Mouse-Brain Topology improved Evolutionary Neural Network for Efficient Reinforcement Learning

Xuan Han[1,2], Kebin Ja[1,3], and Tielin Zhang[2,3]

[1] Beijing University of Technology, Beijing, 100124, China.
[2] Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.
[3] Corresponding authors.

Abstract. The brain structures are key indicators to represent the complexity of many cognitive functions, e.g., visual pathways and memory circuits. Inspired by the topology of the mouse brain provided by the Allen Brain Institute, whereby 213 brain regions are linked as a mesoscale connectome, we propose a mouse-brain topology improved evolutionary neural network (MT-ENN). The MT-ENN model incorporates parts of biologically plausible brain structures after hierarchical clustering, and then is tuned by the evolutionary learning algorithm. Two benchmark Open-AI Mujoco tasks were used to test the performance of the proposed algorithm, and the experimental results showed that the proposed MT-ENN was not only sparser (containing only 61% of all connections), but also performed better than other algorithms, including the ENN using a random network, standard long-short-term memory (LSTM), and multi-layer perception (MLP). We think the biologically plausible structures might contribute more to the further development of artificial neural networks.

Keywords: Evolutionary neural network · Reinforcement learning · Mouse-brain topology.

1 Introduction

The mammalian brain has been studied deeply in the past decades. The brain structures are key indicators for representing the learned knowledge from millions of years of evolution in different animals. Like many genetic evolutionary algorithms in machine learning, the biological brain is the best evolutionary outcome in a natural organism. However, it is still an open question whether new network structures copied from natural brains could contribute to the development of artificial neural networks, whereby the design of structures is usually considered more important than neuronal types.

At present, the whole mouse brain has been widely examined, which contains around 213 brain regions, and the sparseness of the entire brain is about 36% [13]. Many types of research have been given on the topology of the mouse brain, and

one of the most important motivations is copying it to the conventional artificial neural networks (ANNs) for higher performance, or more energy efficiency [14].

Besides network architectures, which are usually designed by nature, network plasticity or learning algorithms are also important for nurturing cognitive functions. The evolutionary algorithm [2] is biologically plausible, which refers to the evolutionary operations for gene coding, population initialization, crossover (or mutation) operations, and other extended operations. The evolutionary computation also uses a global optimization function for self-organizing, adaptive- and self-learning to solve complex tuning problems.

Inspired by the biological networks and evolutionary algorithms, we propose a mouse-brain topology improved evolutionary neural network (MT-ENN), which incorporates biologically plausible brain structures and an evolutionary learning algorithm for efficient learning on two benchmark reinforcement learning (RL) tasks. The dynamic neurons contain 1st-order dynamics of membrane potentials supported by some additional key dynamic parameters. Then these neurons are connected by the copied topology from the biological brain and tuned by a global evolutionary algorithm. The two open-AI gym [3] tasks, e.g., MountainCar-v2 and Half-cheetah-v2, are selected to verify the proposed MT-ENN.

In addition, we also made a three-dimensional (3D) visual reconstruction of the entire network topology after network learning, which is convenient for identifying the connectome of different brain regions and might inspire back to the neuroscience researchers and answer the question that why some topology or brain regions are important for reinforcement learning.

2 Related works

A new continuous-time differential learning was proposed for RL tasks containing continuous dynamics [5]. Then, a hybrid learning framework was proposed by extending this new learning rule with a multi-scale dynamic coding [17]. The topology-focused algorithm was proposed by using a sub-network to replace a previous global network and achieved comparable performance [10], named as the lottery ticket hypothesis [7]. A biological network using C.elegans topology was proposed to achieve higher performance than algorithms using random networks [9], which showed the efficiency of the biological topology.

These algorithms all performed well on RL tasks. However, they seldom use complicated network topology. One of the main motivations of this paper is that the further incorporation of network topologies, especially from smarter animals than C.elegans, will strengthen intelligent algorithms on cognitive functions to handle RL tasks. In addition, we think reducing the size of fully connected networks for better interpretability is at least as important as performance, which will also be further discussed in the following sections.

3 Methods

3.1 The Allen Mouse Brain Atlas

The Allen Institute for Brain Science has provided the public with the whole mouse brain atlas, containing neuronal types and network topology from some standard mouse brains. The network topology contains the directional mapping of 213 brain regions, which are further visualized by Houdini software for a better understanding.

