



## OPEN ACCESS

## EDITED BY

Zhijiang Wan,  
Nanchang University, China

## REVIEWED BY

Changhong Jing,  
Hong Kong Polytechnic University,  
Hong Kong SAR, China  
Yulong Liu,  
Changchun University of Science and  
Technology, China  
Zhang Bin,  
Kanagawa University, Japan

## \*CORRESPONDENCE

Jia Guo  
✉ guojia@hbue.edu.cn

RECEIVED 16 February 2024

ACCEPTED 02 April 2024

PUBLISHED 17 April 2024

## CITATION

Zuo Q, Li R, Shi B, Hong J, Zhu Y, Chen X,  
Wu Y and Guo J (2024) U-shaped  
convolutional transformer GAN with  
multi-resolution consistency loss for restoring  
brain functional time-series and dementia  
diagnosis.

*Front. Comput. Neurosci.* 18:1387004.  
doi: 10.3389/fncom.2024.1387004

## COPYRIGHT

© 2024 Zuo, Li, Shi, Hong, Zhu, Chen, Wu and Guo. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# U-shaped convolutional transformer GAN with multi-resolution consistency loss for restoring brain functional time-series and dementia diagnosis

Qiankun Zuo<sup>1,2,3</sup>, Ruiheng Li<sup>1,2</sup>, Binghua Shi<sup>1,2</sup>, Jin Hong<sup>4</sup>, Yanfei Zhu<sup>5</sup>, Xuhang Chen<sup>6</sup>, Yixian Wu<sup>7</sup> and Jia Guo<sup>1,2,3\*</sup>

<sup>1</sup>Hubei Key Laboratory of Digital Finance Innovation, Hubei University of Economics, Wuhan, Hubei, China, <sup>2</sup>School of Information Engineering, Hubei University of Economics, Wuhan, Hubei, China, <sup>3</sup>Hubei Internet Finance Information Engineering Technology Research Center, Hubei University of Economics, Wuhan, Hubei, China, <sup>4</sup>Medical Research Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China, <sup>5</sup>School of Foreign Languages, Sun Yat-sen University, Guangzhou, China, <sup>6</sup>Faculty of Science and Technology, University of Macau, Taipa, Macao SAR, China, <sup>7</sup>School of Mechanical Engineering, Beijing Institute of Petrochemical Technology, Beijing, China

**Introduction:** The blood oxygen level-dependent (BOLD) signal derived from functional neuroimaging is commonly used in brain network analysis and dementia diagnosis. Missing the BOLD signal may lead to bad performance and misinterpretation of findings when analyzing neurological disease. Few studies have focused on the restoration of brain functional time-series data.

**Methods:** In this paper, a novel *U*-shaped convolutional transformer GAN (UCT-GAN) model is proposed to restore the missing brain functional time-series data. The proposed model leverages the power of generative adversarial networks (GANs) while incorporating a *U*-shaped architecture to effectively capture hierarchical features in the restoration process. Besides, the multi-level temporal-correlated attention and the convolutional sampling in the transformer-based generator are devised to capture the global and local temporal features for the missing time series and associate their long-range relationship with the other brain regions. Furthermore, by introducing multi-resolution consistency loss, the proposed model can promote the learning of diverse temporal patterns and maintain consistency across different temporal resolutions, thus effectively restoring complex brain functional dynamics.

**Results:** We theoretically tested our model on the public Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, and our experiments demonstrate that the proposed model outperforms existing methods in terms of both quantitative metrics and qualitative assessments. The model's ability to preserve the underlying topological structure of the brain functional networks during restoration is a particularly notable achievement.

**Conclusion:** Overall, the proposed model offers a promising solution for restoring brain functional time-series and contributes to the advancement of neuroscience research by providing enhanced tools for disease analysis and interpretation.

## KEYWORDS

hierarchical topological transformer, multi-level temporal-correlated attention, central connectivity perception, time-series restoration, multi-head attention, brain neurological disease

# 1 Introduction

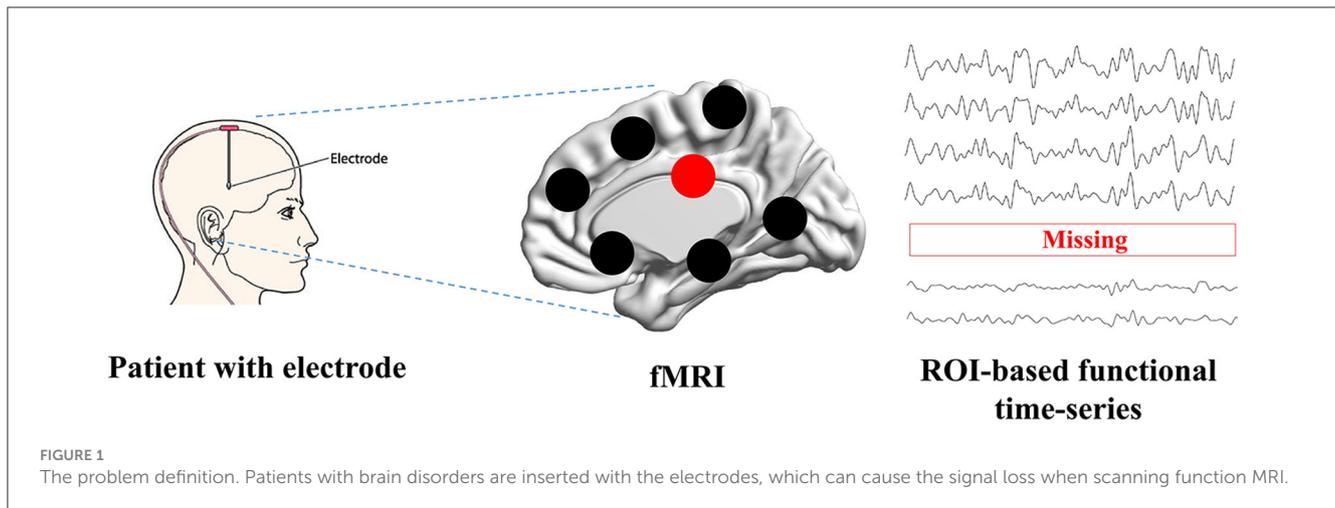
The blood oxygen level-dependent (BOLD) signal derived from functional neuroimaging is commonly used in brain disorder analysis. As a common brain disorder, Alzheimer's disease (AD) is a progressive neurodegenerative condition characterized by cognitive decline, memory impairment, and changes in behavior (Knopman et al., 2021). The exact cause of AD is not fully understood, but it involves the accumulation of abnormal proteins in the brain, particularly beta-amyloid plaques and tau tangles. To treat brain disorders (e.g., AD, Parkinson's disease), deep brain stimulation (DBS) is a possible way to solve the problem of movement disorders (Limousin and Foltynie, 2019; Ríos et al., 2022). It is a neurosurgical procedure that involves the implantation of electrodes into specific regions of the brain to modulate its electrical activity (Leoutsakos et al., 2018). DBS has been investigated as a potential treatment for AD because it offers a way to modulate brain activity in specific areas that are associated with memory and cognition. The electrodes play a critical role in the DBS procedure. These thin, insulated wires are surgically implanted into the brain region of interest. Once in place, they are connected to an implanted pulse generator, which delivers electrical impulses to the brain. These electrical pulses can help regulate the abnormal brain activity associated with certain neurological disorders, potentially improving symptoms (Medtronic, 2020; Alajangi et al., 2022). When it comes to AD, researchers can explore the use of DBS to target brain regions such as the fornix, which is involved in memory and learning. By stimulating these areas, the DBS is able to help improve cognitive function in cognitive patients (Neumann et al., 2023; Siddiqi et al., 2023; Vogel et al., 2023).

Functional Magnetic Resonance Imaging (fMRI) has revolutionized the field of neuroscience, particularly in the study of brain diseases such as Alzheimer's disease (Forouzannezhad et al., 2019; Yin et al., 2022; Zuo et al., 2023). fMRI is a non-invasive neuroimaging technique that provides valuable insights into the functioning of the human brain by measuring blood oxygenation level-dependent (BOLD) signals. fMRI has been confirmed as a reliable instrument to investigate the brain's functional aspects and explore the brain's mechanisms, enabling early detection, understanding cognitive disease progression, and assessing the impact of interventions (Warren and Moustafa, 2023; Yen et al., 2023). Many studies (Wang et al., 2018; Ibrahim et al., 2021; Sendi et al., 2023) have constructed connectivity-based features and analyzed cognitive disease from fMRI. The constructed features in non-Euclidean space can establish relations between distant brain regions, which is superior than the image-based features in Euclidean space (Chen et al., 2023; Wan et al., 2023d). When evaluating the treatment's performance, fMRI allows researchers to monitor and assess changes in brain activity before and after DBS treatment (Boutet et al., 2021; Soleimani et al., 2023). However, fMRI can sometimes be impacted by the presence of implanted electrodes. The metallic components of these electrodes can create artifacts in the MRI images, which may lead to signal loss or distortion in the region of interest (ROI) (Nimbalkar et al., 2019; Luo et al., 2022; Wang X. et al., 2022). These artifacts include signal intensity changes and temporal and spatial variability. Presently, there are no post-processing MRI techniques available

to effectively mitigate these artifacts. Therefore, identifying these specific characteristics is essential for restoring neural activity from artifacts and ensuring the accuracy and validity of fMRI findings in clinical and research settings. Researchers need to address issues related to side effects and fMRI signal loss to further our understanding of the technique's effectiveness in treating this complex and devastating neurodegenerative disease. The possible way to solve this issue is to construct a deep learning model to recover missing signals, as it has achieved complex tasks in medical image analysis (Wang S. et al., 2022; You et al., 2022; Hu et al., 2023; Wan et al., 2023e). As shown in Figure 1, when patients are treated by electrode stimulation, the brain fMRI suffers from signal loss in the stimulated brain regions.

Generative adversarial networks (GANs) have gained prominence in the fields of medical image analysis (Hong et al., 2022) and functional time series reconstruction as a powerful tool for generating synthetic data that closely resemble real-world time series data (Luo et al., 2018, 2019). In addition, Transformer's self-attention mechanism has been successfully applied in medical data analysis (Li et al., 2023; Wan et al., 2023a,b,c). The parallel processing capability and adaptability to various data types make it a versatile tool for time series generation (Tang and Matteson, 2021; Zerveas et al., 2021). Therefore, combining the GAN and transformer can enable the reconstruction of missing time series. Transformer GANs (Generative Adversarial Networks with a Transformer architecture) have been applied to time series reconstruction, offering innovative solutions to various data reconstruction tasks (Wu et al., 2020; Li et al., 2021; Li X. et al., 2022). In many domains, time series data may have missing or incomplete observations. Transformer GANs can be trained to impute the missing data by learning the underlying patterns and relationships in the time series. The generator network creates synthetic data points to fill in the gaps, while the discriminator evaluates the realism of the imputed values. Transformer GANs offer advantages for time series reconstruction due to their ability to capture long-range dependencies and complex patterns and their interpretability through attention mechanisms (Jiang et al., 2021; Zhao et al., 2021). However, these models cannot capture topological relationships at different temporal resolutions, which may degrade the reconstruction performance and the ability to analyze the brain network.

