

A Supervised Contrastive Learning-based Analysis of rs-fMRI Data Captures Gender Differences in Nonlinear Functional Network Coupling*

Reihaneh Hassanzadeh, Vince Calhoun

Abstract— Many studies in neuroscience have focused on interpreting brain activity using functional connectivity (FC). The most widely used approach for measuring FC is based on linear correlation (e.g., the Pearson correlation), where the temporal co-fluctuations between functional brain regions are computed. However, such approaches ignore nonlinear dependencies among regions that might carry distinctive information across groups of subjects. In this study, we offer a deep learning-based approach that also captures nonlinear temporal relationships between brain networks. Our approach consists of two main parts: an encoder that learns domain-specific embeddings of time courses estimated from independent component analysis (ICA) and a similarity metric that measures the similarities between the embeddings. We call such similarities as nonlinear functional relationships between networks. Our findings on a large dataset (including above 11k normal control subjects) suggest that male subjects exhibit stronger nonlinear network-network relationships than female subjects in most cases. Furthermore, we observe that, unlike FC, our approach could capture some intra-network relationships, especially between cognitive control and visual networks, which are significantly different between males and females, suggesting that our approach can provide a complementary interpretation of the functional brain activity to FC.

I. INTRODUCTION

Functional connectivity is a method to measure functional relationships between brain regions that might be spatially distant or anatomically disconnected. It quantifies the statistical dependencies between blood-oxygenation-level-dependent (BOLD) signals generated during a functional magnetic resonance imaging (fMRI) scan. Functional connectivity can be estimated either from time series computed from a seed-based method (called FC) or time courses estimated from an ICA method (called functional network connectivity, FNC) [1]. Both methods have been widely applied to fMRI (especially resting-state fMRI, rs-fMRI) data. Networks estimated from ICA are attractive in terms of reproducibility across sessions and centers as ICA is not biased towards a prior-knowledge selection of seeds [2].

Many rs-fMRI studies have shown functional connectivity varies between female and male subjects. For example, [3] observed that intra-domain networks (especially within default mode and sensory-motor networks) have stronger FNC in males than females, while the opposite was found for inter-domain networks. [4] and [5] reported stronger FC in males

compared with females in default mode networks. Furthermore, [6] reported increased FC in parietal and occipital networks but decreased FC in frontal and temporal networks in men. Also, the authors found that male subjects have stronger FNC between cognitive and sensory networks but weaker FNC between right working memory and attention networks.

Other rs-fMRI studies have used gender differences in functional connectivity to predict gender in new data. For example, [7] predicted gender using FC and highlighted intra-network connections in default mode, frontoparietal and sensorimotor networks as the highest contributing features to the prediction. Also, [8] showed that the connection between the cingulate cortex, medial and lateral frontal cortex, temporoparietal regions, precuneus, and insula to other regions contributed the most to gender classification.

In general, previous gender-related rs-fMRI studies have mostly used FCs that are derived from correlation-based approaches, such as the Pearson correlation, in which only the linear dependencies between time series are considered. While such approaches are simple to implement and fast to execute, they fall short in extracting complex nonlinear temporal patterns that tie together pairs of networks. Hence, in this paper, we propose a deep learning-based model to extract “nonlinear” functional couplings/relationships between ICA-estimated time courses. In light of that, we first adapt a supervised contrastive learning-based framework, which has shown promise in the image processing domain, to the time series data for learning domain-informative embeddings. Then, we measure the similarity of the embeddings, which measures the nonlinear functional network couplings (henceforth, nFNC). Finally, we use nFNC to investigate the role of gender in differentiating the functional relationships between pairs of networks.

II. METHOD

We develop a contrastive learning-based model that maps the input samples, i.e., ICA-estimated time courses, into 128-dimensional embeddings such that samples of the same classes fall near each other in a learnable embedding space, while at the same time, those belonging to other classes are pushed away from each other. The model is designed in a supervised setting, where domain labels, such as auditory, visual, etc., are considered as classes. Our model, in a nutshell, is comprised

*Research supported by NSF #2112455 and NIH # R01MH118695.

Reihaneh Hassanzadeh is with the Computer Science Department, Georgia State University, Atlanta, GA 30302, USA.

Vince Calhoun is with the School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.

Reihaneh Hassanzadeh and Vince Calhoun are with the Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Atlanta, GA 30303, USA.

of three main components, 1. an encoder network that generates normalized embeddings of time course inputs, 2. a projection head where a supervised contrastive loss is applied, and 3. a cosine similarity metric layer which is applied on the normalized outputs from the embedding layer to calculate a nonlinear similarity between the networks.