The overall connectome of the mouse brain is shown in Fig. 1A, where the 213×213 matrix is the connectomes between 213 brain regions. The connectome is sparse, indicating that the bottom-up and top-down connections are hierarchical and highly related to cognitive functions.

3.2 The clustered hierarchical circuits

The connectivity matrix of the mouse brain atlas was then clustered into subclusters for easier analysis. It is a general challenge to find a proper clustering algorithm that could filter out trivial branches but leave out the key functional topology. A special hierarchical clustering method, i.e., Tanimoto coefficient algorithm [12], is selected for this clustering. The algorithm could generate multiple clusters given a directional topology [1]. The detailed procedure of the clustering algorithm could be concluded as the following equations:

$$S(e_{ik}, e_{jk}) = \frac{a_i \cdot a_j}{\mid a_i \mid^2 + \mid a_j \mid^2 - a_i \cdot a_j}, \tag{1}$$

We divide 213 regions of the mouse brain into 71 clusters at a cutting height of 0.8 (relative value designed by experience) based on the links or connections between each two brain regions in both directions as well as weights or intensities of projections (Fig. 1C, D).

3.3 The neuron model

Here, we briefly describe the neuron and synapse model to design neural circuit dynamics [8] by using the following equation.

$$\begin{cases} \dot{V}_i(t) = [I_L + \sum I_C(t)]/C_m \\ I_L(t) = \omega_L[E_L - V_{post}(t)] \\ I_C(t) = \omega_C[E_C - V_{post}(t)]g(t) \\ g(t) = 1/[1 + exp(-\sigma(V_{pre}(t) - \mu))] \end{cases}, \tag{2}$$

where C_m is the membrane capacitance of the neuron, I_C and I_L are the input currents of the chemical channel and leakage channel, respectively. E_C and E_L are the corresponding reversal potentials. $V_{post}(t)$ and $V_{pre}(t)$ are the membrane potentials of post-synapses and pre-synapses, respectively. $g(t)$ is the membrane conductance, defining whether a synapse is excitatory or inhibitory by E_C. ω_C and ω_L stand for the conductances of the chemical channel and leakage channel.

These equations show that the algorithms have a strong ability to model the time series reaching any time step [4, 6].

