Received: 22 December 2016 Accepted: 3 March 2017

A Computational Approach towards the Microscale Mouse Brain Connectome from the Mesoscale

Tielin ZHANG [a], Yi ZENG [a,b,c,1] and Bo XU [a,b,c]

[a] *Institute of Automation, Chinese Academy of Sciences, Beijing, China*
[b] *Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China*
[c] *University of Chinese Academy of Sciences, Beijing, China*

Abstract. The wiring diagram of the mouse brain presents an indispensable foundation for the research on basic and applied neurobiology. It is also essential as a structural foundation for computational simulation of the brain. Different scales of the connectome give us different hints and clues to understand the functions of the nervous system and how they process information. However, compared to the macroscale and most recent mesoscale mouse brain connectome studies, there is no complete whole brain microscale connectome available because of the scalability and accuracy of automatic recognition techniques. Different scales of the connectivity data are comprehensive descriptions of the whole brain at different levels of details. Hence connectivity results from a neighborhood scale may help to predict each other. Here we report a computational approach to bring the mesoscale connectome a step forward towards the microscale from the perspective of neuron, synapse and network motifs distribution by the connectivity data at the mesoscale and some facts from the anatomical experiments at the microscale. These attempts make a step forward towards the efforts of microscale mouse brain connectome given the fact that the detailed microscale connectome results are still far to be produced due to the limitation of current nano-scale 3-D reconstruction techniques. The generated microscale mouse brain will play a key role on the understanding of the behavioral and cognitive processes of the mouse brain. In this paper, the conversion method which could get the approximate number of neurons and synapses in microscale is proposed and tested in sub-regions of Hippocampal Formation (HF), and is generalized to the whole brain. As a step forward towards understanding the microscale connectome, we propose a microscale motif prediction model to generate understanding on the microscale structure of different brain region from network motif perspective. Correlation analysis shows that the predicted motif distribution is very relevant to the real anatomical brain data at microscale.

Keywords. Microscale mouse brain connectome, Brain connectivity atlas, Motif distribution, Synaptic degree distribution analysis

[1]Tielin Zhang and Yi Zeng contributed to the work equally and should be regarded as co-first authors. Corresponding Author: Yi Zeng; E-mail: yi.zeng@ia.ac.cn.

1. Introduction

Identifying the structural architecture of the brain has been one of the most important and challenging tasks for the investigation of the brain and neuroscience. The connectome of the brain is the structural foundation and provides insights for deeper understanding of neural networks and neural functions. It also shows the genetic and evolutionary properties on the brain organization across different species. Small world connection and motif distribution property of the neural networks have been found in brains of different species, ranging from drosophila brain to the human brain [Achard *et al* (2006)]. Experiments have shown that various types of motif support different special functional properties of the network, for example, three-node feed-forward loop motif plays an important role on information abstraction [Mangan and Alon (2003)].

Despite of the fundamental and important role for neuronal connectivity to brain and neuroscience research, the current understanding about them is far to complete. The massive whole brain connectivity can be roughly divided into multiple scales, namely, the macroscale, the mesoscale, and the microscale. The connectivity in different scales describes different principles on the organization of brain building blocks at different levels. The connectome at different scales share some mutual characteristics (e.g. small world phenomenon, similar motif distribution), although their interrelationships are complex.

The macroscale connectome of the brain describes connectivity of different brain building blocks in a coarser level which is usually from one region to another region of the brain. It can be inferred by functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI) which could predict the functional connectivity of the brain by detecting the changes of the blood flow and restricted diffusion of water respectively [Friston *et al* (1998),Assaf and asternak (2008)]. The wiring diagram of the macroscale connectome for the whole brain has been built and has been used on many aspects of neural biological researches and applications.

The microscale connectome of the brain is aimed at describing the connectivity of brain building blocks at the neuron and synapse level. Automatic microscopic reconstruction at the nanometer scale (e.g. stimulated emission depletion, STED for short) is a supporting technique for building the microscale connectome of the brain. Nevertheless, due to the scalability issue for automatic recognition and 3-D reconstruction, there is no microscale whole brain connectome atlas for mammals.

The mesoscale connectome is between the macroscale and the microscale. It is finer than the scale of large brain regions and coarser than the scale of neurons and synapses. The connectivity of the sub-regions can be described at this level. One of the most representative results at this level is the mesoscale connectome of the mouse brain based on the enhanced green fluorescent protein (EGFP) technique [Oh *et al* (2014)].

The microscale connectome is essential and unique since it shows the basic wiring principles at the neuron level. The effort from mesoscale to microscale can be considered as a step forward towards the understanding of the behavioral and cognitive processes in mouse brain before the real and complete data come out through synaptic-level neuroanatomical investigations.

In this paper, we attempt to take a step forward towards the microscale atlas from the mesoscale by computationally combining understandings and data both from the mesoscale and the microscale. The microscale connectome efforts are made from two perspectives: (1) the predicted number of synapses and neurons in each region; (2) the

predicted microscale network motif distributions in target regions. The approach combines the understanding of the mesoscale connectome (which is composed of positive and negative connective strength based on the method of tracing projections between brain regions) and the anatomical experimental results (e.g. the number of neurons and synapses in some identified brain building blocks measured by anatomical processing) at the microscale to predict the number of neurons and synapses in other regions. In addition, anatomical network motif data in 54 brain regions are collected and used to train a microscale motif prediction model which could predict the motif distribution of a specific region at the neuronal and synaptic level. The new generated microscale connectome which contains the number of neurons and synapses, and the microscale motif distribution will be verified by the real anatomical data from future experiments.

Section 2 will introduce the current progress of imaging methods in different scales, and a mesoscale connectome atlas produced by Allen Institute for Brain Science is selected as the basis of our investigation. In Section 3, anatomical data from sub regions of Hippocampal Formation (HF) are selected to validate the reasonability of the proposed conversion method. In Section 4, a model for microscale motif distribution prediction is proposed and validated. Section 5 gives a statistical degree analysis for the predicted microscale atlas. Section 6 gives a brief conclusion for the proposed method.

2. Related Works

With the advancement of new techniques and equipments, attempts of mesoscale and microscale neuron imaging for the mouse brain have been conducted and accelerated our understanding on the structure and function of the neural circuits.

At the microscale, from the structure perspective, electron microscopic equipment only enables relatively small scale reconstruction and observation of the brain anatomy [Sporns *et al* (2005),Osten and Margrie (2013)]. From the function perspective, at this scale, functional calcium imaging methods are used to understand the relationship between cognition and neuron activities, still at very small scale [Stosiek *et al* (2006)]. Large scale (including whole brain scale) structural and functional connectivity at the microscale is one of the most challenging investigations for Brain and Neuroscience research [Kasthuri *et al* (2015)].

