CONTEXTUAL CONVOLUTIONAL NEURAL NETWORK FILTERING IMPROVES EM IMAGE SEGMENTATION

Xundong Wu, Yong Wu, Ligia Toro & Enrico Stefani
Department of Anesthesiology
University of California, Los Angeles
{xundong,wuyong,ltoro,estefani}@ucla.edu

ABSTRACT

We designed a contextual filtering algorithm for improving the quality of image segmentation. The algorithm was applied on the task of building the Membrane Detection Probability Maps (MDPM) for segmenting electron microscopy (EM) images of brain tissues. To achieve this, we executed supervised training of a convolutional neural network to recover the ground-truth label of the masked-out center pixel from patches sampled from an un-refined MDPM. Through this training process the model learns the distribution of the segmentation ground-truth map . By applying this trained network over MDPMs we are able to integrate contextual information to obtain with better spatial consistency in the high-level representation space. By iteratively applying this network over the MDPMs for multiple rounds, we were able to significantly improve the EM image segmentation results.

1 INTRODUCTION

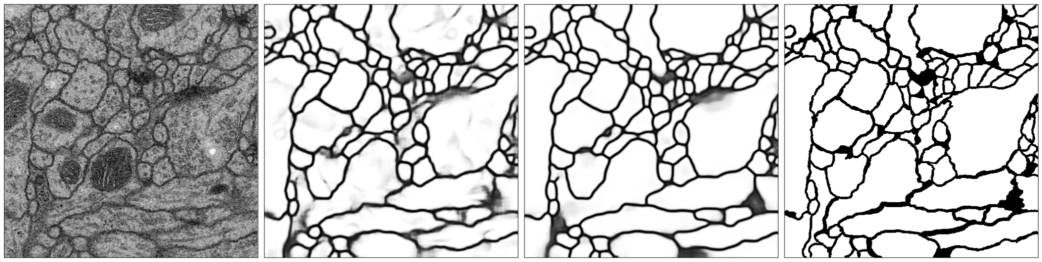
To further expand our understanding on the structure and the working mechanisms of the human brain, it is necessary to map the entire neural connections of the nervous system at the micro-scale level. One main approach is to acquire serial 2-D images of brain tissues at nanometric resolution with serial-section Transmitted Electron Microscopy (ssTEM). Much effort has been made to develop tools to automatically process those images. Previous approaches Ciresan et al. (2012)Jurrus et al. (2010)Jain et al. (2007) use the contextual information surrounding a pixel to assign a probability value of it representing the cell membrane. Applying those detectors to every pixel of an original EM image leads to a pixel wise Membrane Detection Probability Map (MDPM). Post-processing on top of these detection maps is necessary in order to obtain the final region segmentation results. The post-processing methods can be simply, for example, simple thresholding, or creating a smoother probability map by using a median filter Ciresan et al. (2012).

In this work, we simply by training an Iterative Convolutional Neural Network (I-CNN) to recover the masked-out center pixel value of patches sampled from MDPMs, and then iteratively applying this network over the resultant MDPMs, one can obtain a high quality segmentation map both visually and measured by the foreground-restrict Rand score Arganda-Carreras et al. (2015); Jain et al. (2010).

2 RELATED WORKS

Computer vision research on real-world image contour detection and segmentation tasks has come up with many solutions to ensure the consistency of image segmentation. For the EM image segmentation task, the detection of membrane resembles the contour detection problem in general computer vision. The quality of membrane detection can be directly measured by the pixels error: but just like in contour detection, a high quality membrane detection does not guarantee a good segmentation Arbelaez et al. (2011). A small gap in the contour formed by the detected membrane can lead to an incorrect merge of two different regions, or a false section of membrane can incorrectly split one region into two parts.

