Task-guided Generative Adversarial Networks for Synthesizing and Augmenting Structural Connectivity Matrices for Connectivity-Based Prediction

Tatsuya Yamamoto[1], Tomoki Sugiura[1], Tomoyuki Hiroyasu[2], and Satoru Hiwa[2]
Graduate School of Life and Medical Sciences, Doshisha University. [1]
(1-3 Tatara Miyakodani, Kyotanabe, Kyoto 610-0394, Japan)
Department of Biomedical Sciences and Informatics, Doshisha University. [2]
(1-3 Tatara Miyakodani, Kyotanabe, Kyoto 610-0394, Japan)
Corresponding Author: Satoru Hiwa (shiwa@mail.doshisha.ac.jp)

Keywords: Data augmentation; structural connectivity; generative adversarial networks (GANs); connectome-based prediction; generative models

Abstract

Recent machine learning techniques have improved the modeling of complex dependencies between brain connectivity and cognitive/behavioral traits, facilitating connectome-based predictions. However, they typically require large datasets. While large open datasets like the Human Connectome Project have offered significant benefits to connectomics research, collecting such large data remains a challenge due to the financial cost and time. To address this issue, we propose Task-guided GAN II, a novel data augmentation method leveraging generative adversarial networks (GANs) to enhance the sample size from limited datasets for connectome-based prediction tasks. Distinguishing from previous approaches, our method incorporates a task-guided branch within the conventional Wasserstein GAN framework, specifically designed to synthesize structural connectivity matrices. It aims to effectively augment data and improve the prediction accuracy of human cognitive traits by capturing more task-directed features within the data. We evaluated the effectiveness of data augmentation using Task-guided GAN II in predicting fluid intelligence utilizing the NIMH Health Research Volunteer Dataset. Our results demonstrate that data augmentation with Task-guided GAN II not only improves prediction accuracy but also ensures that its latent space effectively captures correlations between structural connectivity and cognitive outcomes. Our method would be beneficial in leveraging small datasets for human connectomics research.

1. Introduction

Connectomics, which represents neural circuits as networks and analyzes their topology through mathematical approaches, has emerged as an essential methodology in cognitive and computational neuroscience [1] [2]. It facilitates a deeper understanding of the complex relationships among brain networks, cognition, behavior, and individual differences in human cognitive traits and behaviors. Additionally, recent advances in machine learning methodologies have employed connectome-based machine learning models using structural and functional connectivity matrices to predict behavior or cognitive traits [3].

Recent studies have revealed that building prediction models requires more than a hundred samples [4], [5] and has also been shown to be improved in accuracy and robustness as the sample size becomes larger, not depending on the choice of regression models [6]. Therefore, researchers have often used large-sample open datasets such as the Human Connectome Project (HCP) Dataset (N = 1200) [7], ABCD Study (N = 11878) [8], UK Biobank (N = 40000+) [9]. These datasets have contributed significantly to developing analytical methods and understanding human connectomics [10]. However, Yeung et al. have reported that most self-recruited samples included around 100 subjects [5]. In the exploratory stage of the research, it is not easy for individual researchers to collect large samples over 100 due to the limitations on time and cost.

One of the best solutions to resolve the small sample size issue is to employ a data augmentation that generates a new sample by manipulating the given small dataset. In traditional data augmentation, geometric transformation has often been utilized in image data, such as rotation or flipping to the original data. However, such manipulations are inapplicable to the matrix-shaped data of brain functional and structural connectivity because their shape was constrained by the order of brain regions. In recent years, data augmentation methods using generative models have been proposed [11], [12], [13], [14], [15], [16], [17]. In these methods, generative augmentation models approximate the distribution of a given dataset and synthesize new samples with similar characteristics to the original dataset. Specifically, generative adversarial networks (GAN) have been increasingly used to data synthesis in numerous application fields with its strong feature of learning the local and global structure of data [18]. In the neuroscience field, Chao Li et al. proposed the structural connectivity augmentation method named BrainNetGAN to build the model for the classification of Alzheimer’s disease [12]. The accuracy of the graph neural network-based classification model of Alzheimer's disease was improved using the structural connectivity data synthesized by their BrainNetGAN. Ruizhe Li et al. also suggested a data augmentation method for T1-weighted image named Task-guided GAN (TG GAN) for the brain age prediction model [13]. They improved the accuracy of the age prediction [13]. In their TG

GAN model, the task-guided branch of a regression model for brain age prediction was incorporated into the GAN architecture. This accelerated the data synthesis to be more task-specific and contributed to improving the prediction task.

There are a lot of studies on data augmentation methods for medical images, but a few studies on brain connectome data augmentation methods for the connectome-based prediction task. The BrainNetGAN, as mentioned above, was developed for the classification task and trained using specific class labels to synthesize conditioned data. On the other hand, the TG GAN was designed with a 3D image input structure specifically for the T1-weighted image synthesis. Neither model is suitable for synthesizing connectivity matrices aimed at prediction tasks. Given the rapidly increasing interest in connectome-based prediction studies, there is an urgent need to develop a data augmentation model tailored to these tasks. Here, we propose a novel data augmentation method, Task-guided GAN II (TG GAN II). It was designed for synthesizing the structural connectivity matrices and employs a task-guided branch to predict human cognitive traits from the structural connectivity. We aim to enhance the accuracy of predicting cognitive traits from structural connectivity by augmenting the original dataset with the data synthesized by our TG GAN II. To assess the effectiveness of our method, we utilize the NIMH Health Research Volunteer Dataset [19] to build the model predicting fluid intelligence based on structural connectivity. The TG GAN II employs the Wasserstein GAN (WGAN) without task-guided branches as its baseline model. It is evaluated against this baseline in the following three perspectives: 1) the similarity of graph features between synthesized and original data, 2) the improvement in prediction performance, and 3) the influence of the added synthetic data volume on prediction performance enhancement. We hope that these investigations will make a substantial contribution toward addressing the sample size challenges in connectome-based prediction tasks.

2. Methods

2.1 Baseline model: Wasserstein GAN with gradient penalty (WGAN-GP)

We employed a Wasserstein GAN with gradient penalty (WGAN-GP) [20] as the baseline model for our proposed TG GAN II. The WGAN-GP incorporated the Wasserstein distance measure into its objective function to improve the stability of model training by suppressing the mode collapse. It consisted of the generator that synthesized new structural connectivity matrices, and the critics assessed the synthesized images by estimating the Wasserstein distance between synthesized and given ones. We used an autoencoder-based generator in our WGAN-GP model to synthesize new images by interpolating the latent feature space (**Figure 1**).

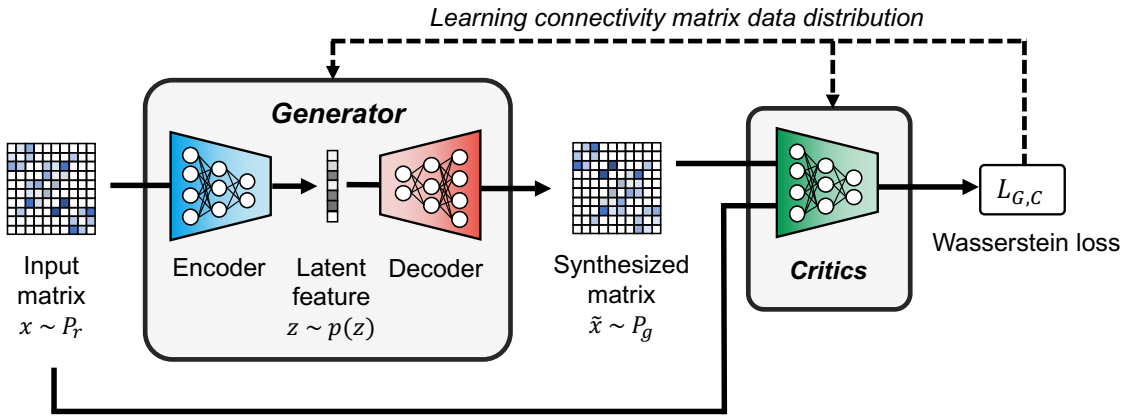

Figure 1. The architecture of the WGAN-GP used for synthesizing connectivity matrices, serving as the baseline model, where P_r and P_g denote the distributions of given and synthesized connectivity matrices, respectively, and x and $\tilde{x}$ are the given input and synthesized matrices, respectively. The encoder transforms the input matrix into latent variables z, while the decoder reconstructs the matrix $\tilde{x}$ from z. The critics aim to estimate the Wasserstein distance between the given and synthesized matrices, while the generator seeks to minimize this distance.

