Adversarial Brain Multiplex Prediction From a Single Network for High-Order Connectional Gender-Specific Brain Mapping

Ahmed Nebli ID1,2 and Islem Rekik ID1 *

1 BASIRA Lab, Faculty of Computer and Informatics, Istanbul Technical University, Istanbul, Turkey
2 National School for Computer Science (ENSI), Mannouba, Tunisia

Abstract. Brain connectivity networks, derived from magnetic resonance imaging (MRI), non-invasively quantify the relationship in function, structure, and morphology between two brain regions of interest (ROIs) and give insights into gender-related connectional differences. However, to the best of our knowledge, studies on gender differences in brain connectivity were limited to investigating pairwise (i.e., low-order) relationship ROIs, overlooking the complex high-order interconnectedness of the brain as a network. A few recent works on neurological disorder diagnosis addressed this limitation by introducing the brain multiplex, which in its shallow form, is composed of a source network intra-layer, a target intra-layer, and a convolutional inter-layer capturing the high-level relationship between both intra-layers. However, brain multiplexes are built from at least two different brain networks, inhibiting its application to connectomic datasets with single brain networks such as functional networks. To fill this gap, we propose the first work on predicting brain multiplexes from a source network to investigate gender differences. Recently, generative adversarial networks (GANs) submerged the field of medical data synthesis. However, although conventional GANs work well on images, they cannot handle brain networks due to their non-Euclidean topological structure. Differently, in this paper, we tap into the nascent field of geometric-GANs (G-GAN) to design a deep multiplex prediction architecture comprising (i) a geometric source to target network translator mimicking a U-Net architecture with skip connections and (ii) a conditional discriminator which classifies predicted target intra-layers by conditioning on the multiplex source intra-layers. Such architecture simultaneously learns the latent source network representation and the deep non-linear mapping from the source to target multiplex intra-layers. Our experiments on a large dataset demonstrated that predicted multiplexes significantly boost gender classification accuracy compared with source networks and identifies both low and high-order gender-specific multiplex connections.

* corresponding author: irekik@itu.edu.tr, http://basira-lab.com, GitHub: http://github.com/basiralab. This work is accepted for publication in the PRedictive Intelligence in MEdicine (PRIME) workshop Springer proceedings in conjunction with MICCAI 2020.
1 Introduction

The brain is a complex interconnected network encoding the connectional fingerprint of each individual and representing a notable biomarker of its gender [1]. In fact, several studies suggest that gender is a result of a natural distinction caused by human genetics and translated into dimorphic cortical connectivities. For instance, [2] showed that males excel in memory and motor tasks while females have better linguistic and emotional processing yet more prone to anxiety and depression [3]. Therefore, an accurate gender classification of brain networks might help better spot gender-biased brain disorders and contribute to more reliable and personalized treatments. Despite the breadth of research on gender differences in brain functional and structural connectivity [1] as well as morphological connectivity [4], existing works are limited to investigating pairwise (i.e., low-order) relationship between ROIs, overlooking not only the complex high-order interconnectedness of the brain as a network but also the topological configuration of brain ROIs connectivities.

To address the limitation of conventional low-order brain network representation, recent pioneering works [5,6] introduced the concept of a brain multiplex, which in its shallow form, is composed of a source network intra-layer, a target intra-layer, and a convolutional inter-layer capturing the high-order relationship between both intra-layers. Basically, a brain multiplex can be viewed as a tensor stacking two brain networks (also called intra-layers) and one inter-layer that encodes the similarity between these two intra-layers. While the intra-layers capture the low-order pairwise relationship between ROIs, the inter-layer models the high-order relationship between both intra-layer networks. In particular, the inter-layer is estimated by convolving one intra-layer with the other. The multiplex representation of brain connectivities boosted the classification accuracy in patients with early mild cognitive impairment [5] and patients with late mild cognitive impairment compared to the low-order brain network as well as their simple concatenation [6]. Although compelling, building a brain multiplex depends on the availability of different intra-layer brain networks [5,6], limiting its applicability to only multi-view (or multi-modal) brain connectomic datasets. One way to circumvent the issue of connectomic data with missing observations is to discard those samples, however this might not be convenient for devising learning-based classification frameworks. Ideally, one would learn how to predict a multiplex from a single intra-layer network (i.e., a source network) to (i) boost classification results, and (ii) discover low-order and high-order connectional biomarkers for the target classification task.

