Research White Paper

Hanbo Chen

My research interest centers on one core theme: understanding the architecture of brain connectome. It is believed that the brain circuitry made up of billions of neurons is where the mind originates. Thus to fully understand brain architectures, a comprehensive investigation on its connectome is the first step. Around this theme, I studied different aspects of the problem such as: brain network alternations in diseased brain, correlations between functional and structural connectivity, brain evolution and development mechanisms, by using various imaging modalities (MRI, fMRI, DTI, neuron tracing, microscopy images, genomic images, etc.) and analyzed different animal models (drosophila, mouse, macaque, chimpanzee, and human). In addition to the many exciting findings on brain connectome, the other aspect of my contribution lies in a set of smart computational methods designed to tackle various challenges associated with **big neuroimaging data**. Notably, "big" not only indicates the size of data or the number of subjects, but also infers to a huge variety of image modalities, which offer rich and complementary views of the brain. My attempts in solving this big data problem are discussed from the following 3 aspects:

- Design computational framework for group-wise analysis between different populations and different species.
- Bridge the gap between different imaging modalities on different resolution scales.
- Develop computational tools for accurate, efficient and high-throughput neuroimaging analysis.

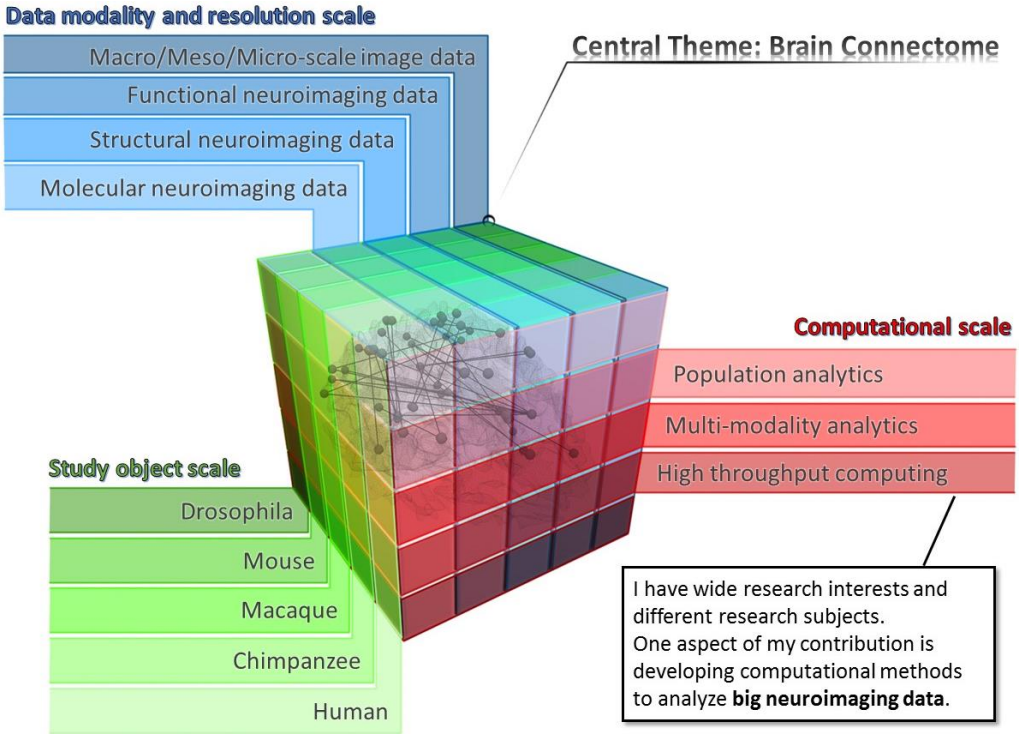

Figure 1. Summary and overall vision of my research topics.

1 GROUP-WISE ANALYSIS BETWEEN DIFFERENT POPULATIONS

Defining reliable, reproducible and accurate brain regions of interest (ROIs) is the first and fundamentally important step in performing group-wise analysis of brain connectome. Traditional approaches rely on registering individual brains to a common template brain with atlas derived from domain knowledge (E.g. Brodmann area) to define common ROIs. However, automatic registration algorithm has limited accuracy in mapping ROIs due to the variability across individuals. Although more precise ROIs can be obtained by manual annotation of experts, it is not applicable for large populations. Moreover, finer parcellation of ROIs beyond current domain knowledge are required for accurate mapping of functional ROIs (E.g. pre-central gyrus should be further parcellated into different functional regions).