The encoder in the generator module extracted features from given connectivity matrices and mapped them into the latent space, and the decoder module synthesized the matrices from the latent variables. We employed convolutional neural networks for brain networks (BrainNetCNN) proposed by Kawahara et al. [21], which consisted of four convolutional layers and two fully connected layers for the encoder. The BrainNetCNN had three convolutional layers specialized for extracting topological features of connectivity matrices: edge-to-edge (E2E), edge-to-node (E2N), and node-to-graph (N2G) convolutions (**Figure 2**). We adopted a simple feed-forward neural network with five fully connected layers for the decoder. Besides, we used the softplus function as an activation function in the final layer because the elemental value of structural connectivity took zero or more. The decoder outputted a vector of the lower triangle part in the connectivity matrix, which was finally matricized. We also utilized the BrainNetCNN architecture with four convolutional layers for the critics.

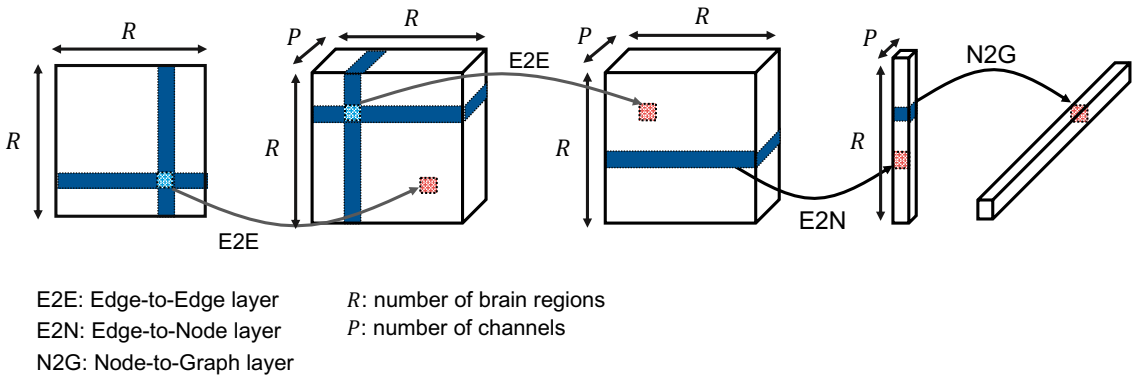

Figure 2. Figure 2. Architecture of BrainNetCNN. This architecture introduces convolution kernels designed for the adjacency matrix of brain networks, a development unique to BrainNetCNN. These kernels are applied to an element's entire row and column, enabling edge-to-edge (E2E) and edge-to-node (E2N) convolution operations. Initially, the connectivity matrix undergoes convolution with one or more E2E kernels, emphasizing the connections between adjacent brain regions. Subsequently, the output from the E2E convolution is processed through an E2N filter, which calculates a weighted sum of edges for each brain region. A node-to-graph (N2G) kernel integrates these weighted node values to produce a single output value, representing an aggregate measure of brain connectivity.

The mathematical description of the loss function for the generator and critics in our WGAN-GP is presented in Equation 1. The critics aimed to maximize this equation to accurately estimate the Wasserstein distance between the synthesized and original matrices. Conversely, the generator was trained to minimize the same equation, bringing the distributions of the synthesized and original matrices closer together.

$$L_{G,C} = \min_{G} \max_{C} \mathbb{E}_{x\sim P_r}[C(x)] - \mathbb{E}_{\tilde{x}\sim P_g}[C(\tilde{x})] - \lambda(\|\nabla_{\hat{x}} C(\hat{x})\|^2 - 1)^2 \quad (1)$$

where P_r is the distribution of given connectivity matrices, and P_g is the distribution of synthesized matrices. $C(x)$ denotes the output of the critics for a given input x, and $\tilde{x}$ represents the output of the decoder $G(z)$ when given a latent variable $z\sim p(z)$ generated in the encoder. The third term in the Equation 1 introduces a gradient penalty for encouraging $C(x)$ to adhere to the 1-Lipschitz continuity, with λ being the coefficient of this penalty.

2.2 Task-guided GAN II

In this study, we have built upon the baseline WGAN-GP model to develop a novel GAN-based connectivity synthesis model named TG GAN II. Our proposed model extends the WGAN-GP by incorporating a task-guided branch specifically designed for prediction tasks. Its objective is to augment the dataset with additional connectivity matrices and their corresponding objective

variables, thereby facilitating the construction of a more accurate cognitive trait prediction model based on brain connectivity.

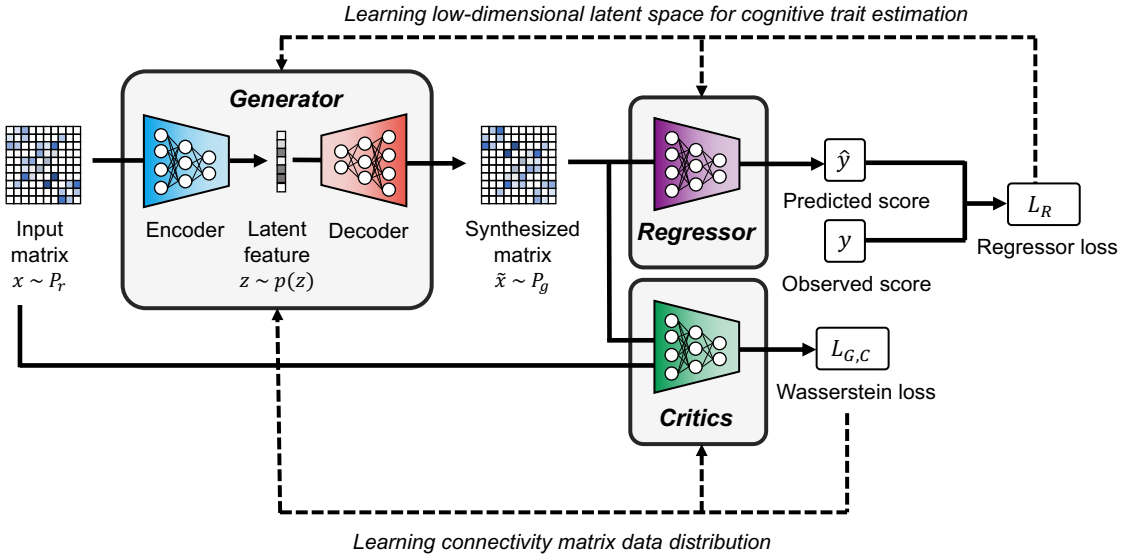

Figure 3. The architecture of Task-guided GAN II for synthesizing connectivity matrices. This model introduces modifications over the WGAN-GP, notably incorporating a task-guided regressor branch and specialized loss function calculations. The regressor branch features four convolutional layers and four fully connected layers, which together output the predicted objective variable based on the synthesized connectivity matrix. The loss function of the WGAN-GP is expanded to include the root mean squared error (RMSE) between the observed and predicted objective variables, facilitating more accurate predictions.