The DBS has emerged as a promising therapeutic approach for Alzheimer's disease (AD), offering potential benefits in alleviating symptoms and modifying disease progression. Although still in the investigational stage, DBS for AD holds promise as a novel intervention aimed at improving cognitive function and quality of life for individuals affected by this devastating neurodegenerative disorder. Among the publicly available datasets, the representative dataset containing brain imaging data for all stages of AD is the Alzheimer's Disease Neuroimaging Initiative (ADNI). Currently, there are no patients implanted with intracortical electrodes for DBS treatment. Our study is the first to theoretically remove some ROI's signals and then utilize our model to recover the removed signals. In this study, we propose a novel U-shaped convolutional transformer GAN (UCT-GAN) model to restore the missing brain functional time-series. First, the fMRI is preprocessed to obtain the ROI-based functional time series. Then, we exclude some ROI's



time-series and treat them as a missing signal. The rest of the ROI-based time-series are sent to the U-shaped topological transformer generator to recover the missing time-series by capturing complex temporal patterns and relationships. Next, the recovered time-series from the generator is sent to the discriminator, consisting of multi-head attention and central connectivity perception, to evaluate the realism of the generated data compared with real fMRI data at different scales. Both spatiotemporal and connectivity features are utilized to constrain the generated missing signals. Finally, we implement a loss function that enforces consistency across different temporal resolutions. This loss encourages the generator to capture diverse temporal scales in the data. When the training reaches Nash equilibrium, the model can recover the missing time-series signal. The main works of this study are as follows:

- The proposed model leverages the power of generative adversarial networks (GANs) while incorporating a U-shaped architecture to effectively capture both global and local features in the restoration process. The temporal characteristics of missing time series can be highly recovered for downstream brain network analysis.
- The multi-level temporal-correlated attention in the transformer-based generator is devised to model the temporal relationship between the missing ROI and other normal ROIs. The topological properties of the missing time-series can be well explored.
- By introducing multi-resolution consistency loss, the proposed model can promote the learning of diverse temporal patterns and maintain consistency across different temporal resolutions, thus effectively restoring complex brain functional dynamics.

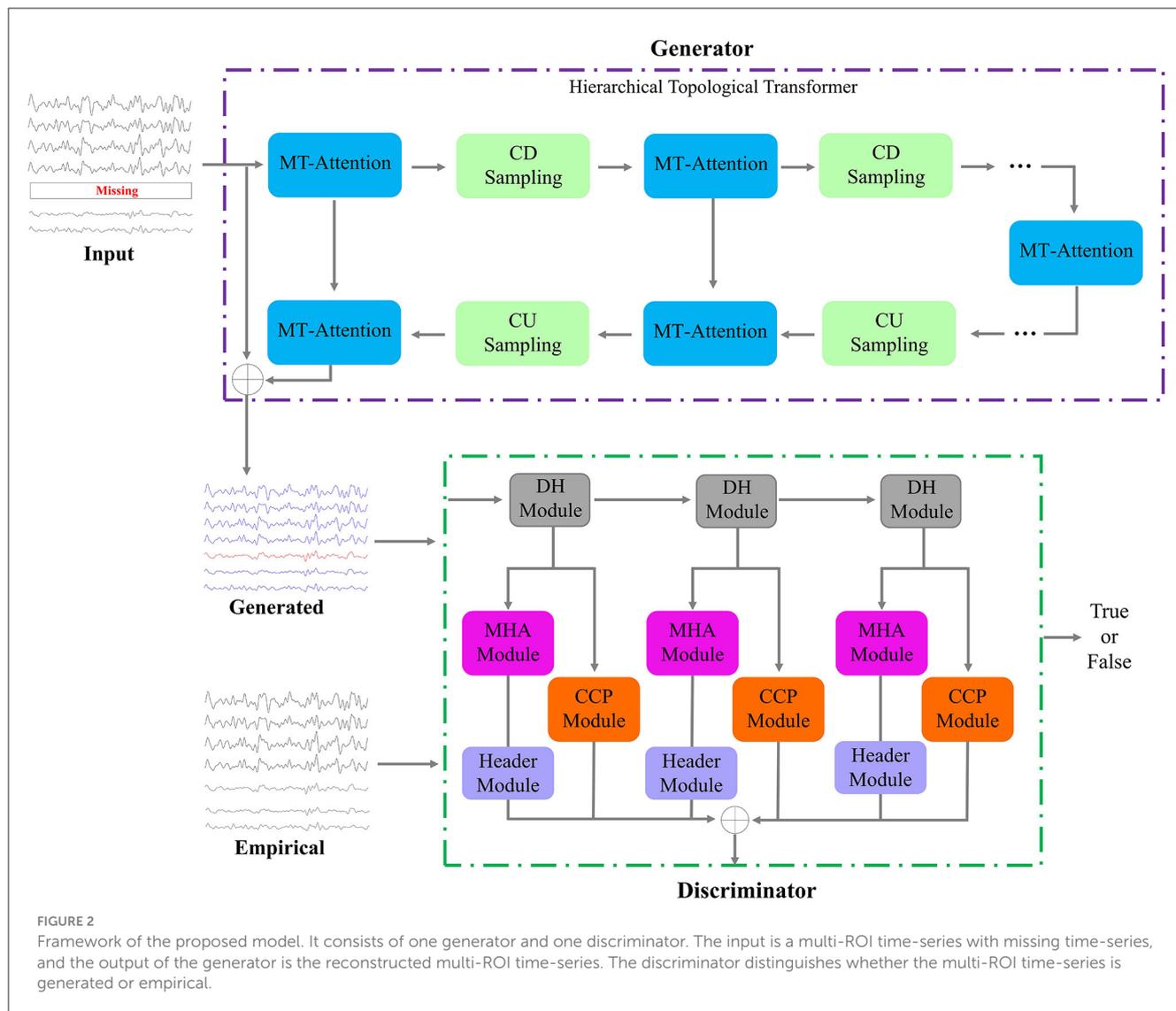
## 2 Related work

Reconstructing time-series data using generative adversarial networks (GANs) is a burgeoning field of research with several

related studies. GANs offer the potential to generate synthetic time-series by capturing statistical and temporal characteristics. The main advantage of GANs is that they can be used to augment existing time-series datasets by generating additional synthetic data. Increasing the data size is particularly valuable when training machine learning models in medical image analysis. Considering the architecture of the generator, we divide the GAN-based models into two groups: recurrent neural network (RNN)-based approaches and transformer-based approaches.

Mogren (2016) combined RNN and GAN to synthesize more realistic continuous sequential from random noise. Similarly, Esteban et al. (2017) embedded the RNN into both the generator and discriminator to synthesize realistic medical time-series signals by introducing label constraints. Meanwhile, Donahue et al. (2018) introduced the WaveGAN model to generate time-series waveforms by applying one-dimensional convolution kernels. To preserve temporal dynamics, Yoon et al. (2019) proposed the TimeGAN framework to project temporal features onto embedded space through supervisory and antagonistic learning and generate realistic time-series signals to preserve temporal correlation between different variables. In addition, Ni et al. (2020) proposed the SigCWGAN to capture the temporal dependencies inherent in joint probability distributions within time-series signals. Nonetheless, RNN-based approaches are challenging for generating long synthetic sequences. This stems from the sequential processing of time steps in time-series data, where recent time steps exert a stronger influence on the generation of subsequent time steps. Therefore, RNNs fail to establish relationships between distant time steps in lengthy sequences.

Reconstructing time-series signals using transformer GANs, which combine the transformer architecture with GAN, has the potential to capture complex temporal relations of long sequential time-series. Li X. et al. (2022) successfully designed the generator and discriminator with transformer to synthesize long-sequence time-series signals. Srinivasan and Knottenbelt (2022) proposed the TST-GAN to solve the problem of errors accumulating over time when synthesizing temporal features. This model can accurately simulate the joint distribution of the entire time-series, and the generated time-series can be used instead



of real data. To synthesize multivariate time-series within the overall distribution, Madane et al. (2022) introduced different conditions into the transformer-based GAN to approximate the joint distribution of multiple time-series. Li Y. et al. (2022) pointed out the weakness of short-term dependencies in transformers and proposed adversarial convolutional transformers (ACTs) to pay attention to local information of time series. They greatly improved the forecasting accuracy of time-series datasets. Xia et al. (2023) also combined the convolutional networks and transformer in the adversarial training to preserve both global and local temporal features in the time-series generation. However, these models fail to capture the hierarchical temporal features and ignore the temporal characteristics of different frequencies, which may hinder synthesis performance during the time-series generation.

Considering the shortcomings of related methods, we incorporated transformer-based networks into a U-shaped architecture to model temporal relationships on both global and local scales. In addition, the restoration process of time-series is to learn complex distribution, where the generative adversarial networks (GANs) show great ability in learning the underlying

patterns and relationships in the time series. Therefore, we try to combine the U-shaped convolutional transformer and GANs to restore the missing brain functional time series for dementia diagnosis.

### 3 Materials and methods

#### 3.1 Data description

The Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset<sup>1</sup> is a comprehensive and widely used resource for studying Alzheimer’s disease and related neurological conditions. It includes a variety of data types, including structural and functional MRI (fMRI) data. In this study, we successfully downloaded about 311 subjects from the ADNI website. The patients scanned with fMRI are distributed among the normal controls (NC), early mild cognitive impairment (EMCI), and late mild cognitive impairment

1 <http://adni.loni.usc.edu/>

(LMCI). The numbers for the three categories are 105, 110, and 96, respectively. The time of repetition (TR) is 3.0 s. The scanning time for each subject is  $\sim 10$  min.

Preprocessing fMRI data typically involve several steps to ensure data quality and prepare it for analysis. We use the routine GREYNA (Wang et al., 2015) software to preprocess the fMRI to construct multi-ROI time-series. The general preprocessing steps (Zuo et al., 2022, 2024) are as follows: convert the DICOM files into NIfTI format for easier handling and compatibility, remove the first 10 volumes, correct for differences in acquisition times between slices to ensure temporal alignment, correct for head motion during scanning, register the fMRI data to a standard anatomical template (e.g., MNI152) to ensure spatial consistency across subjects, apply spatial smoothing to the data to improve the signal-to-noise ratio and compensate for small anatomical differences between subjects, apply temporal filters to remove low-frequency drifts (e.g., high-pass filtering) and to attenuate high-frequency noise (e.g., Gaussian or bandpass filtering), register the fMRI data to the structural MRI data for each subject, and wrap the fMRI volumes into the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) to obtain the time-series of 90 ROIs. At last, the output is the multi-ROI time-series  $S_e$  with the size  $N \times 187$ . In the following experiments, we remove one or more ROI time-series from  $S_e$  and recover the removed time-series through the proposed model.

## 3.2 Architecture

The main framework of this study is shown in Figure 2. The proposed UCT-GAN model consists of a hierarchical topological transformer generator and a multi-resolution relational discriminator. Given fMRI with missing signals on some brain area, after preprocessing, we can obtain the input data of the proposed model. We denoted it as the incomplete multi-ROI time-series signal  $S_m \in \mathbb{R}^{N \times T}$ , where  $N$  is the ROI number and  $T$  is the scanning functional signal length. The transformer-based generator aims to extract hierarchical features to recover the missing ROI-based signal. The multi-resolution discriminator is utilized to constrain the generated time-series ( $S_g \in \mathbb{R}^{N \times T}$ ) as close as the empirical time-series ( $S_e \in \mathbb{R}^{N \times T}$ ). To ensure the generation's good performance, we design three loss functions to optimize the model's parameters, including the generative loss, the discriminative loss, and the multi-resolution consistency loss.

### 3.2.1 Hierarchical topological transformer generator

The generator is a neural network architecture that combines the principles of hierarchical attention mechanisms from transformers with one-dimensional convolutional layers. This architecture can capture both global and local temporal information at different scales and is often used for processing sequences or time series data efficiently and effectively. In the generator, we designed multiple layers of multi-level temporal-correlated attention (MT-Attention) and convolutional sampling to explore hierarchical temporal features. The output is the generated multi-ROI time-series  $S_g$ . Setting  $L$  convolutional down sampling (CDS) layers, there are also  $L$  layers of convolutional

up sampling (CUS) and  $2L + 1$  layers of MT-Attention. The computation process can be expressed by the following formula:

$$S_g = G(S_m) = f(MTA(S_m), CDS(S_m), CUS(S_m))$$

Here, the symbol  $f$  means the calculation processes in the generator.

