ISSN: 2688-8645

Learning the Heuristics of Hub Identification over Brain Networks

Anqi Chen

School of Artificial Intelligence, Zhejiang College of Security Technology, Wenzhou, Zhejiang, China

chenanqicool@gmail.com

Abstract

A plethora of neuroscience studies find that connector hub nodes play a key role in regulating multiple modules and supporting brain functions such as consciousness and cognition, due to its critical topological location in the network. Current approaches mainly rely on the hand-crafted attributes (aka. graph embedding at each node) from the domain knowledge of network neuroscience such as high connectivity degree to identify connector hub nodes. However, simple ranking heuristic based on the pre-defined attributes has limited power to characterize the complex network topology, which often results in less accurate hub identification results. Although graph theory allows us to find connector hubs with a greater mathematical insight, the large scale of brain network often compromises the well-defined optimization into a local and sub-optimal solution. To overcome these limitations, we propose a joint graph embedding and hub identification solution in a reinforcement learning framework to discover the unprecedented heuristics from the existing knowledge of network neuroscience and graph theory, which allows us to outperform the current state-of-the-art hub identification methods. We have achieved more reliable and replicable hub identification results on both simulated and real brain network data, suggesting the high applicability to various network analysis studies in neuroscience and neuroimaging fields.

Keywords

Graph Embedding; Reinforcement Learning; Hub Identification.

1. Introduction

Human brain is a complex and economic system where large portion of brain functions are supported by a small number of hub nodes [1, 2]. For example, a recent study [3] shows that densely connected hub nodes play an important role in integrating and exchanging the global information from different parts of the brain. In light of this, investigation on network organization becomes the key to understand brain function as well as the emerging behavior. Since connector hubs usually link multiple network modules (communities), many efforts have been investigated into the accurate identification of connector hubs since last decade [2, 4-6].

In network neuroscience field, hand-crafted nodal attributes such as connectivity degree are widely used to rank the likelihood of being connector hub nodes, based on the intuition that hub nodes are usually densely connected to non-hub nodes in the network as well as the statistical principles such as small-worldness. In general, the idea is to separate the hub nodes and non-hub nodes based on the embedding vectors (e.g., a set of nodal attributes) on each node that characterizes the topological difference between hub and non-hub nodes. However, it is very difficult to find a universal attribute that works well for all brain networks since provincial hub nodes have similar nodal attributes except most of the links are presented within the same

ISSN: 2688-8645

module [7, 8]. Since the heuristic of selecting connector hub nodes is simply by ranking, it becomes highly empirical to find the cut-off, which undermines replicability. More critically, the embedding vector is not adjusted in selecting hub nodes, despite that the removal of selected connector hub nodes has significantly changed the network topology. Such gross simplification is one of the major reasons for less accurate results of connector hub nodes. In what follows, we refer hub as connector hub unless stated otherwise.

Recently, graph theory has been used to detect network module by clustering the graph embeddings that are derived from the network Laplacian matrix. Although it is efficient to find the hub nodes based on the network communities, the accuracy of hub identification result completely relies on the network community detection which is still computationally challenging. A joint hub identification method has been proposed in [9] by examining the break-down effect after removing a set of hub nodes. Although this approach is supported by the very insightful mathematics theory, the global optimization of all possible combinations of hub nodes is NP-hard for large-scale brain network. Thus, the authors relax the combinatory optimization into a convex and continuous energy function at the expense of sub-optimal solution under the framework of ADMM (Alternating Direction Method of Multipliers) [10].

None of the above state-of-the-art hub identification method uses machine learning technique, since the lack of ground truth regarding the hub locations in the network prohibits the classic supervised learning scheme. In this context, the alternative solution is reinforcement learning of the unprecedented heuristics for hub identification from the insights of the existing algorithms. Specifically, we first construct an incremental exemplar container which consists of existing brain network data as well as multiple hub identification results by various existing algorithms such as sorting-based and module-based methods. Next, we propose a reinforcement learning framework to jointly find (1) the conditional graph embedding vector for each node which dynamically reflects the topological profile in the context of identified hub nodes in the network, and (2) the best heuristic of hub identification to achieve highest reward under the current setting of graph embeddings, which is guided by mathematical principles in graph theory. As the plot in Fig. 1, our reinforcement learning is initialized from the expert heuristics from network neuroscience. As we incrementally augment the exemplar container by adding more and more accurate hub identification results (with the latest learned heuristics), our reinforcement learning model continues to polish the heuristics of identifying the hub nodes, which allows our learning-based method to perform beyond human experts. In the testing stage, we sequentially find connector hub nodes by repeating three steps at a time: (1) updating conditional graph embeddings based on the current hub selection result, (2) calculating the rewards based on the learned heuristics, (3) selecting new node with largest reward as the connector hub node.