Figure 3 illustrates the architecture of TG GAN II. Within this framework, a regressor model for cognitive trait prediction was integrated into the WGAN-GP structure as the task-guided branch to enhance the performance of the prediction task. This branch was designed to generate a latent space that captured the variations in both explanatory and objective variables. It consisted of four convolutional layers and four fully connected layers, culminating in the output of the predicted objective variable from the synthesized connectivity matrix. The loss function for the task-guided regressor of TG GAN II, outlined in Equation 2, was calculated by the root mean squared error (RMSE) between the observed and predicted objective variables, with both the generator and critics optimized to minimize this function:

$$L_R = \min_R \sqrt{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{2}$$

where y_i denotes the observed objective variable and $\hat{y}_i$ is the predicted one. In the training of TG GAN II, both Equations 1 and 2 were optimized simultaneously, resulting in the comprehensive loss function for TG GAN II being represented in Equation 3:

$$\min_{G,R} \max_{C} L_{G,C} + \alpha\, L_R \tag{3}$$

where the coefficient α served as the weight against L_R, the regression loss, balancing the two terms of the loss function.

2.3 Latent space interpolation for connectivity matrices and objective variable synthesis

Using the trained TG GAN II model, we aimed to synthesize new connectivity matrices and their corresponding objective variables by performing the interpolation in the latent space. A single pair of the connectivity matrix and its objective variable was synthesized using two pairs of them (X_i, y_i) and (X_j, y_j) following Equations 4 and 5 referring to the method provided by Li et al. [10]:

$$X_s = G(z_s),\ \ z_s = \varepsilon D(X_i) + (1-\varepsilon) D(X_j) \tag{4}$$

$$y_s = \varepsilon y_i + (1-\varepsilon) y_j \tag{5}$$

where ε is the value between 0 and 1. This procedure is illustrated in **Figure 4**.

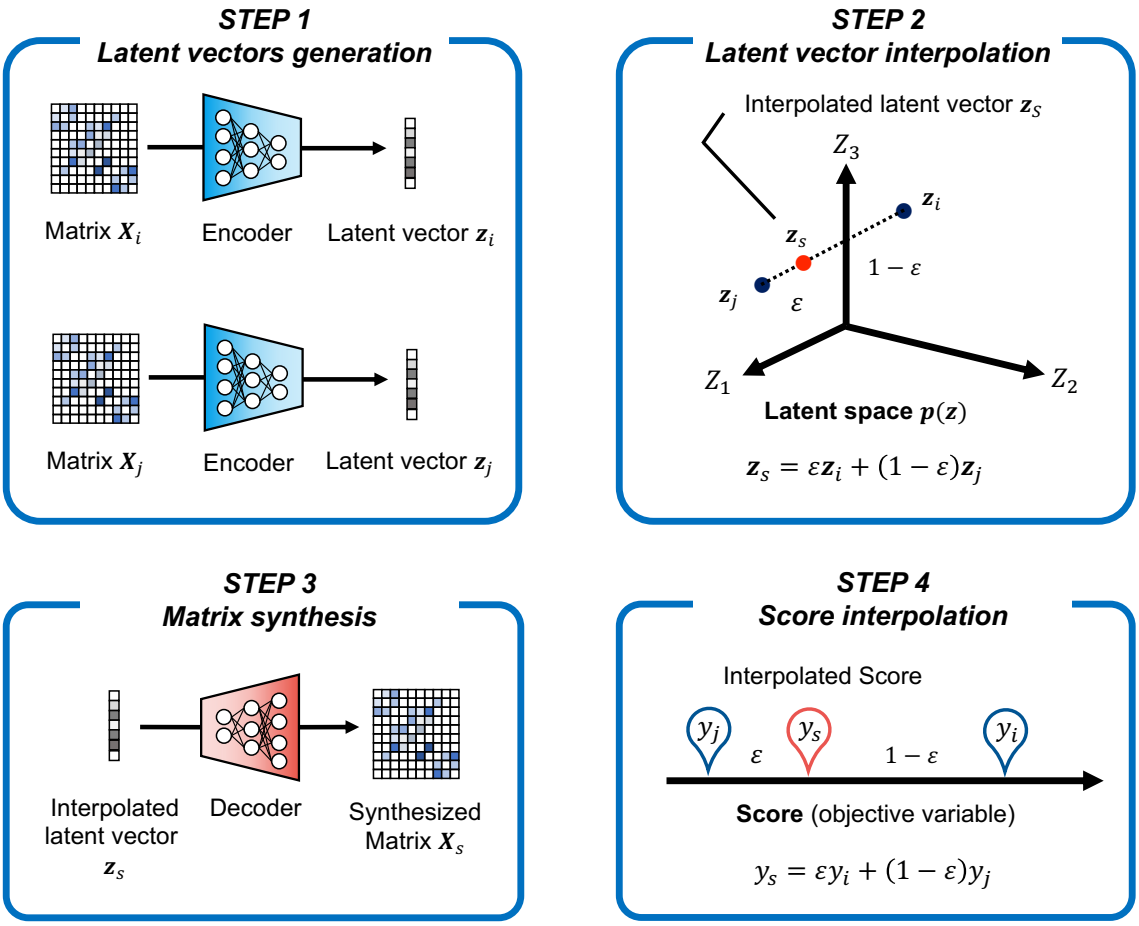

Figure 4. Synthesis of connectivity matrices and their corresponding objective variables through latent vector interpolation. This process involves using two pairs of given connectivity matrices and objective variables, (X_i, y_i) and (X_j, y_j), as inputs. These samples are first mapped to the latent space using the trained generative model. A new set of latent variables is then generated by performing linear interpolation between these mapped points. The interpolated latent variables are fed into the decoder, which reconstructs the connectivity matrices.

2.4 Model training of generative models for data augmentation

2.4.1 Dataset

In this study, we used MRI images and cognitive scores in the NIMH Healthy Research Volunteer Dataset (https://openneuro.org/datasets/ds004215/versions/1.0.1) [19]. We used T1-weighted images, diffusion-weighted images, and the score of NIH Toolbox Cognition Battery [22] (age-adjusted score) as cognitive traits from 108 samples in which T1-weighted image, DWI, and resting fMRI had been obtained without deficiencies. We showed the average and standard variation of these scores in Table. 1. We summed up four task scores in each participant and used

it as a fluid intelligence score. This composed score shows a higher signal-noise ratio than each task score and is less affected by the variability of each task score [23].

Table 1. NIH Toolbox Cognition Battery Score

Domain	Average	SD
Flanker Inhibitory Control and Attention	91.3	12.1
List Sorting Working Memory	103.5	13.0
Picture Sequence Memory	114.9	16.7
Dimensional Change Card Sort	109.6	17.9
Fluid Intelligence	419.2	40.9

2.4.2 MRI data preprocessing

We preprocessed DWI and T1 images using a pre-configured DWI process pipeline, QSIPrep (0.15.4) [24]. The preprocessing process consists of the following six steps: (1) T1-weighted image preprocessing, (2) Denoising DWI using Marchenko-Pastur Principal Component Analysis (MP-PCA) [25], (3) B1 field inhomogeneity correction, (4) Susceptibility distortion correction (TOPUP) [26], (5) Eddy current and head motion correction (EDDY) [27], (6) Registration of DWI into T1-weighted image.

2.4.3 Structural connectivity mapping

We also reconstructed whole brain tractograms and structural connectivity using QSIPrep (0.15.4) [24]. The reconstruction of structural connectivity maps was performed in the following steps. First, the fiber orientation distribution function (fODF) was estimated using the Single-Shell 3-Tissue constrained spherical deconvolution (SS3T CSD) model [28] (the response function was estimated following Dhollander et al. [29]). Second, the whole brain tractogram was estimated to be 10 million streamlines using probabilistic (iFOD2 [30]) and anatomically-constrained tractography (ACT) [31]. Third, we computed streamline weights based on the SIFT2 algorithm that reduced the biases in probabilistic tractography [32]. Finally, the whole brain was parcellated into 116 regions based on automated anatomical labeling (AAL), and the structural connectivity was defined as the sum of SIFT2-weighted streamlines connecting two arbitrary regions divided by the sum of the volumes of those regions.

