Brain Effective Connectivity Learning with Deep Reinforcement Learning

Yilin Lu[1], Jinduo Liu[1*], Junzhong Ji[1], Han Lv[2], Mengdi Huai[3],
[1] Beijing University of Technology, shertheus@gmail.com, {jinduo, jjz01}@bjut.edu.cn
[2]Capital Medical University, chrislvhan@126.com
[3]Iowa State University, mdhuai@iastate.edu

2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 978-1-6654-6819-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/BIBM55620.2022.9995284

***Abstract*—In recent years, using functional magnetic resonance imaging (fMRI) data to infer brain effective connectivity (EC) between different brain regions is an important advanced study in neuroinformatics. However, current methods always perform not well due to the high noise of neuroimaging data. In this paper, we propose an effective connectivity learning method with deep reinforcement learning, called EC-DRL, aiming to more accurately identify the brain effective connectivity from fMRI data. The proposed method is based on the actor-critic algorithm framework, using the encoder-decoder model as the actor network. More specifically, the encoder adopts the Transformer model structure, and the decoder uses a bidirectional long-short-term memory network with an attention mechanism. A large number of experimental results on simulated fMRI data and real-world fMRI data show that EC-DRL can better infer effective connectivity compared to the state-of-the-art methods.**

***Index Terms*—Brain effective connectivity, deep reinforcement learning, encoder-decoder model, bidirectional long-short-term memory network, fMRI time series.**

I. Introduction

Recently, learning brain effective connectivity (EC) which is the interaction of brain regions at the neural level has become a frontier subject. Since it not only is crucial to evaluate brain function but also has a strong relation with neurodegenerative diseases, e.g. Parkinson's disease and Alzheimer's disease [3], [5]. There are substantial efforts on analyzing the causal relations of the brain regions or regions of interest (ROI) with neuroimaging data, e.g., functional magnetic resonance imaging (fMRI) data [6]–[8]. And the effective connectivity between the brain regions can be seen as directed edges in a causal graph (directed graph) where nodes denote brain regions [9]. Hence, learning brain effective connectivity can be turned into a problem that discovering a causal graph from fMRI time series data.

Causal graph inferring has made considerable progress over the past few decades. PC algorithm [12] is a classic constraint-based algorithm that first learned the skeleton of the causal structure through the independence test, and then determines the directions of edges using colliders. And greedy equivalence search (GES) [1] is another widely used score-based method. GES starts with a graph with no edge, instead of beginning with a complete undirected graph, like the PC algorithm. Then GES continuously adds edges to the graph and removes unnecessary edges to maximize a properly defined score that can evaluate the generated graph. The functional Causal Model (FCM) where the functions are assumed to denote the causal mechanisms in variables is an emerging method in casual discovery. FCM does not need strong assumptions and Linear Non-Gaussian Acyclic Model (LiNGAM) is the first FCM method proposed by Shimizu et al [10].

In the last decade, deep learning is one of the fastest-growing fields. And many researchers try to apply deep learning techniques to causal graph inferring. In 2018, Zheng et al [17]. proposed the continuous optimization for structure learning (NO TEARS) method which first turns the causal discovery problem from a combinatorial optimization problem into a continuous optimization problem so that the causal relations can be solved more efficiently. And empirically, the results are close to the optimal solution. And a year later, Yue et al [14] extend Zheng's method to the nonlinear case. Gradient-Based Neural DAG Learning (DRAN-DAG) method proposed by Lachapelle et al, which combines score-based methods with deep learning to infer causal relations. Zhu et al. [18] firstly use reinforcement learning method to solve the causal discovery problem. However, the performance of deep learning methods is not as good as traditional statistic methods on estimating brain effective connectivity.

In this paper, we propose a brain effective connectivity learning method with deep reinforcement learning, named EC-DRL. The new method employs BiLSTM with an attention mechanism [13] as a decoder in the actor to better recover the causal graph from fMRI data, which improves the authenticity of the generated data. We have tested our model on both simulation data and real data, and the experimental results show that the proposed method has certain advantages in performance compared with existing state-of-the-art methods.

II. Methodology

In this section, we put forward our proposed novel model, i.e., EC-DRL, which can estimate brain effective connectivity from fMRI time series data. Specifically, we first give an overview of the proposed EC-DRL, and then describe the details of the main components.