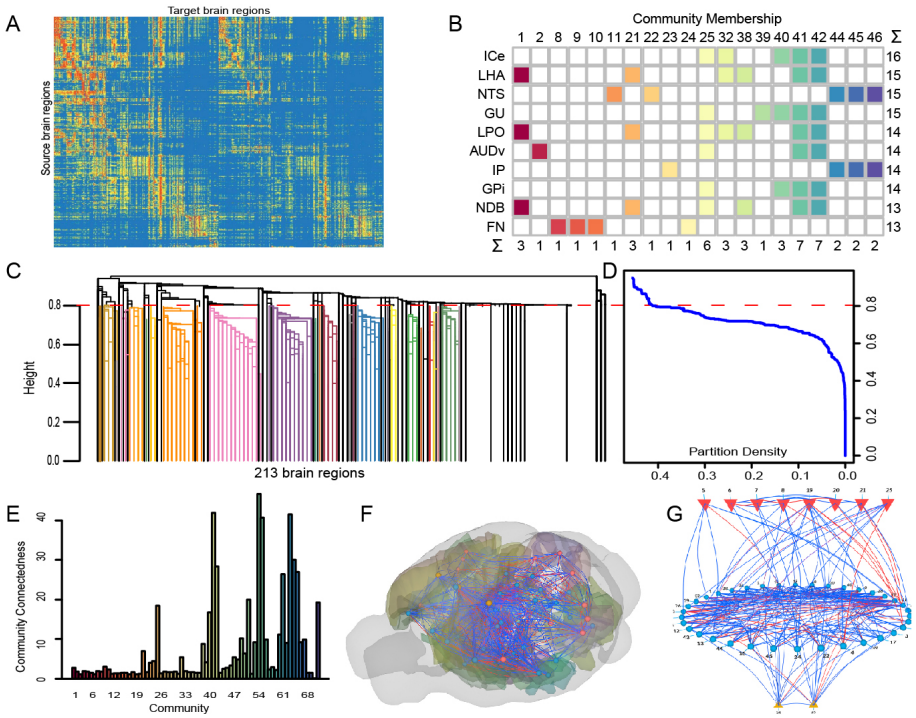

Fig. 1. The generation procedure of brain-inspired topology. (A) The mesoscale connectome of the Allen mouse brain atlas. Each dot is the mapping from a source (y-axis) to a target (x-axis) brain region with strong (red) or weak connectivity (blue). The source data is from the mesoscale Allen mouse brain atlas [13]. (B) An example of clustered topology from 213 brain regions. The most frequent brain regions and the communities share the most nodes of 71 communities with a cutting height of 0.8. (C,D) The hierarchically clustered brain regions. The 213 mouse brain regions are split into 71 clusters at a cutting height of 0.8 (C) and the partition density (D). In order to cut out more clusters with the appropriate cluster size, the cutting height is manually designed. (E) The connectedness of different communities (clusters). The connectedness of individual link communities with cutting at 0.8 including 71 communities. Community-71, Community-65, and Community-61 contain 31, 49, and 46 brain regions with sparse connectivity. (F) Schematic diagram of the connection of 46 brain regions in the 3D brain. The 46 brain regions contain three parts: the red sphere represents the input layer (sensory brain regions); the blue sphere represents the middle layer (hidden brain regions for memory or other cognitive functions); the yellow sphere represents the output layer (motor brain regions). Red lines represent inhibitory connections, and blue lines represent excitatory connections. (G) Schematic diagram of the connection of 46 brain regions in 2D visualization.

3.4 Coping the biological circuits to artificial ones

According to the definition of these 213 brain regions in the Allen mouse brain atlas, we reorganized them into sensation, hidden, and motor brain regions, according to structural projection and physiological function. We also annotate the biological functions of the brain regions at each level of the communities of interest, which makes it possible to compare the hierarchical composition and functional circuitry of the community or groups of interest with the combination of the atlas.

Hence, some different types of topology are selected for the next-step simulation, including the whole 213 brain regions (directly copied from the biological atlas) and some other sub-brain topology (selected from 31, 46, and 49 brain regions based on the hierarchical clustering), which will be further discussed in the experimental sections.

3.5 The network learning

In this paper, a simple search-based algorithm is selected for network reinforcement learning [9], whereby the agent learns to make decisions after observing the current state in an environment and then receives a timely or delayed reward. A fitness function is designed to collect these rewards and guide the direction of the random search. At the beginning of learning, the agent makes random decisions for exploration, and a good decision for a lower fitness function will be kept by saving the current parameters and focusing more on the exploitation. The search-based algorithm ARS can train a network by repeating two training strategies until convergence. First, V_n values are obtained by calculating θ with fitness function f. Then the adaptive search algorithm calculates $E(R_\theta)$ by the average value of the worst K samples among the V_n values.

In addition, two benchmark Mujoco RL tasks are used to test the performance of the proposed algorithm. One important motivation is that the RL tasks are more related to the biological agent's learning procedure, which makes decisions after observation by maximizing its predicted rewards.

4 Experiments

4.1 The clustered brain regions

The 213 brain regions were clustered into multiple graphs, and parts of critical brain regions shared by different communities were recorded in Fig. 1B. It represents a sparse connection between different brain regions, congruent to the biological brain.

The community-61, community-65, and community-71 in 71 clusters were selected for subsequent experiments, given the cutting height of 0.8. It is impressive that the number of nodes in the largest connectedness in all clusters is less than 60 (Fig. 1E), indicating the network is very sparse. The community-61 contains 46 brain region nodes, which is called Circuit-46 (which means the clustering

containing 46 brain regions). Similar to it, we give the names of community-65 and community-71 as Circuit-49 and Circuit-31, respectively. These circuits obtained after clustering will be used for the next-step RL experiments.

4.2 The network topology from biological mouse brain

The Circuit-46 is visualized in a standard 3D mouse brain common coordinate framework (CCF) from the Allen Brain Institute (Fig. 1F), where each point represents a brain region, and each link represents excitatory (blue ones) or inhibitory (red ones) connections between different regions (the total is 819 connections). We set 657 excitatory, and 162 inhibitory connections for Circuit-46, inspired by the biological discovery [16]. The connectivity strength lower than 0.05 was omitted and only left for those larger than the threshold for the ease of visualization (i.e., only 207 excitatory connections and 60 inhibitory connections were left after filtering).

