Structure Mapping Generative Adversarial Network for Multi-View Information Mapping Pattern Mining

Xia-an Bi, *Member, IEEE*, YangJun Huang, Zicheng Yang, Ke Chen, Zhaoxu Xing, Luyun Xu, Xiang Li, Zhengliang Liu, and Tianming Liu, *Senior Member, IEEE*

***Abstract*—Multi-view learning is dedicated to integrating information from different views and improving the generalization performance of models. However, in most current works, learning under different views has significant independency, overlooking common information mapping patterns that exist between these views. This paper proposes a Structure Mapping Generative adversarial network (SM-GAN) framework, which utilizes the consistency and complementarity of multi-view data from the innovative perspective of information mapping. Specifically, based on network-structured multi-view data, a structural information mapping model is proposed to capture hierarchical interaction patterns among views. Subsequently, three different types of graph convolutional operations are designed in SM-GAN based on the model. Compared with regular GAN, we add a structural information mapping module between the encoder and decoder wthin the generator, completing the structural information mapping from the micro-view to the macro-view. This paper conducted sufficient validation experiments using public imaging genetics data in Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. It is shown that SM-GAN outperforms baseline and advanced methods in multi-label classification and evolution prediction tasks.**

***Index Terms*—Multi-view learning, structural information mapping, graph convolution, generative adversarial network.**

Manuscript received 21 October 2022; revised 9 August 2023 and 17 October 2023; accepted 4 November 2023. Date of publication 6 November 2023; date of current version 6 March 2024. This work was supported in part by the National Natural Science Foundation of China under Grants 62072173 and 62372169, in part by the Natural Science Foundation of Hunan Province, China, under Grant 2023JJ30410, in part by the Key Open Project of Key Laboratory of Data Science and Intelligence Education (Hainan Normal University), Ministry of Education, under Grant DSIE202101, in part by the National Key Research and Development Program of China under Grant 2020YFB2104400, in part by the Medical Humanities and Social Sciences Project of Hunan Normal University, and in part by the Innovation and Entrepreneurship Training Program of Hunan Xiangjiang Artificial Intelligence Academy. Recommended for acceptance by Shuiwang Ji. *(Corresponding authors: Xia-an Bi; Luyun Xu; Tianming Liu.)*

Xia-an Bi, YangJun Huang, Zicheng Yang, and Ke Chen are with the College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China (e-mail: bixiaan@hnu.edu.cn; huagyangjunhnnu@163.com; yangzichenghnnu@163.com; chenkehnnu@163.com).

Zhaoxu Xing is with the College of Computer Science and Electronic Engineering, Hunan University, Changsha 410012, China (e-mail: xingzhaoxuhnnu@163.com).

Luyun Xu is with the College of Business, Hunan Normal University, Changsha 410081, China (e-mail: xuluyun@hunnu.edu.cn).

Xiang Li is with the Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114 USA (e-mail: xli60@mgh.harvard.edu).

Zhengliang Liu and Tianming Liu are with the Department of Computer Science and Bioimaging Research Center, University of Georgia, Athens, GA 30602 USA (e-mail: zl18864@uga.edu; tliu@uga.edu).

Digital Object Identifier 10.1109/TPAMI.2023.3330795

I. Introduction

In PRACTICAL applications, obtaining a comprehensive understanding of objects through a single view is often challenging. Multi-view learning is currently a flourishing field in artificial intelligence, with its core idea aimed at integrating insights from multiple views to enhance the model's generalization abilities [1], [2]. After years of development, multi-view learning has been proven effective in various fields, including retrieval systems [3], object recognition [4], [5], disease diagnosis [6], [7], etc. It is generally accepted that the effectiveness of multi-view learning stems from the consistency and complementarity characteristics of multi-view data. On one hand, the views essentially depict the same object from different perspectives, which brings a natural semantic consistency among them. On the other hand, each view contains unique information, so there is complementarity in multi-view data [8].

Multi-view learning aims to model each view and optimize all models, for which several common strategies have been developed. For example, co-training approaches iteratively train classifiers under different views [9], [10]. For co-regularization, efforts are made to enhance the ob-jective function by designing regularization terms [11], [12]. Subspace learning assumes that there exists a common subspace in all views, and its objective is to find the mapping of all sample points in the subspace [13], [14].

Despite the evaluated effectiveness, the current approaches could be further improved by modeling the complex inter-view interactions. In realistic scenarios, e.g., biomedical data analysis, hierarchical multi-view data are common and the association pattern between different levels cannot be mined through simple linear analysis. This phenomenon is initially noticed in the field of natural language processing [15], where case views can be classified into macroscopic views and microscopic views according to the level of expressed information. These two types of views have different implications for intelligent research. In general, microscopic views often provide less significant characteristic information than macroscopic views but usually contain fundamental factors that lead to macro-level manifestations and changes. For instance, for complex disease research, one can investigate microscopic gene mutations from genetic views [16], [17], [18], and also can observe the macroscopic lesions from imaging views [19], [20]. In general, the former is believed to be endogenous to macro-level changes, yet the exact mechanisms remain largely unidentified [21], [22], [23].