A. EC-DRL Architecture

In the proposed method, we propose to integrate the Actor-Critic algorithm into brain effective connectivity learning and

*Corresponding author: Jinduo Liu

Authorized licensed use limited to: University at Buffalo Libraries. Downloaded on May 23,2024 at 05:03:26 UTC from IEEE Xplore. Restrictions apply.

design a novel effective connectivity learning method (i.e., EC-DRL) based on the deep reinforcement learning method. There are three crucial parts in the proposed EC-DRL framework: *actor*, *critic* and *reward*. In the following, we will show the details of the three components.

Actor. The actor component is an encoder-decoder model, which takes noise variables and real fMRI time series data as input and generates directed graphs (graph adjacency matrices). Since the encoder-decoder model can naturally extract contextual information, it is very suitable for processing the fMRI time series data. Below, we will describe the adopted Encoder and Decoder.

- *Encoder.* The encoder based on the Transformer model. In detail, after embedding the inputs by a linear layer, the embedded input will be processed by multiple identical encoder blocks, and each encoder block is made up of a multi-head self-attention layer and a feedforward layer. We believe that multi-head self-attention can perform the task of extracting information from fMRI data very well. Importantly, compared with self-attention, it reduces the dependence on external information and can better capture the internal correlation in the data. Given the fMRI data X, the operation of the encoder blocks are given as follows:

$$\begin{aligned} Q_h &= W^Q {X'}^h + bias^Q, \\ Q_h &= W^K {X'}^h + bias^K, \\ V_h &= W^V {X'}^h + bias^V, \end{aligned} \tag{1}$$

where X' denotes the embedded input and ${X'}^h$ denotes the h-th input after dividing the embedded input X' into H groups. Q_h, Q_h, V_h denote query, key and value of the h-th input respectively. Then self-attention is calculated as follows:

$$SelfAttn_h = softmax(\frac{Q_h {K_h}^t}{\sqrt{D_{K_h}}}) V_h, \tag{2}$$

where D represents the number of elements in the last dimension of K. Then, we can get multi-head attention by concating all H groups of self-attention, which is given as follows:

$$MultiHead = Concat(SelfAttn_1, ..., SelfAttn_H). \tag{3}$$

And the result of multi-head attention will go through a feedforward layer which consist of 2 liner layer and a ReLU activation.

$$Block = ReLU(MultiHeadW_1 + bias_1)W_2 + bias_2. \tag{4}$$

- *Decoder.* The decoder we chose is Bi-directional Long Short-Term Memory (BiLSTM) [16]with an attention layer, since it not only has a strong ability of global information modeling but also can solve the problems of gradient disappearance and gradient explosion in the process of long sequence training. The BiLSTM model is made up of multiple LSTM cells. The process of the BiLSTM model consists of two layers: the forward layer and the backward layer. Specifically, the forward layer, from time step 1 to T, updates the long-term memory and stores the hidden state. And given the encoder output enc_t of the t-th time step, the hidden state can be represented as follows:

$$\overrightarrow{H_t} = f(enc_t {W_1}^{(f)} + \overrightarrow{H_{t-1}} {W_2}^{(f)} + bias^{(f)}), \tag{5}$$

where ${W_1}^{(f)}, {W_1}^{(f)}$ and $bias^{(f)}$ are parameters of the forward layer and function f denotes the LSTM model. The process of the backward layer is same to the process of the forward layer, except the time step is from T to 1:

$$\overleftarrow{H_t} = f(enc_t {W_1}^{(b)} + \overleftarrow{H_{t-1}} {W_2}^{(b)} + bias^{(b)}), \tag{6}$$

where ${W_1}^{(b)}, {W_1}^{(b)}$ and $bias^{(b)}$ are parameters of the backward layer. After concatenating the hidden state of both layers, we will get the hidden state of the BILSTM. The output can be represented by the following formula

$$O_t = [\overrightarrow{H_t}, \overleftarrow{H_t}]W + bias. \tag{7}$$

Critic. The critic we use is a 2-layer feed-forward neural network with a ReLU activation function. The input of the critic network is the encoder output. And the loss function is the mean-square error between its output and the true rewards of the actor network.

In the training, we find that the critic network can be replaced by an exponential moving average, which is a method that estimates the local mean of a variable. In this way, the update of the variables is related to the historical values over a period of time. This approach can improve the robustness of the model. Its mathematical expression is given as

$$V_t = \alpha V_{t-1} + (1 - \alpha) S_t, \tag{8}$$

where $\alpha \in (0, 1)$, and S_t is the value in the t-th time step. This method can reduce a certain running time without affecting the results. The reason is that the critic network is too simple, which means its role is limited. And the exponential moving average can be seen as a simplification of the critic network.