In a typical MDPM such as the one from Ciresan et al. (2012), the detection probability of every pixel is made in isolation. In Ciresan et al. (2012), the authors used a simple median filter with



(a) Original image (b) Original P map (c) Round 8 (d) Ground truth

Figure 1: One sample image processed by I-CNN. Probability map is shown in reverse manner. The darkest pixel value means that the probability of being membrane is 1. Original image is a raw EM image; Original P map is the MDPM output from the base CNN network. Round 8 are the maps processed by I-CNN for 8 rounds. Ground truth is the map labeled by a human expert.

a small radius to smooth out the detection map. However, the limitation of applying a median filter is its isotropy: if one applies a median filter with an increasing radius to the MDPMs the performance will quickly deteriorate as the radius becomes larger. Therefore, an algorithm that can take long-range information and avoid isotropy smoothing was needed. To achieve this goal, we have developed an Iterative Convolutional Neural Network (I-CNN) that significantly improves the definition of boundaries. In Jurrus et al. (2010); Pinheiro & Collobert (2013); Lee et al. (2015); Tu (2008); Tu & Bai (2010), they applied a network that conditioned on both the raw image features and the previous round label map to recursively refine the label maps. Our approach differs from theirs in that the probability maps are refined without directly conditioned on the raw image features.

3 System description and result

The dataset used in this experiment consists of two stacks of EM images used in the ISBI 2012 EM segmentation challenge. One stack has 30 EM images and their corresponding labels for training. The other stack also contain 30 images, and their labels are concealed.

The network we used in this stage is analogous to the convolutional network implemented by Ciresan et al. (2012). To differentiate the base network and the iterative secondary network, we call this base network CNN and the secondary Network I-CNN. All detail about both CNN and I-CNN networks training will be released with the code very soon. Next, we will describe our approach to refine the probability map. For this task, we generated an non-overfitted MDPM training set through cross training.

The pixel-wise probability map generated from the network described in the last section shows high pixel-wise accuracy yet it was short of local consistency in certain areas (see Figure1, 2), which is inconsistent with the spatial continuity of the cell membrane. In the approach described in the last two sections, although contextual information is used to generate the pixel detection probability values, those probability values are generated independently. Here we propose a simple convolutional network (I-CNN) which directly learns the statistics of segmentation maps to significantly improve the segmentation quality. The main difference between I-CNN network and the previous CNN is input. In the CNN network, the input to the network is the raw EM image. In I-CNN, the input image patches are replaced by the patches extracted from MDPM, while the label of the center pixel is masked out.

4 Result

As shown in Figure1, when we iteratively applied the I-CNN to the MDPM, the first thing we notice is that by iteratively refining the probability, we removed the noise in the map. After 8 rounds, the map turned out to be a map with rather clear boundaries as opposed to the fuzzy boundaries in the original MDPM. If we zoom into areas where the CNN was unable to make an affirmative inference about pixels, as shown in Figure2, we can see that the I-CNN is able to integrate the information in

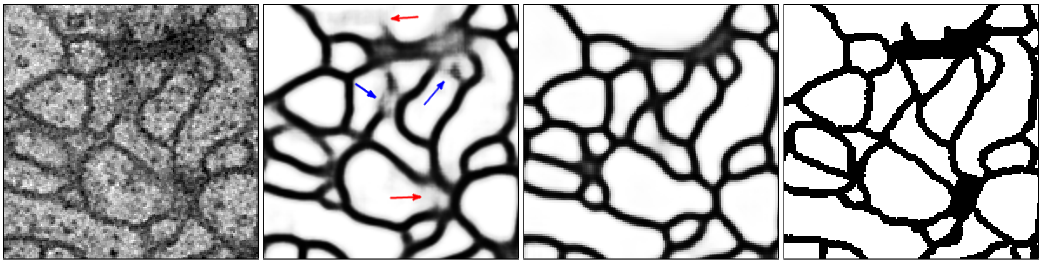


Figure 2: Examples of patches where the model closes gaps and removes uncertain membrane sections. Blue arrows indicate where the model adds or solidifies a section of link between membrane parts; red arrows indicate where the model removes sections of uncertain membrane.

a neighbourhood to recognize (blue arrows) a section of membrane shown with low probability but with good spatial continuity, and eventually label the section with high confidence and closes gaps at the boundary. At the same time, the I-CNN was able to identify noise pixel and areas (Red arrow) that do not appear like a section of membrane and eventually completely removed it.