A. The encoder network

We utilize a bidirectional long short-term memory (LSTM) [9] network to learn from the past and future of a time course. We implement the LSTM in two layers, each of which has a hidden size of 128. That is, the first recurrent layer is stacked over the second layer, which processes the first layer's outputs. We concatenate the outputs of the two layers, which results in 256 output features. Furthermore, we add a dropout on the first layer with a dropout probability equal to 0.2. Finally, the 256 outputs are fed into a 128-dimensional fully connected layer, i.e., the embedding layer, followed by a normalization layer on top of it. In summary, the encoder maps a 450-length time course into normalized 128-dimensional high-level embeddings, which contain the domain information. A contrastive loss function guides the model to encode domain-informative embeddings in a supervised way which is explained in the subsection.

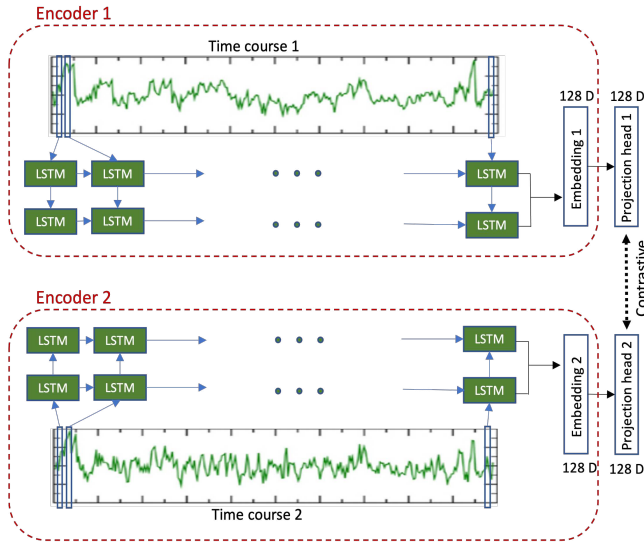


Figure 1. **Proposed model high-level architecture.** The model consists of an LSTM-based encoder that generates normalized domain-specific embeddings of time courses and a projection head where the supervised contrastive loss is applied.

B. The projection layer

We stack a fully connected layer, called the projection layer, on the normalized outputs from the embedding layer. We then apply a supervised contrastive learning loss [10] on the normalized outputs of the projection layer. This loss induces a model for which generated embeddings of same-domain samples remain close to each other and far from others in the embedding space. Hence, the encoder network generates domain-specific embeddings of the input networks.

The contrastive learning loss operates on pair of samples, including multiple positive pairs (i.e., pairs with samples of the

same class) and multiple negative pairs (i.e., pairs with samples of different classes). In our work, we generate positive and negative pairs as follows. For a given input network, to generate a positive pair, we link it with a network of the same domain of a different subject. Likewise, to generate a negative pair, we link the sample with a network of a different domain and a different subject.

A significant benefit of the supervised contrastive loss is that it computes the loss by comparing and contrasting multiple positive pairs and multiple negative pairs¹ per anchor. Whereas other well-known contrastive learning-based losses are based on single-positive or/and single-negative pairs in each batch (e.g., triplet loss [11] and hard negative-mining [12]). This encourages the supervised contrastive loss to estimate a more accurate margin between domains and within networks of the same domain.

C. The similarity metric layer

Once the model is trained, we freeze the encoder network to generate domain-specific network embeddings. Then, we compute the cosine similarity between any two embeddings to quantify how nonlinearly similar their corresponding time courses are. A large value of similarity suggests a strong relationship between the underlying brain networks and conversely.

D. Transferring the knowledge and finetuning

Once the model is pretrained as described, we need to finetune it for the domain classification task. To do that, we add a fully-connected linear classification layer on top of the pre-trained encoder and finetune the overall model while having the weights of the encoder component frozen.

III. EXPERIMENTAL DATA

A. Participants

We evaluate our proposed model on a resting-state fMRI (rs-fMRI) dataset from the UK Biobank study² [13]. The dataset includes 11754 participants with the same age distribution between males and females. The demographic data from the participants are reported in Table 1.

TABLE I. PARTICIPANT DEMOGRAPHICS

	Population Number	Age			
		Mean	SD	Min	Max
All	11754 (100%)	62.56	7.38	45	80
Male	5772 (49%)	63.08	7.54	45	80
Female	5982 (51%)	62.07	7.19	46	80

B. Data acquisition and preprocessing

Rs-fMRI data were acquired during a six-minute scan by a 3-Tesla Siemens Skyra scanner. The preprocessing steps performed by UK Biobank are motion correlation with MCFLIRT, normalization based on grand-mean intensity, Gaussian-weighted high-pass temporal filtering, and EPI and GDC unwarping (for more details, see here³).