In addition, the sensory, hidden, and motor brain regions are visualized in a 2D figure, including 8 sensory areas (red inverted triangles), 36 hidden areas (blue circles), and 2 motor areas (yellow triangles), as shown that in Fig. 1G.

4.3 Analyse of the network Motif

The distribution of all the 13 types of 3-node motifs is shown in Fig. 2A, B, C. The "credible frequency", which is the product of the occurrence frequency and $1 - P$, was proposed to visualize the composition of the network. P is the P-value of each motif in the selected network compared to 1,000 randomly generated networks of the same size. Each generated network was sampled from a uniformly random distribution. The smaller P-value, the less likely it is that a random network will have the same occurrence probability as the corresponding motif, which means higher credibility of the motif exists in the selected input network. Hence, the proposed "credible frequency" is an appropriate index to reveal the basic units of the input mouse brain circuit without the interference of the low-credibility motif.

As shown in Fig. 2A,B,C, we found that compared with other circuits, the distribution of Motif-5 (cross-layer connections) is higher in Circuit-46, which is important for network learning and will be remained for the further analysis of the following sections.

4.4 Results with different network degrees

We conducted an experiment where random connections were given to a MT-ENN with the same number of neurons and synapses as those in the Circuit-46. The synapses' initial polarity was set randomly (excitatory or inhibitory). Only feedforward connections were given from sensory to hidden neurons, same as those in the hidden to motor neurons. The random circuits were then trained on RL tasks whose performances are reported in Fig. 2D,G. We observed that

the performance of MT-ENN using mouse brain topology was higher than that using random connections, and the performance would be more significant for a simple (e.g., MountainCar-v2) instead of hard (e.g., Half-cheetah-v2) tasks. This result on two benchmark RL tasks demonstrates the usefulness of the biological structures.

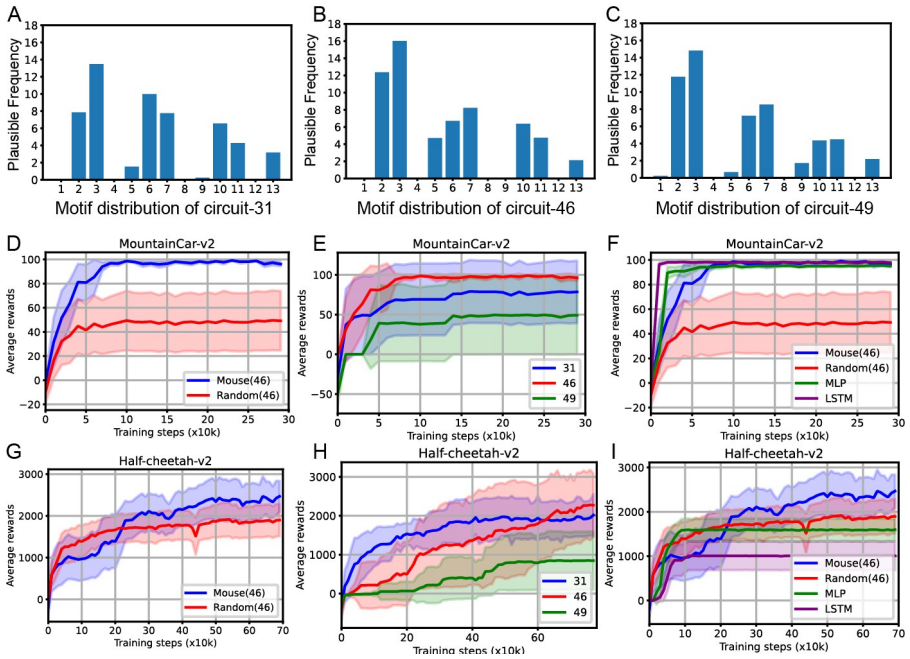

Fig. 2. Performance of different algorithms on reinforcement learning tasks. (A-C) Motif distribution of Circuit-31, Circuit-46, and Circuit-49. The horizontal axis represents 13 different types of motifs, and the vertical axis represents the trusted frequency. The higher the frequency, the higher the proportion of such motif. (D,G) Performance of the MT-ENN using Circuit-46 and random structures on the MountainCar-v2 and Half-cheetah-v2 tasks, respectively. The horizontal axis is the number of training steps, and the vertical axis is the average scores. The shadows represent the area of standard deviation. In the MountainCar-v2 task, a 100 score would be given only after they reached the top. (E,H) Some to those in (D,G) but for the MT-ENNs using Circuit-31, Circuit-46, and Circuit-49. (F,I) Some to those in (D,G) but for the MT-ENNs using Circuit-46, random connections, MLP, and LSTM.