2.4.4 Training setup for generative models

The entire dataset was divided into the discovery dataset (70%; $N = 75$) and the test dataset (30%; $N = 33$). In this division, we sorted the samples by their objective variables and then assigned the samples with a rank of (2nd, 3rd, 5th, 6th, ..., 103rd, 105th, 106th, 108th) to the

discovery dataset and (1st, 4th, 7th, ..., 101st, 104th, 107th) to the test dataset. This method, proposed by Cui et al. [33], was employed to ensure a similar distribution of behavioral scores across datasets and minimize the random bias resulting from the division. The discovery dataset was used to train TG GAN II and WGAN-GP, select the data augmentation model, and construct the prediction model. Conversely, the test dataset validated the prediction model's accuracy. We adopted a hold-out validation scheme for the training and model selection processes in TG GAN II and WGAN GP. Similarly to the discovery-test splitting, we sorted the discovery samples according to their objective variables. Then we assigned samples with a rank of (2nd, 3rd, 5th, ..., 72nd, 74th, 75th) to the training dataset and (1st, 4th, ..., 73rd) to the validation dataset.

The TG GAN II and WGAN-GP models were trained with Adam optimizer with momentum parameters $\beta_1 = 0.9$, $\beta_1 = 0.999$ (learning rate: 0.0001, batch size: 2), and were run for 2000 epochs. We implemented an early stopping strategy to prevent overfitting on the training dataset. The training was stopped if the regressor loss on the validation dataset did not improve for 50 consecutive epochs after reaching the 500th epoch.

The dropout rate p of the encoder, critics, and regressor was selected from a set of values: 0.1, 0.2, 0.3, 0.4, and 0.5. Similarly, the coefficient α was selected from a range between 0.1 and 1.0 in increments of 0.1. For all trained TG GAN II models, we selected the model that minimized the sum of the regressor loss on the training and validation datasets for data augmentation. In the case of WGAN-GP models, which do not include a regressor, we selected models based on having the same dropout rate as the chosen TG GAN II model.

2.4.5 Evaluation of synthesized connectivity matrices based on graph theory metrics

We evaluated the similarity between the structural connectivity matrices synthesized by TG GAN II and WGAN-GP and the original (acquired) matrices within the discovery dataset. First, we synthesized a connectivity matrix for each sample using these models. To quantify the differences between the acquired and synthesized matrices, we subtracted the synthesized matrices from their corresponding acquired matrices for each generative model. Then, we calculated the average difference matrix for each model. These averaged difference matrices were visualized and compared.

In the subsequent analysis, we quantitatively assessed the similarity between the acquired and synthesized matrices using graph theory metrics: connectivity strength, betweenness centrality, and clustering coefficient. For this purpose, we computed the average matrices for both acquired and synthesized connectivity matrices and binarized them according to edge density thresholds,

ranging from 5% to 25% in increments of 5%. We then estimated the distribution of these three graph theory metrics for the acquired and synthesized averaged matrices. The similarity between these distributions was quantified using the Kullback-Leibler (KL) divergence.

2.4.6 Latent space evaluation

We assessed the relevance between the fluid intelligence scores and the latent space generated by TG GAN II and WGAN-GP encoders. Latent variables $\boldsymbol{z}$ were sampled from the discovery dataset. We projected the acquired latent variables onto a two-dimensional space using principal component analysis (PCA) to visualize the latent space. Subsequently, we calculated the Pearson correlation between the first principal component (PC1) scores and the corresponding fluid intelligence scores.

2.5 Performance evaluation on data augmentation for connectome-based prediction

To evaluate the effectiveness of TG GAN II, we built the regression model to predict the fluid intelligence scores from the structural connectivity matrices augmented by two generative models. We used the Ridge regression algorithm [34], which is widely used in the neuroscience field. The prediction performance of the fluid intelligence was then compared between models augmented by TG GAN II and WGAN-GP.

2.5.1 Training setup for a prediction model

To assess the prediction model's efficacy within the discovery dataset, we utilized repeated 5-fold cross-validation (5F CV). Additionally, the L2 regularization parameter in Ridge regression was fine-tuned using 5-fold cross-validation within each fold of the outer 5F CV. The methodology for this repeated nested CV process is outlined in subsequent sections.

Repeated outer 5F CV In the outer 5F CV phase, the discovery dataset was randomly divided into five subsets. Four subsets were merged and used as the training dataset, while the remaining subset served as the validation dataset. We constructed the prediction model using the training dataset and the parameters determined during the inner 5F CV phase. This model was then validated against the validation dataset. This cycle of training and validation was repeated until each subset had been used as the validation dataset once. The prediction accuracy was calculated as the Pearson correlation between the observed and predicted scores. The outer 5F CV was repeated 20 times to avoid the bias resulting from random splitting.

Inner 5F CV and parameter tuning Within each iteration of the outer 5F CV, the L2 regularization parameter for Ridge regression was optimized through an inner 5F CV process. We

selected this parameter from a set of 16 possible values: $\alpha \in \{x \mid x = 2^n, n \in \mathbb{Z}, n \in [-10, 5]\}$.. For the inner 5F CV, the training dataset from the outer loop was randomly split into five subsets. Four subsets were used in each iteration to train the model under each parameter setting, with the fifth subset reserved for validation. This procedure was executed five times, ensuring each subset was utilized as the validation dataset. For each parameter across the inner 5F CV loops, we calculated a Pearson correlation and a mean absolute error (MAE) between the observed and predicted scores, averaging these metrics across all loops. The sum of the mean correlation and the inverse of MAE, standardized for differing scales, served as the measure of prediction accuracy in inner validation [6], [33]. The parameter that yielded the highest accuracy model was selected in the subsequent outer 5F CV.

Evaluating Test-Retest Reliability To evaluate the test-retest reliability of the prediction model built through repeated nested CV, 100 prediction models from the outer 5F CV were validated using the test dataset. The prediction accuracy was defined by Pearson's correlation and RMSE between observed and predicted scores.

2.5.2 Evaluation of prediction performance

Finally, we evaluated the accuracy of the fluid intelligence score prediction model augmented by TG GAN II and WGAN-GP. Data augmentation was applied to the discovery dataset, increasing the structural connectivity matrices and fluid intelligence scores to double (original data + 100%), triple (+ 200%), quadruple (+ 300%), and quintuple (+ 400%) through the latent space interpolation, as detailed in section 2.4.6. Utilizing the original and augmented datasets, we constructed fluid intelligence prediction models using the repeated nested CV method outlined in section 2.5.1. These models were then validated against the test dataset. To compare the prediction accuracy between the data-augmented models and the baseline (non-augmented: original sample size) model, we conducted Welch's *t*-test. For this analysis, outlier detection was performed on the accuracy of 100 prediction models, identifying any values more than three standard deviations from the mean as outliers. Additionally, we applied the Bonferroni correction for multiple comparisons.

3. Results

3.1 Analysis of synthesized connectivity matrices

Figure 5 presents the average connectivity matrices for both the acquired and the synthesized structural matrices using TG GAN II and WGAN-GP, with brain regions aligned according to Yeo's 7 Network definition. In **Figure 5A**, the weighted matrices are displayed, whereas **Figure 5B** showcases those binarized at a 25% edge density threshold. The synthesized matrices from

WGAN-GP exhibit connections that do not exist in the acquired matrices, not observed with TG GAN II. This is evident in both the weighted and binarized connectivity matrices.

Figure 6 displays the average difference in connectivity matrices between the acquired and those synthesized by TG GAN II and WGAN-GP. Each part of the connectivity matrix differentiated the connectomes: intra-hemispherical connections (left hemisphere: A, right hemisphere: C), inter-hemispherical connections (B), cerebrum-cerebellum connections (D), and intra-cerebellar connections (E). The differences were notable in the intra-hemispherical and intra-cerebellar connections. TG GAN II and WGAN-GP tended to estimate weaker connections than those observed in the actual data. Additionally, WGAN-GP was shown to generate more inter-hemispherical and cerebrum-cerebellum connections compared to TG GAN II, indicating that TG GAN II synthesized connectivity matrices similar to the acquired data in these regions than WGAN-GP.

