Deep Active Learning for Myelin Segmentation on Histology Data

Mélanie Lubrano di Scandalea[1,2], Christian S. Perone[1], Mathieu Boudreau[1], Julien Cohen-Adad[1,3]

[1] NeuroPoly Lab, Institute of Biomedical Engineering, Polytechnique Montreal, Montreal, QC, Canada
[2] Télécom Paristech, Paris, France
[3] Functional Neuroimaging Unit, CRIUGM, University of Montreal, Montreal, QC, Canada

Introduction

Context:

Histology data often require segmentation of microstructure (e.g. axons, myelin in spinal cord).

Problem 1:

Manual segmentation is long, tedious, and user-biased.

▶ **Solution**:

Deep learning methods offer highest performance for multi-class segmentation.

Problem 2:

Deep learning requires large amounts of data and labels to train new models (high sample complexity).

▶ **Solution**:

Active Learning [1] enables a learning algorithm to obtain **similar performance** (when compared to pure supervised learning) **with less manual labeling,** by selecting specific data points (uncertainty sampling in this work).

Goal:

Develop an open-source active learning framework for segmenting myelin from histology data based on uncertainty sampling.

Methods

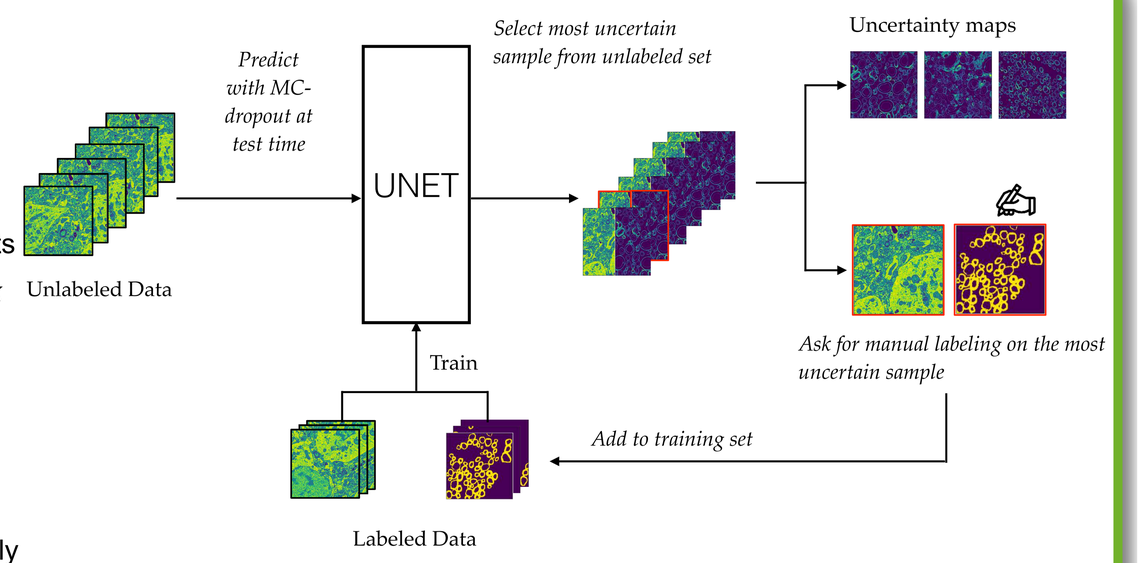
Active Learning framework

Objective:

- Minimize expert's time and effort for data labeling
- Improve a model's performances on small datasets
  - ▶ *Wisely select the samples to train the network*

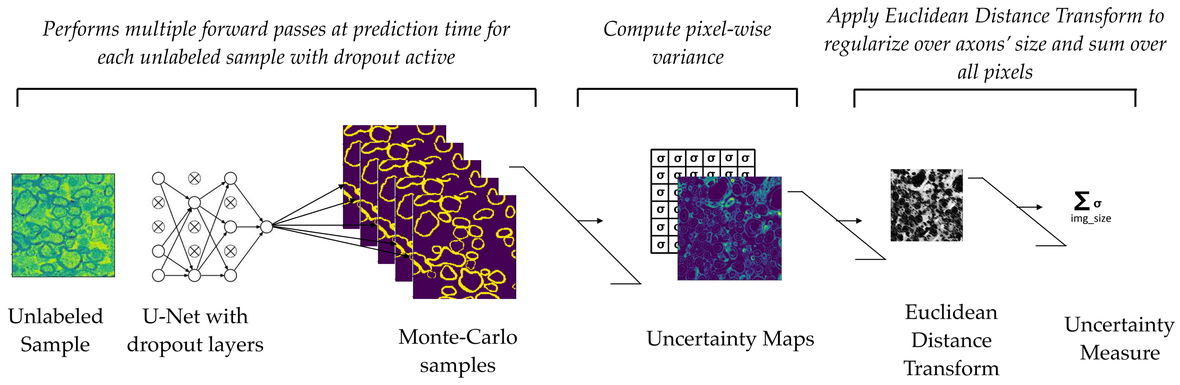
Method: *(Based on CEAL framework [2])*

1. Train U-Net [3] with SMALL initial dataset
2. Use MC-Dropout to compute uncertainty on unlabelled data and RANK it
3. Select MOST UNCERTAIN sample to be manually annotated by an oracle and add to the training set
4. Re-train U-Net from scratch, iterate