We then measured the segmentation result by the Rand score used by the ISBI 2012 challengeArganda-Carreras et al. (2015). For this part of experiment, trained I-CNNs were applied to left-out validation MDPM set; their segmentation error scores were then measured for every round by the Rand score. We observed that by iteratively applying the I-CNNs on the membrane detection map one can dramatically reduce the Rand error of the segmentation result at the beginning of the iteration. This reduction in the Rand error disappears only after about 6 rounds of iterations and afterwards deteriorates. We also applied the I-CNN to the test image stack submitted the result to the ISBI 2012 challenge website obtaining a Rand error score of 0.0263, which is much better than the score of 0.0551 obtained from the original CNN result before refining.

We also applied our secondary network to a set of MDPMs of better qualify from Chen et al. (2016). Even though our network was not trained to process the exact same kind of data, the refine process still managed to significantly reduce the Rand error from 0.0351 to 0.0255. Furthermore, according to a recent update from the organizer of the ISBI 2012 Arganda-Carreras et al. (2015), with their new evaluation method (thinned rand score), our post-processing obtained a thinned Rand score of 0.9765, which is just very small fraction behind 0.9768, the score of Chen et al. (2016). This indicates that our approach performed in par with the water-shed algorithm used by Chen et al. (2016) when measured with the new Rand score.

5 Conclusion and Discussion

The new algorithm presented in this work learns the manifold of membrane morphology distribution; it enforces these constraints through iteration on a MDPM, refining it to fit a membrane morphology distribution learned from the training data. From another perspective, instead of generating a membrane detection probability of every pixel in isolation, we congregated information in the local neighbourhood through applying the I-CNN iteratively to the MDPMs and obtained significantly better consistency in neighbouring pixels. A significant improvement, measured by the Rand error, was achieved over the original MDPM result. It is also important to point out that training with the ground-truth map directly provided no benefit in improving the segmentation quality. It seems to be essential to learn the gradient field that can guide a raw MDPM toward the manifold of the ground-truth maps by train from un-refined MDPM.

Acknowledgments

We would like to thank Dr. Riccardo Olcese for helpful discussions. We are also gratefully for the support of NVIDIA Corporation with the donation of two GPUs. This research was supported by NIH R01HL107418 grant.

References

Björn Andres, Ullrich Köthe, Moritz Helmstaedter, Winfried Denk, and Fred A Hamprecht. Segmentation of sbfsem volume data of neural tissue by hierarchical classification. In *Pattern recognition*, pp. 142–152. Springer, 2008.

Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(5):898–916, 2011.

Ignacio Arganda-Carreras, Srinivas C Turaga, Daniel R Berger, Dan Cireşan, Alessandro Giusti, Luca M Gambardella, Jürgen Schmidhuber, Dmitry Laptev, Sarvesh Dwivedi, Joachim M Buhmann, et al. Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in neuroanatomy*, 9, 2015.

Davi D Bock, Wei-Chung Allen Lee, Aaron M Kerlin, Mark L Andermann, Greg Hood, Arthur W Wetzel, Sergey Yurgenson, Edward R Soucy, Hyon Suk Kim, and R Clay Reid. Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471(7337):177–182, 2011.

Albert Cardona, Stephan Saalfeld, Stephan Preibisch, Benjamin Schmid, Anchi Cheng, Jim Pulokas, Pavel Tomancak, and Volker Hartenstein. An integrated micro-and macroarchitectural analysis of the drosophila brain by computer-assisted serial section electron microscopy. *PLoS biology*, 8 (10):2564, 2010.

hao Chen, Xiaojuan Qi, Jiezhi Cheng, and Phengann Heng. Deep contextual networks for neuronal structure segmentation. In *Thirtieth AAAI Conference on Artificial Inteligence*. AAAI, 2016.