At the mesoscale, several investigations on the mouse brain connectome has already been made. One of them is by the Golgi silver impregnation method [Rakic (2006)], another even more efficient method is neuroanatomical-tracer method which has played an important role on the measurement of connectivity of sensory, motor and other sub systems [Felleman and Essen (1991),Rocklanda and Pandyaa (1979)]. Comparing with the traditional methods which are time consuming, a quicker way is to make the mapping of point-to-point connections between two brain building blocks by the method of anterograde tracers and retrograde tracers [Bohland *et al* (2009)]. This method uses two approaches to generate three-dimensional mouse brain atlas: the first one is the light-sheet fluorescence microscopy for the brain tissue after chemical processing [Maizel *et al* (2011)], and the second one is the integration of microscopy tissue sectioning (e.g. line-scan imaging, or two-photon microscopy [Ragan *et al* (2012)]). The main characteristic of this method is the mechanical removal of the brain tissue after the mosaic imaging of the upper tissue, so that the tissue will always be

on the top of the camera, and high resolution camera or lenses can be used to achieve high standard imaging pictures. Instruments can be serial two-photon tomography [Ragan *et al* (2012)], knife-edge scanning microscopy or fluorescent micro-optical sectioning tomography [Maniadakis and Trahanias (2003),Seress (1988)].

The mesoscale mouse brain connectome atlas from Allen Institute for Brain Science uses EGFP to make the measurement of the projections of axons from 213 mouse brain sub regions (e.g. V1 and V4 in visual cortex, CA1 in hippocampus, and POL in thalamus) covering the whole brain [Oh *et al* (2014)]. The atlas is calculated from 469 injected experiments and has been one of the state of the art atlases [Ragan *et al* (2012)]. There are 213×213 values in which each one stands for the strength of the connectivity among two sub regions in the atlas.

In this paper, we will try to combine mesoscale atlas with some other anatomical experimental results to make a predictive analysis on the neuron, synapse and motif distributions to bring a step forward towards the microscale mouse brain atlas.

3. From Mesoscale to Microscale Connectome

The microscale connectome atlas is able to identify the number of neurons and synapses in each sub region and the level of convergence and divergence for each cell in the network. Even though the detailed microscale atlas has not been acquired by the microscopic reconstruction method, with the high-resolution mesoscale mouse brain atlas and some sufficient and detailed microscale connections of sub-regions (e.g. the approximate number of neurons and synapses for some specific sub regions by anatomical experiments), we can bring the mesoscale connectome a step forward towards the microscale connectome. However, since different sub regions share different cell types and cell densities, our attempt is just a prototype to show the reasonability and importance of conversion method, and with the more real and more detailed anatomical results are found by experiments, a more feasible microscale atlas with higher accuracy will be produced based on the proposed convertion method.

The essential idea that bridges the mesoscale connectome and the microscale connectome is that the EGFP method can obtain projection weights, while they are naturally weights of connections among clusters of neurons that belong to different regions (or the same region for interconnections). Hence, the projection weights at the mesoscale can be converted to the number of synapses at the microscale with the help of more and more anatomical experimental results. In addition, by using different kinds of EGFP methods, different types of connectivity (i.e. excitatory or inhibitory synapses) at mesoscale can be obtained.

In order to make comparison with the prediction results, some special sub regions which already contain partial detailed microscale information by anatomical experiments are selected. Compared to other regions, most of the complex intrinsic wiring diagram of the hippocampus is increasingly refined over more than a hundred years. Hence, we select HF for verification. Here we separate the conversion procedure into four major steps.

3.1. Regrouping the Mesoscale Connections

As the first step, we regroup the mesoscale connections of building blocks which belong to the same upper region in the mouse brain. The data of mesoscale connections in 213 regions of mouse brain are collected. Here we give three examples, as shown in Fig 1, the hippocampus, the visual cortex and the thalamus. (1) illustrates the Sub regions of hippocampal formation and their connectivities, and the right image illustrates the connectivity strength of different sub regions, from top to bottom is CA1, CA2, CA3, DG, LEC, MEC, Sub, respectively; (2) illustrates some part of the sub regions of the visual cortex and their connectivities, for the image on the right, the connectivity strength is among VISal, VISam, VISI, VISp, VISpl, VISpm, from top to bottom respectively. (3) illustrates sub regions of the thalamus and their connectivity, and in the right image, from top to bottom, the sub regions are: AD, Amd, Amv, AV, CL, CM, IMD, LD, LGd, LGv, LH, MGd, MGm, MGv, MH, PT, PVT, RH, SMT, SPA, SPFm, SPFp, VAL, VPL, VPM, VPMpc, respectively.

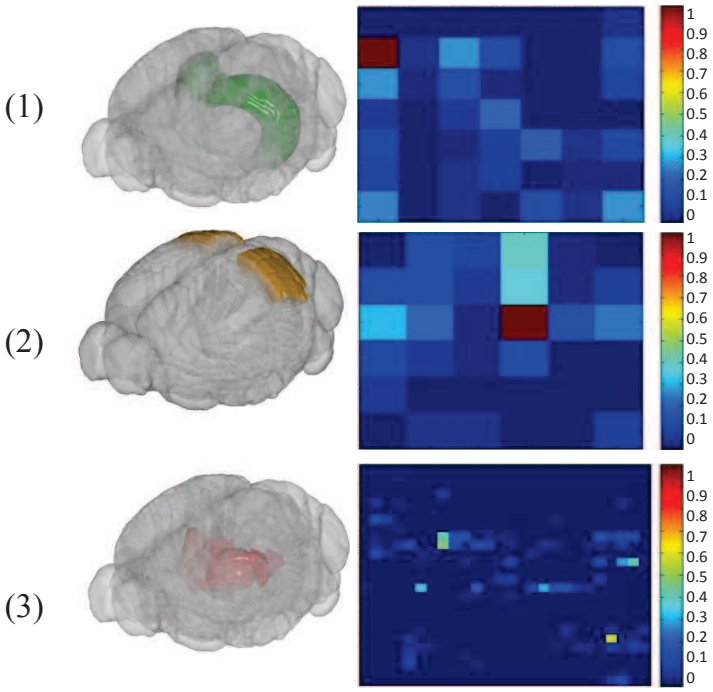


Figure 1. The location and connectivity matrix of sub regions within hippocampal formation (1), part of the visual cortex (2), and thalamus (3) in the mouse brain.