¹The size of negative samples is usually larger than the positive samples.

²This is an academic study for which no ethical approval was required.

³https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf

C. Group-independent component analysis

We apply a fully automated spatially constrained ICA [14] using the NeuroMark framework [15] to our preprocessed data to estimate one hundred reproducible functional networks. Subsequently, we choose fifty-three of these networks based on our empirical filtering policy. ICA estimates both spatial maps and their corresponding time courses of functional networks. In this study, we focus only on the time courses, each having a length of 450 time points. Figure 2 visualizes the location of these networks in the brain and groups them into functional domains.

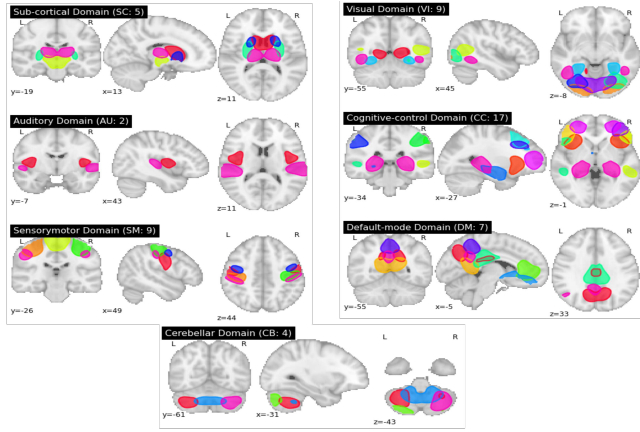


Figure 2. **Brain networks estimated from ICA.** Functional networks are grouped into seven domains: subcortical (SC, five networks), visual (VI, nine networks), auditory (AU, two networks), cognitive control (CC, seventeen networks), sensory-motor (SM, nine networks), default mode (DM, seven networks), and cerebellar (CB, four networks).

IV. EXPERIMENTAL RESULTS

We trained our proposed model for 200 epochs with a learning rate of 0.5 and a batch size of 128. We employed a five-fold cross-validation technique where three folds are considered as the training set, one fold is considered as the validation set, and the remaining fold is used as a test set. We deduplicated these sets to avoid any data leakage. Our model is implemented in PyTorch, and we used NVIDIA Tesla V100 GPUs with a RAM size of 32GB.

We performed a hyperparameter search using a grid search strategy to pick the model that generates the most informative embeddings. The grid search was run on two sets of parameters for the hidden size and the batch size⁴ on the task in the method section part D. As such, we used the cross-entropy loss on the predicted functional domain as a fitness measure for our grid-search selection strategy.

To evaluate the selected best model, for each subject on the test dataset, we first compute a similarity matrix with each cell containing the cosine similarity between the embedding of the corresponding networks, which we call the nonlinear functional network coupling matrix (nFNC, for short). Next, we take the average of the nFNC matrices across female and male samples separately and calculate the difference between the resulting two averaged matrices. Figure 3.A visualizes the difference between the two matrices (the upper triangular matrix) along with the corresponding cells with significant

false discovery rate (FDR) corrected p-values using a two-sample t-test (the lower triangular matrix)⁵. Interestingly, most (76%) of network-network couplings among male subjects exhibit a significantly stronger nonlinear nFNC than those of female subjects. For example, the couplings between the visual and cognitive control networks, between the subcortical and default mode networks, and within the cerebellar networks are stronger in males. This is while, for females, the nFNC matrix shows stronger functional couplings between the subcortical and sensory motor networks, within subcortical networks, and within visual networks.

We furthermore compute the FNC matrix generated from the Pearson correlation between ICA-estimated time courses. Figure 3.B shows the difference between the FNC matrix of males and females, as well as values with significant FDR corrected p-values from a two-sample t-test in the lower triangular matrix. We can observe that 55% of networks in males have significantly stronger FNC than females, this is especially the case for intra-networks in subcortical, sensory-motor, and cerebellar. On the contrary, the SC-AU, SC-SM, AU-CB, and SM-CB connections in females are higher than in males.