Reward. Reward is to measure the matching degree between the brain effective connectivity network and the fMRI data. Therefore the goal of the actor network or the entire network is to maximize the reward. Note that the Bayesian information criterion score (BIC) is widely used in score-based causal discovery methods. It is not only consistent but also locally consistent for its decomposability. Therefore, motivated by this, we choose BIC as our reward function which will be maximized by our actor. To better discover DAGs, we propose to add a acyclicity constraint to the reward. In this work, we use the acyclicity constraint proposed by Zheng et al [17]. Hence, our reward can be represented as follows

$$reward(\mathcal{G}) = -[BIC(\mathcal{G}) + \lambda A(M)], \tag{9}$$

where $\lambda \geq 0$ is a parameter that would be adjusted during training, M is binary adjacency matrix of $\mathcal{G}$, $BIC(\mathcal{G})$ is the

Authorized licensed use limited to: University at Buffalo Libraries. Downloaded on May 23,2024 at 05:03:26 UTC from IEEE Xplore. Restrictions apply.

BIC score of graph $\mathcal{G}$ and $A(M)$ is the acyclicity constraint. And if the fMRI data do not need to guarantee the acyclicity, λ needs to be set always equal to 0.

III. Experiments

In this section, to assess the performance of EC-DRL, we conduct experiments on a public simulated fMRI dataset generated from known ground-truth networks and then compare the results with other state-of-the-art methods. Finally, to demonstrate the application potential of EC-DRL, we apply it to the real-world fMRI data.

A. Data Description

Simulation Dataset. To validate the effectiveness of our proposed method, we use a widely used simulation datasets to compare our method with other state-of-the-art methods. The dataset we use is the Netsim dataset [11] which contains rich, realistic simulated fMRI data for a wide range of underlying networks, experimental protocols and problematic confounds in the data. The data is available at https://www.fmrib.ox.ac.uk/datasets/netsim/index.html. And TableI is the detailed information of the data we used in this paper from the public dataset.

TABLE I: Description of simulated dataset

Sim	Nodes	Session(min)	TR(s)	Noise(%)	HRF(s)	Other factors
1	5	10	3.00	1.0	0.5	
2	5	10	3.00	1.0	0.5	shared inputs
3	5	10	3.00	1.0	0.5	global mean confound
4	10	10	3.00	1.0	0.5	
5	5	2.5	3.00	0.1	0.5	
6	5	5	3.00	0.1	0.5	

Real world fMRI Dataset. We also use the real-world fMRI dataset to evaluate the performance of the proposed method. We chose rhyming task fMRI data from the Sanchez dataset. The rhyming task data is the fMRI data in which nine subjects judged if a pair of visual stimuli rhymed or not. The rhyming task data is available at https://github.com/cabal-cmu/feedback-discovery.

B. Comparison Methods for Evaluation

In order to intuitively illustrate the competitiveness of our EC-DRL, we compare with six other state-of-the-art or classic methods. These methods include: Peter and Clark (PC) [12], greedy equivalence search (GES) [1], Independent component analysis based linear non-gaussian acyclic model (ICA-LiNGAM) [2], continuous optimization for structure learning (No Tears) [17] and reinforcement learning (RL) [18] methods.

We employ the existing implementations from the literature. For ICA-LiNGAM, No Tears, and RL, the implementations we used are from gCastle toolbox proposed by [15]. The code is available at https://github.com/huawei-noah/trustworthyAI. And for PC and GES, the implementations we used are from the causal discovery toolbox [4]. All the hyperparameters are using the default settings. And EC-DRL is using the same hyperparameters as RL.

C. Results on Simulation fMRI Dataset

In the experiments, we choose 6 datasets from the Netsim data and we run each method on each data. We present the results in Table II. Note that when we test, we are using the data of all the subjects concatenated. We compared the learned results to ground-truth networks on the most common graph metrics: structural Hamming distance (SHD), precision, recall, and F1-measure (F1). An algorithm performs well when it gets higher values of precision, recall, and F1 and a lower value of SHD.

From Table II, we can find that GES and PC are not good as other methods. Because they get a relatively high SHD and low precision and recall. And No tears only perform better than GES and PC. ICA-LiNGAM has a good and stable result but is not as better as RL and EC-DRL. And it can be easy to find out that EC-DRL performs better than RL. More than that, it is worth noting that EC-DRL gets good results in all sims while some other methods have extremely poor performance on a few datasets which indicates EC-DRL has good robustness.