In recent years, the in-vogue generative adversarial networks (GANs) submerged the field of medical data synthesis [7] with unprecedented results in learning how to generate a target brain imaging modality from a different modal-
ity (e.g., computed tomography (CT) from MRI or T1-w from T2-w imaging) [7].

Despite the high efficiency of GANs in solving generative-related problems, all these methods focused on generating only images. However, brain connectivities as well as manifolds and graphs are essentially non-Euclidean spaces. To fill this gap in graph deep learning, geometric deep learning has been recently proposed to handle such data [8,9] and propose new graph convolution and deconvolution operations. Given the absence of works on predicting brain multiplex as well as investigating the discriminative potential of these high-order brain connectonental representations for gender classification, we design in this paper a deep brain multiplex prediction (DBMP) architecture comprising (i) a geometric source to target network translator mimicking a U-Net architecture with skip connections and (ii) a conditional discriminator which classifies predicted target intra-layers by conditioning on the multiplex source intra-layers. Such architecture simultaneously learns the latent source network representation and the deep non-linear mapping from the source to target multiplex intra-layers. Taking into account its insensitivity to overtraining and its notable performance in paired classification, we train a support vector machine (SVM) classifier in combination with a feature selection step on the predicted brain multiplexes for gender classification and connectional gender marker discovery. The main innovative contributions of our work on both methodological and clinical levels are listed below.

Methodological advance. We propose the first geometric deep learning on brain multiplex prediction based on G-GAN. This allows to learn the multiplex instead of predefining it for each individual in the population and circumvents the scarcity of multi-view connectomic brain data compared to the breadth of computer vision datasets. Furthermore, to the best of our knowledge, this is the first work adapting geometric deep learning for brain connectome synthesis and classification in general.

Clinical/scientific advances. This is the first attempt to explore gender differences using brain multiplexes and reveal new connectional gender markers on both low and high-order levels. This contribution will enable early and personalized treatments for gender-biased brain disorders.

2 Proposed Deep Brain Multiplex Prediction using G-GAN for Gender Fingerprinting

In this section, we introduce the different steps of deep multiplex prediction framework from a source network for connectional gender fingerprinting. The proposed DBMP framework draws inspiration from [9], which has pioneered the concept of graph or network translator. In typical GAN models, network generation is designed only for learning the distribution of network representations whereas network translation learns not only the latent network representation but also the generic translation mapping from source network to the target network simultaneously [9]. Fig. 1 illustrates the proposed pipeline for DBMP with application to gender fingerprinting comprising the following steps: 1) network translator training, 2) conditional network discriminator learning, 3) linear SVM
Fig. 1: Proposed deep brain multiplex prediction (DBMP) framework from a source network for gender classification and fingerprinting.

(A) Given a set of morphological brain multiplexes, we train a brain network translator to learn a multiplex correspondence from a source network to a target network. This is achieved by training a brain network translator to optimize a multiplex correspondence loss function. The multiplex correspondence loss function encourages the network to produce multiplexes that are similar to the source multiplexes.

(B) For each of the training subjects, the conditional network discriminator inputs two sets of pairs of multiplex layers: \( \{ M^S_k, M^T_k \} \) and \( \{ \hat{M}^S_k, \hat{M}^T_k \} \) \( \forall k \leq N \). The discriminator then outputs a value between 0 and 1 representing the realness of the translator's predicted multiplex \( \hat{M}^T_k \).

(C) We produce the final multiplex \( M^k = \{ M^S_k, \hat{M}^S_k \rightarrow T, \hat{M}^T_k \} \) by stacking the source intra-layer, the predicted intra-layer, and inter-layer which is the result of a convolution of both intra-layers.

(D) We use an infinite feature selection strategy that selects the most gender discriminative brain multiplex connections to train a support vector machine classifier to assign each subject's brain multiplex gender. We display the most discriminative brain multiplex connections which belong to both intra-layers and inter-layers.
classifier training with feature selection, and 4) connectional gender discriminative brain multiplex connections discovery.