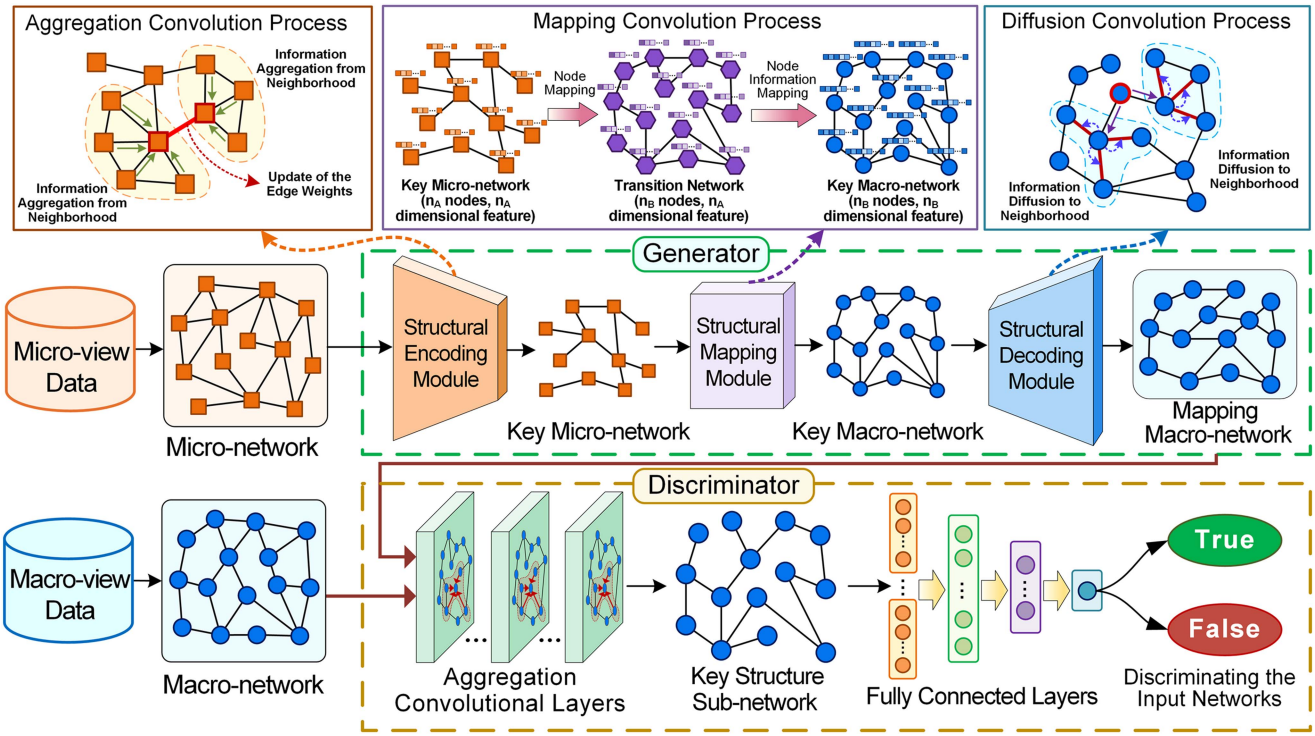


Fig. 1. Flowchart of the proposed SM-GAN approach. In the generator, structural encoding module includes aggregation convolution operations for mining key micro-view information; structural mapping module performs the micro-macro structural information mapping to obtain the key macro-networks; the structural decoding module is to finalize the generated macro-networks. The discriminator judges whether the input macro-networks are real or generated. The discrimination result (True/False) is determined by the multiple aggregation convolutional layers and fully connected layers. After training, SM-GAN will capture the micro-macro nonlinear mapping patterns.

Existing methods are generally good at mining view-specific information, but the knowledge about the multi-view interaction patterns is limited. This paper innovatively explores the association patterns between different views from the perspective of information mapping from micro-views to macro-views, which fully leverages the hierchachial complementary of multi-view data.

Deep learning methods for multi-view analysis have become emerging research interests [2], [24]. In particular, Graph Neural Network (GNN) [25], [26] have developed vigorously due to the increasing tendency of network-based data representations. Graph Convolutional Network (GCN), as a major improvement branch of GNN, can efficiently aggregate and extract key information of the nodes in graphs [27]. Generative Adversarial Network (GAN) is widely used for data generation or augmentation. GAN-based multi-view methods have also been emphasized in recent years [28]. However, these methods generally lack interpretability, limiting their potential applications. On the other hand, most current multi-view methods only focus on sample classification. Yet, in many practical scenarios characterized by long-term evolutionary trends, the ability to predict the future evolutionary pattern of data based on the current state is a task in significant need, especially for preventing certain loss-causing outcomes, e.g., disease deterioration [29], [30], traffic accident [31], etc. This paper designs an interpretable deep learning algorithm, where the convolutional layers are based on mathematical modeling to study the impact of micro-level factors on macro-level manifestations. Our method can be used not only for sample classification but also for predicting the evolutionary trend of the data, which broadens its application prospects.

This paper proposed a general framework for multi-view learning by utilizing the complementary information of multi-view data from the perspective of macro-to-micro scale mapping. First, a network representation is constructed for microscopic and macroscopic views of data, named as "micro-network" and "macro-network". Second, a structural information mapping model is established to investigate the mapping patterns from the micro-network to the macro-network. Finally, convolutional operations are designed based on the model and integrated into the GAN architecture to form SM-GAN, as depicted in Fig. 1. We validate the performance of SM-GAN in tasks of multi-label sample classification and evolution prediction using the imaging genetic data published in the Alzheimer's Disease Neuroimaging Initiative (ADNI).