3.2. Providing Partial Microscale Anatomical Information

Following the first step , we collect three main parts of anatomical experimental results, including hippocampal formation, visual cortex, and some parts of thalamus. As hippocampus formation for example, it is with a six layered architecture that comprises five

distinct sub-regions: the Cornu Ammonis area one to three (CA1-CA3), the Subiculum (Sub) and the Dentate Gyrus (DG). A sketch on the mesoscale connection among different sub regions of HF and EC are shown in Fig 2 (which is integrated and refined from [Cutsuridis *et al* (2010),Treves and Rolls (1992)].). The EC contains two parts (i.e. MEC and LEC) and six layers (i.e. EC-I to EC-VI). Perforant pathway (PP) and trisynaptic pathway are the two main projections from EC to HF. The output regions from HF to EC are CA1 and Sub [Treves and Rolls (1992)].

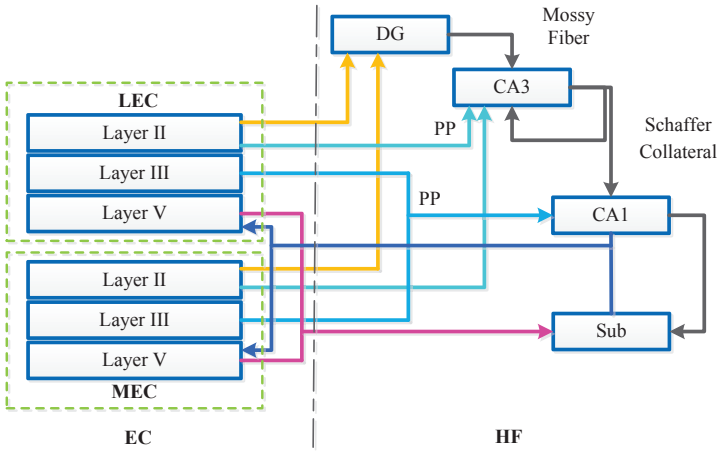


Figure 2. The mesoscale connectivity diagram of the EC-HF network. This image is integrated and refined from [Cutsuridis *et al* (2010),Treves and Rolls (1992)].

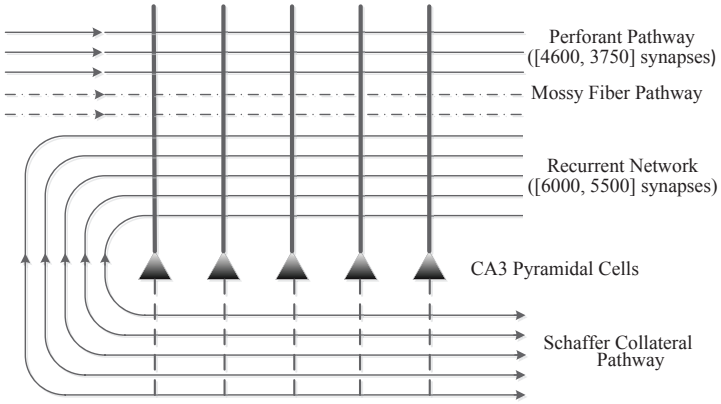


Figure 3. The CA3 synapse connectivity in HF (integrated and refined from [Cutsuridis *et al* (2010), Treves and Rolls (1992),Witter (2010)]).

To illustrate the idea, here we select one of the sub regions (e.g. CA3) as the additional anatomical information. As shown in Fig 3, the brackets in the form of [A, B] denote the average synaptic connections between neurons in two regions. A denotes the number of CA3 synapses innervated by a single cell in EC or CA3. And B denotes the number

of cells that converge on a single CA3 cell. In Fig 3, CA3 receives inputs from EC by perforant pathway and from DG by mossy fibers, and produces outputs to CA1 by Schaffer Collateral pathway. In addition, it also has strong recurrent network which is very different from other sub-regions in HF. The average number of synapses for each cell which converges on CA3 is $3,750$ for EC and $5,500$ for CA3 [Cutsuridis *et al* (2010)]. The average synapse number of a single cell (in EC, DG or CA3, here we do not distinguish the types of cells) which projects to CA3 cells is $4,600$ and $6,000$ respectively [Treves and Rolls (1992),Witter (2010)].

3.3. Converting from the Mesoscale Connectome to the Microscale Connectome

Up to now, there is no detailed atlas about the numbers of neurons and synapses in each sub regions of the whole mouse brain. The basic idea of this step is to provide a concrete method to predict them. In order to make the method more understandable, we introduce it combining with a concrete example.

Table 1 presents the detailed description of the mesoscale connectivity strength in HF and EC. The values in the table are the voxel strength measured by EGFP in paper [Oh *et al* (2014)]. CA3 is one of the sub regions with enough anatomical details about the number of neurons and recurrent-type synapses. So we select it as the sample region in Equation 1. Based on the average number of synapses in sample sub regions, we can calculate the approximate total number of synapses for each region based on the Equation 1.

$$Syn_{sub}(i) = \frac{\sum_{j=1}^{213} W_{i,j}}{\sum_{j=1}^{213}\sum_{s=1}^{N} W_{s,j}} \times \left(\sum_{s=1}^{N} Syn_s \times Neu_s \right) \tag{1}$$

where $Syn_{sub}(i)$ denotes the total number of synapses of the target region i, $W_{i,j}$ denotes the connectivity weight from the region i to region j in mesoscale mouse brain atlas (e.g. some weights of sub regions in HF are shown in Table 1, more weights of other regions can be found in [Oh *et al* (2014)]); $W_{s,j}$ is the value of connectivity weight from sample sub regions to other regions (e.g. $W_{CA3,CA3}$ is the recurrent value and is 0.116, as shown in Table 1); Neu_s is the number of neurons in sample areas (hippocampal formation, visual cortex or thalamus) which is from the anatomical experiments; Syn_s is the average number of synapses of cells in the sample areas (e.g. for CA3 the value is $6,000$, as shown in Fig 3 [Treves and Rolls (1992),Witter (2010)].

Based on the proposed method and the mesoscale connectome atlas in Table 1, we then can obtain the approximate synapse distribution of the whole mouse brain. Notice that this distribution is still the mesoscale distribution of the synapses for each sub regions.

In order to predict the microscale number of neurons and the number of synapses for each neuron in each sub region of mouse brain, two attempts are tried to make the measurement on the number of neurons, as defined in Equation 2 and Equation 3.