**Conventional shallow brain multiplex construction.** Given $N$ training subjects, let $\{M^k_S\}_{k=1}^N$ denote the set of $N$ source intra-layers and $\{M^k_T\}_{k=1}^N$ the set of $N$ target intra-layers. A conventional shallow brain multiplex $M^k$ for subject $k$ is defined as a tensor stacking three layers: a source intra-layer $M^k_S$, a target intra-layer $M^k_T$, and a convolutional inter-layer $M^k_{S\rightarrow T}$ encoding the similarity between the source and the target intra-layers [5]. The inter-layer is typically generated by convolving both multiplex intra-layers as follows [6,5]:

$$M^k_{S\rightarrow T}(a, b) = \sum_p \sum_q M^k_S(p, q)M^k_T(a - p + 1, b - q + 1)$$

Where $a$ and $b$ respectively denote the row and the column of a specific element in the inter-layer while $p$ and $q$ respectively denote the row and column of a specific element in the intra-layers.

**Deep brain multiplex prediction using G-GAN.** For the target prediction task, we assume that each sample is represented by a ground-truth source network intra-layer. Our goal is to learn the target intra-layer from the source, then convolve both to ultimately synthesize a shallow brain multiplex. To do so, inspired by the graph translator model introduced in [9] proposing new graph convolution and deconvolution operations on graphs, we formulate our multiplex prediction task in the spirit of generative adversarial learning. Specifically, given an input ground-truth source intra-layer $M^k_S$ and a random noise, we predict the target intra-layer $\hat{M}^k_T$ by learning a local and global source to target $M^k_{T}$ mapping where ground-truth and predicted target intra-layers are enforced to share sparse patterns. This is achieved via an adversarial training of a network translator learning to generate fake target data from input source data by which it tries to fool the conditional discriminator aiming to learn to tell real and fake samples apart. The main advantage of such architecture is its adversarial training resulting in one network training the other in a bi-directional way.

Given a brain multiplex $M^k$ of subject $k$, let $M^k_S = (V, E, W)$ denote a source intra-layer, represented as a fully-connected (directed or undirected) graph. $V$ is a set of $n$ nodes, $E \subseteq V \times V$ is the set of edges and the set of weights for their corresponding edges are encoded in matrix $W \in \mathbb{R}^{n \times n}$, namely a weighted adjacency matrix. Let $e(i, j) \in E$ denote the edge from the node $v_i \in V$ to the node $v_j \in V$ and $W_{i,j} \in W$ denotes the corresponding weight of the edge $e(i, j)$. Analogously, $M^k_T = (V', E', W')$ denote the target intra-layer in $M^k$. The translation $T_c$ from source to target intra-layer mapping is defined as $T_c : U, S \rightarrow T$, where $U$ refers to the random noise and $S$ and $T$ respectively denote the domains of both source and the target intra-layers. The proposed G-GAN based DBMP framework aims to synthesize a fake target intra-layer $\hat{M}^k_T = T(M^k_S, U)$ that mimics the real target intra-layer $M^k_T$ in its topological and structural properties by minimizing the following adversarial loss function over all $k$ samples:
\[
\mathcal{L}(T_r, D) = \mathbb{E}_{M_S^k, M_T^k} [\log D(M_T^k|M_S^k)] + \mathbb{E}_{M_S^k, U} [\log (1 - D(T_r(M_S^k, U)|M_S^k))],
\]

(2)

where both \(T_r\) and \(D\) are trying to minimize the output of this loss function in an adversarial way. Since we translate the source into the target intra-layer which might lie on different manifolds, we enforce sparsely shared patterns across translated source intra-layer and ground-truth target intra-layer via \(L_1\) regularization which can boost the optimization process. The updated loss function is defined as follows to estimate the mapping translator \(T_r\):

\[
T_r^* = \arg\min_{T_r} \max_D \mathcal{L}(T_r, D) + \mathcal{L}_{L_1}(T_r),
\]

(3)

Where \(\mathcal{L}_{L_1}(T_r) = \mathbb{E}_{M_S^k, M_T^k, U} [||M_T^k - T_r(M_S^k, U)||_1]\). This is solved using ADAM optimizer for learning the translator and discriminator mappings alternatingly.