On the whole, this paper presents a pioneering work about multi-view mapping pattern mining. It is important to emphasize that this paper specifically concentrates on view-level information mapping patterns, while specific feature-level mapping patterns (e.g., the patterns describing how specific genes affect specific brain regions) are not yet considered. The paper employs efficient deep learning to create a model with an extensive set of parameters designed for characterizing view-level mapping patterns. Through experimental results, SM-GAN can effectively and steadily map the original view data to the target view using

the generator, verifying the existence of the view information mapping pattern and the effectiveness of SM-GAN in capturing view-level mapping patterns. This work provides a brand-new perspective of multi-view learning and lays the foundation for further exploration, including the detection of the feature-level mapping patterns (e.g., how certain etiologies like genes regulate the activities of specific brain regions), which is believed to be a valuable task with significant application values.

In summary, our contributions are listed as follows:

1) A structural information mapping model is developed. In this model, multi-view interactions are modeled as a structural information mapping process from micro-networks to macro-networks, fully expoiting the intricate interrelationships among various views.
2) A GAN-based deep learning method is proposed. The convolution operations are designed according to the structural information mapping model. The trained model can extract the information mapping pattern from multi-view data, and the sample classification and evolutionary pattern prediction are based on the extracted patterns.
3) Effectiveness of SM-GAN is validated on a multi-modal biomedical dataset, and the results demonstrate that SM-GAN outperforms other commonly used methods.

II. Relevant Works

A. Traditional Techniques for Multi-View Learning

Leveraging the consistency and complementarity within multi-view data is acknowledged as the key to improving the performance of multi-view learning. For this purpose, the generic and intuitive idea is to train a model for each view and combine the multiple understandings from all views when making decisions [32]. Representative methods under this idea include co-training and multi-kernel learning. The former designs an alternating training process to ensure the consistency of multi-view information. The latter tries to pre-specify the base kernel matrix for each view. Cai et al. [33] considered both global and local cluster structures in co-training, incorporating more multi-view information than trivial methods. Liu et al. [34] proposed a solution for incomplete base kernels by adaptively complementing the missing elements. However, it is worth noting that the definition of the uniformity measure in co-training and the pre-selection process of the kernel are highly subjective, which restricts the flexibility of methods.

To better exploit the consistency of multi-view data, researchers have proposed new ideas. For example, some studies introduce and optimize new regularization terms to maximize consistency. Based on the classical support vector machine algorithm, on the other hand, Ye et al. [12] attempted to introduce a set of double-sided constraint terms to improve the robustness. Subspace learning advocates mapping samples under different views to a common subspace, and the dimensionality of the subspace thus reflects the view consistency. In this direction, canonical correlation analysis (CCA) is a general means [35]. Kumar et al. [36] used supervised information to improve the CCA objective function to obtain a joint multi-view representation of samples in a common subspace, and the validity of its method is verified in several data sets.

These works solved certain problems in multi-view learning and gained fair performance, but the independency of views caused by primary information fusion remains a great obstacle to further improvement. Overall, though good performance is achieved in mining view-specific characteristic information, a proper way to combine the multi-view information is lacking. Little is known about the interaction patterns among views, making it difficult to adapt to realistic scenarios.

B. Multi-View Deep Learning Models

Deep learning is becoming a mainstream research method in order to improve the efficiency of mining multi-view data. As reviewed in the literature [2], common deep model training methods include convolutional neural network (CNN), GCN, GAN, etc., and this trend is still maintained nowadays.

Specifically, CNN, as a traditional deep learning method, is widely applied in image-typed multi-view analysis, such as multi-view 3D object recognition [5] and medical image segmentation [30], [37]. However, most non-image data do not have a grid-like structure, which causes a weak performance of CNN strategy. To adapt to a wider range of multi-view data, the application of GCN is emphasized [31], [38]. The principle of GCN is to transfer and update the node feature information on the network structure, which requires a network-structured data representation as the prerequisite. GAN is a common generative model usually for semi-supervised learning. Overall, the majority of current improvements in GANs within the multi-view domain are focused on improving architectural designs and refining training processes [39], [40], while the methods for encoding and decoding information are still relatively basic, limiting their ability to extract and integrate information of the data features. Utilizing convolutional layers as a component of the generator and discriminator is a promising idea for GAN improvement [41], [42].

Some deep learning methods have noticed the importance of the association patterns between multi-view data. However, due to the black-box nature of deep learning, and the lack of mathematical models for convolution designing, it is difficult to show the adaptability of the trained models to application scenarios. In this context, some studies have attempted to use deep learning models to improve traditional multi-view learning, thereby preserving the interpretability of the method [43], [44]. However, previously described problems such as insufficient view information fusion remain unsolved.

III. Methods

This section divides the views describing the same system into two categories. The macro-view network reflects the explicit manifestations, while the micro-view network delves into the underlying influencing factors. We then study the micro-macro information mapping to explore the factor-manifestation effect pattern. To this end, we first describe a structural information

mapping model in Section III-A, and then Section III-B proposes a deep learning model to implement the model idea. Section III-C demonstrates the application of the method to sample classification and prediction of data evolutionary patterns.