Comparing nFNC (Figure 3.A) to FNC (Figure 3.B) suggests that while in both cases, the corresponding heatmaps show clustered arrangement of high-intensity points, in terms of the domains, they present different gender-based patterns. For example, for males, nFNC shows stronger couplings between subcortical and cognitive control in most cases, while this is the opposite in FNC. Also, intra-subcortical networks have stronger FNC in males, but they have stronger nFNC in females. On the other hand, both FNC and nFNC show similar relationships between the two genders in AU-SM, SM-CB, VI-CB, and VI-DM. Another interesting observation is that our contrastive learning-based approach detects significant gender differences in functional relationships, whereas for FNC this is not the case, especially for VI-CC (see Figure 3.C).

V. DISCUSSION

Overall, our proposed model shows that nonlinear temporal relationships between brain networks that are computed based on a contrastive framework can provide complementary information of brain neural activity to FNC. Interestingly, nFNC recognizes more links between the visual networks and cognitive control network than FNC that are significantly different across the two groups of gender. We observe a greater number of significantly stronger couplings in nFNC than in FNC for males, which suggests that nonlinear relationships are important for evaluating interactions with gender. Furthermore, we found that males have stronger intra-networks FNC in sensory-motor and intra-networks nFNC in default mode, a fact that has been corroborated by other studies as in [3, 4]. It is also interesting to note that both FNC and nFNC find consistent strong patterns of couplings between subcortical and sensory-motor networks and within visual networks (females) and within cerebellar networks (males).

Last but not least, all the aforementioned observations have been reported by a model that is trained without any prior

⁴ including a hidden size of 64, 128, 256, and 512, and a batch size of 128, 256, 512, and 1024.

⁵ Note that the figure shows the difference matrix that is averaged across the five folds and the p-values that are significant in all the five folds

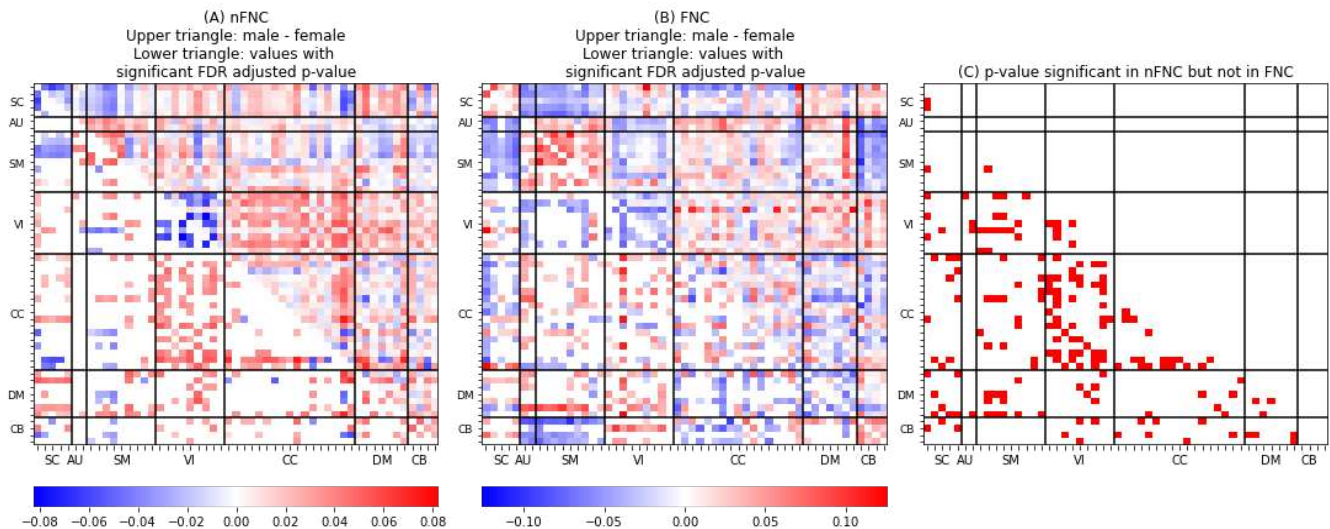


Figure 3. **Brain networks estimated from ICA.** Figure A shows the nonlinear functional relationships between networks computed from a contrastive learning-based model. Figure B shows the functional network connectivity using the Pearson correlation. Values in the lower triangular matrix are their corresponding values in the upper triangular matrix that have a significant FDR-corrected p-value from a two-sample t-test. Figure C shows the cells that have significant p-values in nFNC but not in FNC.

knowledge of gender or any other subject-specific information (e.g., age), and hence, the results are less subject to bias.