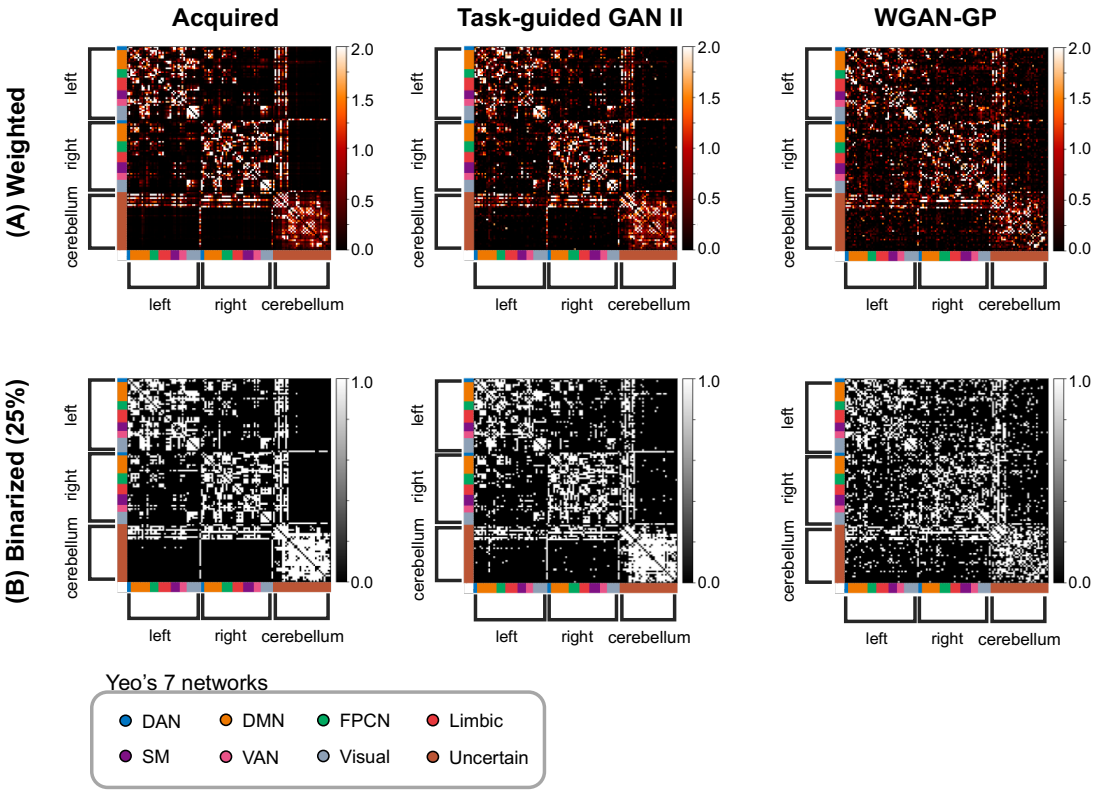

Figure 5. Comparison between acquired and synthesized structural connectivity matrices using Task-guided GAN II and WGAN-GP. (A) displays weighted connectivity matrices, and (B) shows binarized connectivity matrices, where binarization was applied to retain only the top 25% of the strongest connections. The brain regions in matrices were aligned according to Yeo's 7 Network definition.

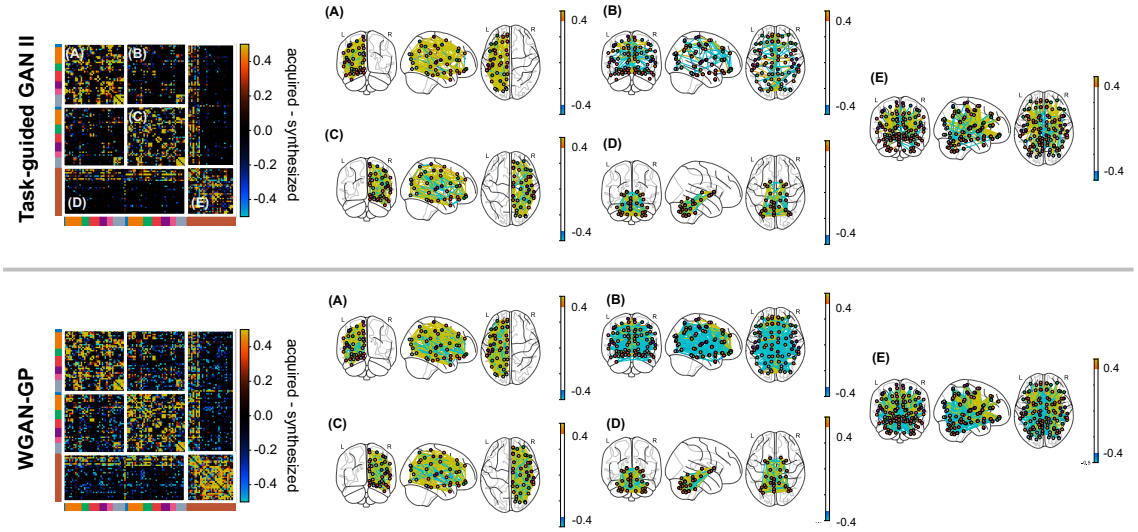

Figure 6. Comparison of differences between acquired and synthesized structural connectivity matrices using Task-guided GAN II and WGAN-GP. This figure illustrates the average differences across connectivity matrices, highlighting the differences between acquired and synthesized connectomes. The connectivity matrix and connectome are segmented to distinguish various connectome components: intra-hemispherical connections (left hemisphere: A, right hemisphere: C), inter-hemispherical connections (B), cerebrum-cerebellum connections (D), and intra-cerebellar connections (E).

Figure 7 indicates a quantitative analysis of the topological features of both the acquired and synthesized connectivity matrices, utilizing graph theory metrics: connectivity strength, betweenness centrality, and clustering coefficient. The distribution of each graph metric is displayed as a density plot, along with the similarity between the distributions of the acquired and synthesized matrices estimated using KL-divergence. Notably, the KL-divergence values and density plots indicate no difference in the distribution of connectivity strength between matrices synthesized by TG GAN II and those by WGAN-GP. Conversely, the distributions of betweenness centrality and clustering coefficient suggested that TG GAN II produced matrices with topological characteristics more closely aligned to those of the acquired matrices than WGAN-GP. This is evidenced by smaller KL-divergence values and the similarity of density plots across all edge density thresholds.