A. Structural Information Mapping Model

For a sample described by macro-view A and micro-view B, data, the structural information can be presented by two matrices $\boldsymbol{S}_A \in \mathbb{R}^{n_A \times \ell_A}$ and $\boldsymbol{S}_B \in \mathbb{R}^{n_B \times \ell_B}$, where n_A and n_B correspond to the feature number under a particular view, and ℓ_A and ℓ_B are the uniform length of features. Introducing an indicator t (initialized as 0) to count information aggregation times, the model elements listed as follows.

1. *Micro-network:* $G_A = \langle \boldsymbol{V}_A, \boldsymbol{S}_A^{(t)} \rangle$, where $\boldsymbol{V}_A$ is the vertex set of the micro-network. This network is constructed based on micro-view feature matrix. $\boldsymbol{S}_{A_i}^{(t)}$ is the i-th row vector of matrix $\boldsymbol{S}_A^{(t)}$, which corresponds to a micro-view feature.
2. *Macro-network:* $G_B = \langle \boldsymbol{V}_B, \boldsymbol{S}_B^{(t)} \rangle$. Similarly, $\boldsymbol{V}_B$ and $\boldsymbol{W}_B^{(t)}$ are the vertex set and weight matrix, and $\boldsymbol{S}_{B_i}^{(t)}$ corresponds to the i-th macro-view feature.
3. Key micro-network: $G_A^K = \langle \boldsymbol{V}_A, \boldsymbol{S}_A^{(T)} \rangle$, where T is the number of structural information aggregating, $\boldsymbol{S}_A^{(T)}$ and $\boldsymbol{W}_A^{(T)}$ denote the updated structural information matrix.
4. Key macro-network: $G_B^K = \langle \boldsymbol{V}_B, \boldsymbol{S}_{BM}^{(0)} \rangle$. $\boldsymbol{V}_B$ is the vertex set and $\boldsymbol{S}_{BM}^{(0)}$ is the mapped structural information matrix. This network is obtained through the structure information mapping
5. Mapping macro-network: $G_B^{Map} = \langle \boldsymbol{V}_B, \boldsymbol{S}_{BM}^{(u)} \rangle$, where $\boldsymbol{S}_{BM}^{(u)}$ denotes the structure information matrix obtained from the u times of structural information diffusion. After T times of information diffusion in total, $\boldsymbol{S}_{BM}^{(T)}$ is finalized as the structural information mapping result.

For a micro-network $\boldsymbol{S}_i^{(0)}$, the information mapping model is presented as follows:

$$\begin{cases} \boldsymbol{S}_{A_i}^{(t)} = \alpha \sum_{k=1}^{n_\mathrm{A}} \boldsymbol{S}_{A_k}^{(t-1)} + \boldsymbol{S}_{A_i}^{(t-1)}, i \in [1, n_\mathrm{A}]; t \in [1, T] & (1) \\ G_\mathrm{A}^K = Graph\left(\boldsymbol{S}_A^{(T)}\right) & (2) \\ \boldsymbol{S}_{BM}^{(0)} = Mapping\left(\boldsymbol{S}_A^{(T)}\right) & (3) \\ \boldsymbol{S}_{BM_j}^{(u)} = \beta \sum_{k=1}^{n_\mathrm{B}} \boldsymbol{S}_{BM_k}^{(u)} + \boldsymbol{S}_{BM_j}^{(u-1)}, j \in [1, n_\mathrm{B}]; \\ \quad u \in [1, T] & (4) \\ G_\mathrm{B}^{Map} = Graph\left(\boldsymbol{S}_{BM}^{(T)}\right) & (5) \end{cases}$$

where (1) denotes that each vertex in Micro-network aggregates the structural information from its neighbors. (2) means the key micro-network is constructed based on the updated structural information. (3) outlines the process of the structural information mapping, where the key macro-network is obtained. (4) presents information diffusion, where the vertices in the key macro-network diffuse the information to their neighborhood. (5) denotes the generation of the mapping macro-network through the diffusion of structural information.

B. Design of the SM-GAN

The model describes the updating and mapping of the structural information matrices of micro-networks and macro-networks. It's essential to note that the structural information matrix is an abstract concept, with each row describing the structural information of a certain node (i.e., feature). The specific information can be represented in different ways, such as node degree, node centrality, etc. In this section, we will proceed to construct a deep learning network named SM-GAN based on this model. In specific, we choose the weight matrix of the network as the structural information matrix, where each row represents the connection strength between a node and all other nodes. We introduce $\boldsymbol{W}_A \in \mathbb{R}^{n_A \times n_A}$ and $\boldsymbol{W}_B \in \mathbb{R}^{n_B \times n_B}$ to respectively represent the correlation weight matrix of the micro-network and macro-network.

1) Generator: We improve the GAN architecture by introducing a structural information mapping module within the traditional encoder-decoder structure. This addition aims to capture the effect pattern from the micro-view to the macro-view. In SM-GAN, the generator is designed with three modules of information encoding, information mapping, and information decoding. These are achieved through by aggregation convolution (AC), mapping convolution (MC), and diffusion convolution (DC) layers, respectively.