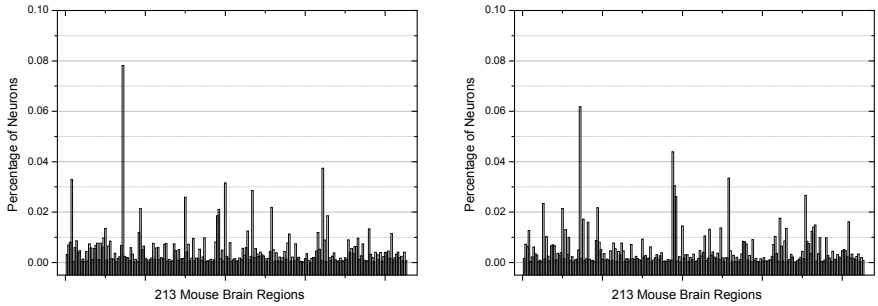
$$N_1(i) = \frac{Syn_{sub}(i)}{\sum_{i=1}^{213} Syn_{sub(i)}} \times Neu_{All} \tag{2}$$

$$N_2(i) = \frac{voxel(i)}{\sum_{i=1}^{213} voxel(i)} \times Neu_{All} \tag{3}$$

Table 1. The connectivity weights of EGFP-labeled axons between the sub regions of HF and EC (The values are extracted from [Oh *et al* (2014)]).

	CA1	CA2	CA3	DG	LEC	MEC	Sub
CA1	0	0	0	0	0.0132	0.0072	0.039
CA2	0.846	0.030	0.237	0.092	0	0	0.100
CA3	0.233	0.013	0.116	0.024	0.001	0.004	0.077
DG	0.047	0.008	0.051	0.172	0.001	0.004	0
LEC	0.086	0.003	0.013	0.058	0.169	0.033	0.080
MEC	0.073	0.001	0.031	0.059	0	0.013	0.022
Sub	0.221	0.012	0.028	0	0.116	0.027	0.228

Equation 2 corresponds to the first method, and it is based on the idea that the number of synapses in each sub region is the sum of the synapse of all kinds of neurons. $Syn_{sub}(i)$ denotes the approximate total number of synapses in the sub regions of mouse brain. Neu_{All} denotes the total number of neurons. Equation 3 presents the second method, and it obtains the number of neurons by the proportion voxel size of the target sub regions.



(a) Results based on the synapse method, as defined in Equation 2

(b) Results based on the voxel method, as defined in Equation 3

Figure 4. The distribution of percentage of neurons in 213 mouse brain regions by the two methods.

Fig 4(a) and Fig 4(b) present the percentages of numbers of neurons in 213 sub regions of the mouse brain by two different methods. The cosine similarity of them is 0.828 which shows the high consistency of the two methods. By using this approach, we obtain the percentages on the number of neurons in each sub region at whole brain scale.

3.4. Verifying the Conversion Results

Comparing the predicted data with the realistic data from anatomical experiments is a reasonable attempt to verify the conversion results.

3.4.1. The verification from the sub regions in HF

For the anatomical results, as Table 2 shows, some quantitative data about the number of neurons, the ratio of excitatory and inhibitory synapses at microscale in HF has been measured [Insausti *et al* (1998)].

Table 2. The approximate number of neurons in HF [Maniadakis and Trahanias (2003), Cutsuridis *et al* (2010),Boss *et al* (1985)]

Region Names	Number of Neurons	Excitatory Synapses	Inhibitory Synapses
CA1	390,000	95%	5%
CA3	250,000	88%	12%
DG	500,000	66%	34%

For the prediction results, the number of neurons and the number of synapses in sub regions of HF and EC based on the two methods are shown in Table 3 and Table 4. Comparing with the anatomical experimental results from Table 2, the predicted number of neurons in each sub region of HF by two different methods are generally consistent to the anatomical ones.

Table 3. The predicted number of synapses in hippocampus by method one

Region	Number of Neurons (10^3)	Number of Synapses (10^3)		
		CA1	CA3	DG
CA1	378.3	2.8	0.0	0.0
CA3	330.0	12.0	6.0	1.2
DG	495.4	2.4	2.6	8.9

Table 4. The predicted number of synapses in hippocampus by method two

Region	Number of Neurons (10^3)	Number of Synapses (10^3)		
		CA1	CA3	DG
CA1	355.5	2.8	0.0	0.0
CA3	280.2	10.1	5.1	1.0
DG	585.2	6.8	7.4	25.5

Note that due to the technical limitation of EGFP, many of the connectivity strength within or among sub regions are not visible (i.e. with the value 0). Hence, the average number of synapses for each neuron in specific region cannot be calculated based on the proposed method. While some of the blank results can be refined based on domain knowledge. For example, in the mammalian brain, almost all the regions are with interconnections among neurons within the specific region. Hence, for those values marked as 0 for self connections, the average number of synapses for this region (e.g. CA1 in Table 3 and Table 4) is assigned an average value of synapses calculated based on the data of whole mouse brain (namely, 2,800).

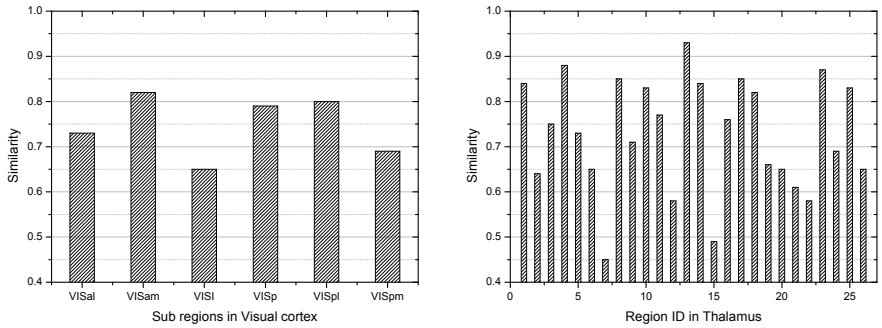
Using the same methods, we generate the number of average synapse for neurons in all the 213 regions of the mouse brain. The over all predicted synapse number for the whole mouse brain is approximately at the same order of magnitude compared to the real house mouse brain. The predicted number of synapses in different regions have validated the reasonability of the proposed computational approach to convert the mesoscale connectivity weights to the microscale synapse distribution.

3.4.2. The Verification from Other Sample Regions

Since the number of neurons in each sub regions of the mouse brain has not yet been measured by the anatomical experiments, in order to verify the predicted results of the neurons and the synapses, we select the predicted number of neurons and synapses in the sample regions to make comparison with biological experimental evidence. The similarity is calculated based on Equation 4.

$$Sim(i) = \frac{|Pre(i) - Ana(i)|}{Ana(i)} \tag{4}$$

The similarity comparisons of predicted and anatomical numbers of neurons are shown in Fig 5(a) and Fig 5(b). From the similarity results, we can conclude that most of the results are above 50%, this shows that the two results are weakly consistent with each other. Because this attempt just uses the anatomical numbers of neurons in three sample regions, we think that with more anatomical information added into the conversion methods, a higher accuracy will be achieved.