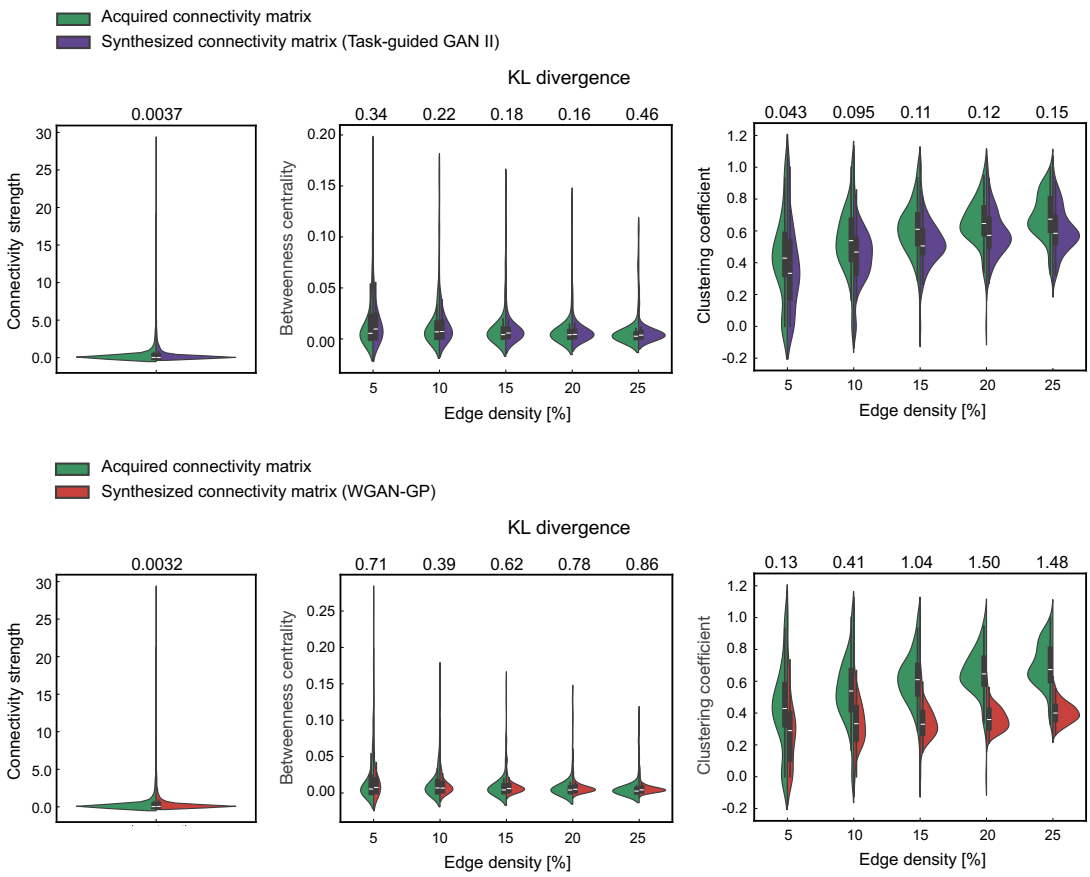

Figure 7. Distribution of graph theory metrics (connectivity strength, betweenness centrality, clustering coefficient) across acquired and synthesized connectivity matrices via TG GAN II and WGAN-GP under various threshold settings. The upper section displays results from Task-guided GAN II, while the lower section indicates WGAN-GP results. The similarity between the distributions of graph metrics is quantified using the Kullback-Leibler divergence.

3.2 Latent space analysis

Figure 8 illustrates the latent spaces of each generative model, projected into two-dimensional space using PCA, with each sample annotated by its corresponding fluid intelligence score. The left side of **Figure 8** displays the PCA-projected latent spaces, while the right side shows the correlation between the first principal component (PC1) and the fluid intelligence scores. The correlation is higher in TG GAN II (r = 0.68, p < .05) than WGAN-GP (r = 0.18, p > .05), indicating that the latent space in TG GAN II is more closely associated with fluid intelligence scores. These results demonstrated that TG GAN II effectively captured the variance in fluid intelligence scores within its latent space.

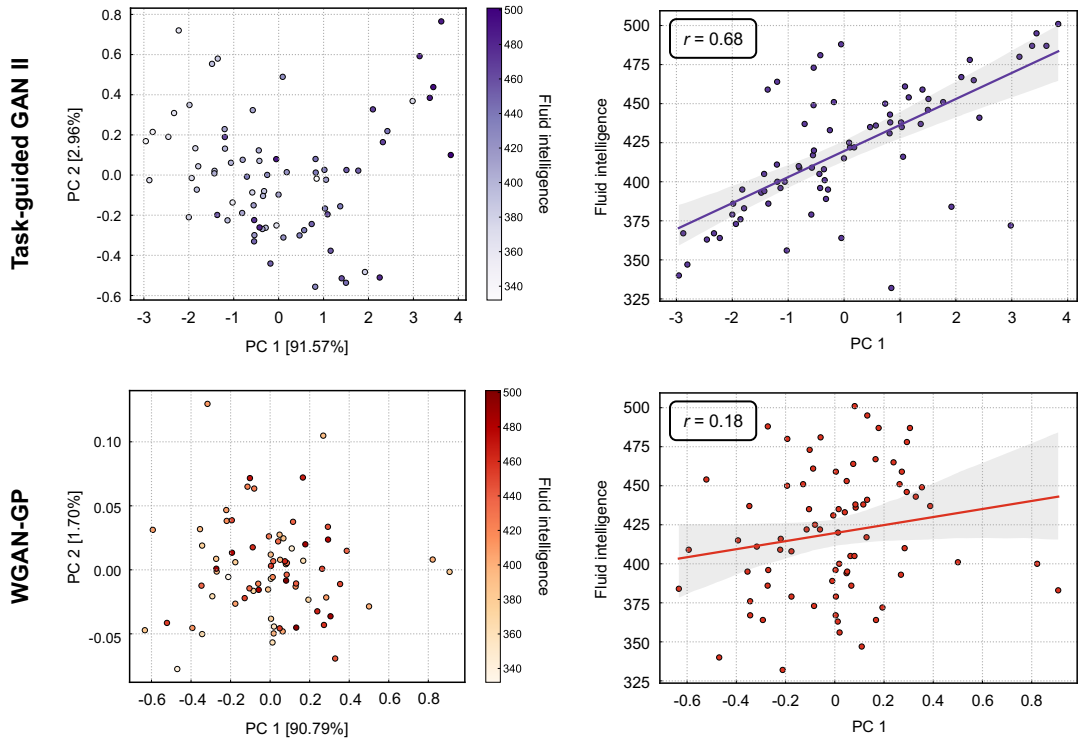

Figure 8. Visualization of the latent space through principal component analysis (PCA) on the left, and the correlation between the first principal component (PC 1) and the fluid intelligence score on the right.

3.3 Prediction performance

Figure 9 and **Figure** 10 illustrate the accuracy of the prediction models using Pearson's correlation between observed and predicted fluid intelligence scores, as well as RMSE, for data augmented by TG GAN II and WGAN-GP, respectively. The term 'baseline' in each figure refers to models developed solely with acquired samples without data augmentation. Each data point represents a model constructed through the repeated nested CV method. The results in **Figure 9** revealed that data augmentation with TG GAN II significantly enhanced the correlation between observed and predicted scores beyond the baseline ($p < .01$, Bonferroni-corrected), whereas augmentation with WGAN-GP did not improve. Regarding RMSE, augmentation with neither TG GAN II nor WGAN-GP led to improvements. Also, data augmentation with WGAN-GP resulted in worse RMSE than the baseline, whereas the RMSE for models augmented with TG GAN II remained consistent with the baseline (**Figure 10**). These results were consistent also in the summary of the accuracy of data-augmented and baseline prediction models indicated in **Table 2**. Notably, increasing the number of synthesized samples did not lead to further improvements in prediction accuracy, even with TG GAN II.

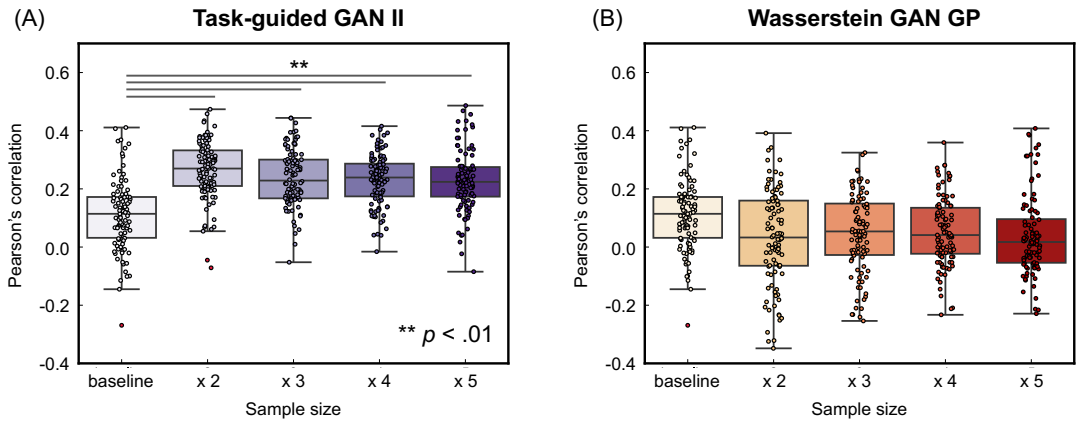

Figure 9. Accuracy of the prediction model, evaluated by Pearson's correlation between observed and predicted scores. (A) indicates data augmentation results using Task-guided GAN II, while (B) presents results from employing Wasserstein GAN GP.