Multi-level temporal-correlated attention (MTA) is an attention mechanism designed to capture dependencies and patterns at multiple levels of the characteristics of temporal sequences. This attention mechanism is especially useful for modeling time-series relationships between different ROIs. As shown in Figure 3, assuming the input of MT-Attention is the multi-ROI temporal feature  $F^i$  with the size  $2C \times N \times T/2$ . We first split it into  $2C$  slices, where each slice is sent to the level-topological computing (LTC) network to learn temporal dependencies between ROIs. For instance, some slices may represent lower levels and can be used to capture short-term dependencies within the sequence, while other slices may indicate higher levels and can be used to capture long-term dependencies. This multi-level structure allows MTA to consider different levels of temporal dynamics when reconstructing missing signals. Each slice is passed through the norm layer, linear projection (LP), splitting, attention map (AM), merge, dropout, norm, LP, and dropout. The output is the updated multi-ROI temporal feature  $F^{i+1}$  with the size of  $F^i$ . The whole computation can be determined by

$$F^{i+1} = MTA(F^i) = LTC(F_1^i) || LTC(F_2^i) || \dots || LTC(F_{2C}^i)$$

where  $F^i$  means the input of the  $i$ -th module in the generator.  $||$  indicates the concatenation operation.  $F^{i+1}$  is the output of the  $i$ -th module in the generator. After separating the  $2C$  channels of  $F^i$ , each channel component is represented as  $F_j^i$ . Here,  $j$  is in the range of  $1 - 2C$ . These components are computed by the attention network and feedforward transform (FFT). The formula can be expressed as follows:

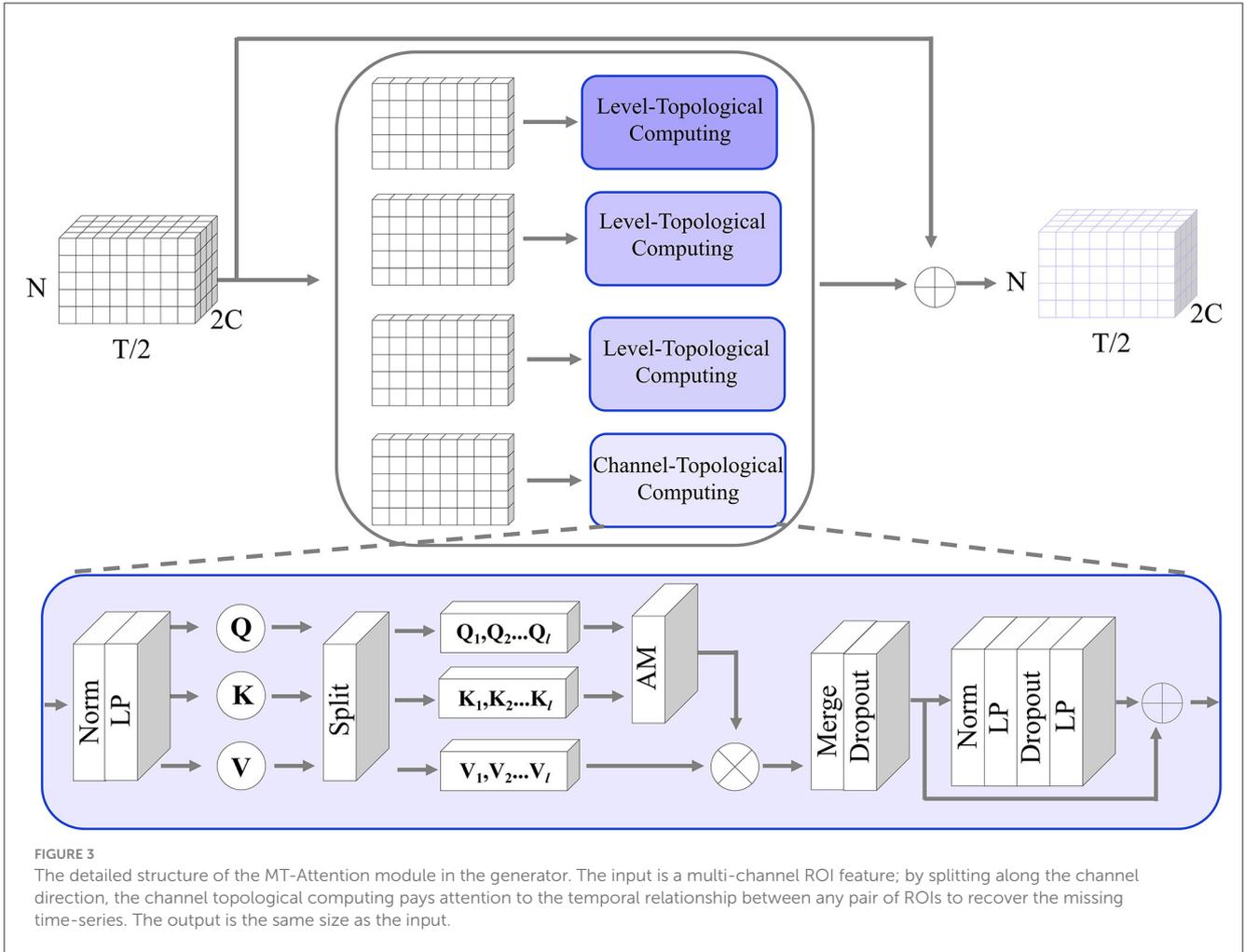
$$LTC(F_j^i) = Attention(F_j^i) + FFT(F_j^i), \quad j = \{1, 2, \dots, 2C\}$$

In the attention network, the norm is applied to the temporal feature to stabilize the training process. The LP layer is used to learn temporal attention matrices  $Q, K, V$ . We applied  $l$  heads to the attention matrices and computed an attention map for each head. The attentioned heads are then merged by a LP layer and a dropout layer. Here is the computing formula:

$$Attention(F_j^i) = Dropout(LP(Att(Q_1, K_1, V_1) || Att(Q_2, K_2, V_2) || \dots || Att(Q_l, K_l, V_l)))$$

$$Att(Q_l, K_l, V_l) = \text{softmax} \left( \frac{Q_l K_l^T}{\sqrt{T/(2l)}} \right) V_l$$

$$Q = (Q_1 || Q_2 || \dots || Q_l)$$



$$K = (K_1 || K_2 || \dots || K_l)$$

$$V = (V_1 || V_2 || \dots || V_l)$$

$$Q = LP(Norm(F_j^i)), K = LP(Norm(F_j^i)), V = LP(Norm(F_j^i))$$

In the feedforward transform network, it consists of Norm, LP, and dropout layers. They are used to provide a non-linear transformation to the intermediate representation produced by the self-attention mechanism. The LP layer projects the input temporal features from a lower-dimensional space to a higher-dimensional space, introducing some non-linearity in the process. The dropout layer aims to make the mapping weights more sparse for robust learning. The second linear layer then projects the result back to the original dimension. The computation formula is defined as follows:

$$FFT(F_j^i) = Attention(F_j^i) + LP(Dropout(LP(Norm(Attention(F_j^i))))))$$

Convolutional sampling is utilized to reduce or increase temporal dimensions, including convolutional down (CD) sampling and convolutional up (CU) sampling. In the generator, the CD sampling halves the dimension along the temporal direction while doubles the channels. The CU sampling doubles

the dimension along the temporal direction while halving the channels.

For the CD sampling, one-dimensional convolutional kernels are applied to the multi-ROI temporal features. 1D convolution is used to capture local patterns or features within the time series data. By moving the filter across the sequence, it can detect changes, peaks, valleys, and other patterns within the temporal features. To reduce the dimensions, we set the stride step 2 and the doubled channels. For example, the input incomplete multi-ROI time-series signal  $S_m$  is sent to the generator. We treated it as the multi-ROI temporal features  $F_1$  with the size  $C \times N \times T$  ( $C = 1$ ). After passing the MT-Attention module, the output  $F_1$  has the same size. Then, going through the CD sampling, the output  $F_2$  changed the size to  $2C \times N \times T/2$ . For the CU sampling, we adopted transposed convolution to increase temporal dimensions.

### 3.2.2 Multi-resolution discriminator

The multi-resolution discriminator aims to distinguish between generated multi-ROI time-series and empirical multi-ROI time-series. The discriminator's feedback can help optimize the generator. When the discriminator can easily distinguish between a true and false sample, it provides feedback to the generator

to improve its generation capabilities. The generator then adjusts its parameters to produce a sample that is more similar to the true one.

The structure of the discriminator consists of three dimension halving (DH) modules, three multi-head attention (MHA) modules, three header modules, and three central connectivity perception (CCP) modules. The generated/empirical multi-ROI time-series are first passed through the DH modules. For each DH module, the time-series dimension is halved but the channels are unchanged, which is different from the CD sampling in the generator. Through the three DH modules, the input multi-ROI time-series (e.g.,  $\mathbf{S}_g$ ) are resampled into three samples:

$$\mathbf{R}_1 = DH(\mathbf{S}_g)$$

$$\mathbf{R}_2 = DH(\mathbf{R}_1)$$

$$\mathbf{R}_3 = DH(\mathbf{R}_2)$$

where  $\mathbf{R}_i, i = \{1, 2, 3\}$  represents the high frequency signals (with the size  $N \times T/2$ ), middle frequency signals (with the size  $N \times T/4$ ), and low frequency signals (with the size  $N \times T/8$ ), respectively.

The resampled sample is sent to two branches: MHA and CCP. The former is used to capture the temporal dynamics and learn to measure temporal consistency; the latter is used to compute the consistency of missing-signal ROI-related connections. Both of them can contribute to the consistency measurement between the generated and empirical samples. Combining them can make the generated samples more realistic than the empirical samples. The MHA is the same as the transformer network with  $l$  heads. The header module transmits the attentioned temporal features into one scalar (1 means true, 0 means false). The CCP module first transforms the resampled sample into a connectivity matrix and then selects the missing-signal ROI-related connections. A one-layer LP is used to transmit the connectivity features into one scalar. The detailed computing steps are defined as follows:

$$o_1 = \text{Header}(\text{MHA}(\mathbf{R}_1)) + \text{CCP}(\mathbf{R}_1)$$

$$o_2 = \text{Header}(\text{MHA}(\mathbf{R}_2)) + \text{CCP}(\mathbf{R}_2)$$

$$o_3 = \text{Header}(\text{MHA}(\mathbf{R}_3)) + \text{CCP}(\mathbf{R}_3)$$

$$o_m = \frac{o_1 + o_2 + o_3}{6}$$

where  $o_i, i = \{1, 2, 3\}$  is the scalar.  $o_m$  is the final output score. After the model converges, the value of  $o_m$  approaches 0.5.

### 3.3 Hybrid loss functions

The adversarial loss, also known as the discriminator loss or the GAN loss, is a key component of a generative adversarial network (GAN). It quantifies how well the discriminator can distinguish

between real and generated data. The goal of the generator is to minimize this loss, while the discriminator aims to maximize it. The adversarial loss is typically defined as a binary cross-entropy loss. The formulation of the adversarial loss is as follows:

$$\mathcal{L}_D = \|\mathbf{D}(\mathbf{S}_e) - 1\|^2 + \|\mathbf{D}(\mathbf{G}(\mathbf{S}_m))\|^2$$

$$\mathcal{L}_G = \|\mathbf{D}(\mathbf{G}(\mathbf{S}_m)) - 1\|^2$$

In addition, to keep the generated time-series as precisely similar as the empirical time-series, we introduced the multi-resolution consistency loss  $\mathcal{L}_{MRC}$ . It contains the reconstruction loss, the cross-correlation loss, and the topological loss at different temporal resolutions. The reconstruction error is a metric used to quantify the local dissimilarity between empirical time-series and generated time-series, and the cross-correlation loss can measure the overall temporal patterns between generated and empirical time-series. The topological loss computes the connectivity difference between the generated and empirical time-series. Here, we use the temporal mean absolute error (TMAE) and mean cross-correlation coefficient (MCC) to compute the loss functions. The multi-resolution consistency loss is defined as follows:

$$\begin{aligned} \mathcal{L}_{MRC} = & \sum_{k=0}^3 TMAE(DH^k(\mathbf{S}_g), DH^k(\mathbf{S}_e)) \\ & + \sum_{k=0}^3 MCC(DH^k(\mathbf{S}_g), DH^k(\mathbf{S}_e)) \\ & + \sum_{k=0}^3 TMAE(PCC(\mathbf{S}_g), PCC(\mathbf{S}_e)) \end{aligned}$$

$$TMAE(\mathbf{S}_g, \mathbf{S}_e) = \frac{1}{NT} \sum_{i=1}^N \sum_{j=1}^T |\mathbf{S}_{g,ij} - \mathbf{S}_{e,ij}|$$

$$MCC(\mathbf{S}_g, \mathbf{S}_e) = \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=1}^T (\mathbf{S}_{g,ij} - \mathbf{S}_{\bar{g},i})(\mathbf{S}_{e,ij} - \mathbf{S}_{\bar{e},i})}{\sqrt{\sum_{j=1}^T (\mathbf{S}_{g,ij} - \mathbf{S}_{\bar{g},i})^2} \sqrt{\sum_{j=1}^T (\mathbf{S}_{e,ij} - \mathbf{S}_{\bar{e},i})^2}}$$

where the  $DH^k$  means stacking  $k$  DH layers.  $\mathbf{S}_{\bar{g},i}$  means averaging the  $i$ -th ROI time-series for  $\mathbf{S}_g$ . In summary, the total loss of the proposed UCT-GAN can be optimized by the following loss functions:

$$\mathcal{L}_{all} = \mathcal{L}_G + \mathcal{L}_D + \alpha \mathcal{L}_{MRC}$$

The detailed training pseudo-code is shown in [Algorithm 1](#).