TABLE II: Netsim Simulation fMRI Dataset result

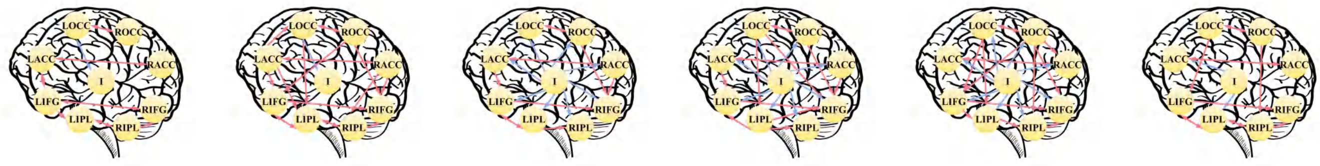
Sim	Metrics	PC	GES	ICA-LiNGAM	No Tears	RL	EC-DRL
1	SHD	4	6	1	2	1	1
	Precision	0.33	0.45	1	0.6	0.8	0.8
	Recall	0.4	1.0	0.8	0.6	0.8	0.8
	F1	0.36	0.63	0.89	0.6	0.89	0.89
2	SHD	11	10	2	1	1	1
	Precision	0.31	0.29	0.75	1	0.8	0.8
	Recall	1.0	0.8	0.6	0.8	0.8	0.8
	F1	0.48	0.42	0.67	0.89	0.8	0.8
3	SHD	13	10	0	1	0	0
	Precision	0.28	0.29	1.0	1.0	1.0	1.0
	Recall	1.0	0.8	1.0	0.8	1.0	1.0
	F1	0.43	0.42	1.0	0.89	1.0	1.0
4	SHD	5	6	1	7	6	1
	Precision	0.73	0.69	1.0	0.45	0.54	1.0
	Recall	0.92	0.92	0.80	0.42	0.5	0.92
	F1	0.81	0.79	0.89	0.43	0.52	0.96
5	SHD	3	3	0	1	1	0
	Precision	0.63	0.63	1.0	1.0	0.8	1.0
	Recall	1.0	1.0	1.0	0.8	0.8	1.0
	F1	0.77	0.77	1.0	0.89	0.8	1.0
6	SHD	3	3	1	1	2	0
	Precision	0.63	0.63	1.0	0.8	0.6	1.0
	Recall	1.0	1.0	1.0	0.8	0.6	1.0
	F1	0.77	0.77	1.0	0.8	0.6	1.0

In summary, the proposed model EC-DRL performs better than the five comparison methods on the simulation data. However, EC-DRL has high precision, but the outputs have a relatively small number of edges. The reward function can be adjusted in future work to produce more edges. Next, We discuss its performance on the real fMRI data in the following section.

D. Results on Real fMRI Dataset.

In this section, we will use real-world data to test the performance of EC-DRL. In detail, we run EC-DRL and other methods on the fMRI data of all subjects separately. If 40% of the generated graphs have a certain directed edge then

Authorized licensed use limited to: University at Buffalo Libraries. Downloaded on May 23,2024 at 05:03:26 UTC from IEEE Xplore. Restrictions apply.



(a) PC (b) GES (c) LiNGAM (d) No Tears (e) RL (f) EC-DRL

Fig. 1: The Brain effective connectivity network structure learned by 6 methods on rhyming task fMRI data.

we assume there is effective connectivity between the brain regions.

The rhyming task fMRI data included an Input variable built by convolving the rhyming task boxcar model with a canonical hemodynamic response function. And that means the edges from the Input variable must feedforward into the regions of interest, and no edge should point backward into the Input variable. Therefore, we can use the Input variable as a simple but gold standard for the accuracy of the results of methods. The results are shown in Fig1

From Fig1, we can easily find that there is an edge in the results of PC and GES pointing from ROIs to the Input variable ($LOCC \rightarrow I$). And that means their result may not reliable enough. ED-DRL discovers 2 edges that point out from the Input variable. LiNGAM, No Tears and RL discovers 8, and PC discovers 1. And from Fig1 we also can find that 4 of the methods get the results of $I \rightarrow LIFG$ and $I \rightarrow LIPL$, which indicates it is very likely that these two edges exist. And And some edges were discovered by all the methods i.e. $ROCC \rightarrow LOCC$, $RACC \rightarrow LACC$. This phenomenon may represent that the right hemisphere of brain regions always activated earlier than the left hemisphere of brain regions under this experiment.