(a) The result in Visual Cortex (b) The result in Thalamus

Figure 5. The similarity comparison of predicted and anatomical number of neurons

4. Region Specific Microscale Motif Distribution Prediction

Most of the networks which could be represented as graphs can use network motif to describe their network properties [Sporns and Kotter (2004)]. Three-node network motifs are commonly used for analyzing complex networks [Mangan and Alon (2003)]. Although until now, there are no detailed microscale network structures for each sub regions, investigations on the motif distribution at the synaptic level could also help to get deeper understanding on the network structures of the brain at the microscale.

Since the motif distribution at different scales share similarities to a certain degree, in this paper, we try to establish the link between the mesoscale connectome and microscale motif distribution.

A three layered neural network is built to make the motif type prediction at the microscale, as shown in Fig 6. The input layer contains 213 neurons (corresponding to the 213 mouse brain regions at the mesoscale) and each cell receives the inputs from one of

the 213 connectivity strengths with the specific predicted region. The hidden layer contains 500 neurons and the output layer contains 13 neurons which correspond to the 13 types of three-node motifs. The output of the network is the motif distribution prediction results to a specific region.

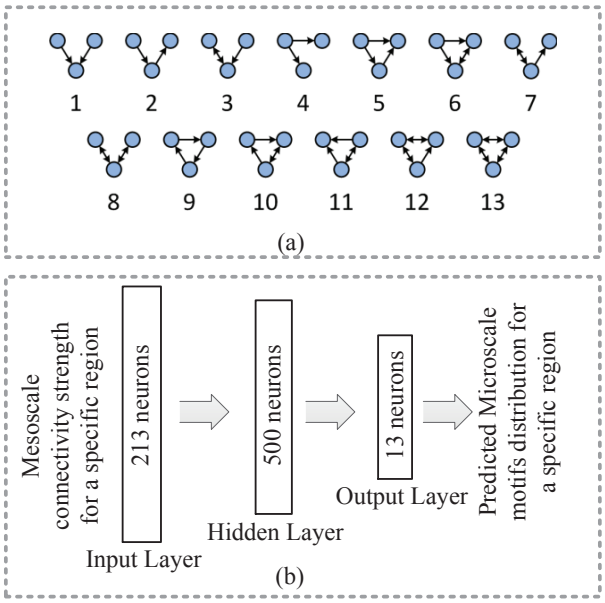


Figure 6. The 13 types of motifs (a) and the three layered classifier for motif type distribution analysis (b)

The network are trained by the data from two sources: (1) The mesoscale connectivity strength of CA1, CA3 to the 213 regions of the mouse brain as inputs and microscale motifs distributions in mouse by the network structure prediction method as outputs [Zhang *et al* (2015)]; (2) The mesoscale connectivity strength of 46 regions from both cats cortex and macaque cortex to the 213 corresponding regions as inputs [Sporns *et al* (2007)] and network motif distribution of cats cortex and macaque cortex as outputs (We use data from the same region of other mammalian brains since they generally share many structural similarities) [Sporns *et al* (2007)]. On one hand, there are no ground truth for microscale connectome and motif distribution for cat and monkey brain. On the other hand, the mesoscale motif distribution provides partial evidence for microscale since mesoscale connections are also established by specific connections among neurons from different regions, and the motifs from mesoscale and microscale are similar to some extent, as shown in Fig 7. Hence, it is rational to borrow them as a version of partial microscale connectome, and to use them as outputs of the network for training the model.

We use the Matlab deep learning toolbox to develop the three layered classifier with the setting of iteration time 100, learning rate 1.0. Different with the traditional training procedure which set the target class as 1 and other classes as 0, we give the motif distribution values directly into the output layers as the right classification results. Fig 8 is the description of the percentage distribution of different types of motifs in CA1. From the result of Fig 8, we could find that the motifs are clustered more on type No.3, No.7,

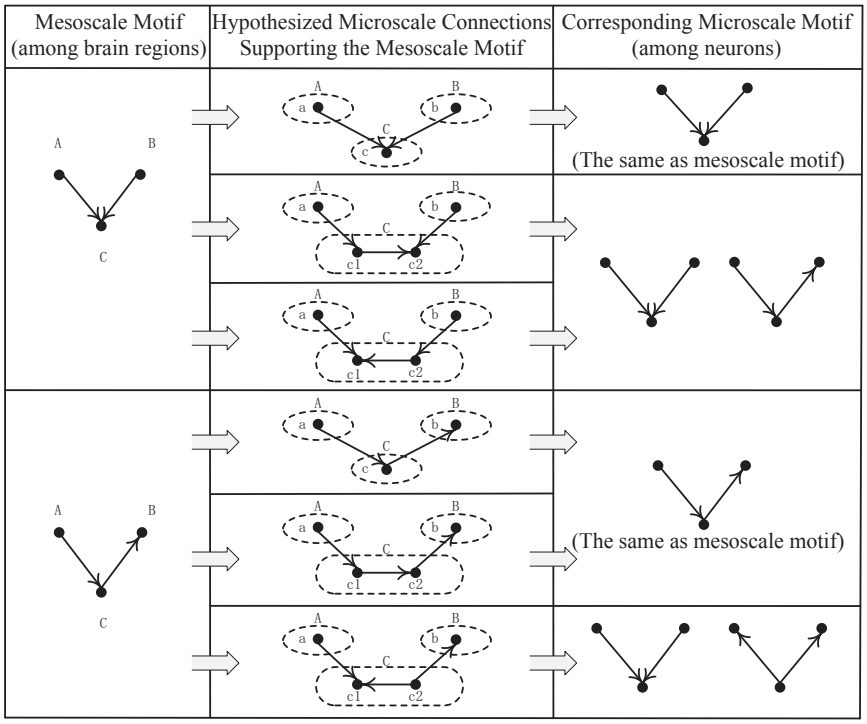


Figure 7. From mesoscale to microscale motifs

No.8, No.12 and No.13 which can be considered as representative motifs for describing the characteristics of CA1. Fig 9 presents the anatomical motif distribution results of the 46 regions in cats and macaque cortex for training the three layered classifier.