## 4 Experiments and results

### 4.1 Model settings and evaluating metrics

The UCT-GAN model is trained on Windows 11 using the pytorch deep learning framework to reconstruct the incomplete

```

Input:  $S_m$ : incomplete multi-ROI time series;
          $O$ : the number of iteration steps for
         updating the generator;
          $T$ : the number of iteration steps for updating
         the discriminator;
          $L$ : the number of convolutional down sampling
         (CDS) layers.
Output:  $S_g$ : the reconstructed multi-ROI time-series
1: for  $i = 1, 2, \dots, O$  do
2:   for  $j = 1, 2, \dots, T$  do
3:     denote the input  $S_m$  as  $F^0$ .
4:     for  $k = 1, 2, \dots, L$  do
5:       compute the attentioned feature:  $F^k$ 
6:        $MTA(F^{k-1})$ 
7:       compute the down sampled feature:  $F^k$ 
8:        $CDS(F^k)$ 
9:     end for
10:    compute the attentioned feature:  $\hat{F}^k$ 
11:     $MTA(\hat{F}^k)$ 
12:    compute the up sampled feature:  $\hat{F}^{k+1}$ 
13:     $CUS(\hat{F}^{k+1})$ 
14:    compute the attentioned feature:  $\hat{F}^k$ 
15:     $MTA(\hat{F}^k) + F^k$ 
16:  end for
17:  compute the reconstructed multi-ROI
18:  time-series:  $S_g = F^0 + \hat{F}^k$ 
19:  compute the loss function:  $\mathcal{L}_D = \|\mathbf{D}(S_e) - 1\|^2 +$ 
20:   $\|\mathbf{D}(S_g)\|^2$ 
21:  update the discriminator  $\mathbf{D}$  by propagating
22:  the gradient  $-\nabla \mathcal{L}_D^j$ 
23: end for
24: compute the combined loss function  $\mathcal{L}$ 
25:  $\mathcal{L}_G + \alpha \mathcal{L}_{MRC}$ 
26: update the generator by back-propagating the
27: gradient  $-\nabla \mathcal{L}^i$ 
28: end for

```

Algorithm 1. Optimizing the UCT-GAN model.

multi-ROI time-series. The parameter  $L$  is studied in the range of 1–10 to find the optimal value. In addition, the hyperparameter  $\alpha$  in the loss functions is investigated to determine the best weighting of the multi-resolution consistency loss. During the training, we first train the discriminator and then train the generator. The learning rate for the generator and the discriminator is set at  $3.e - 4$  and  $1.e - 4$ , respectively. The Adam was used to train the models with a batch size of 16. Overall, 10-fold cross verification is adopted to evaluate our model's reconstruction performance.

Measuring the similarity between generated and empirical time-series data is a crucial step in evaluating the performance of the proposed model. Three metrics can be used for this purpose, including mean absolute error (MAE), root mean square error (RMSE), coefficient of determination of the prediction (R2) (Ma et al., 2021), and dynamic time warping (DTW) (Philips et al., 2022). MAE measures the average absolute difference between the values of the generated and empirical time-series. It is calculated by

taking the absolute difference between each corresponding pair of points in the generated and empirical time-series, summing these differences, and then dividing by the total number of data points. MAE is sensitive to outliers and provides a straightforward measure of the magnitude of errors. The formula is defined as follows:

$$MAE = \frac{1}{N_m} \sum_{i=1}^{N_m} \sum_{j=1}^T |s_{\text{generated}}(ij) - s_{\text{empirical}}(ij)|$$

RMSE calculates the square root of the average of the squared differences between the generated and empirical time-series. It provides a measure of the magnitude of errors and gives higher weight to larger errors because of the squaring operation. The formula is defined as follows:

$$RMSE = \frac{1}{N_m} \sum_{i=1}^{N_m} \sqrt{\frac{1}{T} \sum_{j=1}^T (s_{\text{generated}}(ij) - s_{\text{empirical}}(ij))^2}$$

where the  $N_m$  means the number of missing time-series ROIs.  $s_{\text{generated}}(ij)$  is the  $i$ -th and  $j$ -th element in the  $S_g$ , and the  $s_{\text{empirical}}(ij)$  is the  $i$ -th and  $j$ -th element in the  $S_e$ . The R2 measures how the reconstructed time-series linearly regresses the empirical time-series, and large value indicates good reconstruction performance. The DTW measures the distance between reconstructed and empirical time-series, where small values indicate good reconstruction performance.

## 4.2 Parameter analysis

The generator is important for reconstructing missing ROI time-series. To explore the optimal MT-Attention layer number, we studied 10 values of  $L$  to determine the best value. We treated the left amygdala as the missing time-series ROI. The MAE is calculated by measuring the difference between the reconstructed time-series and the empirical time-series. As shown in Figure 4, the MAE changes as the  $L$  increases. The best value of  $L$  is 5. The smaller value of  $L$  with a large MAE may be the result of model underfitting, while the larger value of  $L$  with a large MAE may be the result of overfitting.

The proposed multi-resolution consistency loss can guarantee the model's good reconstruction performance. We investigated the optimal importance of  $L_{MRC}$  in the hybrid loss functions. As shown in Figure 5, we chose the value of  $\alpha$  from 0.0 to 1.0. The 0.0 means the  $L_{MRC}$  is removed from the total loss. As the value of  $\alpha$  increases, the MAE shows a downward trend. This indicates the importance of the proposed multi-resolution consistency loss in reconstructing the missing signals. The best value of  $\alpha$  is achieved at 0.9. In the downstream tasks, we will evaluate the model's performance using the optimal  $L$  and  $\alpha$ .

## 4.3 Reconstruction performance

We adopted the above settings and continued to investigate the time-series reconstruction performance of the left amygdala. We presented the training details about the reconstructing processes.

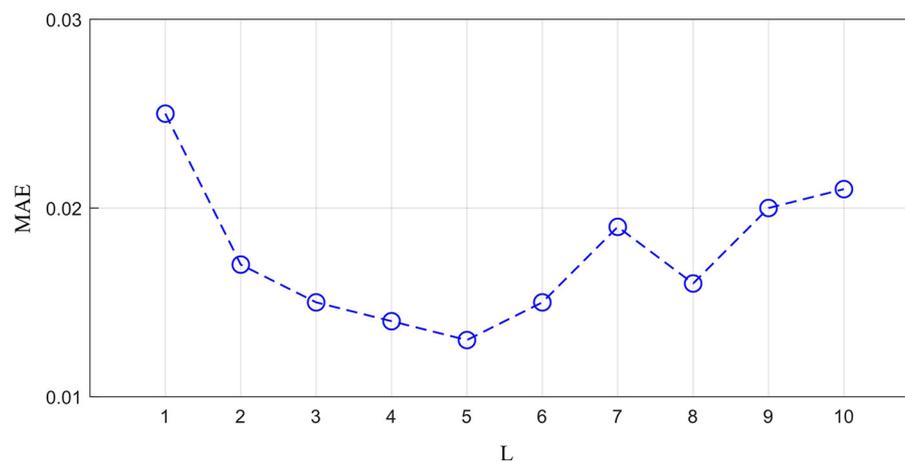


FIGURE 4  
The impact of layer number  $L$  on the generator.

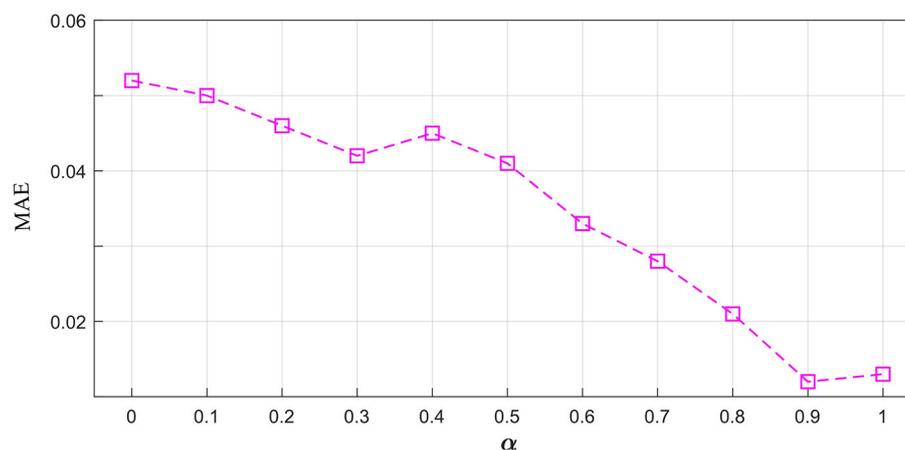


FIGURE 5  
The influence of parameter  $\alpha$  on the hybrid loss functions.

As shown in [Figure 6](#), we initialized the missing signal as Gaussian noise at the 0 epoch, and then the Gaussian noise is getting closer to the empirical signal as the epoch arrives at 500. The right column shows the frequency spectrum of the left column. The frequency spectrum is computed using the Fast Fourier transform, which converts the left time-series into individual spectral components and thereby provides frequency information about the it. At 0 epoch, the frequency information between empirical and reconstructed time-series is very different, and as the epoch increases, the frequency information difference gradually decreases; at the final epoch, the frequency information between the two signals is almost the same, indicating good reconstruction result. Furthermore, we quantitatively evaluate the functional connectivity using the PCC computed by empirical and generated time series. [Figure 7](#) shows that the larger difference is the missing-signal ROI-related connections in the right column. The maximum PCC change is lower than 0.05, which has little influence on brain network analysis.