Facing this challenge, my current lab directed by Prof. Tianming Liu has made great progress on defining brain ROIs [1] in the past 5 years and a good portion of my research contributes to this trend. Together with my colleague Dr. Dajiang Zhu and Dr. Kaiming Li, we reproducibly and accurately identified and predicted brain ROIs with consistent structural connectivity across individuals [2], [3]. Specifically, we defined a set of accurate and efficient descriptors for brain connectivity profiling. Then based on a data-driven exhaustive search, regions with common connectivity profiles were identified. For instance, in my works, 12 consistent white matter hubs were reproducibly identified [2] and pre-central/post-central gyri were consistently parcellated into multiple sections [4]. I also included different metrics to further refine ROIs including: functional

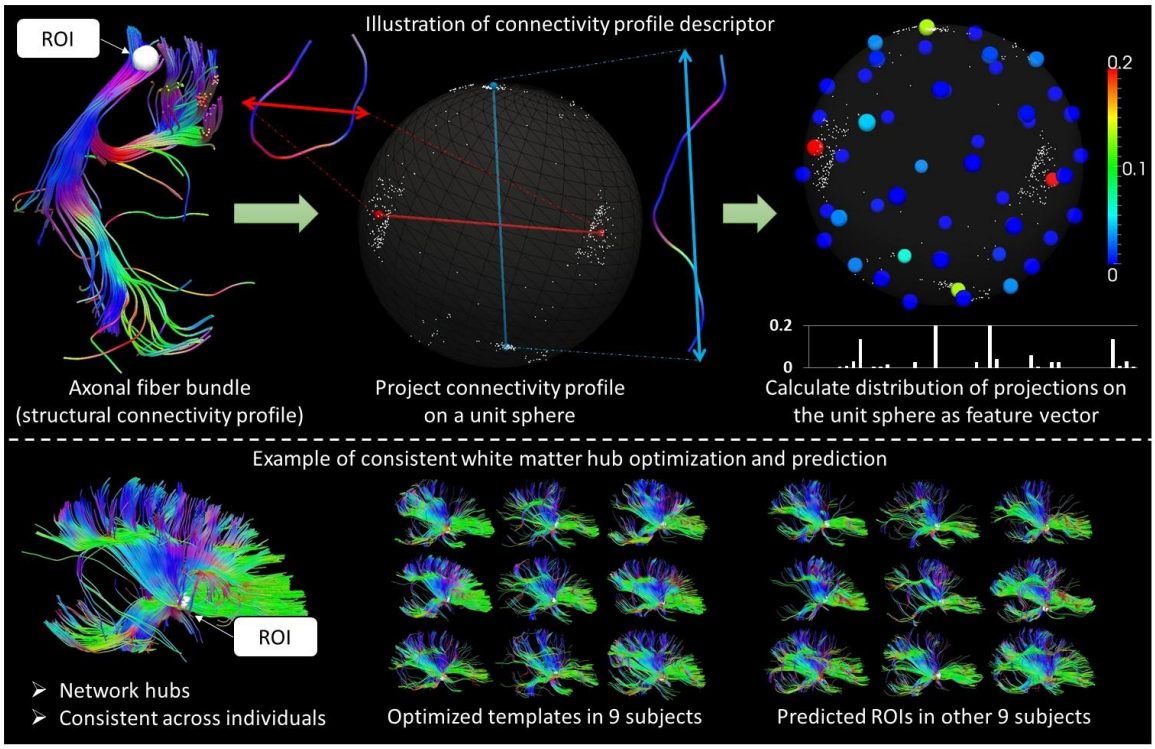

Figure 2. Illustration of connectivity profile descriptor for ROI detection and an example of consistent white matter hub identified as brain ROI [2].

connectivity profile [5], inter-regional connectivity profile [6], and global distance constrain [7]. These ROI systems will serve as a GPS map of brains. In the future, the similar approach can be extended to identify ROIs that are not only consistent across individuals but also consistent across species.

According to the 'fingerprint' concept suggested by Passingham that each brain's cytoarchitectonic area has a unique set of extrinsic inputs and outputs that largely determines the functions that each brain area performs, those ROIs with consistent structural connectivity are ideal for both functional and structural connectome studies. With the identified ROIs as foundations, I studied the regularity and variability of both functional and structural connectome across populations in both healthy brains and diseased brains. Specifically, I adopted a novel algorithm called multi-view spectral clustering to fuse functional connectome and structural connectome across individuals to identify common brain network modules [8]. Based on this framework, I performed longitudinal and group-wise analysis on mild traumatic brain injury subjects and found that after brain injury decreased inter-regional structural connectivity will be complemented by increased functional connectivity [9], [10]. In the future, the method can be widely applied on larger database analysis which includes many more subjects of variety brain diseases across different imaging sites for a deep understanding of abnormal connectome in mental disorders.