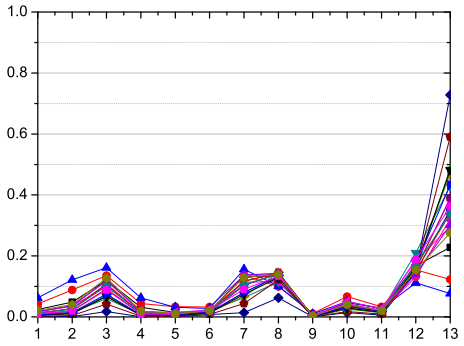


Figure 8. The 20 groups of motif distribution in CA1 of the mouse brain

As Fig 6 shows, after we obtained the anatomical results and predicted results, we use Pearson correlation to judge the similarity of these two results. The results are mi-

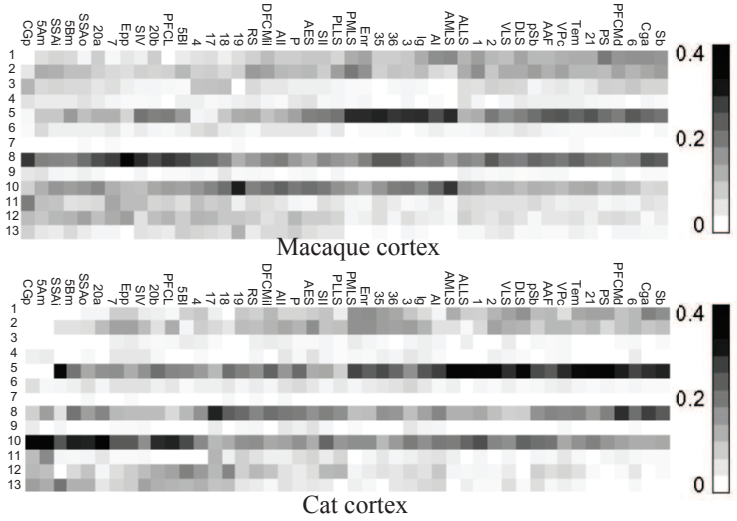


Figure 9. The motif distributions of different regions in cortexes of cat and macaque (x and y coordinates present the 46 sub regions and corresponding frequency of the 13 types of motifs respectively), extracted from [Sporns *et al* (2007)].

croscale motif distribution of 48 regions (CA1, CA3 in the mouse brain, and another 46 brain regions in cat and macaque brain) distributed in 112 groups of data (i.e. 20 groups for CA1 and CA3 respectively in mouse brain, 92 groups from cats and macaque cortex). If we consider the mesoscale motif distribution from cat and macaque brain as a possible version of the microscale, after data cross validation (90% for training and 10% for prediction, with 10 times), the accuracy is 91.6%, which indicates that the proposed method works well for the prediction of the motif distribution.

With the high accuracy of the prediction model, we apply it to the whole mouse brain, and generated a prediction of the motif distribution for each of the 213 regions. As shown in Fig 10, we could get predicted different motif distributions (y axis) in the whole mouse brain regions (x axis). This result will give us tips on the analysis and construction of structural and functional whole mouse brain model.

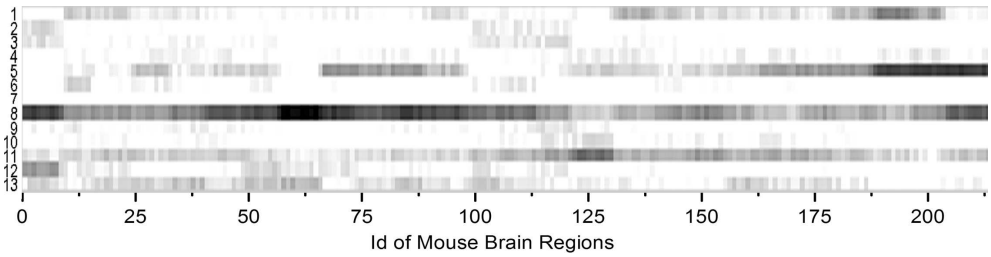


Figure 10. The predicted microscale motif distribution of the 213 regions in the mouse brain.

Here we introduce by far the largest microscale cortical neuron connectome on the mouse brain (to the best of our knowledge) from [Lee *et al* (2016)] to validate the prediction on the mouse brain. Due to the scalability issue for synaptic level reconstruction,

the largest microscale mouse brain cortical neuron connectome reported in literature by far (to the best of our knowledge) is a connectome containing 201 neurons and 1,278 synapses from V1 of the mouse brain [Lee *et al* (2016)]. Although the connectome of the 201 neurons are interconnections within these neurons, it still partially reflects the structural properties of the most realistic microscale connections. In the motif distribution analysis, [Lee *et al* (2016)] reported that there are only 4 types of three-node motifs (out of 13 types over all) which has been observed in the connectome (which is rational since the connectome is only on how the 201 neurons are interconnected), namely, motif No.1, No.2, No.4, and No.5 in Fig 6. The order and frequency of the motif distribution is No.4(1918) >No.2(430) >No.1(347) >No.5(32). While the predicted motif distribution order and ratio on V1 from the proposed model follows the sequence of No.4(0.17) >No.2(0.072) >No.1(0.04) >No.5(0.01). It indicates from the order perspective, that the predicted results are consistent with the real anatomical data. This provides an initial validation on the precision of the microscale motif prediction model. Here we want to extend the discussion why the predicted values are consistent with the real data from the order perspective, while still with inconsistency from the motif frequency ratio perspective. The main reason can be that the real anatomical data are interconnected motifs within the 201 neurons from a specific part of V1 (which is somewhat partial and local, due to current technology bottleneck), while our predicted results describe the motif distribution of the whole V1 area. If more comprehensive microscale connectome are available in the future, the predicted results and the real anatomic data can be compared in more details.

Although the proposed microscale whole brain predictive motif distribution model is still in preliminary phase, it gives us a perspective and sketch on the analysis and construction of structural and functional whole mouse brain model at the synaptic level.

5. Distribution Analysis for the Predicted Microscale Connectome

Since the microscale connectome focuses on synaptic connections within and among different brain regions. It is natural and essential to investigate on the distribution of synaptic connections since it reveal the structural characteristics of the mouse brain at the microscale. Based on the generated microscale connectome, the distribution analysis results of synapses in both long-range and short-range projections (synaptic connections within the same region) are shown in Fig 11 to Fig 14.

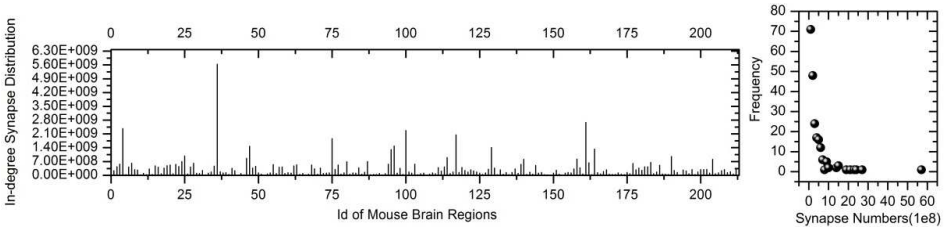


Figure 11. In-degree distribution of long-range projections among different brain regions.