To compare the reconstruction performance using different models, we chose the six competing models: (1) C-RNN-GAN ([Mogren, 2016](#)), (2) RCGAN ([Esteban et al., 2017](#)), (3) waveGAN ([Donahue et al., 2018](#)), (4) TimeGAN ([Yoon et al., 2019](#)), (5) SigCWGAN ([Ni et al., 2020](#)), and (6) TCGAN ([Xia et al., 2023](#)). The input is the incomplete multi-ROI time-series with only one ROI time-series removed. We compared the reconstructed missing signal by computing the three metrics: MAE, RMSE, R2, and DTW. We have randomly split the dataset into 10 folds 10 times. For each method, we calculate the mean and standard deviation for the four metrics. As shown in [Table 1](#), the GAN-based models show inferior performance than the transformer-based models. The possible reason is that the transformer benefits from the relationship modeling ability. Among these methods, the proposed model combining the transformer and GAN achieves the best reconstruction performance in terms of MAE (0.010), RMSE (0.015), R2(0.998), and DTW (1.872). To prove the effectiveness of the reconstructed time-series, we constructed

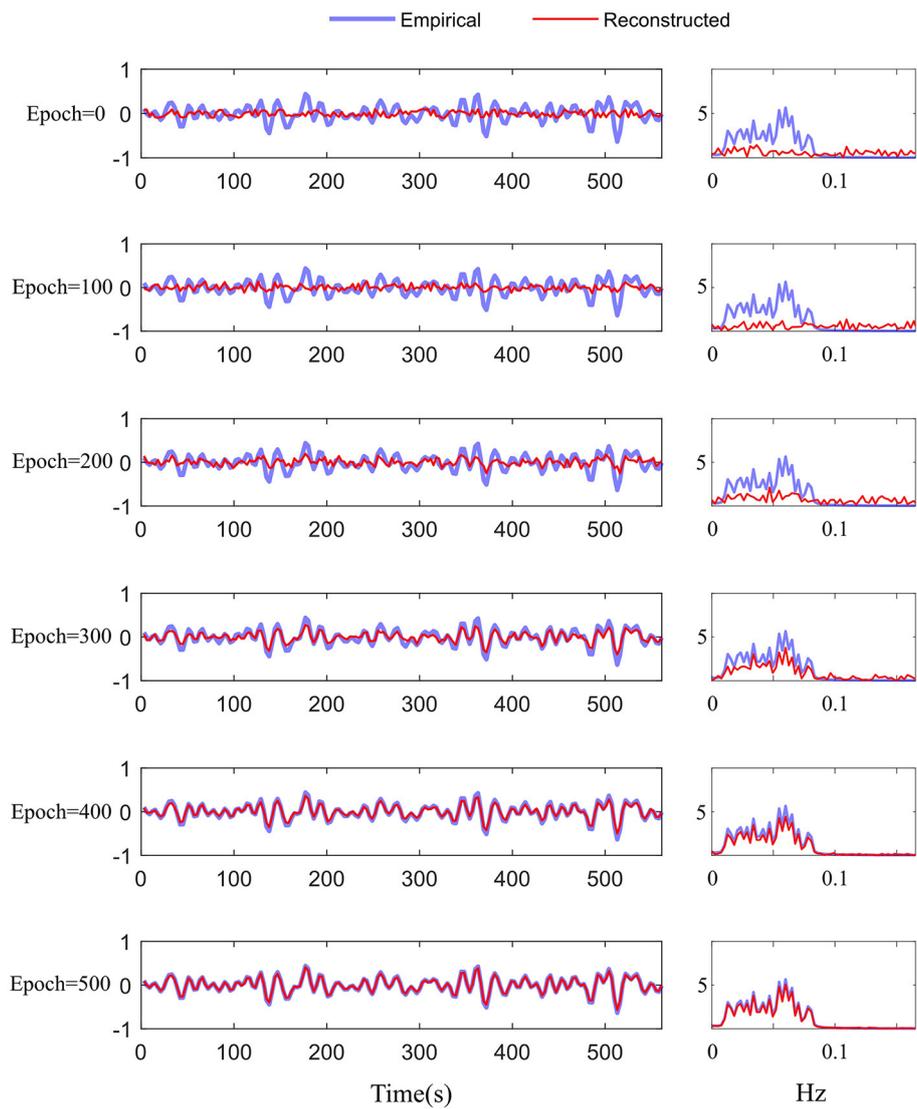


FIGURE 6 Training details of the difference between the empirical and reconstructed time-series in the time and frequency domains from 0 to 500 epochs.

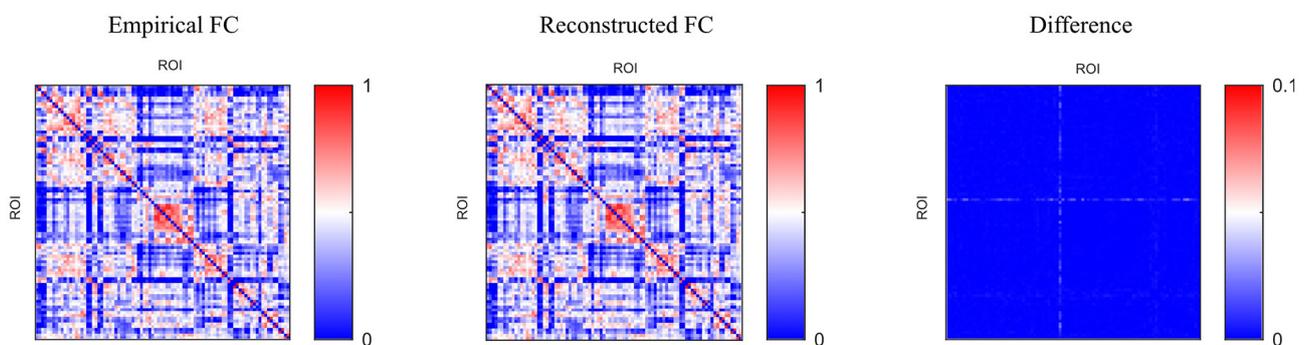


FIGURE 7 Comparison of functional connectivity using different methods.

functional connectivity (FC) from empirical and reconstructed time-series, respectively. The constructed FC was then sent to the BrainNetCNN (Kawahara et al., 2017) classifier to compute four metrics (i.e., ACC, SEN, SPE, and AUC) for NC versus LMCI. The classification results are shown in Figure 8, and there is no significant difference between the four metrics. The p-values of ACC, SEN, SPE, and AUC between the two methods are 0.894, 0.756, 0.703, and 0.358, respectively. The p-values are larger than

0.05, indicating that the proposed method can achieve the same classification performance as the empirical method. Furthermore, we investigated different methods' robustness to the noise, and we added Gaussian noise (1 or 5%) to the training and testing data. Then, we calculated the mean and standard deviation values for the seven methods. As shown in Table 2, our model achieves the best classification performance among the seven methods with a small deviation error, indicating our model's effectiveness and robustness.

To assess the reconstruction performance of other ROIs, we iteratively removed one ROI time-series from the preprocessed empirical multi-ROI time-series and trained them with our model. The reconstructed ROI time-series is compared with the empirical ROI time-series by computing the MAE value. The results are shown in Figure 9. For each ROI, we displayed the MAE. The mean MAE value of all ROIs is 0.010, with an averaged standard error of 0.003. A large MAE of ROI may indicate that this ROI has high degrees. The removal of this ROI fails to model long-range relationships in the generator. To display all the ROI reconstruction performance, Figure 10 shows the correlation between empirical and generated time-series. The left sub-figure displays the correlation between empirical and generated time-series. We remove one ROI time-series and reconstruct it using other ROIs' time-series. For the convenience of display, we utilize the time-series of all ROIs from one subject to plot correlation between reconstructed and empirical BOLD signal. The right sub-figure shows the correlation between empirical and removed ROI-related connections. After reconstructing the removed ROI time-series, we compute the Pearson correlation coefficient between the

TABLE 1 The reconstruction performance using different models.

| Method    | MAE ↓                | RMSE ↓               | R2 ↑                 | DTW ↓                |
|-----------|----------------------|----------------------|----------------------|----------------------|
| C-RNN-GAN | 0.030 ± 0.017        | 0.036 ± 0.026        | 0.960 ± 0.014        | 5.642 ± 0.513        |
| RCGAN     | 0.026 ± 0.015        | 0.033 ± 0.023        | 0.966 ± 0.010        | 4.835 ± 0.449        |
| waveGAN   | 0.022 ± 0.014        | 0.029 ± 0.021        | 0.971 ± 0.008        | 3.578 ± 0.416        |
| TimeGAN   | 0.016 ± 0.012        | 0.023 ± 0.018        | 0.987 ± 0.006        | 3.043 ± 0.372        |
| SigCWGAN  | 0.014 ± 0.011        | 0.020 ± 0.015        | 0.989 ± 0.005        | 2.644 ± 0.301        |
| TCGAN     | 0.013 ± 0.010        | 0.019 ± 0.014        | 0.992 ± 0.003        | 2.302 ± 0.221        |
| Proposed  | <b>0.010 ± 0.006</b> | <b>0.015 ± 0.011</b> | <b>0.998 ± 0.001</b> | <b>1.864 ± 0.247</b> |

The best results are in bold.

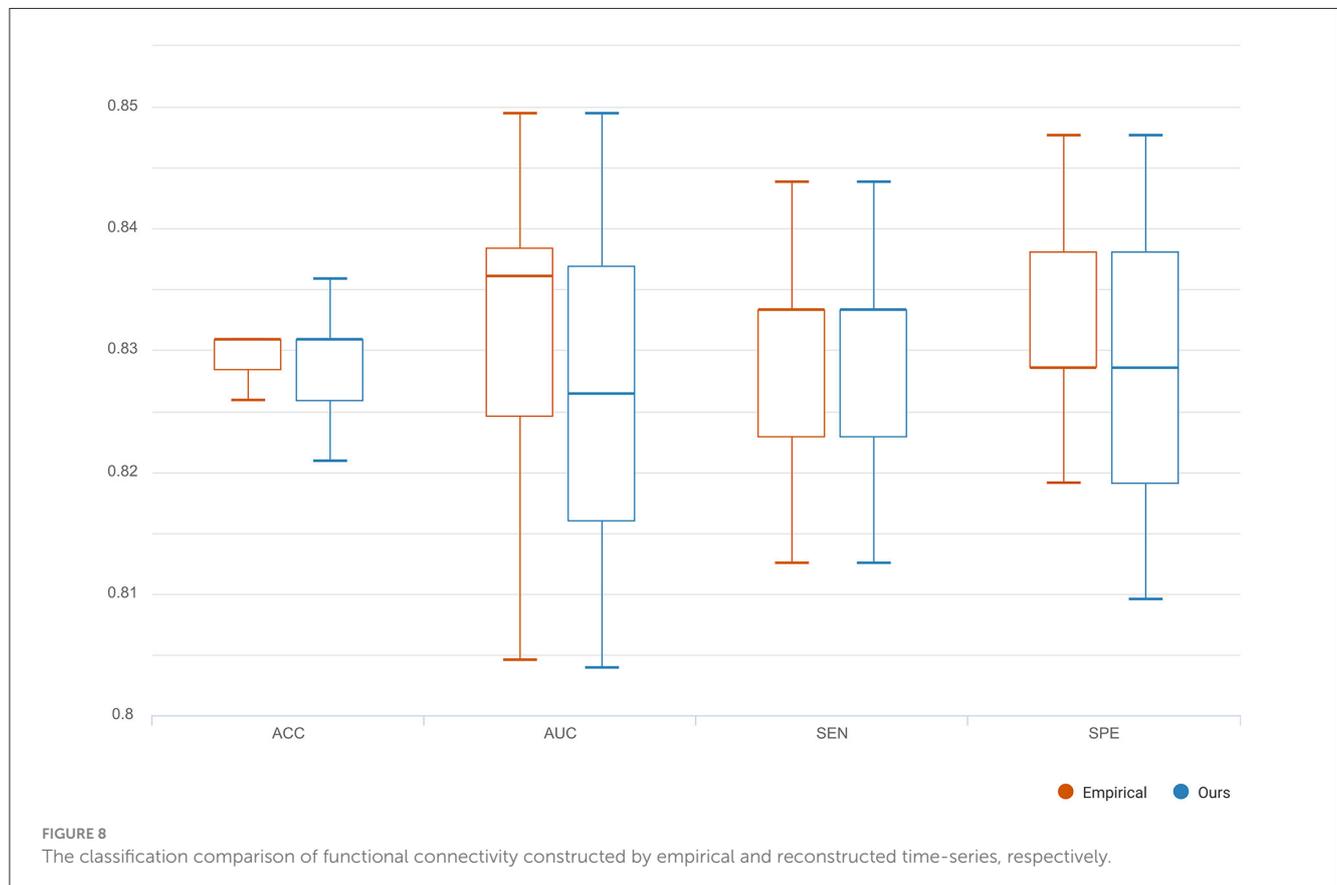
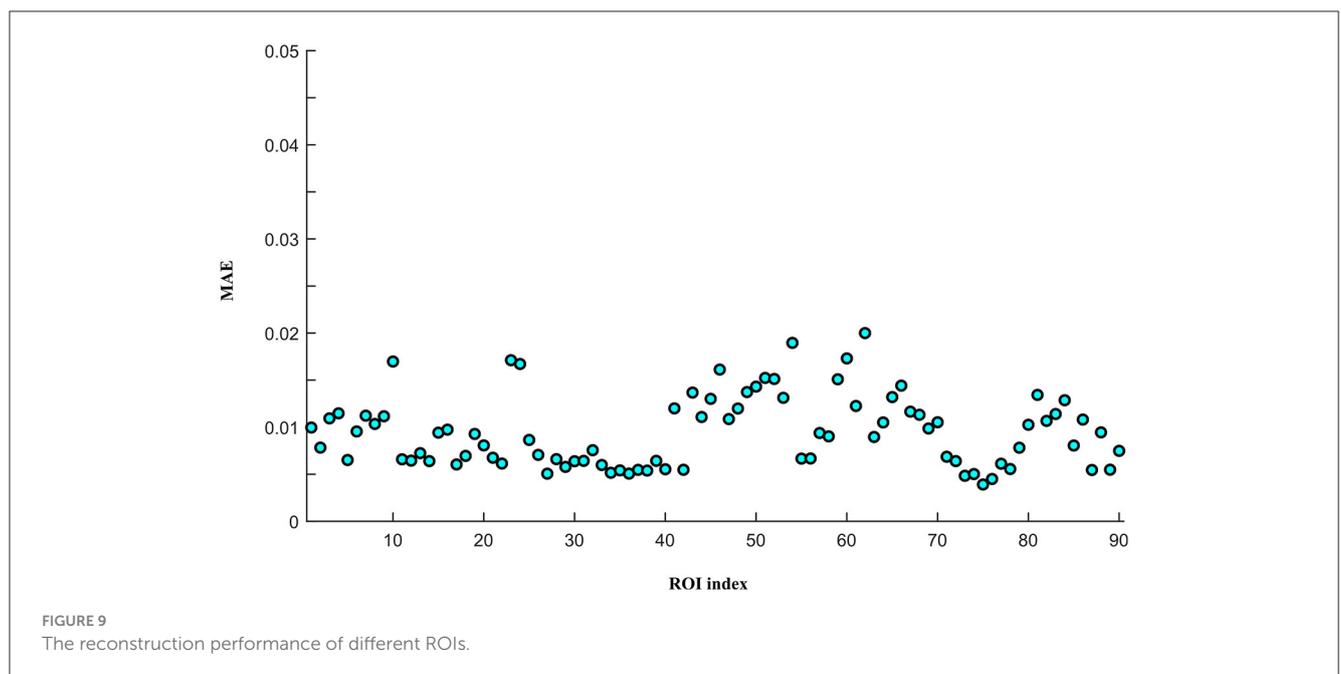