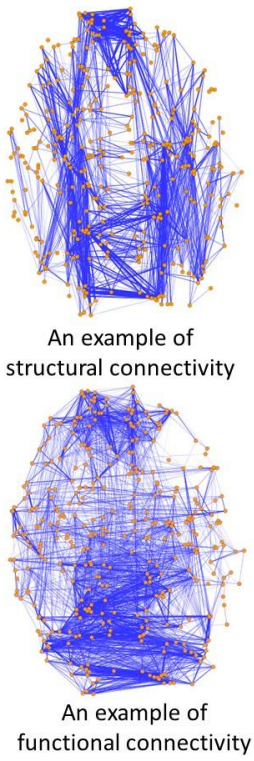

Figure 3. examples of brain connectome based on our ROI system [8].

Besides model driven approaches, machine learning algorithms are also powerful tools in identifying brain regions of interest [11], [12]. Specifically, in collaboration with Yu Zhao (PhD student I supervised), we have developed a large-scale computing framework to study functional network alternations in Autism patient brains [13]. By applying sparse representations and dictionary learning on resting-state fMRI images, brain was decomposed into hundreds of functional networks. After pair-wisely comparing decomposed networks across 178 subjects, we identified 144 group-wisely common functional networks. Further analysis on these networks showed abnormality in both inner-network connectivity and inter-network interactions. Traditional fMRI analysis only focused on a few brain networks such as default mode networks. The novel techniques proposed in this work may revolute future fMRI studies on brain functional connectome such that a more comprehensive understanding of brain functional networks than before is possible. In another work with Yujie Li (PhD student I supervised), we found that machine learning approach can also identify ROIs from gene expression data [14]. By applying the data-driven sparse coding framework on genome-wide in situ hybridization image data set released by Allen Mouse Brain Atlas, we identified transcriptome organization of mouse brain that displayed robust regional specific molecular signatures and corresponded to canonical neuroanatomical

subdivisions consisting of discrete neuronal subtypes. In the future, similar methods can be applied on different imaging modalities.

Moreover, it is also important to correlate those ROIs obtained from different imaging modalities. Traditionally, brain ROIs are defined and predicted based on a single image modality (e.g. manually annotated on averaged MRI templates). My colleagues have showed that by using structural connectivity consistency as constrain, we can more accurately identify brain ROIs computed from task fMRI data. My work further showed that by including the predictability of functional connectivity from structural connectivity into object function, the brain ROIs initially obtained based on structural information can be optimized to preserve better anatomical consistency across individuals [5]. In the future, it is necessary and much more meaningful to define and optimize brain ROIs by include joint information from multiple modalities such as functional connectivity (fMRI), structural connectivity (DTI), anatomical information (MRI, Nissle stain), and gene expression data.