**Source to target intra-layer translator.** As displayed in Fig 1–A, the geometric translator acts as a U-Net with skip connections based on graph convolution in the three encoding layers and deconvolution in the three decoding layers. Specifically, the encoder network comprises two edge-to-edge convolutional layers and one edge-to-node convolutional layer, while the decoder network is composed of a node-to-edge deconvolutional layer and two edge-to-edge deconvolutional layers.

**Conditional network discriminator (Fig. 1–C).** The conditional discriminator is trained by minimizing the sum of two losses: the loss between the target intra-layers \(\{M_T^k\}_{k=1}^N\) and the source \(\{M_S^k\}_{k=1}^N\) and the loss between the predicted target multiplex intra-layers \(\{\hat{M}_T^k\}_{k=1}^N\) and ground-truth target intra-layers \(\{M_T^k\}_{k=1}^N\). In order to compute these losses separately, the conditional network discriminator inputs for training sets of pairs of multiplex layers: \(\{M_S^k, M_T^k\}_{k=1}^N\) and \(\{M_T^k, \hat{M}_T^k\}_{k=1}^N\) to distinguish the pair including the ground truth \(M_T^k\) from the predicted one \(\hat{M}_T^k\) by the translator for each training subject \(k\). The discriminator architecture comprises four stacked layers organized in the following order: two edge-to-edge convolutional layers, one edge-to-node convolutional layer, and fully connected layer followed by a softmax layer to output real or fake label for each input target intra-layer.

**Brain multiplex prediction from source network and classification.** Given the predicted target intra-layer for a testing subject \(tst\), we synthesize its brain multiplex \(\hat{M}_{tst} = \{M_{tst}^S, M_{tst}^{S \rightarrow T}, \hat{M}_{tst}^T\}\). Using 2-fold cross-validation, we train and test the proposed G-GAN DBMP model along a linear SVM classifier to label the multiplex \(\hat{M}\) into male or female. The training of the SVM classifier is proceeded by a feature selection step using the training samples. We particularly use infinite feature selection (IFS) [10] to select the most reproducible features distinguishing between male and female brain multiplexes.

**Gender-related brain multiplex connectivity discovery.** Following feature selection, we identify the top \(n_f\) gender-discriminative features. We design
a feature scoring algorithm by quantifying feature reproducibility across validation folds depending on the occurrence of each selected feature and its weight given by the IFS algorithm. Indeed, the more frequently a feature appears in the top $n_f$ set, the higher its score is.

3 Results and Discussion

**Fig. 2:** Average classification accuracy using linear SVM classifier and multiplex prediction error. (A) We report the mean absolute error (MAE) between the predicted and ground truth target intra-layers by (i) the proposed G-GAN method and (ii) KNN algorithm. (B) Average classification accuracy using: (i) a single source intra-layer network, (ii) the predicted multiplex, and (iii) the ground-truth multiplex [5]. (C) Predicted multiplex classification accuracy while defining the multiplex inter-layer as: (i) the learned G-GAN inter-layer preceding relu, (ii) the learned G-GAN inter-layer following relu, and (iii) the convolutional layer between both intra-layers as in [5].

**Evaluation dataset and method parameters.** We evaluated our proposed framework on a dataset of 400 healthy subjects (226 females with mean age = (21.6 ± 0.8) and 174 males with mean age = (21.6 ± 0.9) from the Brain Genomics Superstruct Project [11]. T1-weighted images were acquired using a 1.2mm isotropic resolution. Test-retest reliability was established with a correlations range from 0.75 for the estimated cortical thickness of the right medial prefrontal cortex to 0.99 for the estimated intracranial volume. We used FreeSurfer processing pipeline to reconstruct the left and right cortical hemispheres. Then we parcellated each hemisphere into 35 regions using Desikan-Killiany atlas. For each subject, we created 4 morphological brain networks (MBN) of size 35×35 as introduced in [12,6], derived from the following cortical attributes respectively: cortical thickness, sulcal depth, mean average curvature, and maximum principal curvature.