The AC operation is designed to realize the aggregation of the structure information in the micro-network, which includes three main steps. First, the structural information of the vertices is updated based on the structural information aggregation:

$$\boldsymbol{E}^{(t)} = \left(\boldsymbol{A}^{(t)} + \boldsymbol{I}^{(t)}\right) \boldsymbol{W}_\mathrm{A}^{(t)} \odot \boldsymbol{P}^{(t)}, \tag{6}$$

where $\boldsymbol{E}^{(t)} \in \mathbb{R}^{n_A \times n_A}$ represents the node-to-edge information aggregation matrix, $\boldsymbol{A}^{(t)} \in \mathbb{R}^{n_A \times n_A}$ is the adjacency matrix, $\boldsymbol{I}^{(t)} \in \mathbb{R}^{n_A \times n_A}$ is an n_A-order identity diagonal matrix, $\boldsymbol{W}_\mathrm{A}^{(t)}$ denotes the edge weight matrix representing the structural information in the micro-network, $\boldsymbol{P}^{(t)} \in \mathbb{R}^{n_A \times n_A}$ denotes the matrix of learnable coefficients in the t-th structural information aggregation, and operator $\odot$ denotes the Hadamard product between two matrices.

Second, the change of weight matrix is calculated as:

$$\Delta \boldsymbol{W}_c^{(t)} = \frac{1}{n_A - 1} \boldsymbol{F} \left(\sum_{q=1, p \neq q}^{n_A} \boldsymbol{E}_{pq}^{(t)} \right) \tag{7}$$

$$\Delta \boldsymbol{W}^{(t)} = \Delta \boldsymbol{W}_c^{(t)} + \left(\Delta \boldsymbol{W}_c^{(t)}\right)^\mathrm{T} \tag{8}$$

where $\boldsymbol{F} \in \{0, 1\}^{n_A \times n_A}$ is an n_A-order matrix with diagonal elements being all 0 while others are all 1, $\boldsymbol{E}_{pq}^{(t)}$ represents the information that vertex p can propagate to edge between p and q, and $\Delta \boldsymbol{W}^{(t)}$ is the weight change matrix. These two equations describe that edges transmit the information to all connected nodes, which, in other words, is an edge-to-node information propagation process.

Third, the weight matrix is updated as follows:

$$\boldsymbol{W}^{(t+1)} = \alpha \Delta \boldsymbol{W}^{(t)} + \boldsymbol{W}^{(t)} \tag{9}$$

where α is the ratio parameter that affects the degree of weight update. Setting the number of AC layers as T, after the in total T times of AC operations, the result network $\boldsymbol{W}_A^{(T)}$ is finalized as the key micro-network.

The MC layers are designed to execute the mapping of structural information from micro-view to macro-view, where the key macro- network will be obtained based on the key micro-network. MC operation can be divided into the following two steps. First, the information mapping matrix $\boldsymbol{W'} \in \mathbb{R}^{n_B \times n_B}$ is calculated by:

$$\boldsymbol{W'} = \boldsymbol{M}^V \boldsymbol{W}_A^{(T)} \boldsymbol{M}^S \tag{10}$$

where parameter matrix $\boldsymbol{M}^V \in \mathbb{R}^{n_B * n_A}$ is for node mapping and $\boldsymbol{M}^S \in \mathbb{R}^{n_A * n_B}$ is for structural information mapping. The former maps n_A vertices in micro-networks to n_B vertices in macro-networks, and the latter maps the n_A-dimension micro-view features to n_B-dimension macro-view features. Both $\boldsymbol{M}^V$ and $\boldsymbol{M}^S$ are learnable.

Then, the mapped weight matrix of the key macro-network is finalized as:

$$\boldsymbol{W}_{BM}^{(0)} = \frac{1}{2}\left(\boldsymbol{W'} + \boldsymbol{W'}^{\mathrm{T}}\right). \tag{11}$$

The DC layers are designed in the information decoding module to realize the structural information diffusion based on the key macro-network. This process encompasses three primary steps. First, the change of vertex structural information is calculated based on the structural information diffusion in the mapped macro-network:

$$\boldsymbol{D}^{(u)} = \boldsymbol{W}_{BM}^{(u)}\left(\boldsymbol{B}^{(u)} + \boldsymbol{I}^{(u)}\right) \odot \boldsymbol{Q}^{(u)} \tag{12}$$

where $\boldsymbol{W}_{BM}^{(u)} \in \mathbb{R}^{n_B * n_B}$ denotes the updated weight matrix after t-th information diffusion, $\boldsymbol{B}^{(u)} \in \mathbb{R}^{n_B * n_B}$ is the adjacency matrix of the key macro-network, $\boldsymbol{I}^{(u)} \in \mathbb{R}^{n_B * n_B}$ is an n_B-order identity matrix, and $\boldsymbol{Q}^{(u)} \in \mathbb{R}^{n_B * n_B}$ is the trainable parameter matrix.

Second, the change of edge weights $\Delta \boldsymbol{W}_{BM}^{(u)}$ is calculated accordingly:

$$\Delta {\boldsymbol{W}'_c}^{(u)} = \frac{1}{n_B - 1}\left(\sum_{p=1, p\neq q}^{n_B} \boldsymbol{D}_{pq}^{(u)}\right)\boldsymbol{F'}, \tag{13}$$

$$\Delta \boldsymbol{W}_{BM}^{(u)} = \Delta {\boldsymbol{W}'_c}^{(u)} + \left(\Delta {\boldsymbol{W}'_c}^{(u)}\right)^T, \tag{14}$$

where $\boldsymbol{F'} \in \{0,1\}^{n_B * n_B}$ denotes an n_B-order square matrix with diagonal all 0 elements while others are filled with the value 1, $\boldsymbol{D}_{pq}^{(u)}$ indicates the information that node p can diffuse to the edge between p and q. These two equations describe the edge-to-node information diffusion process.