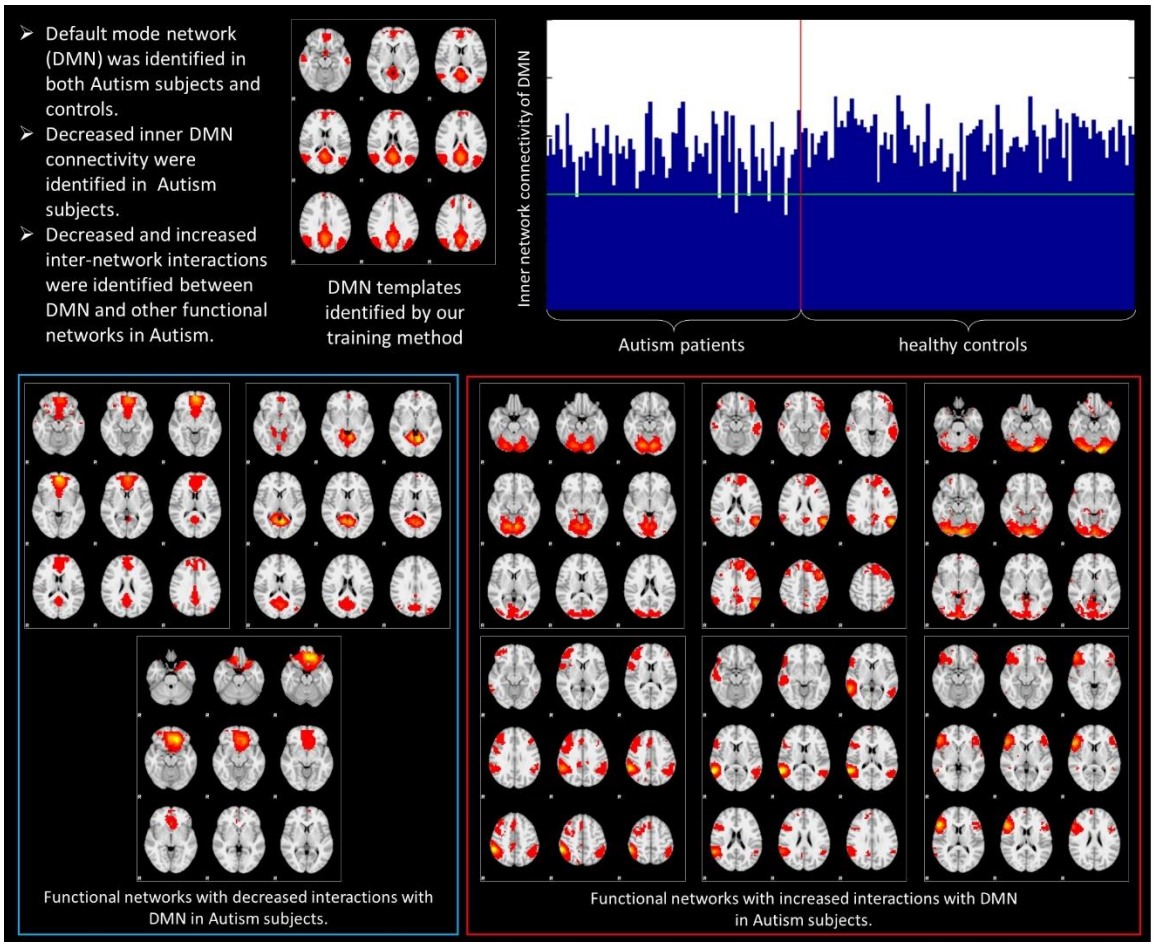

Figure 4. Functional network alternations related to default mode network in Autism subjects identified based on our novel data-driven approach [13].

2 MULTI-SCALE, MULTI-MODEL, ACROSS SPECIES NEUROIMAGING ANALYSIS

2.1 COMPARE FUNCTIONAL AND STRUCTURAL CONNECTOME

If we view brain as a computer, then the structural (anatomical) connectome is the hardware of brain and the functional connectome is the software in brain. Most of the previous studies on mental disorders largely rely on a single imaging modality and thus only focused on one aspect of the brain. However, a comprehensive understanding of brain requires viewing brain from both aspects. Based on the group-wise analysis approach introduced in the previous section, it has been found that structural connectivity is more consistent across individuals in comparison with the functional connectivity [8]. This finding is expected given that the functional connectivity may dynamically alter between different brain states. By fusing the connectome across individuals, we can reproducibly identify brain network clusters for both functional and structural connectivity among different groups of subjects. This framework offered new insight on the joint analysis of brain connectome and has been successfully applied on mTBI studies [9]. In the future, the method will be extended to the analysis of different brain disorders as well as joint analyses across multiple brain diseases.

2.2 CROSS VALIDATE FINDINGS BETWEEN DIFFERENT RESOLUTION SCALES

Advanced neuroimaging techniques allow researchers to investigate brain on different scales (molecular-scale, micro-scale, meso-scale, and macro-scale) which offer complementary pictures of the brain. Coarser scale images enable a global and population-wise view of the brain, while finer scale images carry more details and can be used to validate findings on coarser scales. For instance, diffusion imaging and tractography have been widely used to investigate structural