In the experiments, we have evaluated our reinforcement learning based hub identification method on both simulated and real brain network data. Compared with existing hub identification methods, we achieve more accurate results in terms of reliability and replicability, which indicates the wide application of our reinforcement learning approach in network analyses studies.

Fig. 1. shows the reinforcement learning scenario is grounded on the fundamental of network neuroscience and graph theory. Essentially, we train the neural network to progressively polish heuristics for hub identification from an incremental exemplar container which consists of real network data and the hub identification results using the latest learned heuristics. Under the hood, our proposed neural network architecture has two major components, i.e., graph embedding network (in blue box) and reinforcement heuristic learning network (in red box). In the testing stage, we sequentially identify a set of hub nodes by dynamically updating the graph embeddings (via graph embedding network) and selecting new hub node with largest reward (via reinforcement learning network).

ISSN: 2688-8645

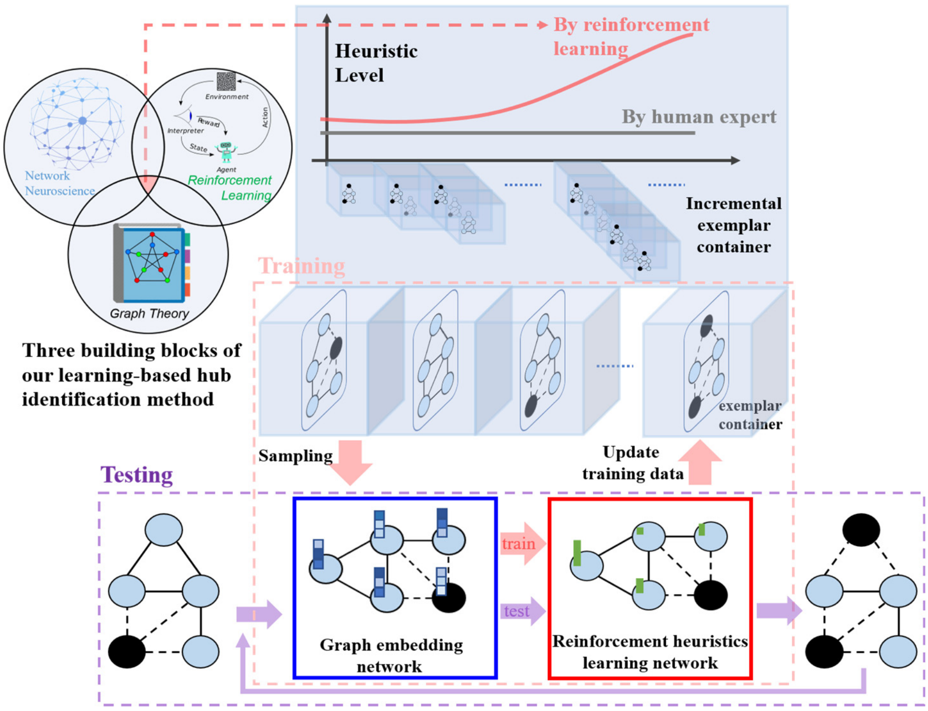

Fig 1. The overview of our learning-based hub identification method.