Third, the edge weights are updated based on the weight change of the edges:

$$\boldsymbol{W}_{BM}^{(u+1)} = \beta \Delta \boldsymbol{W}_{BM}^{(u)} + \boldsymbol{W}_{BM}^{(u)}, \tag{15}$$

where β is the ratio parameter affecting the degree of weight update. Setting the number of DC layers as T, after the total T times of DC operations, the resulting network $\boldsymbol{W}_A^{(T)}$ is finalized as the key micro-network.

Interpretability has long been a key requirement for enhancing the credibility of deep learning models, and the design of our generator provides a notable improvement in interpretability. Based on mathematical modeling, AC, MC, and DC operations are designed. Through AC operations, the representative structural information is encoded as the key micro-network. The MC layers then facilitate the mapping of multi-view structural information, capturing the correlation pattern between the micro and macro views. Subsequently, the DC operations decode the key macroscopic structural information, generating macro-networks closely aligned with the real macro-networks. These three closely-cooperated modules clearly present the information transmission and mapping in the multi-view networks, reflecting the regulation pathway of microscopic features on macroscopic manifestations, and thus providing new horizons for further research.

2) Discriminator: The discriminator discriminates the real macro-networks with macro-networks reconstructed by the generator. Comprising multiple AC layers and fully connected layers, the discriminator's input is denoted as $\boldsymbol{W}_D^{(0)} \in \{\boldsymbol{W}_A^{(T)}, \boldsymbol{W}_B\}$, where $\boldsymbol{W}_B$ represents the weight matrix of the real macro-networks. The output $\boldsymbol{W}_D^{(T_D)}$ is obtained after T_D times of aggregation convolution, which is described as:

$$\Delta \boldsymbol{E}_D^{(n-1)} = \sum_{q=1, p\neq q}^{n_B}\left(\boldsymbol{U}_D^{(n-1)} \boldsymbol{W}_D^{(n-1)} \odot \boldsymbol{R}_D^{(n-1)}\right)_{pq}, \tag{16}$$

$$\Delta \boldsymbol{W}_D^{(n-1)} = \frac{1}{n_B - 1}\left(\boldsymbol{F}_D \Delta \boldsymbol{E}_D^{(n-1)} + \left(\boldsymbol{F}_B \Delta \boldsymbol{E}_D^{(n-1)}\right)^{\mathrm{T}}\right), \tag{17}$$

$$\boldsymbol{W}_D^{(n)} = \alpha \Delta \boldsymbol{W}_D^{(n-1)} + \Delta \boldsymbol{W}_D^{(n-1)}, \tag{18}$$

where $\boldsymbol{U}_D^{(n-1)} \in \mathbb{R}^{n_B * n_B}$ denotes the aggregation matrix of the n-order aggregation, $\boldsymbol{W}_D^{(n-1)} \in \mathbb{R}^{n_B * n_B}$ denotes the weight matrix after the (n-1)-order aggregation convolution, $\boldsymbol{R}_D^{(n-1)} \in \mathbb{R}^{n_B * n_B}$ is the learnable matrix, $\Delta \boldsymbol{E}_D^{(n-1)} \in \mathbb{R}^{n_B * n_B}$ is the information change matrix of the edges in the network, $\boldsymbol{F}_D \in \mathbb{R}^{n_B * n_B}$ represents the edge-to-node information transmission matrix, and $\Delta \boldsymbol{W}_D^{(n-1)} \in \mathbb{R}^{n_B * n_B}$ denotes the change of the weight matrix, and α affects the degree of weight update.

The fully connected layers are for outputting the discrimination result of the network based on the matrix $\boldsymbol{W}_D^{(T_D)}$ finalized by information aggregation:

$$\boldsymbol{X}^{(t+1)} = \begin{cases} ReLU\left(\boldsymbol{W}^{(t)}\boldsymbol{X}^{(t)} + \boldsymbol{b}^{(t)}\right), & \text{if } t \neq T_F - 1 \\ Softmax\left(\boldsymbol{W}^{(t)}\boldsymbol{X}^{(t)} + \boldsymbol{b}^{(t)}\right), & \text{if } t = T_F - 1 \end{cases}, \tag{19}$$

where the input of the fully connected layer is $\boldsymbol{X}^{(0)} = Ran(\boldsymbol{W}_D^{(T_D)})$, $Ran(\cdot)$ is the row-by-row tiling function; T_F is the layer number, and learnable $\boldsymbol{W}^{(t)}$ and $\boldsymbol{b}^{(t)}$ are the weight matrix and bias of the t-th layer.

