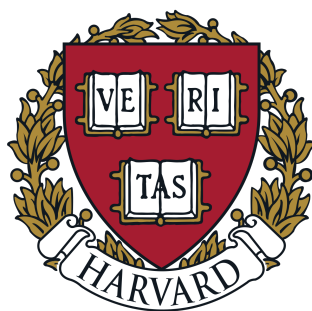# Summer Intern Work Report

Chen Xupeng

Supervisor:
Prof. JEFF LICHTMAN

July 17, 2018

Contents

1 Detailed progress and codes

I have create a category in my website to record more **Detailed Progress in Website**.
Also I have put all my codes in my github including four tasks: **NMJ, Synapse prediction, Synaptic partner, Synapse cluster**. **Project Codes**

2 NMJ

NMJ Detail progress in website

2.1 Dense segmentation

Since the new data is still to be processed, I spent several days doing dense segmentation work both for study and future training.

I have done 25 sections dense segmentation in W12-W14 for 4 days(7.5-7.8), it includes dense segmentation of Axons, Schwann cell, and Schwann cell nucleus.

I have written codes to use python to visualize animation of the 25 segments (segmen animation). Since the computational work use more python codes know, I also shared my animating code with others. (plot segment script)

I also read several articles Marco and Yaron recommened, including previous NMJ work, some segmentation and Connectome processing pipeline papers.

2.2 Future work

We also discuss a lot about the future plan of the project. Since it is more challenging than other tasks, it seems we are a little slow in progress. We have worked with Marco to find a way to label the ROI and use a script to extract coordinates of the bounding box. We have labeled one mask, later we will test Adi's align results and generate more.

- Manually create ROI region for bundles and NMJ for alignment - Do segmentation and **statistical analysis** work on some NMJs(concerning our limited staying time, it seems there isn't enough time to wait for all NMJs' alignment and segmentation results to analyze)

3 Synapse Prediction

Synapse Prediction Detail progress in website
I have spent the leisure time doing some computational work. Currently I am working with donglai and zudi on Synapse project. We would like to include several datasets including CREMI, JWR, Cerebellum data and incorporate the whole pipeline. I will help to refine the deep learning model to predict synapse better(getting higher in CREMI leader board) and

design models for synaptic partner project. Also we would like to do some analysis work if we have a much bigger datasets with hundreds of thousands of synapse to analyze.

3.1 Dataset visualization

I have visualized the CREMI datasets including all images and annotations. (visualization cremi1) (visualization cremi2) (visualization cremi3) I also helped siyan with server and coding environment and visualization work.

3.2 Alignment script

I learned how to align and process the CREMI data. Previously we use a matlab code to do alignment and transformation. Since the whole project is based on python, I study and rewrite the script in python. (cremi shift visualization) (cremi align in python)

3.3 Data augmentation

I finished zudi's assignment for me to incorporate data augmentation into the synapse prediction problem. Funkey group from Janelia is now top at the CREMI leader board. I have studied their prediction method paper and their codes in github.

It is hard to call the API of funkey's data augmentation function since it is based on python2 and Caffe. I spent two days study their codes and compare with Sebastian group's data processing codes since these two have the best data augmentation pipeline.

I rewrite the four kinds of data augmentation methods combining Funkey and Sebastian's codes.(gunpowder) (gunpowder on cremi)

- simple augmentation

- intensity augment

- ElasticAugment

- DefectAugment

Then I incorporate the data augmentation methods into synapse prediction model for further use. (augmentation on cremi)

3.4 Model test and redesign

I tried zudi's synapse prediction model using 3D U-net, since it is based on pytorch, I spent some time to read the pytorch documentation. We have discussed about the model improvement issue and I have tried to use other architecture for better prediction results. (synapse prediction model), (improvement on U net)

4 Synaptic Partner

Synaptic Partner Detail progress in website

We have discussed and agreed to try to build a model for CREMI task3: identify synaptic partner.

To get familiar with task3, I tried to read 2 state of art paper about task3. I summarize and read the 2nd one's codes using 3D U-net to propose potential partner and using 3D CNN to prune.

I also study and rewrite the metrics calculation codes written by CREMI organization. It is a little complex in synaptic partner identification challenge, so rewriting the codes helps me understand what we need to do in challenge three. (study synaptic partner)

5 Synapse Cluster

Synapse Cluste Detail progress in website I did some data visualization for cerebellum and CREMI data which will be used in clustering work. Also implement some computer vision algorithm for feature extraction.(visualize cerebellum)

I am considering to use deep learning based clustering methods, to do feature selection and clustering simultaneously. Also it is essential to consider model interpretability.

6 References

[1]Tapia, Juan C., et al. "Pervasive synaptic branch removal in the mammalian neuromuscular system at birth." Neuron 74.5 (2012): 816-829.

[2]Kasthuri, Narayanan, et al. "Saturated reconstruction of a volume of neocortex." Cell 162.3 (2015): 648-661.

[3]Matveev, Alexander, et al. "A multicore path to connectomics-on-demand." ACM SIGPLAN Notices. Vol. 52. No. 8. ACM, 2017.

[4]Pape, Constantin, et al. "Solving Large Multicut Problems for Connectomics via Domain Decomposition." ICCV Workshops. 2017.

[5]Heinrich, Larissa, et al. "Synaptic Cleft Segmentation in Non-Isotropic Volume Electron Microscopy of the Complete Drosophila Brain." arXiv preprint arXiv:1805.02718 (2018).

[6]Beier, Thorsten, et al. "Multicut brings automated neurite segmentation closer to human performance." Nature Methods 14.2 (2017): 101.

[7]Lee, Kisuk, et al. "Superhuman accuracy on the SNEMI3D connectomics challenge." arXiv preprint arXiv:1706.00120 (2017).

[8]Ding, Jiarui, Anne Condon, and Sohrab P. Shah. "Interpretable dimensionality reduction of single cell transcriptome data with deep generative models." Nature communications 9.1 (2018): 2002.