2. Method

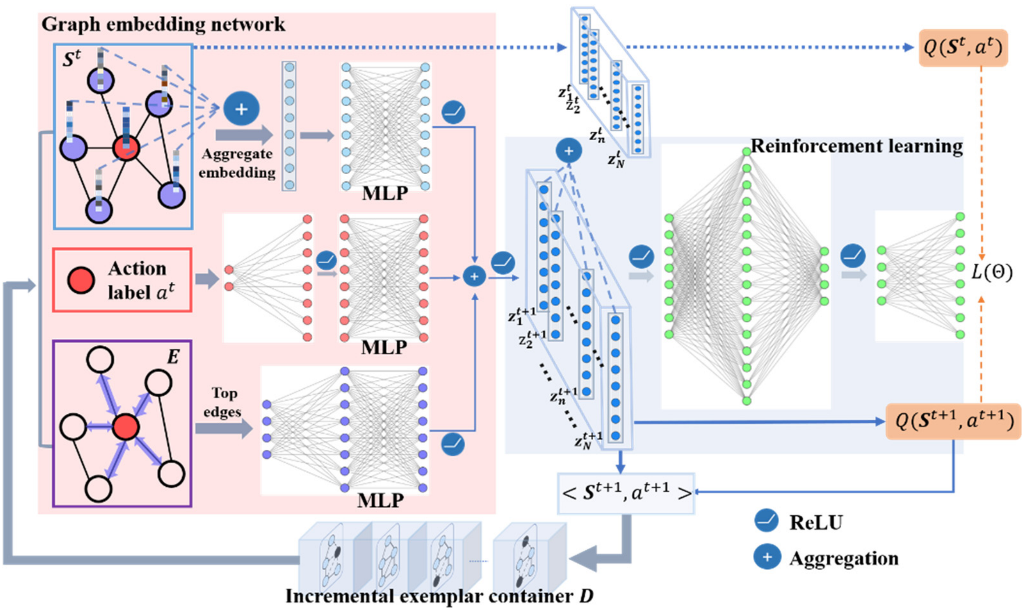

Fig 2. The network architecture of our learning-based hub identification method.

It is common to encode connectivity information in the brain network (with nodes) into a graph structure $\mathcal{G} = (V, E)$, where $V = \{v_n | n = 1, \ldots, N\}$ the node set and E is a $N \times N$ adjacency matrix, respectively. We use vector z_n denote the embedding vector at each node v_n. Since we dynamically update the embedding vector in the sequential hub identification, the z_n becomes a time-dependent embedding vector z_n^t where the variable t indicates the stage of identifying the t^{th} hub node. Accordingly, $S^t = \{z_n^t | n = 1, \ldots, N\}$ is called state since the assemble of embedding vectors is related to the current hub identification result. a^t is a N-length action

ISSN: 2688-8645

vector, where each element is binary ('0' for non-hub and '1' for hub node) and describes the action of selecting the t^{th} hub node based on the existing $(t-1)$ selected hub nodes.

The overview of our first-ever learning-based hub identification method is shown in Fig. 2. In general, there are three key components: (1) an incremental exemplar container, (2) a graph embedding network, and (3) a reinforcement heuristics learning network. We explain each component as follows.

Incremental exemplar container. The exemplar container D^t consists of a set of duals $\langle S_m^t, a_m^t \rangle$ $(m = 1, \dots, M)$ from total M training network data. In the beginning, we set all zeros to each S_m^0 and a_m^0. After selecting the first hub node for all training samples, we update the embedding vectors based on the current graph embedding network and check in $\{S_m^1\}$ and $\{a_m^1\}$ into the container, and so on. Note, such progressive increment of exemplar container D^t allow us to learn a full spectrum of hub identification heuristics using one neural network even we select hub nodes in a sequential manner.

Graph embedding network. The input to our graph embedding network includes (1) previous embedding vector z_n^t at each node v_n, (2) the selected or not-selected result at v_n from previous action vector a^t, and (3) the connectivity profile at v_n from the underlying adjacency matrix E. We deploy three multi-layer perception (MLP) network followed by a ReLU unit to learn the non-linear mapping which eventually generate a new conditional graph embedding vectors z_n^{t+1}. To be robust to spurious links in the network, we use the average embedding vector over the directly connected nodes as the input to MLP for z_n^t, and we only allow top four strongest links (instead of N links) at each node as the input to MLP for connectivity profile.

Reinforcement heuristics learning network. Our reinforcement learning network consists of four key learning elements: (1) state S^t where we assume the conditional graph embeddings S^t describes the current network topology after t nodes have been identified as hub nodes, (2) action a^t which describes the procedure of selecting a new hub node based on existing $(t-1)$ hubs, (3) reward function R which quantify the reward of identifying one node as the hub node, and (4) terminal state where we define the ending point as total K hubs has been found.