Dmitri B Chklovskii, Shiv Vitaladevuni, and Louis K Scheffer. Semi-automated reconstruction of neural circuits using electron microscopy. *Current opinion in neurobiology*, 20(5):667–675, 2010.

Dan Ciresan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pp. 2843–2851, 2012.

Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In *Computer Vision–ECCV 2014*, pp. 184–199. Springer, 2014.

Clement Farabet, Camille Couprie, Laurent Najman, and Yann LeCun. Scene parsing with multiscale feature learning, purity trees, and optimal covers. *arXiv preprint arXiv:1202.2160*, 2012.

Jan Funke, Björn Andres, Fred A Hamprecht, Albert Cardona, and Matthew Cook. Multi-hypothesis crf-segmentation of neural tissue in anisotropic em volumes. *Arxiv preprint*, 2011.

Stephen Gould, Richard Fulton, and Daphne Koller. Decomposing a scene into geometric and semantically consistent regions. In *Computer Vision, 2009 IEEE 12th International Conference on*, pp. 1–8. IEEE, 2009.

Viren Jain and Sebastian Seung. Natural image denoising with convolutional networks. In *Advances in Neural Information Processing Systems*, pp. 769–776, 2009.

Viren Jain, Joseph F Murray, Fabian Roth, Srinivas Turaga, Valentin Zhigulin, Kevin L Briggman, Moritz N Helmstaedter, Winfried Denk, and H Sebastian Seung. Supervised learning of image restoration with convolutional networks. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pp. 1–8. IEEE, 2007.

Viren Jain, Benjamin Bollmann, Mark Richardson, Daniel R Berger, Moritz N Helmstaedter, Kevin L Briggman, Winfried Denk, Jared B Bowden, John M Mendenhall, Wickliffe C Abraham, et al. Boundary learning by optimization with topological constraints. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 2488–2495. IEEE, 2010.

Elizabeth Jurrus, Antonio RC Paiva, Shigeki Watanabe, James R Anderson, Bryan W Jones, Ross T Whitaker, Erik M Jorgensen, Robert E Marc, and Tolga Tasdizen. Detection of neuron membranes in electron microscopy images using a serial neural network architecture. *Medical image analysis*, 14(6):770–783, 2010.

Kisuk Lee, Aleksandar Zlateski, Ashwin Vishwanathan, and H Sebastian Seung. Recursive training of 2d-3d convolutional networks for neuronal boundary detection. *arXiv preprint arXiv:1508.04843*, 2015.

Ting Liu, Elizabeth Jurrus, Mojtaba Seyedhosseini, Mark Ellisman, and Tolga Tasdizen. Watershed merge tree classification for electron microscopy image segmentation. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pp. 133–137. IEEE, 2012.

Ting Liu, Mojtaba Seyedhosseini, Mark Ellisman, and Tolga Tasdizen. Watershed merge forest classification for electron microscopy image stack segmentation. In *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pp. 4069–4073. IEEE, 2013.

Pedro HO Pinheiro and Ronan Collobert. Recurrent convolutional neural networks for scene parsing. *arXiv preprint arXiv:1306.2795*, 2013.

Chris Russell, Pushmeet Kohli, Philip HS Torr, et al. Associative hierarchical crfs for object class image segmentation. In *Computer Vision, 2009 IEEE 12th International Conference on*, pp. 739–746. IEEE, 2009.

Olaf Sporns, Giulio Tononi, and Rolf Kötter. The human connectome: a structural description of the human brain. *PLoS Comput Biol*, 1(4):e42, 2005.

Joseph Tighe and Svetlana Lazebnik. Superparsing: scalable nonparametric image parsing with superpixels. In *Computer Vision–ECCV 2010*, pp. 352–365. Springer, 2010.

Zhuowen Tu. Auto-context and its application to high-level vision tasks. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8. IEEE, 2008.

Zhuowen Tu and Xiang Bai. Auto-context and its application to high-level vision tasks and 3d brain image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(10): 1744–1757, 2010.