**Method evaluation and comparison methods.** Due to the absence of state-of-the-art methods focusing on learning how to predict brain multiplexes using G-GAN as well as investigating how gender differences manifest in brain
Fig. 3: Discovery of the top gender discriminative multiplex connectivities. We display the top 10 most discriminative multiplex connections for $M_1$ ($S =$ maximum principal curvature MBN and $T =$ mean cortical thickness MBN) and $M_2$ ($S =$ mean average curvature and $T =$ mean cortical thickness).

multiplexes, we resorted to comparing our framework against widely used methods such as KNN. This is a valid approach adopted by high-impact publication venues such as [13], where innovative proposed frameworks are benchmarked against the most commonly used techniques. KNN is a widely used algorithm due to its robustness to noisy data. Basically, for each testing subject with a source network, we first identify its top $K$ most similar training source brain networks. Next, we retrieve their corresponding target networks, which we average to predict the target testing intra-layer. To evaluate the potential of the predicted multiplex in boosting gender classification, we benchmarked our predicted multiplexes by G-GAN against classification using: (i) solely source intra-layer (i.e., from a single network), and (ii) ground truth shallow multiplexes where multiplexes were generated using two MBNs and one convolutional layer between them as in [5,14].

Fig 3–A shows the mean absolute error (MAE) between the predicted target intra-layers and the ground truth ones. For KNN, we report the average performance when varying the number of top $K$ selected neighbors from 2 to 10. Notably, G-GAN predicted multiplexes significantly outperforms the baseline KNN method. Fig 3–B displays classification results averaged across multiplexes and number of selected features $n_f \in \{310, \ldots, 350\}$ using three methods: source MBN, predicted multiplex by G-GAN, and ground-truth multiplexes. The predicted multiplexes significantly boosted the classification results in comparison to solely using the ground-truth source MBN. We also notice that it is quite close to the ground-truth multiplex. This classification experiments present an indirect method for evaluating the quality and reliability of the predicted multiplexes.
Insights into learning the multiplex inter-layer. Instead of computing the inter-layer by convolving both intra-layers, we extracted the learned final inter-layer in the encoding part of the translator: before and after the rectified linear unit (ReLU) activation function. Fig 3-C displays the average classification accuracy using the learned embedding of the source intra-layer as its gets translated into the target intra-layer (i.e., a hybrid inter-layer bridging the gap between the source and target intra-layers) and the conventional convolutional inter-layer [5]. Although, this experiment is considered as the first attempt to learn the multiplex inter-layer in the state-of-the-art, the convolutional predefined inter-layer still achieved the best classification result. This might be explained that the learned inter-layer is learned without any supervision by the sample class labels.

Gender multiplex fingerprint discovery. As a proof of concept, Fig ?? displays results from two multiplexes $M_1$ and $M_2$, where we report the top 10 most discriminative multiplex connectional biomarkers for distinguishing between male and female morphological brain networks. These connectional fingerprints were captured on both low and high-order levels. For instance, in $M_1$ ($S =$ maximum principal curvature MBN and $T =$ mean cortical thickness MBN), we found that the most gender-discriminative connectivity links (entorhinal cortex ↔ pericalcarine cortex) in both intra-layers. In fact, the entorhinal cortex is known to be correlated with spatial memory processing, where [15] showed that male participants outperform female participants. Besides, the pericalcarine cortex controlling visual processing and spatial navigation was identified as a gender discriminative region in [16]. While in the multiplex $M_2$ ($S =$ mean average curvature and $T =$ mean cortical thickness), we found that the most discriminative connectivity included the (pars triangularis ↔ pericalcarine cortex). For instance, the pars triangularis is known to be related to language and speech processing, which is in line with the observation that females are better at language processing than males as reported in [17]. This hallmark connectivity was found in the convolutional inter-layer. We also notice other relevant brain multiplex connectivities distributed across intra-layers. In our future work, we will investigate the reproducibility of the discovered multiplex biomarkers fingerprinting gender using different perturbation strategies of the training set while tracking the reproducibility of the top ranked features [18].

4 Conclusion

In this paper, we presented the first framework for brain multiplex synthesis from a source brain network with application to gender classification. Specifically, we rooted our method in geometric generative adversarial network (G-GAN) nicely preserving local and global brain connectional patterns, where a network translator and discriminator are adversarially trained. The predicted multiplexes significantly boosted the classification results compared to using single source networks. More importantly, we identified for the first time both low-order and high-order morphological brain connectivities encoding gender differences using
large GSP dataset. There are many possible future directions yet to explore, such as learning discriminative multiplex inter-layers by integrating a second conditional discriminator by class labels.