The reward function R steers the reinforcement learning of hub identification heuristics. Here, we follow the object function in [6] to quantify the reward by the break-down effect after the removal of hub nodes out of the brain network. Specifically, we first turn the action vector a^t into a diagonal matrix A^t. Then we calculate the degraded adjacency matrix $E^t = (A^t)^T E A^t$ which reflects the removal of hub nodes by zeroing corresponding rows and columns. Since each connector hub node links multiple modules, the total Eigenvalue of the underlying Laplacian matrix (after removing the connector hub nodes) is expected to smaller than the counterpart total Eigenvalue by including any non-hub nodes. Thus, we define reward function as $R(S^t, a^t)$ which returns inverse of total Eigenvalue from the degraded Laplacian matrix.

In classic reinforcement learning, the best action at state S^t is taken by the highest return $Q(S^t, a^t) \in \mathbb{R}$ based on not only the reward $R(S^t, a^t)$ but also the future return $Q(S^{t+1}, a^{t+1})$, which forms a Markov decision process (MDP). Hence, we essentially learn a Q-function $Q^*(S^t, a^t)$ by maximizing the following statistical expectation in a recursive way as:

$$Q^*(S^t, a^t) = \mathbb{E}[R(S^t, a^t) + \gamma \max_{a^{t+1}} Q^*(S^{t+1}, a^{t+1})] \quad (1)$$

where γ is called discount rate[11]. Since the number of brain networks is almost infinite in the real application, the infinite stetting of $\{S^t\}$ prohibits to learn Q^* using the conventional Q-learning [12]. Instead, we use deep Q-network (DQN) [13] to establish the mapping $Q^*(S^t, a^t, \Theta)$ from the given state S^t and action a^t, where the solution space is parameterized by network parameters Θ. The network architecture is shown in the blue box of Fig. 2, where the input is S^t and output is the value of Q-function $Q(S^t, a^t)$ for each action a^t. The loss function $L(\Theta)$ is designed to minimize the squared difference between the left and right side of Eq. 1. To do so, we first substitute S^t with S^{t+1} to train the DQN and update the network parameters from Θ^- (old) to Θ (latest). Since we can track the selected hubs from S^t to S^{t+1} for each instance, we can obtain the Q-function $Q^*(S^t, a^t, \Theta)$ with the latest network parameter Θ, which approximates the left side of Eq. 1. Similarly, we calculate the Q-function $Q^*(S^{t+1}, a^{t+1}, \Theta^-)$ using the old network parameters Θ^-, which approximates the third term in Eq. 1, i.e., the future return at state S^{t+1}. Thus, the DQN squared error loss function $L(\Theta)$ is defined as:

$$L(\Theta) = \left\| R(S^t, a^t) + \gamma \max_{a^{t+1}} Q^*(S^{t+1}, a^{t+1}; \Theta^-) - Q^*(S^t, a^t, \Theta) \right\|^2 \tag{2}$$

Stochastic gradient descent is used to backpropagation in training DQN. We set discount rate $\gamma = 1$. Mini-batch size is set to 128. We obtain the new network parameter Θ after 1,000 updates from the old parameters Θ^-. Once the training of DQN converges, it is straightforward to determine the best action a^t (selecting t^{th} hub node) based on stage S^t. Then we check in the dual instance $< S^t, a^t >$ into the exemplar container to keep polishing the heuristics of hub identification.

Testing stage—hub identification for the unseen network. We identify hub nodes for the unseen brain network in a sequential manner. In the beginning, we initialize graph embedding S^0 with zeros. Thus, the first identified connector hub node is completely based on the reward function which is the graph spectrum of the input network. From the second hub node, we first update the graph embedding vector at each node based on the selected hub nodes and then take the best action (selecting next hub) based on the output of DQN. We repeat these steps until K hubs have been identified.

3. Experiments

3.1. Simulated Network Data

First, we synthesize 500 networks as the training dataset, each with 34 nodes (same as the karate club data [16]), where the each link is the binary result of random guess. Then we apply the trained neural network to identify hub nodes on the karate club data which is considered having four communities in many network science literatures [17]. Based on the literature review, we set . From Fig. 3(a)-(d), we show the hub identification results on karate club data by sorting-based method, module-based method, optimization-based method, and our learning-based method, respectively. It is clear that our learning-based method achieves the most accurate hub identification result since the removal of six hub nodes (as well as their links) completely break down into four disconnected communities.