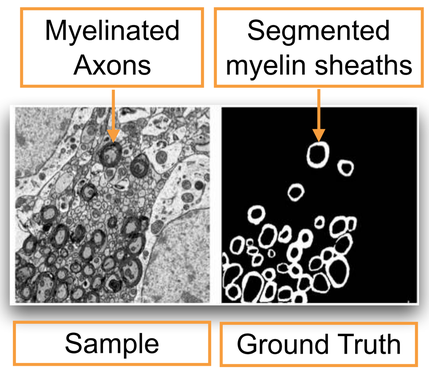
Monte Carlo Dropout

- Easy-to-use, scalable technique to evaluate uncertainty [4] [5]
- Keep dropout layers active at PREDICTION TIME
- Sample predictions to approximate posterior distribution
- Compute pixel-wise standard deviation for multiple predictions.



Results

I. Dataset



Dataset samples

Mouse brain histology observed with Transmission Electron Microscopy (TEM)

II. Myelin segmentation - TEM Dataset [6]

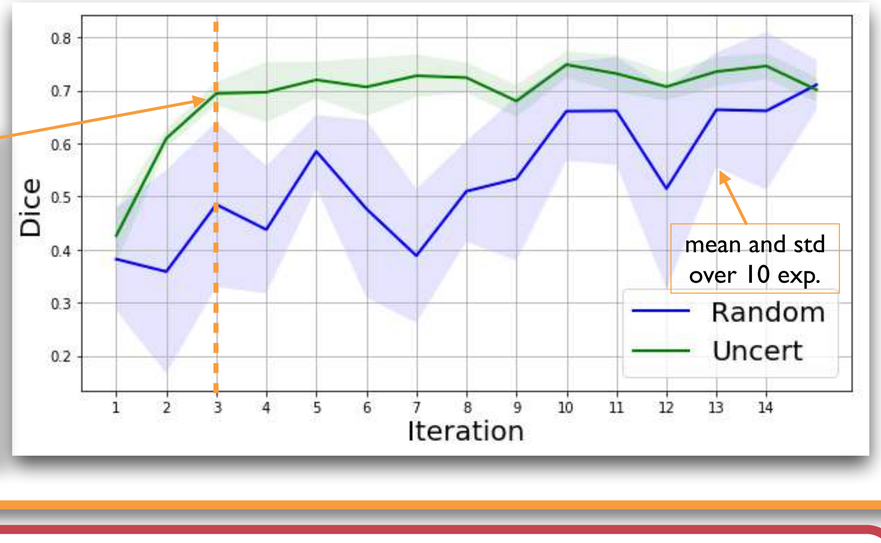
Experimental setting:

- **Initial training set size**: 5 patches (512x512 pixels)
- Adding the **most uncertain sample** (among 26 unlabeled ones) after each iteration
- **Baseline**: Random sampling
- Experiments are repeated 10 times, then averaged
- **Evaluation** with Dice coefficient (the higher the better):

$$\text{Dice} = \frac{2(A \cap B)}{|A| + |B|}$$

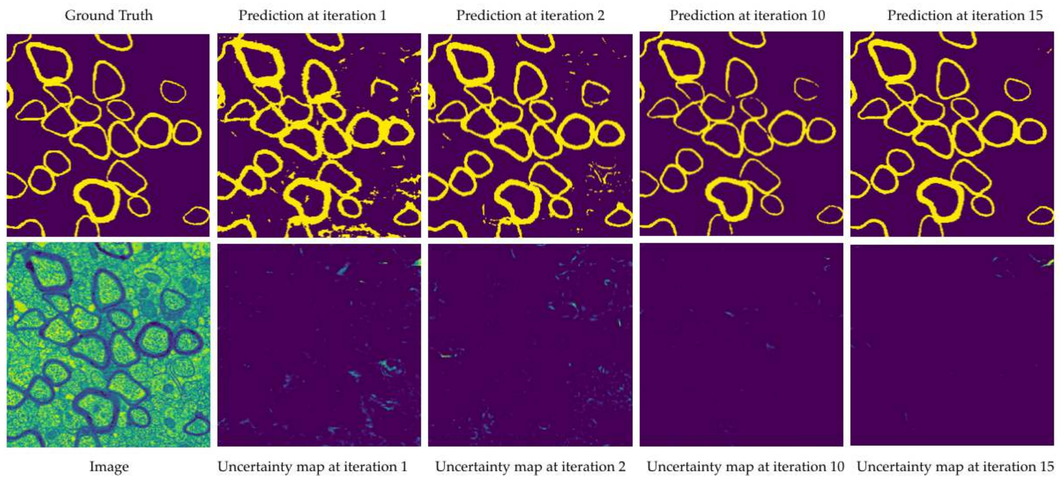
Results: Significant reduction in annotation effort for similar performance.

