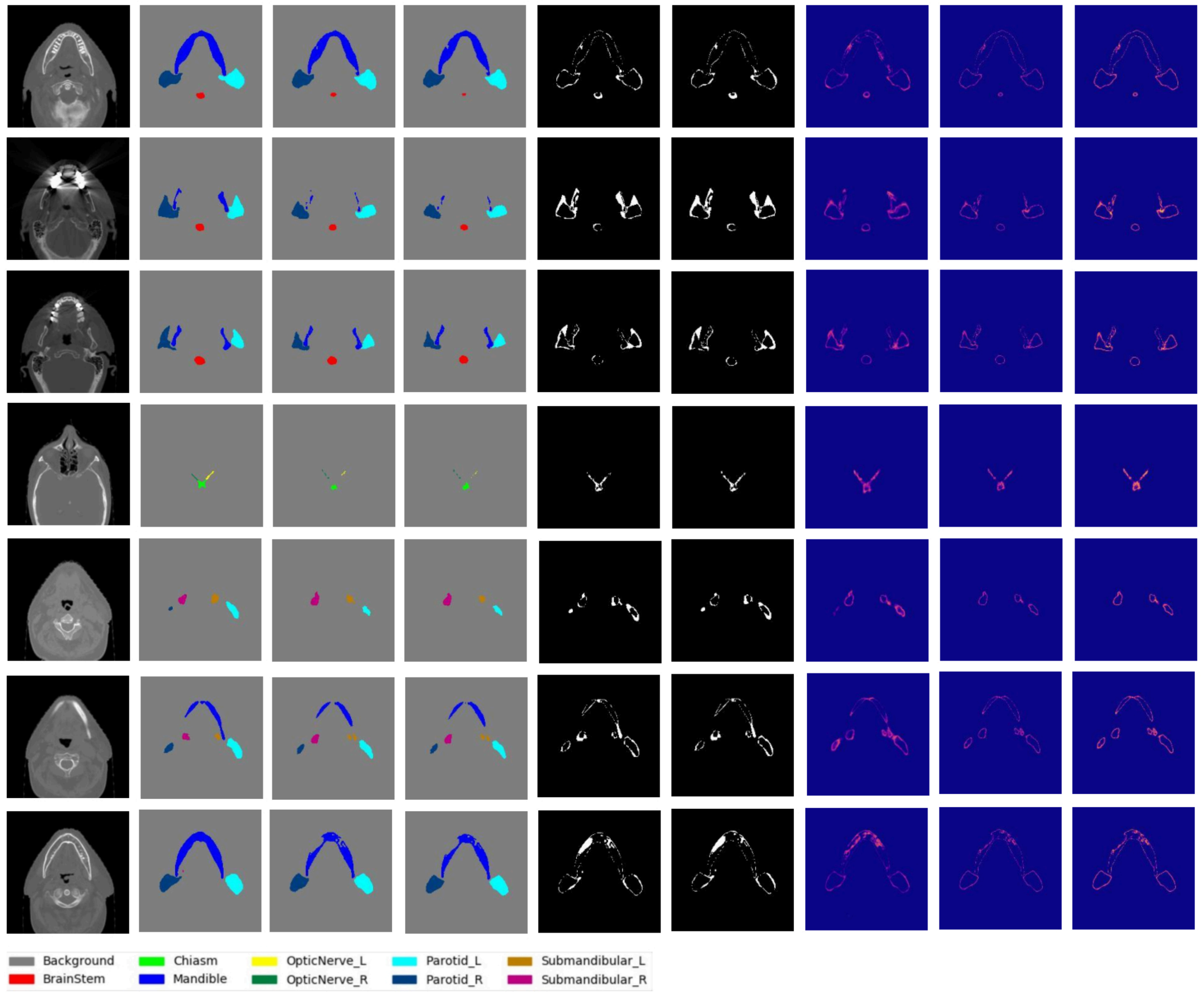
where $\mathbb{C}$ is a set of k classes , $c \in \mathbb{C}$, and p_c is the softmax output of class c. For the case of multi-label segmentation, each class is not correlated to the other classes, and the diagonal components contain values much higher than the others. Uncertainty can be approximated as trace of the estimated multivariate covariance matrix (TCV)

$$\sum_{c \in \mathbb{C}} p_c(1-p_c)$$

VE is represented as

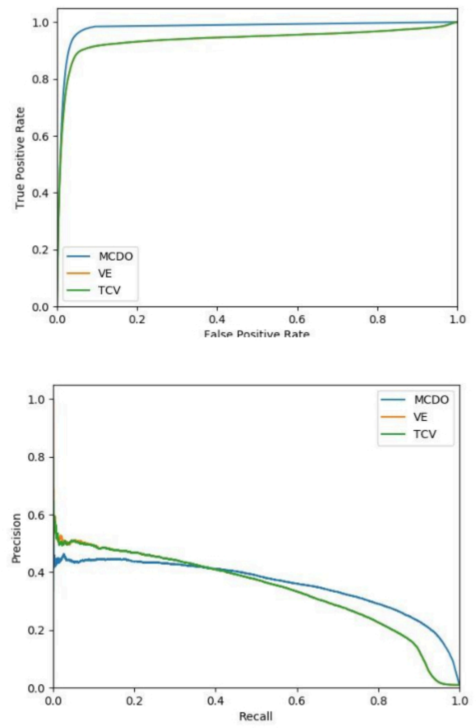
$$-\sum_{c \in \mathbb{C}} p_c \log(p_c)$$

RESULTS



The columns from left to right: Sample image, Ground-truth label, MC multi-sample based label prediction, One-shot label prediction, MC prediction error map, One-shot error map, MCDO uncertainty estimation, One-shot TCV uncertainty estimation, and One-shot VE uncertainty estimation. The images are sampled from the same head and neck volume.

For the segmentation experiment, 48 head and neck volumes were randomly split into the Training (27 volumes, 1295 slices), validation (9 volumes, 476 slices) and testing (12 volumes, 602 slices).



From top to bottom: The AUROC and Precision-Recall chart of a head and neck volume. It can be seen that the general shape of the 2 One-shot approaches are similar to MCDO approach.

The benchmark MCDO and the two experimented uncertainty indicators all corroborate with the error in labeling, with AUC >0.9. However, the required MCDO process induces both performance dependence on the number of process trajectories as well as the time cost. While the MCDO approach with 30-time inference sampling for proper statistics costs a nominal 280 milliseconds, the one-shot approach takes 12.8 milliseconds, much more efficient and amicable to online applications.

CONCLUSIONS

The proposed approach is able to achieve comparable error prediction power as the popular MCDO implementation with a much smaller computational cost. This advantage is desirable in real-time, online applications, or interactive procedures. It is also compatible to any error-driven segmentation refinement procedure.

REFERENCES

1Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding **Kendall A et a**l.. *Proceedings of the British Machine Vision Conference (BMVC), pages 57.1-57.12. BMVA Press, September 2017.*

2Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network **Cai S et al.**. *Quant Imaging Med Surg. 2020;10(6):1275-1285. doi:10.21037/qims-19-1090*

3Dropout as a bayesian approximation: Representing model uncertainty in deep learning **Gal Y et al.**. *International conference on machine learning, pp. 1050-1059. 2016.*

CONTACT INFORMATION

Email: ymin@mednet.ucla.edu