3) Model Training: The objective function of the adversarial training is as:

$$\min_G \max_D V(G,D) = \min_G \max_D \left\{ \mathbb{E}_{x\sim P_x(x)} \left[\log D(x)\right] + \mathbb{E}_{z\sim P_z(z)} \left[1 - \log\left(D\left(G\left(z\right)\right)\right)\right] \right\} \tag{20}$$

where x is the real macro-network, z is the real micro-network, and $P_x(x)$ and $P_z(z)$ denote the respective distributions of x and z. $G(*)$ signifies the macro-network generation based on the structural information mapping, which corresponds to the operations described by (7)–(16). $D(*)$ corresponds to the discriminator depicted by (17)–(20), whose result indicates the possibility that the discriminator judges the input network as real. The training of SM-GAN unfolds in two steps. First, the discriminator is optimized by solving:

$$\max_D \left\{ \mathbb{E}_{x\sim P_x(x)} \left[\log D(x)\right] + \mathbb{E}_{z\sim P_z(z)} \left[1 - \log D\left(G\left(z\right)\right)\right] \right\}. \tag{21}$$

During the training discriminator, the generator is fitted. Then, the generator is optimized by solving:

$$\min_G \left\{ \mathbb{E}_{z\sim P_z(z)} \left[1 - \log D\left(G\left(z\right)\right)\right] \right\}. \tag{22}$$

In other words, for the training of the generator, the aim is to generate networks that are challenging for the discriminator to distinguish. According to the idea of GAN, the adversarial training practice brings a competitive relationship between the generator and discriminator, which can improve the performance of the converged model [39].

The optimization of SM-GAN relies on backpropagation and Adam algorithm. After each epoch of optimization, we evaluate the model performance using the test set, which is divided into N categories $C_1, \ldots, C_N$ accoding to the need of the task. The entire training algorithm of SM-GAN is outlined in Algorithm 1.

C. The Applications of SM-GAN

1) Multi-Label Classification: The versatility of SM-GAN shines through in various application scenarios, with the flexibility to adapt to different input networks. In instances where the input micro-networks and macro-networks belong to subjects in the same category, SM-GAN will extract the discriminant features among different label categories. The macro-network output from the generator is regarded as the reconstructed macro-networks derived from the micro-view data. In that case, after training, the generator will obtain the ability of automatically generating the macro-networks with different labels, while the discriminator can identify the generated macro-networks.

2) Data Evolution Prediction: In many application scenarios, the collected data have a long-term evolutionary trend. Predicting future evolution from current time point holds significant application value. To implement the data evolution prediction modeling, we divide the multi-view dataset into the early dataset (representing the current state) and the late dataset (representing the subsequent states) based on the data timestamp. Subsequently, the micro-network, serving as the input of the generator, is constructed from the early dataset, while the macro-network as the input of the discriminator is constructed from the late dataset. This approach allows SM-GAN to capture the dynamic evolutionary pattern of the data in the training process. In this case, the mapping macro-networks obtained by the generator are interpreted as the prediction of the data evolution, and the following steps need to be carried out.

First, a similarity matrix S between the weight matrices of the generated and real macro-network is calculated, whose elements are defined as:

$$\boldsymbol{S}_{rc} = \begin{cases} 1, \text{ if } |\boldsymbol{W}_{C_{rc}} - \boldsymbol{W}_{F_{rc}}| \leq 0.1 \\ 0, \text{ if } |\boldsymbol{W}_{C_{rc}} - \boldsymbol{W}_{F_{rc}}| > 0.1 \end{cases} \quad r, c \in [1, N_2], \tag{23}$$

where $\boldsymbol{W}_C$ denotes the weight matrix of the generated macro-network and $\boldsymbol{W}_F$ is that of the real macro-network. Then, the similarity $Q \in [0,1]$ is obtained accordingly:

$$Q = \frac{1}{n_B\left(n_B - 1\right)} \left(\sum_{r=1}^{n_B} \sum_{c=1}^{n_B} \boldsymbol{S}_{rc} \right). \tag{24}$$

This similarity allows us to predict the evolutionary trend of the data. When Q exceeds a certain threshold, the current sample is expected to undergo further evolution, aligning with the state presented in the late dataset. For this paper, we set the threshold as 0.5.

IV. Experiments and Discussions

A. Dataset and Preprocessing

The ADNI dataset stands as a comprehensive multi-view dataset dedicated to facilitating the research of Alzheimer's disease (AD) and related diseases. The data were collected according to a strict standard to ensure the data homology. We used the neuroimaging data and genetic data for the experiment of this work, which can correspond to the macroscopic view and microscopic view. We validated the performance of SM-GAN for the tasks of multi-label classification and prediction of Alzheimer's Disease (AD) progression at the patient level, distinguishing among early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI), and AD. EMCI and LMCI present the prodromal stage of AD, posing a challenging task due to their similaraities in patient characterizations such as imaging data. The data collection was approved and supported by the relevant institutions, and the multi-view data in this paper had been authorized by ANDI.

For multi-label classification, we selected 233 subjects with AD, 197 with EMCI, and 203 with LMCI. For the prediction task, we conducted further sample selection. From EMCI data, we selected 33 subjects having deterioration to the next stage (i.e., LMCI) within the follow-up 5 years. These subjects were named processing EMCI (pEMCI). Then, another 33 subjects with stable conditions were selected and named stable EMCI (sEMCI). Similarly, from the LMCI data, 48 stable LMCI (sLMCI) patients and 48 progressive LMCI (pLMCI) patients were selected. pLMCI subjects deteriorate to AD within 5 years. In such settings, pEMCI and pLMCI are the early datasets described in Section III-C-2, and LMCI and AD are their corresponding late datasets. To facilitate model training, sample

Algorithm 1: The Overall Algorithm of SM-GAN.