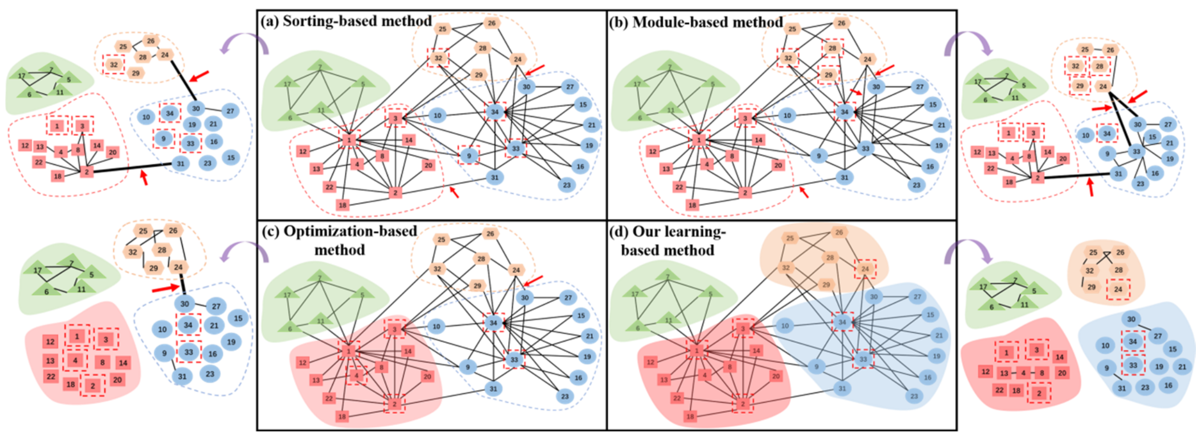

Fig 3. The identification results of four methods on karate club data by sorting-based (a), module-based (b), optimization-based (c), and our learning-based method (d).

ISSN: 2688-8645

3.2. Evaluating Hub Identification Results on Structural Networks

The training dataset consists of 502 structural networks constructed from the neuroimaging data in ADNI database, where we use Destrieux atlas [18] to parcellate cortical surface into 148 regions and measure the structural connectivity strength via fiber tractography [19]. In the left panel of Fig. 4, we display the whole procedure of hub identification by sorting-based method (using human heuristics) and our learning-based method (learning heuristics) on the bottom and top, respectively. The curves of reward function show that our learning-based method leads to more reasonable result than sorting-based method, which is also manifested by the spatial locations of hubs. It is worth noting that the optimization-based method uses the same reward function as the energy function. Since it is computational hard to achieve the global optimum, the compromised local solution renders the final energy function substantially lower than our learning-based method, which lead to sub-optimal hub identification result (designated by the black dotted arrow).