4.5 Results with different number of brain regions

We conducted a series of experiments to test the influence of the number of brain regions, including MT-ENNs using Circuit-31, Circuit-46, and Circuit-49, respectively. As shown in Fig. 2E,H, the MT-ENN using Circuit-46 is better than

those using the other two structures. This result confirmed our hypothesis that Circuit-46 could be a winning ticket, containing better performance and lower sparsity. Hence, we selected Circuit-46 as MT-ENN's main biological structure basis for the subsequent experiments.

Motif distributions of Circuit-46 were calculated in Fig. 2B, where the Motif-5 occupied a higher proportion than that in Circuit-31 (Fig. 2A) and Circuit-49 (Fig. 2C). Hence, we speculated that a proper proportion of cross-layer connections might play a significant role in RL tasks.

4.6 Results with other SOTA algorithms

We further tested the performance of our MT-ENN with other SOTA algorithms, including long short-term memory (LSTM) and multi-layer perception (MLP). We keep the comparisons fair by using the same number of neurons, linear mapping functions, and learning algorithms. Other key parameters were also kept learnable. We chose the same number of cells (neurons) for the LSTM and MLP networks, equal to the size of the MT-ENN. The LSTM and MLP networks are fully connected, and the MT-ENN achieves 61% network sparsity. The result is shown in Table 1 and Fig. 2F,I.

As shown in Table 1, for the MountainCar-v2 task, our MT-ENN reached the score (mean reward) of 99.14, higher than other algorithms tuned with backpropagation. Given a more complex task, e.g., the Half-cheetah-v2 task, our MT-ENN reached a much higher performance (2,468 scores) than the state-of-the-art MLP (1,601 scores) and LSTM (1,009 scores). These results showed the efficiency of the biologically plausible brain circuits.

Table 1. The performance comparisons of the proposed MT-ENN with other SOTA algorithms on MountainCar-v2 and Half-cheetah-v2 tasks.

Tasks	Architectures	Learning rules	Performance	Sparsity
MountainCar-v2	LSTM [11]	BPTT	98.98±0.59	0%
	MLP [15]	PPO	95.5±1.5	0%
	Random	Search	49.59±49.59	61%
	MT-ENN (Ours)	**Search**	**99.14±0.12**	**61%**
Half-cheetah-v2	LTSM [11]	BPTT	1009.51±641.95	0%
	MLP [15]	PPO	1601.05±506.50	0%
	Random	Search	1917.40±819.39	61%
	MT-ENN (Ours)	**Search**	**2468.18±962.36**	**61%**

5 Discussion

Selectively copying biological structures into artificial neural networks is a shortcut for efficiently designing neural networks. In this paper, 213 mouse brain regions were clustered and analyzed to generate some sub-graph topology for the network design of MT-ENN.

The 3D morphology helps us learn more about the neuron types [19], sparseness, and connectome of different brain regions during analysis. The comparisons of different sub-graph topologies will be extended for further discussion [18]. Combined with the biological understandings, these clustered results will help us select more topology that satisfies biological plausibility and efficiency.

Interestingly, the clustered sub topology circuits are also very similar to the biological circuits. As clustered Circuit-31, Circuit-46, and Circuit-49, for example, they all contain sensory, hidden (for memory), and motor brain regions. For the topology with more types of sensation areas, it will be obvious that they could contribute more to the multisensory competition, which has been verified in Circuit-31, which has more sensory neurons ratio but poorer performance than Circuit-46. Hence, the topology structure of Circuit-31 may not well meet the needs of signal integration processing, resulting in the weakening of learning.

The sensory regions in Circuit-46 are mainly auditory, and there are more middle regions for integration. The performance of Circuit-46 is the best in the experiment, which gives us hints and inspiration that Circuit-46 might be more biologically plausible for dynamic information processing. In addition, more brain regions in Circuit-46 are related to the cognitive functions that need more global connections, including neuromodulatory pathways such as 5-HT projections from the CLI region and brain regions related to spatial navigation like the nucleus and thalamus, which corresponds to the current experimental pattern. This series of factors may be why Circuit-46 performs better than other circuits.