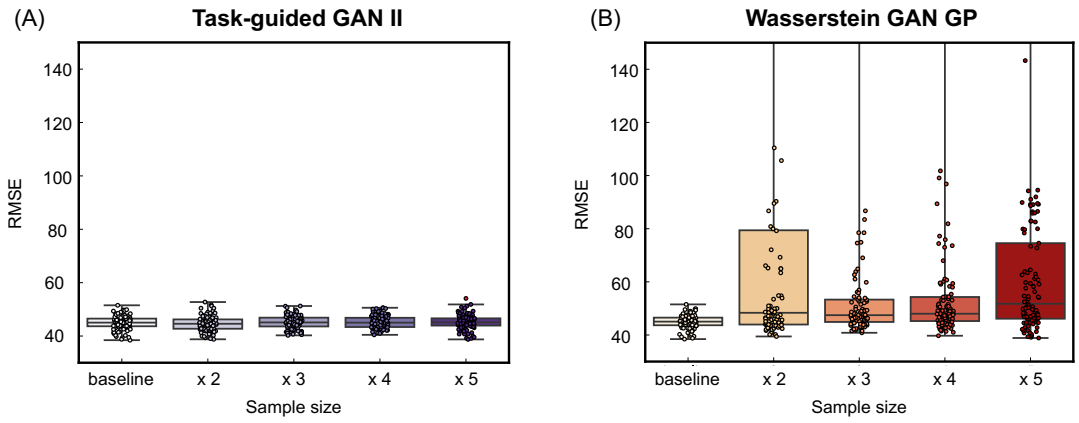

Figure 10. Accuracy of the prediction model, evaluated by the root mean squared error (RMSE) between observed and predicted scores. (A) indicates data augmentation results using Task-guided GAN II, while (B) presents results from employing Wasserstein GAN GP.

Table 2. Summary of the prediction accuracy.

Augmentation method	Training sample size	Pearson's correlation	RMSE
Baseline	75	0.11 ± 0.12	44.8 ± 2.53
TG GAN II (Proposed)	149 (baseline + 100%)	0.26 ± 0.10	44.6 ± 2.77
	223 (+ 200%)	0.24 ± 0.10	45.1 ± 2.48
	297 (+ 300%)	0.23 ± 0.09	45.3 ± 2.34
	371 (+ 400%)	0.23 ± 0.11	45.2 ± 2.75
WGAN GP	149 (+ 100%)	0.030 ± 0.16	470 ± 958
	223 (+ 200%)	0.050 ± 0.14	301 ± 892
	297 (+ 300%)	0.050 ± 0.12	177 ± 685
	371 (+ 400%)	0.040 ± 0.14	214 ± 813

4. Discussion

4.1 The relationship between graph structure and prediction accuracy

In this study, we introduced TG GAN II, a novel method for augmenting brain connectivity data, to enhance the accuracy of predictive models for human cognitive traits and behavior. The analysis of the fluid intelligence prediction model, constructed using the NIMH Healthy Research Volunteer Dataset, showed that TG GAN II could synthesize structural connectivity matrices with features more closely aligned with the actual data than its baseline model, WGAN-GP. Furthermore, data augmentation with TG GAN II resulted in improved prediction accuracy.

Upon examining the average synthesized and acquired matrices, it was observed that WGAN-GP tended to overgenerate inter-hemispherical connections compared to TG GAN II, which inaccurately reflected the fewer connections present in the actual data. Zimmermann et al. have suggested that intra-hemispherical structural connectivity is crucial for capturing the variance in cognitive traits, including fluid intelligence [35]—a notion supported by such connections in the acquired data. Unlike WGAN-GP, TG GAN II effectively minimized inter-hemispherical connections while enhancing intra-hemispherical connections, thereby contributing to the enhanced predictive accuracy of the model.

Furthermore, a graph theoretical analysis of both acquired and synthesized connectivity matrices showed that TG GAN II could produce structural connectivity matrices with topological features more closely matching those of the actual data than WGAN-GP. Garai et al. highlighted that network patterns or topological features in structural connectivity possess predictive power for

human cognitive traits [36]. While Litwińczuk observed no consistent advantage in utilizing graph theory measures over connectivity values for explaining and predicting cognitive functions in healthy and typical domains, they noted instances where nodal graph theory metrics of the structural network outperformed raw connectivity models in predictive ability [37]. These prior studies support that the enhanced prediction accuracy observed with matrices synthesized by TG GAN II can be attributed to their graph theoretical features, which are more akin to those of the acquired matrices.

Consequently, our results suggest that the task-guided branch, implemented as the regression model in TG GAN II, could enhance the synthesis of relevant connectomic features for fluid intelligence prediction while suppressing ineffective features.

4.2 Latent space structure

The previous studies, the foundational work on Task-guided GANs [13] and BrainNetGAN [12], have demonstrated the efficacy of task-guided branches in enhancing task performance—specifically, improving age prediction accuracy from T1-weighted images and classification accuracy for Alzheimer's disease, respectively.

In our study, by visualizing the embedded latent space using PCA, we have shown that the task-guided branch not only improved the task performance but also formed a latent space where latent variables were significantly correlated with fluid intelligence scores (**Figure 8**). This visualization in two-dimensional space provides a novel insight absent in prior research, highlighting that our approach enables a deeper understanding of how input variables are integrated with outcome variables within the latent space of generative models. This finding emphasizes the unique contribution of our method to the field, demonstrating that the task-guided branch is effective in forming a latent space that bridges input variables and outcome variables.

4.3 Limitations and future work

The demonstrated efficacy of our TG GAN II in connectome-based prediction tasks suggested future directions for further investigations despite some limitations that need discussion.

While our evaluation focused on synthesizing structural connectomes, the versatile framework of TG GAN II is not confined to this application alone; it holds the potential for enhancing functional connectome-based predictions through data augmentation. In future work, we will explore whether TG GAN II's augmentation capabilities can indeed improve the accuracy of predictions based on functional connectomes.

Moreover, the initial application of our method was conducted on a relatively small dataset, a common scenario in studies where data augmentation aims to tackle the challenges posed by limited sample sizes. Consequently, the scalability and performance of TG GAN II in contexts involving larger datasets remain to be fully explored. Future studies are encouraged to extend the application of our method to more extensive datasets.

5. Acknowledgements

None to declare.

6. References

[1] O. Sporns, G. Tononi, and R. Kötter, "The Human Connectome: A Structural Description of the Human Brain," *PLOS Computational Biology*, vol. 1, no. 4, p. e42, Sep. 2005, doi: 10.1371/journal.pcbi.0010042.

[2] O. Sporns, "Structure and function of complex brain networks," *Dialogues in Clinical Neuroscience*, vol. 15, no. 3, pp. 247–262, Sep. 2013, doi: 10.31887/DCNS.2013.15.3/osporns.

[3] X. Shen *et al.*, "Using connectome-based predictive modeling to predict individual behavior from brain connectivity," *Nat Protoc*, vol. 12, no. 3, Art. no. 3, Mar. 2017, doi: 10.1038/nprot.2016.178.

[4] D. Scheinost *et al.*, "Ten simple rules for predictive modeling of individual differences in neuroimaging," *NeuroImage*, vol. 193, pp. 35–45, Jun. 2019, doi: 10.1016/j.neuroimage.2019.02.057.

[5] A. W. K. Yeung, S. More, J. Wu, and S. B. Eickhoff, "Reporting details of neuroimaging studies on individual traits prediction: A literature survey," *NeuroImage*, vol. 256, p. 119275, Aug. 2022, doi: 10.1016/j.neuroimage.2022.119275.

[6] R. M. Birn *et al.*, "The effect of scan length on the reliability of resting-state fMRI connectivity estimates," *NeuroImage*, vol. 83, pp. 550–558, Dec. 2013, doi: 10.1016/j.neuroimage.2013.05.099.