VI. CONCLUSION AND FUTURE WORK

In this paper, we propose a deep learning-based model to extract the nonlinear relationship between brain networks. We show that our model can learn domain-informative embeddings of the brain networks by comparing and contrasting the inter- and intra-domain networks. We use a similarity metric on the embeddings to measure the extent of similarity, which result in a map of nonlinear functional network couplings (nFNC). We observe that the differences between nFNC of males and females vary differently across the functional domains. Furthermore, our results reveal that most network-network couplings (76%) have a significantly stronger nFNC than in male subjects. Finally, a comparison between nFNC and FNC suggests that nonlinear relationships between networks can provide some complementary information to the linear correlations between networks.

As a future work, we would like to apply the proposed approach to other datasets and investigate the differences between groups of patients vs. controls as well as young subjects vs. old ones to evaluate the importance of capturing nonlinear relationships. We are also interested in using the nFNC matrices for the prediction of age, gender, and diagnosis.

REFERENCES

- [1] H. Lv *et al.*, "Resting-State Functional MRI: Everything That Nonexperts Have Always Wanted to Know," *AJNR. American journal of neuroradiology*, vol. 39, no. 8, pp. 1390–1399, Aug. 2018, doi: 10.3174/AJNR.A5527.
- [2] L. Wu, A. Caprihan, J. Bustillo, A. Mayer, and V. Calhoun, "An approach to directly link ICA and seed-based functional connectivity: application to schizophrenia HHS Public Access," *Neuroimage*, vol. 179, pp. 448–470, 2018, doi: 10.1016/j.neuroimage.2018.06.024.
- [3] E. A. Allen *et al.*, "A baseline for the multivariate comparison of resting-state networks," *Frontiers in Systems Neuroscience*, vol. 0, no. FEBRUARY 2011, p. 2, Feb. 2011, doi: 10.3389/FNSYS.2011.00002/BIBTEX.
- [4] R. L. Bluhm *et al.*, "Default mode network connectivity: effects of age, sex, and analytic approach," *Neuroreport*, vol. 19, no. 8, pp. 887–891, May 2008, doi: 10.1097/WNR.0B013E328300EBBF.
- [5] B. B. Biswal *et al.*, "Toward discovery science of human brain function," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4734–4739, Mar. 2010, doi: 10.1073/PNAS.0911855107.
- [6] M. Filippi, P. Valsasina, P. Misci, A. Falini, G. Comi, and M. A. Rocca, "The organization of intrinsic brain activity differs between genders: a resting-state fMRI study in a large cohort of young healthy subjects," *Human brain mapping*, vol. 34, no. 6, pp. 1330–1343, Jun. 2013, doi: 10.1002/HBM.21514.
- [7] C. Zhang *et al.*, "Functional connectivity predicts gender: Evidence for gender differences in resting brain connectivity," doi: 10.1002/hbm.23950.
- [8] S. Weis, K. R. Patil, F. Hoffstaedter, A. Nostro, B. T. T. Yeo, and S. B. Eickhoff, "Sex Classification by Resting State Brain Connectivity," *Cerebral Cortex*, vol. 30, no. 2, pp. 824–835, Mar. 2020, doi: 10.1093/CERCOR/BHZ129.
- [9] A. Graves, "Long Short-Term Memory," pp. 37–45, 2012, doi: 10.1007/978-3-642-24797-2_4.
- [10] P. Khosla *et al.*, "Supervised Contrastive Learning," *Advances in Neural Information Processing Systems*, vol. 2020-December, Apr. 2020, Accessed: Jan. 20, 2022. [Online]. Available: <https://arxiv.org/abs/2004.11362v5>
- [11] F. Schroff and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering".
- [12] Y. Tian, D. Krishnan, G. Research, and P. Isola, "Contrastive Multiview Coding", Accessed: Jan. 22, 2022. [Online]. Available: <http://github.com/HobbitLong/CMC/>.
- [13] K. L. Miller *et al.*, "Multimodal population brain imaging in the UK Biobank prospective epidemiological study," *Nature Neuroscience* 2016 19:11, vol. 19, no. 11, pp. 1523–1536, Sep. 2016, doi: 10.1038/nn.4393.
- [14] Q. H. Lin, J. Liu, Y. R. Zheng, H. Liang, and V. D. Calhoun, "Semiblind spatial ICA of fMRI using spatial constraints," *Human brain mapping*, vol. 31, no. 7, pp. 1076–1088, 2010, doi: 10.1002/HBM.20919.
- [15] Y. Du *et al.*, "NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders," *NeuroImage: Clinical*, vol. 28, p. 102375, Jan. 2020, doi: 10.1016/J.NICL.2020.102375.