TABLE 2 The reconstruction performance using different models for different noise levels.

| Noise level | Method    | MAE ↓                | RMSE ↓               | R2 ↑                 | DTW ↓                |
|-------------|-----------|----------------------|----------------------|----------------------|----------------------|
| 1%          | C-RNN-GAN | 0.032 ± 0.018        | 0.038 ± 0.027        | 0.951 ± 0.014        | 5.813 ± 0.503        |
|             | RCGAN     | 0.029 ± 0.014        | 0.036 ± 0.022        | 0.954 ± 0.011        | 4.911 ± 0.432        |
|             | waveGAN   | 0.023 ± 0.013        | 0.030 ± 0.022        | 0.970 ± 0.009        | 3.415 ± 0.413        |
|             | TimeGAN   | 0.017 ± 0.012        | 0.025 ± 0.019        | 0.976 ± 0.007        | 3.112 ± 0.367        |
|             | SigCWGAN  | 0.015 ± 0.012        | 0.021 ± 0.016        | 0.978 ± 0.006        | 2.697 ± 0.304        |
|             | TCGAN     | 0.013 ± 0.011        | 0.019 ± 0.016        | 0.991 ± 0.005        | 2.303 ± 0.231        |
| 5%          | Proposed  | <b>0.010 ± 0.007</b> | <b>0.015 ± 0.012</b> | <b>0.998 ± 0.001</b> | <b>1.867 ± 0.249</b> |
|             | C-RNN-GAN | 0.037 ± 0.021        | 0.041 ± 0.029        | 0.941 ± 0.028        | 6.219 ± 0.627        |
|             | RCGAN     | 0.032 ± 0.019        | 0.039 ± 0.027        | 0.948 ± 0.014        | 5.237 ± 0.507        |
|             | waveGAN   | 0.027 ± 0.016        | 0.033 ± 0.025        | 0.964 ± 0.010        | 3.772 ± 0.486        |
|             | TimeGAN   | 0.019 ± 0.014        | 0.028 ± 0.021        | 0.971 ± 0.008        | 3.348 ± 0.371        |
|             | SigCWGAN  | 0.017 ± 0.013        | 0.024 ± 0.018        | 0.970 ± 0.008        | 2.713 ± 0.319        |
|             | TCGAN     | 0.015 ± 0.012        | 0.021 ± 0.017        | 0.985 ± 0.007        | 2.516 ± 0.229        |
|             | Proposed  | <b>0.011 ± 0.006</b> | <b>0.016 ± 0.012</b> | <b>0.996 ± 0.002</b> | <b>1.896 ± 0.252</b> |

The best results are in bold.



removed ROI and other ROIs. All the computed connections from the corresponding subject are denoted as reconstructed functional connectivity (FC), which is compared with empirical FC. The two subfigures demonstrate our model's reconstruction ability. The reconstructed signals can be applied to the downstream brain network analysis.

#### 4.4 Ablation study

To investigate the influence of the generator and the loss function on the reconstruction performance, we design four

variants of the proposed model. (1) UCT by removing the discriminator from the UCT-GAN model. (2) UCT-GAN without hierarchical topological transformer (MSETD w/o HTT). In the generator, we removed the CD sampling and CU sampling and only kept one MT-attention block. (3) UCT-GAN without the multi-resolution consistency loss (MSETD w/o MRC). In the discriminator, we removed two DH modules and (4) the proposed UCT-GAN model. For each variant, we compute the mean value of MAE, RMSE, R2, and DTW. The results are shown in Table 3. Removing the hierarchical structure or the discriminator greatly reduces the time-series reconstruction performance, which shows the effectiveness and necessity of the proposed model in time-series restoration. The multi-resolution consistency also lowers

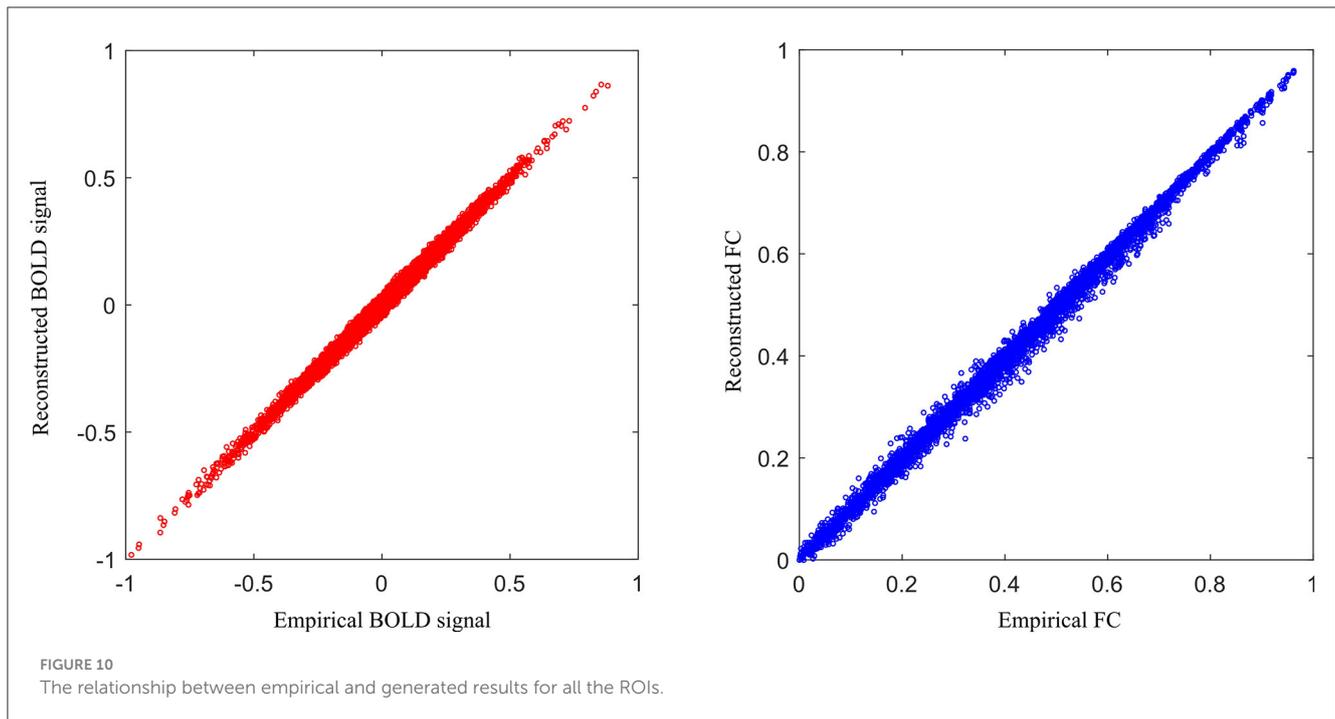


TABLE 3 Influence of different model's module on the reconstruction performance.

| Method          | MAE          | RMSE         | R2           | DTW          |
|-----------------|--------------|--------------|--------------|--------------|
| UCT             | 0.025        | 0.031        | 0.952        | 3.981        |
| UCT-GAN w/o HTT | 0.021        | 0.028        | 0.968        | 3.415        |
| UCT-GAN w/o MRC | 0.018        | 0.025        | 0.974        | 3.153        |
| <b>UCT-GAN</b>  | <b>0.010</b> | <b>0.015</b> | <b>0.998</b> | <b>1.864</b> |

The best results are in bold.

the model's reconstruction performance to some extent. All of them contribute a lot to time-series reconstruction performance. It indicates that the U-shaped generative architecture and multi-resolution consistency loss capture the spatial and temporal characteristics, thus effectively restoring complex brain functional dynamics.