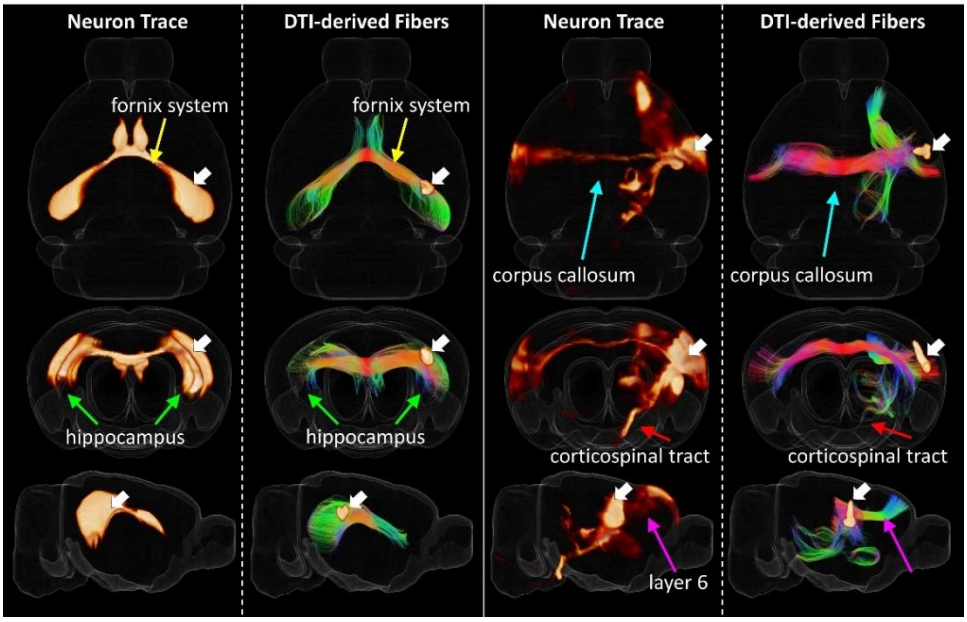

Figure 5. Two examples of comparing neuron trace and DTI-derived axonal fiber bundles [15].

connectome on macro-scale. Despite its capability in none-invasively capturing whole brain structural networks, the technique has been questioned for its limitations and biases for a long time. By comparing whole brain connectome derived from DTI tractography with the connectome derived from meso-scale neuron tracing study released by Allen Mouse Connectivity Atlas, my recent study has showed that a limited agreement between two scales can be achieved with gains in reliability and sensitivity that depend on the selection of optimal tractography parameters and parcellation schemes [15]. In the future, this framework can be applied in designing better DTI acquisition parameters and tractography models.

2.3 LINK FINDINGS ACROSS SPECIES

The rationale of studying and comparing neuroimaging data across species are mainly two-fold: 1) the differences across species allow us to investigate how brain evolves and how human-specific brain functions emerge; 2) the common brain mechanisms and anatomical regions across species allow us to investigate human brain based on animal models. In my previous studies in collaboration with Dr. Tuo Zhang, by analyzing MRI and DTI data across macaque, chimpanzee, and human brain, we found that the complexity of wiring patterns and folding patterns jointly evolves across species [16], [17]. Further investigations showed that fiber orientation coincides with the bending direction of gyral crest. This observation is consistent across the brains of three species we analyzed and support the theory that the wiring progress may regulate cortical convolution during brain development. In another study in collaboration with Xiao Li (graduate student I supervised), we analyzed the gyral folding patterns across these three species [18] and quantified both preserved and evolved patterns. In the future, this comparison can be extended to more animal models such as mouse and rat brain. And we should also include more imaging modalities into analysis such as functional data or molecular data.