Circuit-49 has many olfactory circuits in the sensory regions, which puts greater pressure on information integration and may not match the current experiments. We find that Circuit-49 contains more regions projected with the hippocampus, but the hippocampus regions do not belong to Circuit-49, which may be one of the reasons for poor performance.

The experimental results have verified our hypothesis that mouse brain topology can improve evolutionary neural networks for efficient RL. In this paper, the brain region structure obtained by the joint action of mouse brain clustering and biological trust is better than the commonly used LSTM and MLP in RL tasks. In addition, the mouse brain structure also showed advantages in terms of sparsity. In the future, more biologically credible principles can be borrowed from biological networks and applied to neural networks to achieve better integration of neuroscience and artificial intelligence and promote each other.

Acknowledgements This work was supported in part by the Shanghai Municipal Science and Technology Major Project (2021SHZDZX), the Youth Innovation Promotion Association CAS, the National Natural Science Foundation of China under Grants No. 82171992 and 81871394. More information about the source code and introduction of brain topologies can be found at https://github.com/thomasaimondy/ENN.

References

1. Ahn, Y.Y., Bagrow, J.P., Lehmann, S.: Link communities reveal multiscale complexity in networks. nature **466**(7307), 761–764 (2010)
2. Bäck, T., Schwefel, H.P.: An overview of evolutionary algorithms for parameter optimization. Evolutionary computation **1**(1), 1–23 (1993)
3. Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., Zaremba, W.: Openai gym. arXiv preprint arXiv:1606.01540 (2016)
4. Chen, R.T., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. arXiv preprint arXiv:1806.07366 (2018)
5. Doya, K.: Reinforcement learning in continuous time and space. Neural computation **12**(1), 219–245 (2000)
6. Dupont, E., Doucet, A., Teh, Y.W.: Augmented neural odes. arXiv preprint arXiv:1904.01681 (2019)
7. Frankle, J., Carbin, M.: The lottery ticket hypothesis: Finding sparse, trainable neural networks. arXiv preprint arXiv:1803.03635 (2018)
8. Hasani, R., Lechner, M., Amini, A., Rus, D., Grosu, R.: Liquid time-constant networks. arXiv preprint arXiv:2006.04439 (2020)
9. Hasani, R., Lechner, M., Amini, A., Rus, D., Grosu, R.: A natural lottery ticket winner: Reinforcement learning with ordinary neural circuits. In: International Conference on Machine Learning. pp. 4082–4093. PMLR (2020)
10. Hinton, G., Vinyals, O., Dean, J.: Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531 (2015)
11. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**(8), 1735–1780 (1997)
12. Kalinka, A.T., Tomancak, P.: linkcomm: an r package for the generation, visualization, and analysis of link communities in networks of arbitrary size and type. Bioinformatics **27**(14), 2011–2012 (2011)
13. Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al.: A mesoscale connectome of the mouse brain. Nature **508**(7495), 207–214 (2014)
14. Rubinov, M., Ypma, R.J., Watson, C., Bullmore, E.T.: Wiring cost and topological participation of the mouse brain connectome. Proceedings of the National Academy of Sciences **112**(32), 10032–10037 (2015)
15. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347 (2017)
16. Wildenberg, G.A., Rosen, M.R., Lundell, J., Paukner, D., Freedman, D.J., Kasthuri, N.: Primate neuronal connections are sparse in cortex as compared to mouse. Cell Reports **36**(11), 109709 (2021)
17. Zhang, D., Zhang, T., Jia, S., Xu, B.: Multiscale dynamic coding improved spiking actor network for reinforcement learning. In: Thirty-Sixth AAAI Conference on Artificial Intelligence (2022)
18. Zhang, T., Zeng, Y., Xu, B.: A computational approach towards the microscale mouse brain connectome from the mesoscale. J Integr Neurosci **16**(3), 291–306 (2017)
19. Zhang, T., Zeng, Y., Zhang, Y., Zhang, X., Shi, M., Tang, L., Zhang, D., Xu, B.: Neuron type classification in rat brain based on integrative convolutional and tree-based recurrent neural networks. Sci Rep **11**(1), 7291 (2021)