[7] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, and K. Ugurbil, "The WU-Minn Human Connectome Project: An overview," *NeuroImage*, vol. 80, pp. 62–79, Oct. 2013, doi: 10.1016/j.neuroimage.2013.05.041.

[8] N. D. Volkow *et al.*, "The conception of the ABCD study: From substance use to a broad NIH collaboration," *Developmental Cognitive Neuroscience*, vol. 32, pp. 4–7, Aug. 2018, doi: 10.1016/j.dcn.2017.10.002.

[9] C. Sudlow *et al.*, "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age," *PLOS Medicine*, vol. 12, no. 3,

p. e1001779, Mar. 2015, doi: 10.1371/journal.pmed.1001779.

[10] A. R. Laird, "Large, open datasets for human connectomics research: Considerations for reproducible and responsible data use," *NeuroImage*, vol. 244, p. 118579, Dec. 2021, doi: 10.1016/j.neuroimage.2021.118579.

[11] L. K. Avberšek and G. Repovš, "Deep learning in neuroimaging data analysis: Applications, challenges, and solutions," *Frontiers in Neuroimaging*, vol. 1, 2022, Accessed: Jul. 18, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnimg.2022.981642

[12] C. Li, Y. Wei, X. Chen, and C.-B. Schonlieb, "BrainNetGAN: Data augmentation of brain connectivity using generative adversarial network for dementia classification." arXiv, Jul. 23, 2021. doi: 10.48550/arXiv.2103.08494.

[13] R. Li, M. Bastiani, D. Auer, C. Wagner, and X. Chen, "Image Augmentation Using a Task Guided Generative Adversarial Network for Age Estimation on Brain MRI." arXiv, Aug. 03, 2021. doi: 10.48550/arXiv.2108.01659.

[14] B. Barile, A. Marzullo, C. Stamile, F. Durand-Dubief, and D. Sappey-Marinier, "Data augmentation using generative adversarial neural networks on brain structural connectivity in multiple sclerosis," *Computer Methods and Programs in Biomedicine*, vol. 206, p. 106113, Jul. 2021, doi: 10.1016/j.cmpb.2021.106113.

[15] M. S. Meor Yahaya and J. Teo, "Data augmentation using generative adversarial networks for images and biomarkers in medicine and neuroscience," *Frontiers in Applied Mathematics and Statistics*, vol. 9, 2023, Accessed: May 21, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fams.2023.1162760

[16] Q. Zuo *et al.*, "Hemisphere-Separated Cross-Connectome Aggregating Learning via VAE-GAN for Brain Structural Connectivity Synthesis," *IEEE Access*, vol. 11, pp. 48493–48505, 2023, doi: 10.1109/ACCESS.2023.3276989.

[17] Y.-F. Tan, C.-M. Ting, F. Noman, R. C.-W. Phan, and H. Ombao, "Graph-Regularized Manifold-Aware Conditional Wasserstein GAN for Brain Functional Connectivity Generation." arXiv, Dec. 10, 2022. doi: 10.48550/arXiv.2212.05316.

[18] I. J. Goodfellow *et al.*, "Generative Adversarial Networks." arXiv, Jun. 10, 2014. doi: 10.48550/arXiv.1406.2661.

[19] A. C. Nugent *et al.*, "The NIMH intramural healthy volunteer dataset: A comprehensive MEG, MRI, and behavioral resource," *Sci Data*, vol. 9, no. 1, Art. no. 1, Aug. 2022, doi: 10.1038/s41597-022-01623-9.

[20] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved Training of Wasserstein GANs," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017. Accessed: Feb. 08, 2024. [Online]. Available:

https://proceedings.neurips.cc/paper_files/paper/2017/hash/892c3b1c6dccd52936e27cbd0ff683d6-Abstract.html

[21] J. Kawahara *et al.*, "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017, doi: 10.1016/j.neuroimage.2016.09.046.

[22] R. C. Gershon, M. V. Wagster, H. C. Hendrie, N. A. Fox, K. F. Cook, and C. J. Nowinski, "NIH Toolbox for Assessment of Neurological and Behavioral Function," *Neurology*, vol. 80, no. 11 Supplement 3, pp. S2–S6, Mar. 2013, doi: 10.1212/WNL.0b013e3182872e5f.

[23] R. K. Heaton *et al.*, "Reliability and Validity of Composite Scores from the NIH Toolbox Cognition Battery in Adults," *Journal of the International Neuropsychological Society*, vol. 20, no. 6, pp. 588–598, Jul. 2014, doi: 10.1017/S1355617714000241.

[24] M. Cieslak *et al.*, "QSIPrep: an integrative platform for preprocessing and reconstructing diffusion MRI data," *Nat Methods*, vol. 18, no. 7, pp. 775–778, Jul. 2021, doi: 10.1038/s41592-021-01185-5.

[25] J. Veraart, D. S. Novikov, D. Christiaens, B. Ades-aron, J. Sijbers, and E. Fieremans, "Denoising of diffusion MRI using random matrix theory," *Neuroimage*, vol. 142, pp. 394–406, Nov. 2016, doi: 10.1016/j.neuroimage.2016.08.016.

[26] J. L. R. Andersson, S. Skare, and J. Ashburner, "How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging," *Neuroimage*, vol. 20, no. 2, pp. 870–888, Oct. 2003, doi: 10.1016/S1053-8119(03)00336-7.

[27] J. L. R. Andersson and S. N. Sotiropoulos, "An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging," *Neuroimage*, vol. 125, pp. 1063–1078, Jan. 2016, doi: 10.1016/j.neuroimage.2015.10.019.

[28] T. Dhollander, R. Mito, and A. Connelly, "Single-Shell 3-Tissue CSD (SS3T-CSD) modelling of developing HCP (dHCP) diffusion MRI data," Jun. 2019.

[29] T. Dhollander, D. Raffelt, and A. Connelly, "Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image," Sep. 2016.

[30] J.-D. Tournier, F. Calamante, and A. Connelly, "Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions".

[31] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly, "Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information," *NeuroImage*, vol. 62, no. 3, pp. 1924–1938, Sep. 2012, doi: 10.1016/j.neuroimage.2012.06.005.

[32] R. E. Smith, J.-D. Tournier, F. Calamante, and A. Connelly, "SIFT2: Enabling dense

quantitative assessment of brain white matter connectivity using streamlines tractography," *Neuroimage*, vol. 119, pp. 338–351, Oct. 2015, doi: 10.1016/j.neuroimage.2015.06.092.

[33] Z. Cui, M. Su, L. Li, H. Shu, and G. Gong, "Individualized Prediction of Reading Comprehension Ability Using Gray Matter Volume," *Cereb Cortex*, vol. 28, no. 5, pp. 1656–1672, May 2018, doi: 10.1093/cercor/bhx061.

[34] A. E. Hoerl and R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, Feb. 1970, doi: 10.1080/00401706.1970.10488634.

[35] J. Zimmermann, J. D. Griffiths, and A. R. McIntosh, "Unique Mapping of Structural and Functional Connectivity on Cognition," *J. Neurosci.*, vol. 38, no. 45, pp. 9658–9667, Nov. 2018, doi: 10.1523/JNEUROSCI.0900-18.2018.

[36] S. Garai, F. Xu, D. A. Duong-Tran, Y. Zhao, and L. Shen, "Mining Correlation between Fluid Intelligence and Whole-brain Large Scale Structural Connectivity," *AMIA Jt Summits Transl Sci Proc*, vol. 2023, pp. 225–233, 2023.

[37] M. C. Litwińczuk, N. Muhlert, N. Trujillo-Barreto, and A. Woollams, "Using graph theory as a common language to combine neural structure and function in models of healthy cognitive performance," *Hum Brain Mapp*, vol. 44, no. 8, pp. 3007–3022, Jun. 2023, doi: 10.1002/hbm.26258.

7. Supporting Information

The source code for the methods used in this paper will be made available at https://github.com/MIS-Lab-Doshisha/tg-gan2.