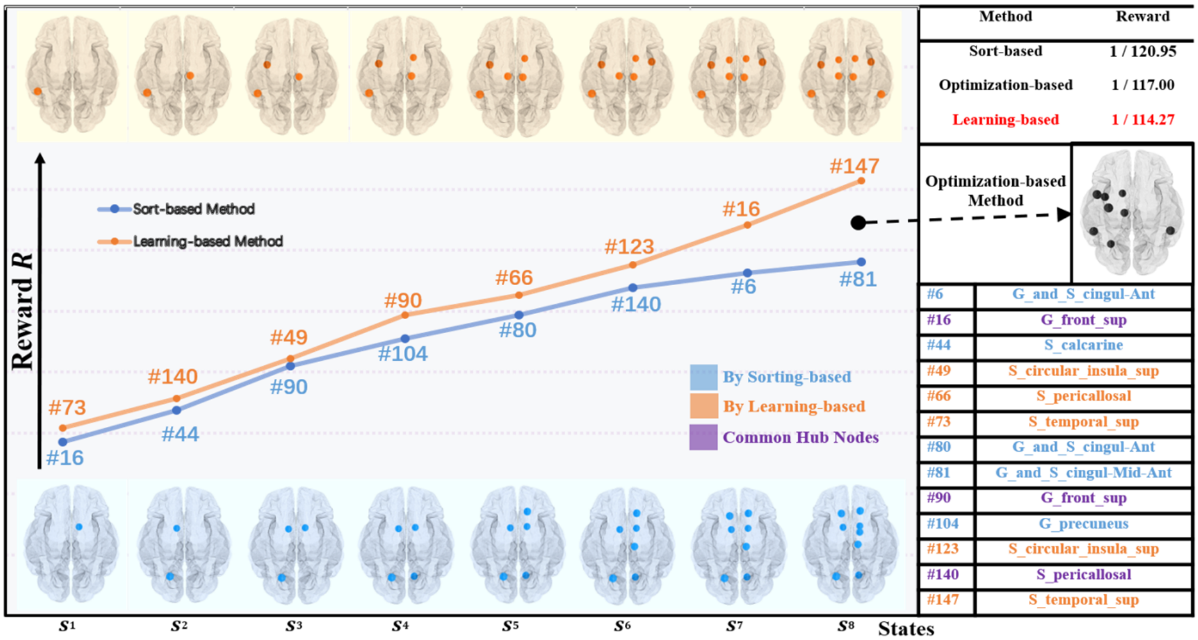

Fig 4. Visualization of hub identification process on individual structural network.

Furthermore, we specifically evaluate the replicability of hub identification result. Since it is not common to expect significant network difference in such short time period, we evaluate the replicability performance of hub identification results between baseline and 6-month follow-up networks, in a test-retest manner. The top ten most frequently selected hub nodes that are both presented in baseline and follow-up networks by above four methods are displayed in Fig. 5, where the node size reflects the frequency. The average agreement of common hub nodes in total 119 test-retest pairs are 35.50% by sorting-based method, 50.42% by module-based method, 56.09% by optimization-based method, and 66.39% by our learning-based method, respectively. We show the count histogram of each network node being both identified in each longitudinal pair in the left panel of Fig.5 (a)-(d). As indicated by the entropy value of the histogram, our reinforcement learning based hub identification method achieves most consistent (higher replicability) result in the replicability test.

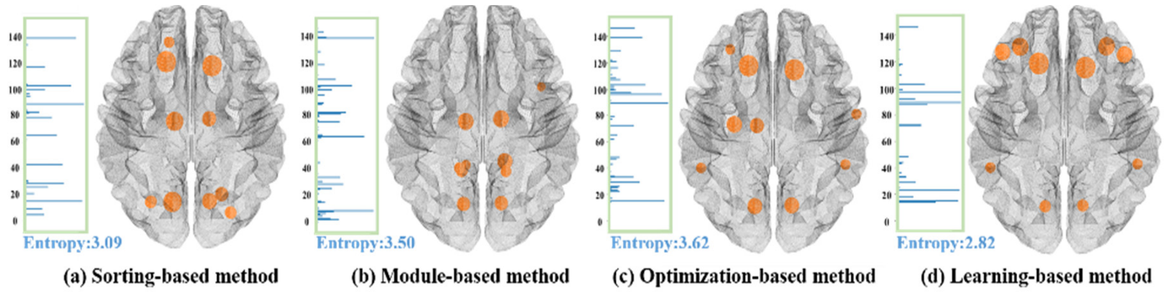

Fig 5. Replicability test by sorting-based method (a), module-based method (b), optimization-based method (c) and our learning-based method (d).

we did a significance test for cortical thickness(p-value<0.05) for structural networks, The hub nodes we detected can also distinguish AD from NC well, as show in Tabel 1.

Table 1. Specific brain regions used to distinguish between NC and AD

Learning-based	NC/AD	Optimization-based	NC/AD	Sorting-based	NC/AD	Module-based	NC/AD
15	×	16	×	16	×	3	×
16	×	30	✓	27	×	10	×
24	✓	34	✓	30	✓	30	✓
30	✓	73	✓	44	✓	35	✓
73	✓	90	✓	66	×	66	×
89	✓	97	✓	80	✓	77	×
90	✓	98	✓	90	✓	83	✓
98	✓	104	✓	101	✓	84	✓
104	✓	140	×	104	✓	104	✓
147	✓	147	✓	147	✓	140	✓

4. Conclusion

In this paper, we propose a first-ever learning-based connector hub identification method for network neuroscience. Specifically, we propose to learn the heuristics of hub identification under the reinforcement learning framework, where the graph theory is used to steer the reinforcement learning. The hub selection procedure is dynamic, that is, we update the attribute vector of each node to reflect the topology change due to the removal of hub nodes in the network. More accurate and replicable hub identification results have been achieved by our novel learning-based method, compared to existing state-of-the-art methods, which shows great applicability of our method in neuroscience and neuroimaging field.