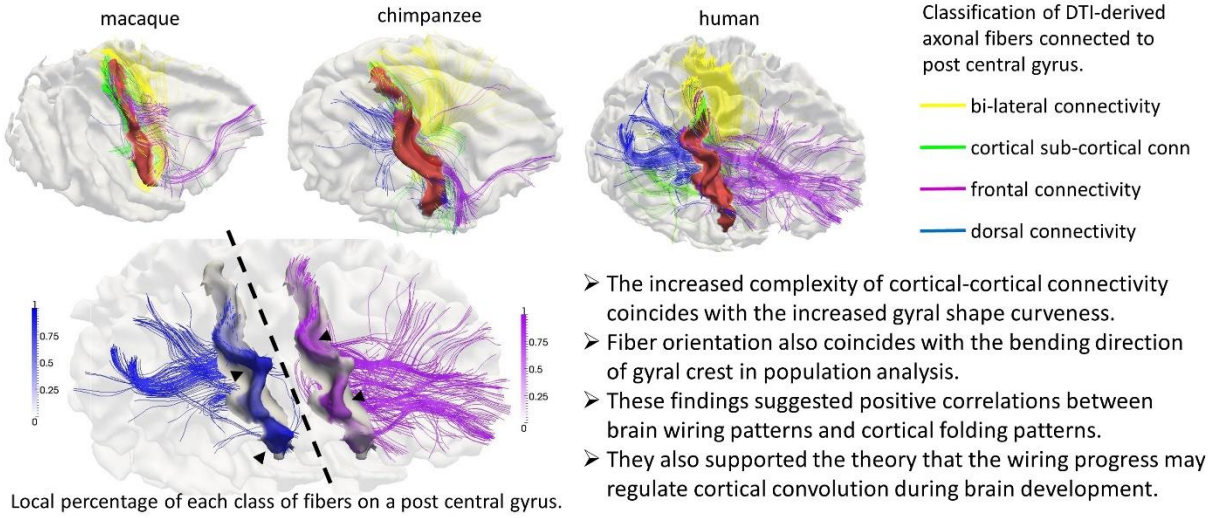

Figure 6. Joint comparisons between cortical folding pattern (MRI) and structural wiring pattern (DTI) among human, chimpanzee, and macaque brains [16].

3 HIGH-THROUGHPUT COMPUTING FOR NEUROIMAGING ANALYSIS

Thanks to the development of bio-imaging techniques, acquiring large population of neuroimaging data in good quality is more feasible than ever. And thanks to the trend of open source in neuroscience field, more and more neuroimaging data are now shared online or between labs. Data availability is no longer the major reason that slows the findings in neuroscience field. Instead, the lack of efficient computational pipelines is becoming the bottleneck. Despite the previously mentioned approaches from the perspective of analytical strategy, the throughput of the computational tool is also very important. To increase the throughput, two features are very important for tool design: 1) 'smart'; 2) high-performance computing.

3.1 SMART

Today, Google cars can already handle complicated road conditions and drive automatically on the road without human intervention. However, in most biology labs, scientists still have to repeat their experiments again and again to acquire data. I share the same perspective as Dr. Hanchuan