In summary, the results on real-world data show that the new method EC-DRL can provide a reliable perspective for the analysis of effective connectivity in task data, and EC-DRL may bring some inspiration to researchers.

IV. Conclusion

In this paper, we proposed a new model to estimate brain effective connectivity networks from fMRI time series data with the deep reinforcement learning technique, called EC-DRL. EC-DRL uses Actor-Critic as the basic frame and employs an encoder-decoder model as the actor to better extract information from data. Experimental results on both simulated and real-world data demonstrate the efficacy of our proposed framework.

V. Acknowledgements

This work was partly supported by NSFC Research Program (62106009, 62276010, 62171297), in part by R&D Program of Beijing Municipal Education Commission (KM202210005030, KZ202210005009), in part by China Postdoctoral Science Foundation (2021M690269), in part by Beijing Postdoctoral Research Foundation, in part by Chaoyang District Postdoctoral Research Foundation.

References

[1] D. M. Chickering, "Optimal structure identification with greedy search," *Journal of machine learning research*, vol. 3, pp. 507–554, 2002.

[2] P. Hoyer, D. Janzing, J. M. Mooij, J. Peters, and B. Schölkopf, "Nonlinear causal discovery with additive noise models," *Advances in neural information processing systems*, vol. 21, 2008.

[3] J. Ji, J. Liu, L. Han, and F. Wang, "Estimating effective connectivity by recurrent generative adversarial networks," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3326–3336, 2021.

[4] D. Kalainathan, O. Goudet, and R. Dutta, "Causal discovery toolbox: Uncovering causal relationships in python." *J. Mach. Learn. Res.*, vol. 21, no. 37, pp. 1–5, 2020.

[5] J. Liu, J. Ji, X. Jia, and A. Zhang, "Learning brain effective connectivity network structure using ant colony optimization combining with voxel activation information," *IEEE journal of biomedical and health informatics*, vol. 24, no. 7, pp. 2028–2040, 2019.

[6] J. Liu, J. Ji, G. Xun, L. Yao, M. Huai, and A. Zhang, "Ec-gan: inferring brain effective connectivity via generative adversarial networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, pp. 4852–4859.

[7] J. Liu, J. Ji, G. Xun, and A. Zhang, "Inferring effective connectivity networks from fmri time series with a temporal entropy-score," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 10, pp. 5993–6006, 2022.

[8] J. Liu, J. Ji, L. Yao, and A. Zhang, "Estimating brain effective connectivity in fmri data by non-stationary dynamic bayesian networks," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2016, pp. 834–839.

[9] R. Sanchez-Romero, J. D. Ramsey, K. Zhang, M. R. Glymour, B. Huang, and C. Glymour, "Estimating feedforward and feedback effective connections from fmri time series: Assessments of statistical methods," *Network Neuroscience*, vol. 3, no. 2, pp. 274–306, 2019.

[10] S. Shimizu, P. O. Hoyer, A. Hyvärinen, A. Kerminen, and M. Jordan, "A linear non-gaussian acyclic model for causal discovery." *Journal of Machine Learning Research*, vol. 7, no. 10, 2006.

[11] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich, "Network modelling methods for fmri," *Neuroimage*, vol. 54, no. 2, pp. 875–891, 2011.

[12] P. Spirtes, C. N. Glymour, R. Scheines, and D. Heckerman, *Causation, prediction, and search*. MIT press, 2000.

[13] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[14] Y. Yu, J. Chen, T. Gao, and M. Yu, "Dag-gnn: Dag structure learning with graph neural networks," in *International Conference on Machine Learning*. PMLR, 2019, pp. 7154–7163.

[15] K. Zhang, S. Zhu, M. Kalander, I. Ng, J. Ye, Z. Chen, and L. Pan, "gcastle: A python toolbox for causal discovery," *arXiv preprint arXiv:2111.15155*, 2021.

[16] S. Zhang, D. Zheng, X. Hu, and M. Yang, "Bidirectional long short-term memory networks for relation classification," in *Proceedings of the 29th Pacific Asia conference on language, information and computation*, 2015, pp. 73–78.

[17] X. Zheng, B. Aragam, P. K. Ravikumar, and E. P. Xing, "Dags with no tears: Continuous optimization for structure learning," *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[18] S. Zhu, I. Ng, and Z. Chen, "Causal discovery with reinforcement learning," *arXiv preprint arXiv:1906.04477*, 2019.

Authorized licensed use limited to: University at Buffalo Libraries. Downloaded on May 23,2024 at 05:03:26 UTC from IEEE Xplore. Restrictions apply.