Fig 11(left) is an analysis on the in-degree synaptic distribution of long-range projections for the mouse brain (including 213 regions). The in-degree describes the number of

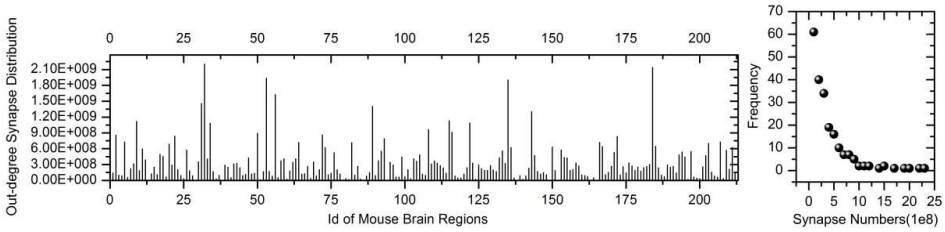


Figure 12. Out-degree distribution of long-range projections among different brain regions.

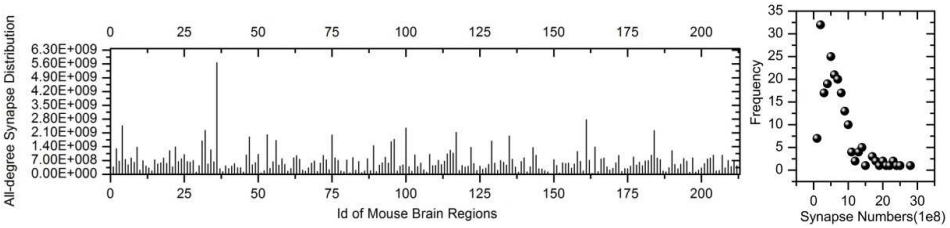


Figure 13. Degree distribution of long-range projections among different brain regions.

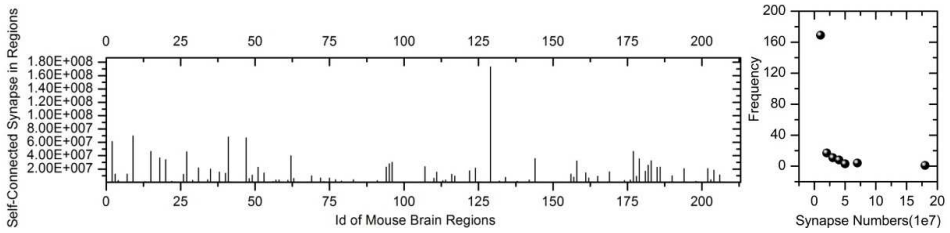


Figure 14. Synaptic degree distribution of projections within the same brain region.

synapses in all the other brain regions which converges on the same mouse brain region. The connections within sub regions are also calculated by the motif distribution and the number of neurons, as shown in Fig 14. Notice that only some regions are with relatively very large number of in-degree synaptic connections, and for most of the regions, the numbers of in-degree synaptic connections are relatively low. As shown in Fig 11(right), the log-log plot shows that the synaptic in-degree distribution of long range projections follows the power-law distribution.

Similar with the in-degree distribution, as shown in Fig 12(left), the synaptic out-degrees (projections which diverge from the same brain region) for different brain regions are to some extent bigger than the corresponding in-degree. Nevertheless the distribution also follows the power-law distribution (as shown in Fig 12(right)). In addition, if we combine the in-degree and out-degree together, the result shows the overall synaptic degree distribution characteristics among different brain regions, as shown in Fig 13.

6. Conclusion

In this paper, a computational approach for constructing an initial microscale mouse brain connectome from the neuron, synapse and neuron motif distribution perspective is proposed. The mesoscale connectome data and some microscale anatomical data are combined together to produce a step forward towards a microscale connectome through predictions on the distributions of neurons, synapses, and microscale neuron motifs. Degree distribution analysis is conducted from various perspectives to explore the characteristics of the connectome at the microscale. The proposed approach should also have a potential to be used for producing microscale connectome for other mammalian brain.

The structural connectome of the brain is a basis for the functional connectome. Structural connectome is definitely not the only reason which decides the function of the brain. Nevertheless, it provides the physical structural basis for information transmission among different building blocks at multiple scales. Hence, it is essential as structural support for cognitive function. We believe the relationship between structural connectome and functional connectome will become much clear as long as both of these connectomes become more complete. This is the reason why this paper aims to bridge the gap between mesoscale and microscale structural connectome of the mouse brain, to support future research on the comparative study on structural and function connectome at multiple scales.

In the future, we would plan to combine other upcoming concrete experimental data to support refined version of this study and make the microscale connectome more realistic, feasible and useful in various disciplines and application domains (e.g. brain simulation and brain-inspired intelligence models [Liu *et al* (2016)]).

7. Acknowledgement

This study was funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB02060007), and Beijing Municipal Commission of Science and Technology (Z161100000216124).

References

[Achard *et al* (2006)] S. Achard, R. Salvador, B. Whitcher, J. Suckling & E. Bullmore. (2006) *A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs*, The Journal of Neuroscience, vol. 26, no. 1, pp. 63–72.

[Mangan and Alon (2003)] S. Mangan & U. Alon. (2003) *Structure and function of the feed-forward loop network motif*, Proceedings of the National Academy of Sciences, vol.100, no. 21, pp. 11980–11985.

[Friston *et al* (1998)] K.J. Friston, F. PC, J. O & L.R. Turner. (1998) *Event-related fMRI: characterizing differential responses*, Neuroimage, vol. 7, no. 1, pp. 30–40.

[Assaf and asternak (2008)] Y. Assaf & O. Pasternak. (2008) *Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review*, Journal of Molecular Neuroscience, vol. 34, no. 1, pp. 51–61.

[Oh *et al* (2014)] S.W. Oh, J.A. Harris, L. Ng, B. Winslow, N. Cain, S. Mihalas, Q. Wang, C. Lau, L. Kuan, A.M. Henry, M.T. Mortrud, B. Ouellette, T.N. Nguyen, S.A. Sorensen, C.R. Slaughterbeck, W. Wakeman, Y. Li, D. Feng, A. Ho, E. Nicholas, K.E. Hirokawa, P. Bohn, K.M. Joines, H. Peng, M.J. Hawrylycz, J.W. Phillips, J.G. Hohmann, Paul Wohnoutka, C.R. Gerfen, C. Koch, A. Bernard, C. Dang, A.R. Jones & H. Zeng. (2014) *A mesoscale connectome of the mouse brain*, Nature, vol. 508, no. 7495, pp. 207–214.