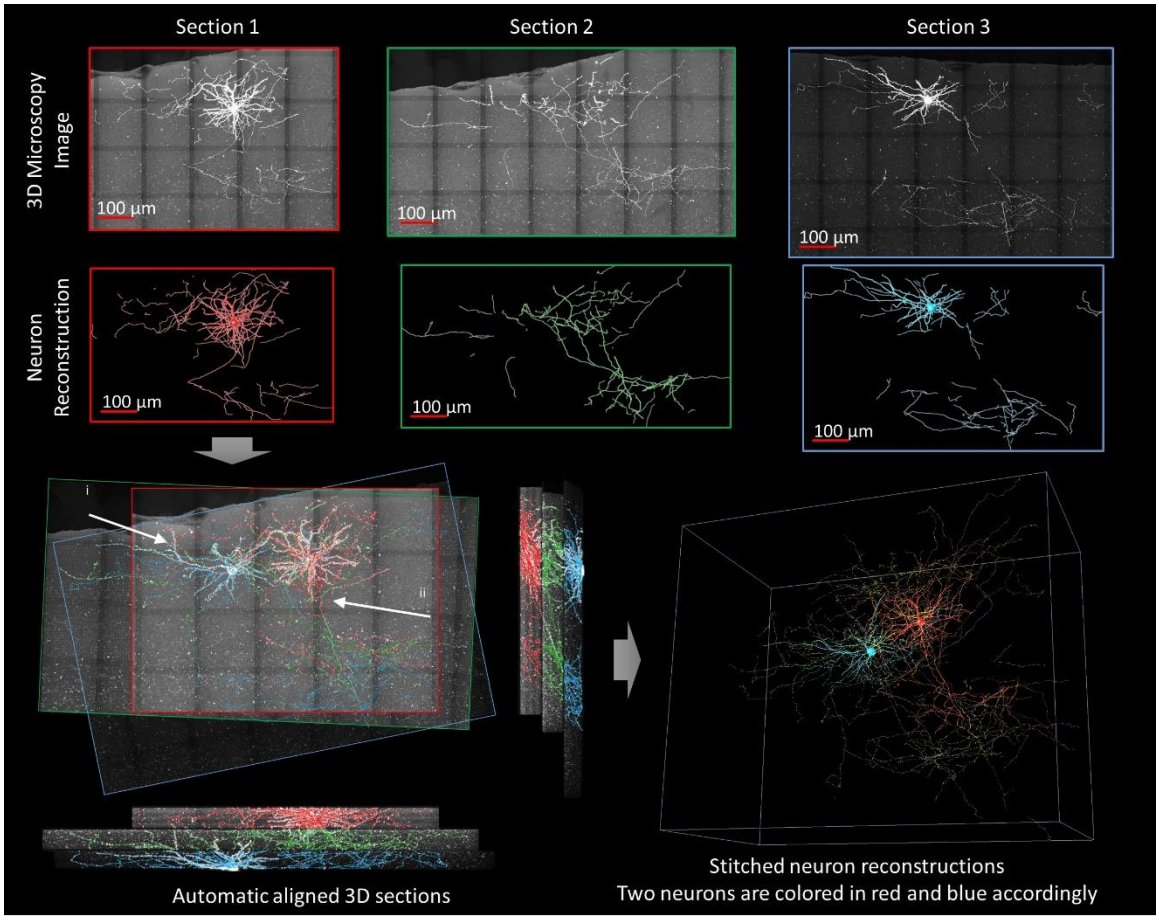

Figure 7. Illustration of neuron stitching problem and the stitching results of NeuronStitcher [19]. Three continuous sections containing 2 neurons in mouse brain are shown as examples.

Peng (my advisor and collaborator) that we should build smarter tools to free those great minds from labor works so they can focus on more intelligent activities. For instance, in collaboration with Dr. Peng, I developed NeuronStitcher which is the first tool that allows neuroscientists to automaticly/semi-automaticly stitch neuron fragments from sliced brain tissues [19]. Due to the limited imaging depth of microscopic, to study complete neuron morphology, researchers need to 1) physically section brains, 2) image such sections under microscope, and 3) stack and align image sections one by one to generate a very big image volume. Before NeuronStitcher, step 3 can only be manually accomplished for confocal microscopic sections. Yet an accurate alignment is extremely difficult. The tool provided an efficient solution to this problem and has been integrated in the IVSCC pipeline of Allen Institute for Brain Science. In another work, I developed one of the first fully automatic software based on machine learning algorithms to reconstruct 3D digital representation of neuron morphology [20]. SmartTracing is a very powerful tool in that it is self-adaptive and does not need human intervention such as parameter tuning to generate results.

3.2 HIGH-PERFORMANCE COMPUTING

Advanced computing framework such as network clusters, computer clouds, or graphic processing unit (GPU) are usually needed to speed up computations for big data. In my recent work in collaboration with Yu Zhao (PhD student I supervised), we successfully reduced the running time of pairwise comparison between 35600 3-dimensional images of brain functional networks from 1280 hours to 3 days [13]. Our computational framework was implemented on Apache Spark, which provides a fast and general engine for big data processing. To our best knowledge, this is one of the earliest works in applying SPARK scalable computing framework in processing fMRI images. In the future, we will complete this high-performance fMRI analysis framework and release it to the field.

REFERENCES:

[1] T. Liu, “A few thoughts on brain ROIs.,” *Brain Imaging Behav.*, vol. 5, no. 3, pp. 189–202, Sep. 2011.

[2] H. Chen, T. Zhang, and T. Liu, “Identifying group-wise consistent white matter landmarks via novel fiber shape descriptor.,” *MICCAI, LNCS*, vol. 8149, pp. 66–73, Jan. 2013.

[3] D. Zhu, K. Li, L. Guo, X. Jiang, T. Zhang, D. Zhang, H. Chen, F. Deng, C. Faraco, C. Jin, C.-Y. Wee, Y. Yuan, P. Lv, Y. Yin, X. X. Hu, L. Duan, X. X. Hu, J. Han, L. Wang, D. Shen, L. S. Miller, L. Li, and T. Liu, “DICCCOL: dense individualized and common connectivity-based cortical landmarks.,” *Cereb. Cortex*, vol. 23, no. 4, pp. 786–800, Apr. 2013.

[4] H. Chen, X. Cai, D. Zhu, F. Nie, T. Liu, and H. Huang, “Group-wise consistent parcellation of gyri via adaptive multi-view spectral clustering of fiber shapes.,” *MICCAI, LNCS*, vol. 7511, pp. 271–9, Jan. 2012.

[5] H. Chen, K. Li, D. Zhu, and T. Liu, “Identifying consistent brain networks via maximizing predictability of functional connectome from structural connectome,” in *ISBI*, 2013, pp. 978–981.

[6] H. Chen, K. Li, D. Zhu, L. Guo, X. Jiang, and T. Liu, “Group-wise optimization and individualized prediction of structural connectomes,” in *ISBI*, 2014, pp. 742–745.

[7] Y. Yuan, H. Chen, J. Lu, T. Zhang, and T. Liu, “Distance Networks for Morphological Profiling and Characterization of DICCCOL Landmarks,” *MICCAI, LNCS*, vol. 9350, pp. 380–7, 2015.