5 Supplementary material

We provide two supplementary items for reproducible and open science:

1. A 6-min YouTube video explaining how our prediction framework works on BASIRA YouTube channel at https://youtu.be/iTjPtC4BULc.
2. An improved version of the adversarial brain multiplex generation code is available on GitHub at https://github.com/basiralab/ABMT.

6 Acknowledgement

This project has been funded by the 2232 International Fellowship for Outstanding Researchers Program of TUBITAK (Project No:118C288, http://basira-lab.com/reprime/) supporting I. Rekik. However, all scientific contributions made in this project are owned and approved solely by the authors.

References

1. Gong, G., He, Y., Evans, A.C.: Brain connectivity: gender makes a difference. The Neuroscientist 17 (2011) 575–591
2. Shirao, N., Okamoto, Y., Okada, G., Ueda, K., Yamawaki, S.: Gender differences in brain activity toward unpleasant linguistic stimuli concerning interpersonal relationships: an fmri study. European archives of psychiatry and clinical neuroscience 255 (2005) 327–333
3. Saunders, P.A., Copeland, J., Dewey, M., Gilmore, C., Larkin, B., Phaterpekar, H., Scott, A.: The prevalence of dementia, depression and neurosis in later life: the liverpool mrc-alpha study. International Journal of Epidemiology 22 (1993) 838–847
4. Nebli, A., Rekik, I.: Gender differences in cortical morphological networks. Brain Imaging and Behavior (2019) 1–9
5. Raeppe, R., Lisowska, A., Rekik, I., Alzheimer’s Disease Neuroimaging Initiative and others: Joint correlational and discriminative ensemble classifier learning for dementia stratification using shallow brain multiplexes. MICCAI (2018) 599–607
6. Mahjoub, I., Mahjoub, M.A., Rekik, I.: Brain multiplexes reveal morphological connectional biomarkers fingerprinting late brain dementia states. Scientific reports 8 (2018) 4103
7. Yi, X., Walia, E., Babyn, P.: Generative adversarial network in medical imaging: A review. arXiv preprint arXiv:1809.07294 (2018)
8. Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P.: Geometric deep learning: going beyond euclidean data. IEEE Signal Processing Magazine 34 (2017) 18–42
9. Guo, X., Wu, L., Zhao, L.: Deep graph translation. arXiv preprint arXiv:1805.09980 (2018)
10. Roffo, G., Melzi, S., Cristani, M.: Infinite feature selection. Proceedings of the IEEE International Conference on Computer Vision (2015) 4202–4210

11. Buckner, R., Hollinshead, M., Holmes, A., Brohawn, D., Fagerness, J., O’Keefe, T., Roffman, J.: The brain genomics superstruct project. Harvard Dataverse Network (2012)

12. Soussia, M., Rekik, I.: High-order connectomic manifold learning for autistic brain state identification. International Workshop on Connectomics in Neuroimaging (2017) 51–59

13. Samusik, N., Good, Z., Spitzer, M.H., Davis, K.L., Nolan, G.P.: Automated mapping of phenotype space with single-cell data. Nature methods 13 (2016) 493

14. Lisowska, A., Rekik, I., Initiative, A.D.N., et al.: Pairing-based ensemble classifier learning using convolutional brain multiplexes and multi-view brain networks for early dementia diagnosis. International Workshop on Connectomics in Neuroimaging (2017) 42–50

15. Cherney, I.D., Brabec, C.M., Runco, D.V.: Mapping out spatial ability: sex differences in way-finding navigation. Perceptual and motor skills 107 (2008) 747–760

16. Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Hakonarson, H., Gur, R.E., Gur, R.C., Verma, R.: Sex differences in the structural connectome of the human brain. Proceedings of the National Academy of Sciences 111 (2014) 823–828

17. Bourne, V.J.: Lateralised processing of positive facial emotion: Sex differences in strength of hemispheric dominance. Neuropsychologia 43 (2005) 953–956

18. Georges, N., Mhiri, I., Rekik, I., Alzheimer’s Disease Neuroimaging Initiative and others: Identifying the best data-driven feature selection method for boosting reproducibility in classification tasks. Pattern Recognition 101 (2020) 107183