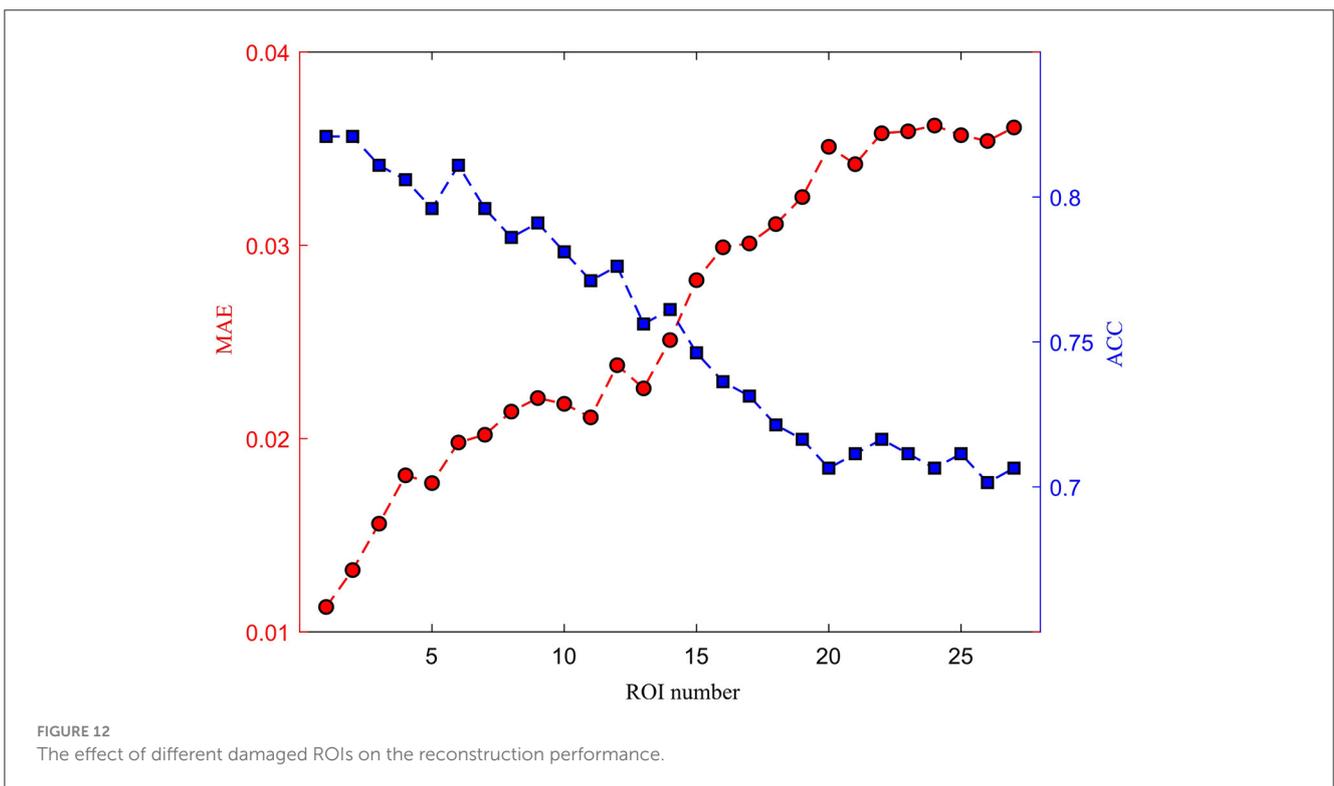
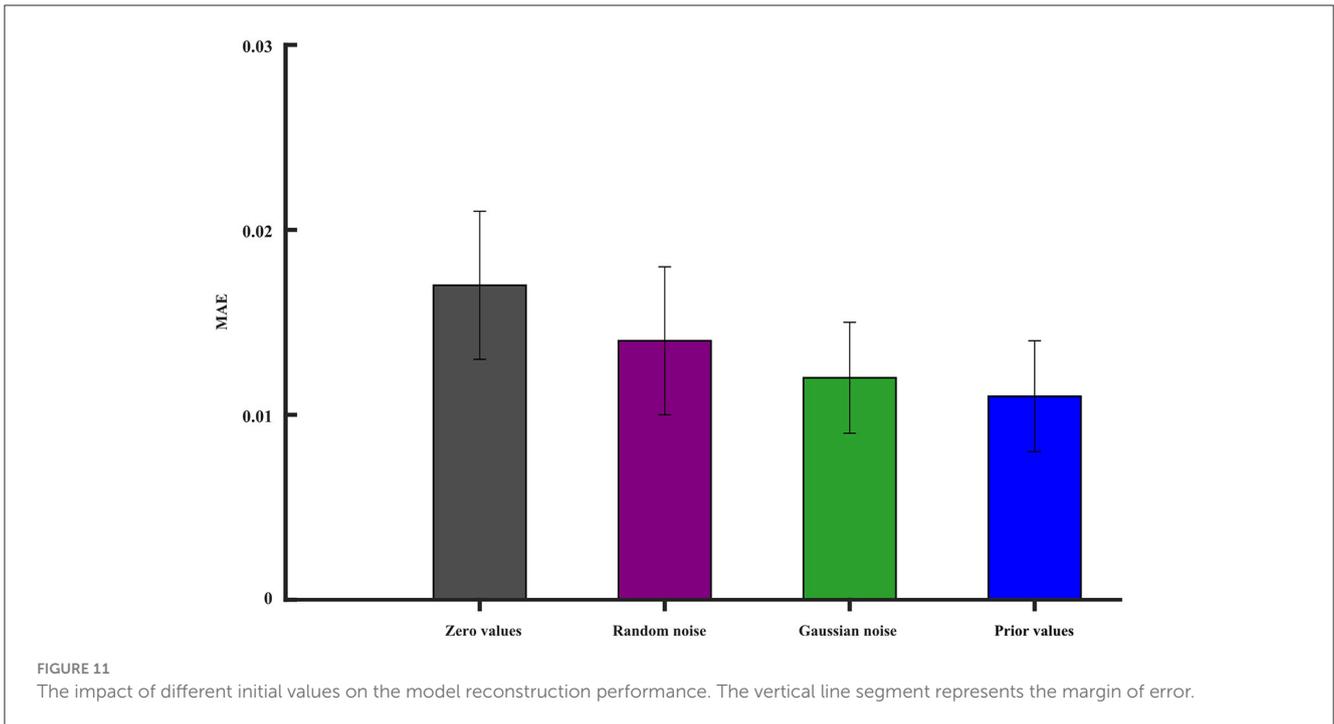
## 5 Discussion

In deep learning, neural networks are often non-convex and have multiple local minima. The choice of initial values can influence whether the optimization algorithm gets stuck in a poor local minimum or finds a more optimal solution. We investigate the iterative initial values during the training. As we know, the proper initial iterative values tend to find the optimal solution of the model. There are many strategies that are used to initialize the model's parameters' weights. We still study the condition when one ROI signal is missing. The missing signal is replaced by (1) zero values, (2) random noise, (3) Gaussian noise, and (4) prior values, which represent the averaged values of other ROI time series. All the initial values are forced into the range of 0 – 1. The MAE is used to evaluate the reconstruction performance. Figure 11

gives the best initial strategy of using the prior values. The prior value strategy can mitigate the risk of convergence to suboptimal solutions.

The medical treatment using the DBS usually involves implanting a device into the brain to alleviate symptoms of various neurological disorders. The intersection of the fornix and stria terminalis in the brain may be the optimal area for DBS treatment. The stria terminalis serves as a major output pathway of the amygdala. Therefore, we investigated the amygdala for potential clinical applications. The damaged signals in the amygdala may also influence the adjacent ROIs, such as the up, down, left, and right brain areas. We cumulatively removed one ROI signal from the scanned fMRI and evaluated the reconstruction performance. As shown in Figure 12, as the number of missing brain regions increases, the MAE gradually increases and the ACC correspondingly decreases. This shows that the reconstruction ability is greatly reduced.

The proposed model combines the U-shaped convolutional transformer and GANs to restore the missing brain functional time series. By restoring complex brain functional dynamics, the proposed model can achieve the same classification results as the empirical method. More missing signals can greatly reduce the reconstruction performance and disease prediction. No more than two ROI missing signals probably have little influence on the dementia diagnosis and brain network analysis. Though the proposed model can achieve good restoration performance, there are two limitations. One limitation is that the studied ROI may have a larger volume than that of the real distortion brain region. In the future, we will try more fined atlas to investigate the BOLD signal distortion, since more fined ROIs can better describe the signal distortion and help precisely reconstruct the missing signals for improving disease analysis. Another one is that the proposed model is tested theoretically with small subjects. In the next study, we will



validate our model on a larger dataset, such as the UK Biobank dataset (<https://www.ukbiobank.ac.uk/>).

## 6 Conclusion

This study proposes a novel U-shaped convolutional transformer GAN (UCT-GAN) model to restore the missing brain

functional time-series data. By leveraging generative adversarial networks (GANs) and the U-shaped transformer architecture, the proposed UCT-GAN can effectively capture hierarchical features in the restoration process. It should be stressed that the multi-level temporal-correlated attention and the convolutional sampling in the generator capture the long-range and local temporal features of the missing signal and associate their relationship with the effective signal. We also designed a multi-resolution consistency

loss to learn diverse temporal patterns and maintain consistency across different temporal resolutions. We theoretically tested our model on the public Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, and our experiments demonstrate superior reconstruction performance with other competing methods in terms of quantitative metrics. The proposed model offers a new solution for restoring brain functional time-series data, driving forward the field of neuroscience research through the provision of enhanced tools for data analysis and interpretation.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: <https://adni.loni.usc.edu>.

## Author contributions

QZ: Conceptualization, Data curation, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing. RL: Data curation, Software, Visualization, Writing – review & editing. BS: Data curation, Software, Visualization, Writing – review & editing. JH: Conceptualization, Methodology, Software, Writing – review & editing. YZ: Investigation, Resources, Validation, Writing – review & editing. XC: Formal analysis, Resources, Validation, Writing – review & editing. YW: Formal analysis, Investigation, Writing – review & editing. JG: Funding

acquisition, Project administration, Supervision, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This paper was supported by the Natural Science Foundation of Hubei Province (2023AFB004 and 2023AFB003).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Alajangi, H. K., Kaur, M., Sharma, A., Rana, S., Thakur, S., Chatterjee, M., et al. (2022). Blood-brain barrier: emerging trends on transport models and new-age strategies for therapeutics intervention against neurological disorders. *Mol. Brain* 15, 1–28. doi: 10.1186/s13041-022-00937-4
- Boutet, A., Madhavan, R., Elias, G. J., Joel, S. E., Gramer, R., Ranjan, M., et al. (2021). Predicting optimal deep brain stimulation parameters for Parkinson's disease using functional MRI and machine learning. *Nat. Commun.* 12:3043. doi: 10.1038/s41467-021-23311-9
- Chen, L., Zhou, Y., Gao, S., Li, M., Tan, H., Wan, Z., et al. (2023). Ara-net: an attention-aware retinal atrophy segmentation network coping with fundus images. *Front. Neurosci.* 17:1174937. doi: 10.3389/fnins.2023.1174937
- Donahue, C., McAuley, J., and Puckette, M. (2018). Adversarial audio synthesis. *arXiv*. [Preprint]. arXiv:1802.04208. doi: 10.48550/arXiv.1802.04208
- Esteban, C., Hyland, S. L., and Rätsch, G. (2017). Real-valued (medical) time series generation with recurrent conditional gans. *arXiv*. [Preprint]. arXiv:1706.02633. doi: 10.48550/arXiv.1706.02633
- Forouzaneshad, P., Abbaspour, A., Fang, C., Cabrerizo, M., Loewenstein, D., Duara, R., et al. (2019). A survey on applications and analysis methods of functional magnetic resonance imaging for Alzheimer's disease. *J. Neurosci. Methods* 317, 121–140. doi: 10.1016/j.jneumeth.2018.12.012
- Hong, J., Zhang, Y.-D., and Chen, W. (2022). Source-free unsupervised domain adaptation for cross-modality abdominal multi-organ segmentation. *Knowl.-Based Syst.* 250:109155. doi: 10.1016/j.knsys.2022.109155
- Hu, B., Zhan, C., Tang, B., Wang, B., Lei, B., Wang, S.-Q., et al. (2023). 3-D brain reconstruction by hierarchical shape-perception network from a single incomplete image. *IEEE Trans. Neural Netw. Learn. Syst.* 1–13. doi: 10.1109/TNNLS.2023.3266819
- Ibrahim, B., Suppiah, S., Ibrahim, N., Mohamad, M., Hassan, H. A., Nasser, N. S., et al. (2021). Diagnostic power of resting-state fMRI for detection of network connectivity in Alzheimer's disease and mild cognitive impairment: a systematic review. *Hum. Brain Mapp.* 42, 2941–2968. doi: 10.1002/hbm.25369
- Jiang, Y., Chang, S., and Wang, Z. (2021). Transgan: two pure transformers can make one strong Gan, and that can scale up. *Adv. Neural Inf. Process. Syst.* 34, 14745–14758.
- Kawahara, J., Brown, C. J., Miller, S. P., Booth, B. G., Chau, V., Grunau, R. E., et al. (2017). Brainnetcnn: convolutional neural networks for brain networks; towards predicting neurodevelopment. *Neuroimage* 146, 1038–1049. doi: 10.1016/j.neuroimage.2016.09.046
- Knopman, D. S., Amieva, H., Petersen, R. C., Chételat, G., Holtzman, D. M., Hyman, B. T., et al. (2021). Alzheimer disease. *Nat. Rev. Dis. Primers* 7:33. doi: 10.1038/s41572-021-00269-y
- Leoutsakos, J.-M. S., Yan, H., Anderson, W. S., Asaad, W. F., Baltuch, G., Burke, A., et al. (2018). Deep brain stimulation targeting the fornix for mild Alzheimer dementia (the advance trial): a two year follow-up including results of delayed activation. *J. Alzheimers Dis.* 64, 597–606. doi: 10.3233/JAD-180121
- Li, M., Liu, S., Wang, Z., Li, X., Yan, Z., Zhu, R., et al. (2023). Myopiadetr: end-to-end pathological myopia detection based on transformer using 2D fundus images. *Front. Neurosci.* 17:1130609. doi: 10.3389/fnins.2023.1130609
- Li, X., Metsis, V., Wang, H., and Ngu, A. H. H. (2022). "TTS-GAN: a transformer-based time-series generative adversarial network," in *International Conference on Artificial Intelligence in Medicine* (Berlin: Springer), 133–143. doi: 10.1007/978-3-031-09342-5\_13
- Li, Y., Peng, X., Zhang, J., Li, Z., and Wen, M. (2021). DCT-GAN: dilated convolutional transformer-based GAN for time series anomaly detection. *IEEE Trans. Knowl. Data Eng.* 35, 3632–3644. doi: 10.1109/TKDE.2021.3130234
- Li, Y., Wang, H., Li, J., Liu, C., and Tan, J. (2022). "Act: adversarial convolutional transformer for time series forecasting," in *2022 International Joint Conference on Neural Networks (IJCNN)* (Padua: IEEE), 1–8. doi: 10.1109/IJCNN55064.2022.9892791
- Limousin, P., and Foltynie, T. (2019). Long-term outcomes of deep brain stimulation in Parkinson disease. *Nat. Rev. Neurol.* 15, 234–242. doi: 10.1038/s41582-019-0145-9
- Luo, B., Qiu, C., Chang, L., Lu, Y., Dong, W., Liu, D., et al. (2022). Altered brain network centrality in Parkinson's disease patients after deep brain stimulation:

- a functional MRI study using a voxel-wise degree centrality approach. *J. Neurosurg.* 138, 1712–1719. doi: 10.3171/2022.9.JNS221640
- Luo, Y., Cai, X., Zhang, Y., Xu, J., and Xiaojie, Y. (2018). "Multivariate time series imputation with generative adversarial networks," in *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)* (Montreal, QC).
- Luo, Y., Zhang, Y., Cai, X., and Yuan, X. (2019). "E2GAN: end-to-end generative adversarial network for multivariate time series analysis and lstm\_bilstm method. *IEEE Trans. Intell. Transp. Syst.* 23, 5615–5624. doi: 10.1109/TITS.2021.3055258
- Madane, A., Dilmi, M.-d., Forest, F., Azzag, H., Lebbah, M., and Lacaille, J. (2022). Transformer-based conditional generative adversarial network for multivariate time series generation. *arXiv*. [Preprint]. arXiv:2210.02089. doi: 10.48550/arXiv.2210.02089
- Medtronic, P. (2020). Green light for deep brain stimulator incorporating neurofeedback. *Nat. Biotechnol.* 38, 1014–1015. doi: 10.1038/s41587-020-0664-3
- Mogren, O. (2016). Continuous recurrent neural networks with adversarial training. *arXiv*. [Preprint]. arXiv:1611.09904. doi: 10.48550/arXiv.1611.09904
- Neumann, W.-J., Horn, A., and Kühn, A. A. (2023). Insights and opportunities for deep brain stimulation as a brain circuit intervention. *Trends Neurosci.* 46, 472–487. doi: 10.1016/j.tins.2023.03.009
- Ni, H., Szpruch, L., Wiese, M., Liao, S., and Xiao, B. (2020). Conditional sigwasserstein gans for time series generation. *arXiv*. [Preprint]. arXiv:2006.05421. doi: 10.48550/arXiv.2006.05421
- Nimbalkar, S., Fuhrer, E., Silva, P., Nguyen, T., Sereno, M., Kassegne, S., et al. (2019). Glassy carbon microelectrodes minimize induced voltages, mechanical vibrations, and artifacts in magnetic resonance imaging. *Microsyst. Nanoeng.* 5:61. doi: 10.1038/s41378-019-0106-x
- Philips, R. T., Torrisi, S. J., Gorka, A. X., Grillon, C., and Ernst, M. (2022). Dynamic time warping identifies functionally distinct fMRI resting state cortical networks specific to VTA and SNC: a proof of concept. *Cereb. Cortex* 32, 1142–1151. doi: 10.1093/cercor/bhab273
- Rios, A. S., Oxenford, S., Neudorfer, C., Butenko, K., Li, N., Rajamani, N., et al. (2022). Optimal deep brain stimulation sites and networks for stimulation of the fornix in Alzheimer's disease. *Nat. Commun.* 13:7707. doi: 10.1038/s41467-022-34510-3
- Sendi, M. S., Zendehehrouh, E., Fu, Z., Liu, J., Du, Y., Mormino, E., et al. (2023). Disrupted dynamic functional network connectivity among cognitive control networks in the progression of Alzheimer's disease. *Brain Connect.* 13, 334–343. doi: 10.1089/brain.2020.0847
- Siddiqi, S. H., Khosravani, S., Rolston, J. D., and Fox, M. D. (2023). The future of brain circuit-targeted therapeutics. *Neuropsychopharmacology* 49, 179–188. doi: 10.1038/s41386-023-01670-9
- Soleimani, G., Nitsche, M. A., Bergmann, T. O., Towhidkhal, F., Violante, I. R., Lorenz, R., et al. (2023). Closing the loop between brain and electrical stimulation: towards precision neuromodulation treatments. *Transl. Psychiatry* 13:279. doi: 10.1038/s41398-023-02565-5
- Srinivasan, P., and Knottenbelt, W. J. (2022). Time-series transformer generative adversarial networks. *arXiv*. [Preprint]. arXiv:2205.11164. doi: 10.48550/arXiv.2205.11164
- Tang, B., and Matteson, D. S. (2021). Probabilistic transformer for time series analysis. *Adv. Neural Inf. Process. Syst.* 34, 23592–23608. Available online at: [https://papers.nips.cc/paper\\_files/paper/2021/file/c68bd9055776bf38d8fc43ced283678-Paper.pdf](https://papers.nips.cc/paper_files/paper/2021/file/c68bd9055776bf38d8fc43ced283678-Paper.pdf)
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Vogel, J. W., Corriveau-Lecavalier, N., Franzmeier, N., Pereira, J. B., Brown, J. A., Maass, A., et al. (2023). Connectome-based modelling of neurodegenerative diseases: towards precision medicine and mechanistic insight. *Natu. Rev. Neurosci.* 24, 620–639. doi: 10.1038/s41583-023-00731-8
- Wan, Z., Cheng, W., Li, M., Zhu, R., and Duan, W. (2023a). GDNET-EEG: an attention-aware deep neural network based on group depth-wise convolution for ssvep stimulation frequency recognition. *Front. Neurosci.* 17:1160040. doi: 10.3389/fnins.2023.1160040
- Wan, Z., Li, M., Liu, S., Huang, J., Tan, H., Duan, W., et al. (2023b). Eegformer: a transformer-based brain activity classification method using eeg signal. *Front. Neurosci.* 17:1148855. doi: 10.3389/fnins.2023.1148855
- Wan, Z., Li, M., Wang, Z., Tan, H., Li, W., Yu, L., et al. (2023c). CellT-Net: a composite transformer method for 2-d cell instance segmentation. *IEEE J. Biomed. Health Inform.* 28, 730–741. doi: 10.1109/JBHI.2023.3265006
- Wan, Z., Liu, S., Ding, F., Li, M., Srivastava, G., Yu, K., et al. (2023d). C2bnet: a deep learning architecture with coupled composite backbone for parasitic egg detection in microscopic images. *IEEE J. Biomed. Health Inform.* 1–13. doi: 10.1109/JBHI.2023.3318604
- Wan, Z., Wan, J., Cheng, W., Yu, J., Yan, Y., Tan, H., et al. (2023e). A wireless sensor system for diabetic retinopathy grading using mobilevit-plus and resnet-based hybrid deep learning framework. *Appl. Sci.* 13:6569. doi: 10.3390/app13116569
- Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., He, Y., et al. (2015). Gretna: a graph theoretical network analysis toolbox for imaging connectomics. *Front. Hum. Neurosci.* 9:386. doi: 10.3389/fnhum.2015.00386
- Wang, S., Chen, Z., You, S., Wang, B., Shen, Y., Lei, B., et al. (2022). Brain stroke lesion segmentation using consistent perception generative adversarial network. *Neural Comput. Appl.* 34, 8657–8669. doi: 10.1007/s00521-021-06816-8
- Wang, X., Wang, M., Sheng, H., Zhu, L., Zhu, J., Zhang, H., et al. (2022). Subdural neural interfaces for long-term electrical recording, optical microscopy and magnetic resonance imaging. *Biomaterials* 281:121352. doi: 10.1016/j.biomaterials.2021.121352
- Wang, Z., Zheng, Y., Zhu, D. C., Bozoki, A. C., and Li, T. (2018). Classification of Alzheimer's disease, mild cognitive impairment and normal control subjects using resting-state fMRI based network connectivity analysis. *IEEE J. Transl. Eng. Health Med.* 6, 1–9. doi: 10.1109/JTEHM.2018.2874887
- Warren, S. L., and Moustafa, A. A. (2023). Functional magnetic resonance imaging, deep learning, and Alzheimer's disease: a systematic review. *J. Neuroimaging* 33, 5–18. doi: 10.1111/jon.13063
- Wu, S., Xiao, X., Ding, Q., Zhao, P., Wei, Y., Huang, J., et al. (2020). Adversarial sparse transformer for time series forecasting. *Adv. Neural Inf. Process. Syst.* 33, 17105–17115. Available online at: [https://proceedings.neurips.cc/paper\\_files/paper/2020/file/c6b8c8d762da15fa8dbbdfb6af9e260-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2020/file/c6b8c8d762da15fa8dbbdfb6af9e260-Paper.pdf)
- Xia, Y., Xu, Y., Chen, P., Zhang, J., and Zhang, Y. (2023). Generative adversarial network with transformer generator for boosting eeg classification. *Biomed. Signal Process. Control.* 80:104276. doi: 10.1016/j.bspc.2022.104276
- Yen, C., Lin, C.-L., and Chiang, M.-C. (2023). Exploring the frontiers of neuroimaging: a review of recent advances in understanding brain functioning and disorders. *Life* 13:1472. doi: 10.3390/life13071472
- Yin, W., Li, L., and Wu, F.-X. (2022). Deep learning for brain disorder diagnosis based on fMRI images. *Neurocomputing* 469, 332–345. doi: 10.1016/j.neucom.2020.05.113
- Yoon, J., Jarrett, D., and Van der Schaar, M. (2019). "Time-series generative adversarial networks," in *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)* (Vancouver, BC).
- You, S., Lei, B., Wang, S., Chui, C. K., Cheung, A. C., Liu, Y., et al. (2022). Fine perceptive GANS for brain MR image super-resolution in wavelet domain. *IEEE Trans. Neural Netw. Learn. Syst.* 34, 8802–8814. doi: 10.1109/TNNLS.2022.3153088
- Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., and Eickhoff, C. (2021). "A transformer-based framework for multivariate time series representation learning," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (New York, NY: ACM), 2114–2124. doi: 10.1145/3447548.3467401
- Zhao, L., Zhang, Z., Chen, T., Metaxas, D., and Zhang, H. (2021). Improved transformer for high-resolution gans. *Adv. Neural Inf. Process. Syst.*, 34, 18367–18380. Available online at: [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/98dce83da57b0395e163467c9dae521b-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/98dce83da57b0395e163467c9dae521b-Paper.pdf)
- Zuo, Q., Lu, L., Wang, L., Zuo, J., and Ouyang, T. (2022). Constructing brain functional network by adversarial temporal-spatial aligned transformer for early ad analysis. *Front. Neurosci.* 16:1087176. doi: 10.3389/fnins.2022.1087176
- Zuo, Q., Wu, H., Chen, C. P., Lei, B., and Wang, S. (2024). Prior-guided adversarial learning with hypergraph for predicting abnormal connections in Alzheimer's disease. *IEEE Trans. Cybern.* 1–14. doi: 10.1109/TCYB.2023.3344641
- Zuo, Q., Zhong, N., Pan, Y., Wu, H., Lei, B., Wang, S., et al. (2023). Brain structure-function fusing representation learning using adversarial decomposed-vae for analyzing MCI. *IEEE Trans. Neural Syst. Rehabil. Eng.* 31, 4017–4028. doi: 10.1109/TNSRE.2023.3323432