[8] H. Chen, K. Li, D. Zhu, X. Jiang, Y. Yuan, P. Lv, T. Zhang, L. Guo, D. Shen, and T. Liu, “Inferring group-wise consistent multimodal brain networks via multi-view spectral clustering.,” *IEEE Trans. Med. Imaging*, vol. 32, no. 9, pp. 1576–86, Sep. 2013.

[9] H. Chen, A. Iraji, X. Jiang, J. Lv, Z. Kou, and T. Liu, “Longitudinal Analysis of Brain Recovery After Mild Traumatic Brain Injury Based on Group-wise Consistent Brain Network Clusters,” *MICCAI, LNCS*, vol. 9350, pp. 194–201, 2015.

[10] A. Iraji, H. Chen, N. Wiseman, R. D. Welch, B. J. O’Neil, E. M. Haacke, T. Liu, and Z. Kou, “Compensation through Functional Hyperconnectivity: A Longitudinal Connectome Assessment of Mild Traumatic Brain Injury,” *Neural Plast.*, 2016.

[11] J. Lv, X. Jiang, X. Li, D. Zhu, S. Zhang, S. Zhao, H. Chen, T. Zhang, X. Hu, J. Han, J. Ye, L. Guo, and T. Liu, “Holistic atlases of functional networks and interactions reveal reciprocal organizational architecture of cortical function.,” *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1120–31, Apr. 2015.

[12] J. Lv, X. Jiang, X. Li, D. Zhu, H. Chen, T. Zhang, S. Zhang, X. Hu, J. Han, H. Huang, J. Zhang, L. Guo, and T. Liu, “Sparse representation of whole-brain fMRI signals for identification of functional networks.,” *Med. Image Anal.*, vol. 20, no. 1, pp. 112–34, Feb. 2015.

[13] Y. Zhao, H. Chen, Y. Li, J. Lv, X. Jiang, X. Li, F. Ge, T. Zhang, S. Zhang, B. Ge, C. Lyu, S. Zhao, J. Han, L. Guo, and T. Liu, “Connectome-scale Group-wise Consistent Resting-state Network Analysis in Autism Spectrum Disorder,” *Hum. Brain Mapp.*

[14] Y. Li, H. Chen, X. Jiang, X. Li, J. Lv, H. Peng, J. Z. Tsien, and T. Liu, “Comprehensive genoarchitecture of adult mouse brain using sparse coding,” *Brain Struct. Funct.*

[15] H. Chen, T. Liu, Y. Zhao, T. Zhang, Y. Li, M. Li, H. Zhang, H. Kuang, L. Guo, J. Z. Tsien, and T. Liu, “Optimization of large-scale mouse brain connectome via joint evaluation of DTI and neuron tracing data.,” *Neuroimage*, vol. 115, pp. 202–13, Jul. 2015.

[16] H. Chen, T. Zhang, L. Guo, K. Li, X. Yu, L. Li, X. X. Hu, J. Han, and T. Liu, “Coevolution of gyral folding and structural connection patterns in primate brains.,” *Cereb. Cortex*, vol. 23, no. 5, pp. 1208–17, May 2013.

[17] T. Zhang, H. Chen, L. Guo, K. Li, L. Li, S. Zhang, D. Shen, X. Hu, and T. Liu, “Characterization of U-shape Streamline Fibers: Methods and Applications,” *Med. Image Anal.*, Apr. 2014.

[18] X. Li, H. Chen, T. Zhang, X. Yu, X. Jiang, K. Li, L. Li, M. J. Razavi, X. Wang, X. Hu, J. Han, L. Guo, X. Hu, and T. Liu, “Evolutionarily-Preserved and Species-Specific Gyral Folding Patterns across Primate Brains,” *Cereb. cortex*.

[19] H. Chen, D. M. Iascone, N. M. da Costa, E. S. Lein, T. Liu, and H. Peng, “Fast Assembling of Neuron Fragments in Serial 3D Sections,” *Nat. Methods*.

[20] H. Chen, H. Xiao, T. Liu, and H. Peng, “SmartTracing: self-learning-based Neuron reconstruction,” *Brain Informatics*, vol. 2, no. 3, pp. 135–144, Aug. 2015.

[21] D. Zhang, L. Guo, D. Zhu, K. Li, L. Li, H. Chen, Q. Zhao, X. Hu, and T. Liu, “Diffusion tensor imaging reveals evolution of primate brain architectures,” *Brain Struct. Funct.*, vol. 218, no. 6, pp. 1429–1450, Nov. 2013.