Acknowledgments

This paper was supported by 2024 Basic Research Project of the Wenzhou Municipal Science and Technology Bureau (Project No. Y20240831).

References

[1] Sporns, O. (2011). Networks of the brain. MIT Press.

[2] Sporns, O., Honey, C. J., & Kötter, R. (2007). Identification and classification of hubs in brain networks. PLOS ONE, 2(10), e1049. https://doi.org/10.1371/journal.pone.0001049.

[3] van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. Trends in Cognitive Sciences, 17(12), 683–696. https://doi.org/10.1016/j.tics.2013.09.012.

[4] Crossley, N. A., Mechelli, A., Scott, J., Carletti, F., Fox, P. T., McGuire, P., & Bullmore, E. T. (2014). The hubs of the human connectome are generally implicated in the anatomy of brain disorders. Brain, 137(8), 2382–2395. https://doi.org/10.1093/brain/awu132.

[5] Jiao, Z., Xia, Z., Cai, M., Zou, L., Xiang, J., & Wang, S. (2018). Hub recognition for brain functional networks by using multiple-feature combination. Computers & Electrical Engineering, 69, 740–752. https://doi.org/10.1016/j.compeleceng.2018.02.014.

[6] Yang, D., Yan, C., Nie, F., Zhu, X., Turja, A., Zsembik, L., Styner, M., & Wu, G. (2019). Joint identification of network hub nodes by multivariate graph interference. In 22nd International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Shenzhen, China.

[7] Papo, D., Buldú, J. M., Boccaletti, S., & Bullmore, E. T. (2014). Complex network theory and the brain. Philosophical Transactions of the Royal Society B: Biological Sciences, 369, 20130520. https: //doi.org/10.1098/rstb.2013.0520.

[8] Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. NeuroImage, 52(3), 1059–1069. https://doi.org/10.1016/ j.neuroimage. 2009. 10.003.

[9] Nie, F., Wang, X., & Huang, H. (2014). Clustering and projected clustering with adaptive neighbors. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 977–986). https://doi.org/10.1145/2623330.2623377.

[10] Shi, W., Ling, Q., Yuan, K., Wu, G., & Yin, W. (2014). On the linear convergence of the ADMM in decentralized consensus optimization. IEEE Transactions on Signal Processing, 62(7), 1750–1761. https://doi.org/10.1109/TSP.2014.2304432.

[11] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT Press.

[12] Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. Machine Learning, 8(3), 279–292. https:// doi.org/10.1007/BF00992698.

[13] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., & Ostrovski, G. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533. https://doi.org/10.1038/nature14236.

[14] Tijms, B. M., Wink, A. M., de Haan, W., van der Flier, W. M., Stam, C. J., Scheltens, P., & Barkhof, F. (2013). Alzheimer's disease: Connecting findings from graph theoretical studies of brain networks. Neurobiology of Aging, 34(8), 2023–2036. https://doi.org/10.1016/j.neurobiolaging.2013.02.010.

[15] Newman, M. E. J. (2006). Modularity and community structure in networks. Proceedings of the National Academy of Sciences, 103(23), 8577–8582. https://doi.org/10.1073/pnas.0601602103.

[16] Porter, M. A., Onnela, J.-P., & Mucha, P. J. (2009). Communities in networks. Notices of the AMS, 56(9), 1082–1097.

[17] Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. Proceedings of the National Academy of Sciences, 99(12), 7821–7826. https://doi.org/ 10. 1073/ pnas. 122653799.

[18] Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage, 53(1), 1–15. https://doi.org/ 10. 1016/ j.neuroimage.2010.06.010.

[19] Geng, X., Gouttard, S., Sharma, A., Gu, H., Styner, M., Lin, W., Gerig, G., & Gilmore, J. (2012). Quantitative tract-based white matter development from birth to age 2 years. NeuroImage, 61(3), 542–557. https://doi.org/10.1016/j.neuroimage.2012.03.062.

[20] Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage, 15(1), 273–289. https://doi.org/ 10. 1006/ nimg.2001.0978.