After only 3 manual patches, the "Uncert" reaches best performance (versus ~15 for the "Random" selection).

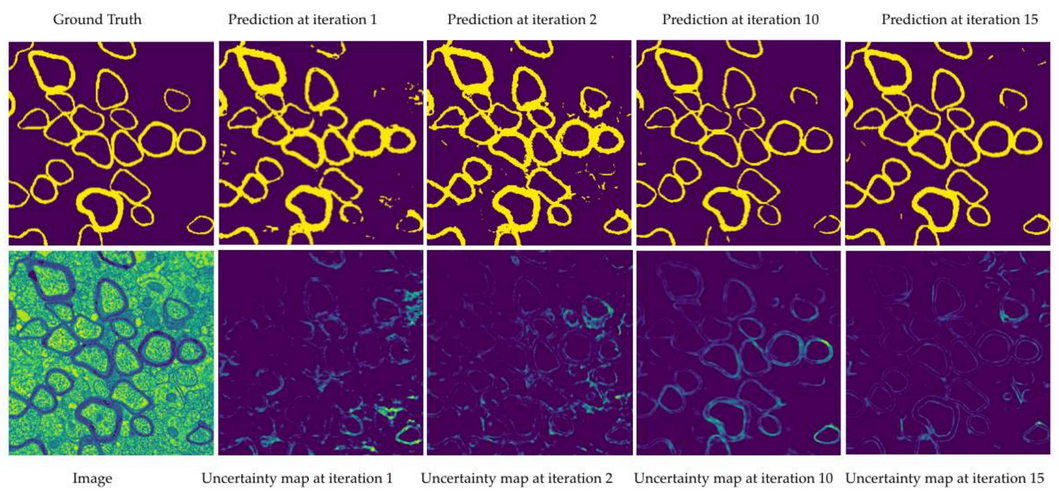


III. Uncertainty Maps

- Each pixel = *variance* of MC-samples' predictions
- Choice of *loss function* impact the *aspect* of the uncertainty maps
- Highlights areas and features on which the model tends to fail
- **Weighted cross-entropy**: uncertainties about the object borders (here: myelin thickness)
- **Dice loss**: uncertainties about other brain microstructures



Evolution of Uncertainty Map when U-Net trained with **Dice Loss**



Evolution of Uncertainty Map when U-Net trained with **Weighted Binary Cross Entropy**

Conclusion

- Active learning can converge to **high performances** with **less data**
- Although it was shown that MC Dropout uncertainty measure [5] doesn't concentrate with more data, the uncertainty metric was proved to yield good results in the active learning context.
- Uncertainty estimation is linked to the **loss function**
- Code is available as open source: https://github.com/neuropoly/deep_active_learning

Future work:

- Explore uncertainty maps benefits
- Explore other uncertainty measurements (such as anchored ensembles)
- Implement method in open-source software package *AxonDeepSeg* [7]

Acknowledgements

We thank Shawn Mikula for sharing the histology data. Funded by the Canada Research Chair in Quantitative Magnetic Resonance Imaging [950-230815], the Canadian Institute of Health Research [CIHR FDN-143263], the Canada Foundation for Innovation [32454, 34824], the Fonds de Recherche du Québec - Santé [28826], the Fonds de Recherche du Québec - Nature et Technologies [2015-PR-182754], the Natural Sciences and Engineering Research Council of Canada [435897-2013], the Canada First Research Excellence Fund (IVADO and TransMedTech) and the Quebec BioImaging Network [5886].

Contacts: Mélanie Lubrano di Scandalea - lubrano.mel@gmail.com

References

[1] Tong, S. (2001). Active learning: theory and applications. **[2]** Blanch, Marc Gorriz. 2017. "Active Deep Learning for Medical Imaging Segmentation." Edited by Xavier Giro-I-Nieto: Axel Carlier and Emmanuel Faure. Universitat Politecnica de Catalunya (UPC). **[3]** Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation.**[4]** Gal, Y. (2016). Uncertainty in Deep Learning. PhD thesis, University of Cambridge. **[5]** Osband, I. (2016). Risk versus uncertainty in deep learning: Bayes, bootstrap and the dangers of dropout. In Proceedings of the NIPS* 2016 Workshop on Bayesian Deep Learning. **[6]** Mikula, S. and Denk, W. (2015). High-resolution whole- brain staining for electron microscopic circuit reconstruction. *Nat. Methods*, 12(6):541–546. **[7]** Aldo Zaimi, Maxime Wabartha, Victor Herman, Pierre-Louis Antonsanti, Christian S. Perone, and Julien Cohen-Adad. 2018. "AxonDeepSeg: Automatic Axon and Myelin Segmentation from Microscopy Data Using Convolutional Neural Networks." Scientific Reports 8 (1): 3816.