[Sporns *et al* (2005)] O. Sporns, G. Tononi & R. Kotter. (2005) *The human connectome: a structural description of the human brain*, PLoS Computational Biology, vol. 1, no. 4, pp. e42.

[Osten and Margrie (2013)] P. Osten & T.W. Margrie. (2013) *Mapping brain circuitry with a light microscope*, Nature Methods, vol. 10, no. 6, pp. 515–523.

[Stosiek *et al* (2006)] C. Stosiek, O. Garaschuk, K. Holthoff, & A. Konnerth. (2006) *In vivo two-photon calcium imaging of neuronal networks*, Proceedings of the National Academy of Sciences, vol. 100, no. 12, pp. 7319–7324.

[Rakic (2006)] P. Rakic. (2006) *A century of progress in cortico-neurogenesis: from silver impregnation to genetic engineering*, Cerebral Cortex, vol. 16, no. suppl 1, pp. 3–17.

[Felleman and Essen (1991)] D.J. Felleman & D.C.V. Essen. (1991) *Distributed hierarchical processing in the primate cerebral cortex*, Cerebral Cortex, vol. 1, no. 1, pp. 1–47.

[Rocklanda and Pandyaa (1979)] K.S. Rocklanda & D.N. Pandyaa. (1979) *Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey*, Brain Research, vol. 179, no. 1, pp. 3–20.

[Bohland *et al* (2009)] J.W. Bohland, C. Wu, H. Barbas, H. Bokil, M. Bota, H.C. Breiter, H.T. Cline, J.C. Doyle, P.J. Freed, R.J. Greenspan, S.N. Haber, M. Hawrylycz, D.G. Herrera, C.C. Hilgetag, Z.J. Huang, A. Jones, E.G. Jones, H.J. Karten, D. Kleinfeld, R. Ketter, H.A. Lester, J.M. Lin, B.D. Mensh, S. Mikula, J. Panksepp, J.L. Price, J. Safdieh, C.B. Saper, N.D. Schiff, J.D. Schmahmann, B.W. Stillman, K. Svoboda, L.W. Swanson, A.W. Toga, D.C.V. Essen, J.D. Watson & P.P. Mitra. (2009) *A proposal for a coordinated effort for the determination of brain wide neuroanatomical connectivity in model organisms at a mesoscopic scale*, PLoS Computational Biology, vol. 5, no. 3, pp. e1000334, .

[Maizel *et al* (2011)] A. Maizel, D.V Wangenheim, F. Federici, J. Haseloff & E.H.K. Stelzer. *High-resolution live imaging of plant growth in near physiological bright conditions using light sheet fluorescence microscopy*, The Plant Journal, vol. 68, no. 2, pp. 377–385.

[Ragan *et al* (2012)] T. Ragan, L.R. Kadiri, K.U Venkataraju, K. Bahlmann, J. Sutin, J. Taranda, I.A. Carreras, Y. Kim, H.S. Seung & P. Osten. (2012) *Serial two-photon tomography for automated ex vivo mouse brain imaging*, Nature Methods, vol. 9, no. 3, pp. 255–258.

[Maniadakis and Trahanias (2003)] M. Maniadakis & P. Trahanias. (2003) *A computational model of neocortical-hippocampal cooperation and its application to self-localization*, Advances in Artificial Life, Springer, pp. 183–190.

[Seress (1988)] L. Seress. (1998) *Interspecies comparison of the hippocampal formation shows increased emphasis on the region superior in the Ammon's horn of the human brain*, Journal fur Hirnforsch, vol. 29, pp. 335–340.

[Cutsuridis *et al* (2010)] V. Cutsuridis, B. Graham, S. Cobb & I. Vida. (2010) *Hippocampal microcircuits: a computational modeler's resource book*, Springer.

[Treves and Rolls (1992)] A. Treves & E.T. Rolls. (1992) *Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network*, Hippocampus, vol. 2, no. 2, pp. 189–199.

[Witter (2010)] M.P. Witter. (2010) *Connectivity of the hippocampus*, Hippocampal Microcircuits: A Computational Modeler's Resource Book, Springer, pp. 5–26.

[Insausti *et al* (1998)] A.M. Insausti, M. Megias, D. Crespo, L.M. Cruz-Orivec, M. Dierssena, TF. Vallinaa, R. Insaustid & J. Floreza. (1998) *Hippocampal volume and neuronal number in Ts65Dn mice: a murine model of Down syndrome*, Neuroscience Letters, vol. 253, no. 3, pp. 175–178.

[Boss *et al* (1985)] B.D. Boss, G.M. Peterson, W.M. Cowan. (1985) *On the number of neurons in the dentate gyrus of the rat*, Brain Research, vol. 338, no. 1, pp. 144–150.

[Sporns and Kotter (2004)] O. Sporns & R. Kotter. (2004) *Motifs in brain networks*,PLoS Biology, vol. 2, no. 11, pp. e369.

[Zhang *et al* (2015)] T. Zhang, Y. Zeng, B. Xu. (2015) *Biological Neural Network Structure and Spike Activity Prediction Based on Multi-Neuron Spike Train Data*, International Journal of Intelligence Science, vol. 5, no. 02, pp. 102.

[Sporns *et al* (2007)] O. Sporns, C.J. Honey, R. Kotter. (2007) *Identification and classification of hubs in brain networks*, PloS One, vol. 2, no. 10, pp. e1049.

[Kasthuri *et al* (2015)] N. Kasthuri, K.J. Hayworth, D.R. Berger, R.L. Schalek, J.A. Conchello, S. Knowles-Barley, D. Lee, A. Vázquez-Reina, V. Kaynig, T.R. Jones, M. Roberts, J.L. Morgan, J.C. Tapia, H.S. Seung, W.G. Roncal, J.T. Vogelstein, R. Burns, D.L. Sussman, C.E. Priebe, H. Pfister, J.W. Lichtman. (2015) *Saturated reconstruction of a volume of neocortex*, Cell, vol. 162, no. 3, pp. 648–661.

[Lee *et al* (2016)] W.C.A. Lee, V. Bonin, M. Reed, B.J. Graham, G. Hood, K. Glattfelder, R.C. Reid. (2016).

Anatomy and function of an excitatory network in the visual cortex. Nature, vol. 532, no. 7599, pp. 370–374.

[Liu *et al* (2016)] X. Liu, Y. Zeng, T.L. Zhang, B. Xu. (2016). *Parallel Brain Simulator: A Multi-scale and Parallel Brain-Inspired Neural Network Modeling and Simulation Platform*. Cognitive Computation, vol. 8, no. 5, pp. 967–981.