Input: Train set $D^{\text{Train}} = \{\boldsymbol{W}_{\text{A}}^{(0)}, \boldsymbol{W}_{\text{B}}^{(0)}\}$; Test set $D^{\text{Test}} = \{\boldsymbol{W}_{\text{A}}^{\text{C}_1}, \ldots, \boldsymbol{W}_{\text{B}}^{\text{C}_N}\}$.
Initialize: Number of convolutional layers n.
Repeat:
  for $t = 1, T$ **do**
    Obtain key gene network $\boldsymbol{W}_{\text{A}}^{(t_1)}$ using $\boldsymbol{W}_{\text{A}}^{(t_1-1)}$ by iterating (6)–(9);
  end for
  Obtain key brain network $\boldsymbol{W}_{\text{BM}}^{(0)}$ using $\boldsymbol{W}_{\text{A}}^{(T)}$ using (10)–(11);
  for $t_2 = 1, \ldots, T$ **do**
    Generate the brain network $\boldsymbol{W}_{\text{BM}}^{(t_2)}$ using $\boldsymbol{W}_{\text{BM}}^{(t_2-1)}$ by iterating (12)–(15);
  end for
  $D \leftarrow \{\boldsymbol{W}_{\text{BM}}^{(T)}, \boldsymbol{W}_{\text{B}}^{(0)}\}$
  For $\boldsymbol{W}_{\text{D}} \in D$ **do**
    Discriminate the category of network $\boldsymbol{W}_{\text{D}}$ through (16)–(19);
  end for
  Optimize the parameters of SM-GAN by solving (20)–(22);
  Input $\{\boldsymbol{W}_{\text{A}}^{\text{C}_1}, \ldots, \boldsymbol{W}_{\text{A}}^{\text{C}_N}\}$ to generator to generate brain networks $\{\boldsymbol{W}_{\text{B}}^{\text{C}_1}, \ldots, \boldsymbol{W}_{\text{B}}^{\text{C}_N}\}$
  Identify $\{\boldsymbol{W}_{\text{B}}^{\text{C}_1}, \ldots, \boldsymbol{W}_{\text{B}}^{\text{C}_N}\}$ and calculate the accuracy Acc;
Until the convergence measured by Acc is satisfied.

numbers were balanced across categories. Table I describes the information of the subjects. Age/gender difference and their significance are evaluated among all subjects. The results showed no significant difference of these two factors.

Each patient has single nucleotide polymorphism (SNP) data as micro-view and functional magnetic resonance imaging (fMRI) data as macro-view. The pre-experimental preparations include the preprocessing of the raw data and the construction of the networks under the two views. For data preprocessing, our objective is to get serialized multi-view features. SNP data are filtered through minor allele frequency (MAF), call rates, Hardy-Weinberg genetic balance, etc., and the candidate 45 genes are settled. This process is executed using the PLINK tool. The genes selected in this paper have been indicated in prior works to play a significant or potential role in Alzheimer's Disease (AD) and other similar brain disorders, affirming the validity of the gene screening process [45], [46], [47]. fMRI data is preprocessed by the DPARSF tool. The focus of fMRI preprocessing is to improve the data quality, including head motion correction, spatial smoothing, time domain filtering, registration and standardization, etc. The Anatomical Automatic Labeling (AAL) template is employed to identify 116 brain regions and 116 serialized features were extracted as macro-view features.

We construct network-based representations of the data using the data features obtained through preprocessing. Microscopic and macroscopic features were separately used to construct micro-networks and macro-networks. The connection weights were calculated using the Pearson correlation coefficient, which was also a validated and effective method [48], [49], [50]:

$$Pear_{v_1,v_2} = \frac{l\sum \boldsymbol{v}_1\boldsymbol{v}_2 - \sum \boldsymbol{v}_1 \sum \boldsymbol{v}_2}{\sqrt{l\sum \boldsymbol{v}_1^2 - (\sum \boldsymbol{v}_1)^2}\sqrt{l\sum \boldsymbol{v}_2^2 - (\sum \boldsymbol{v}_2)^2}} \tag{25}$$

where $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ represent the feature sequences, and l is the sequence length. Note that edges are not presented between every pair of vertices. In contrast, we set an edge weight threshold $\theta \in [0, 1]$, and only edges whose weight was greater than θ were retained. θ is a hyperparameter that needs to be manually adjusted, and its determination is described in the experiments of Section IV-C. The micro-networks in this paper can reveal the association of gene expressions, while the macro-networks illustrate the functional connections between brain regions.

B. Performance in Classification and Evolution Prediction

The multi-label classification performance was evaluated using accuracy (ACC), precision (PRE), recall (REC), and F1-Score (F1). The prediction task was evaluated by ACC, REC, specificity (SPE), and balance accuracy (BAC). These evaluation metrics are formulized as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{26}$$

$$REC = \frac{TP}{TP + FN} \tag{27}$$

$$SPE = \frac{TN}{TN + FP} \tag{28}$$

$$BAC = \frac{SEN + SPE}{2} \tag{29}$$

$$PRE = \frac{TP}{TP + FP}, \tag{30}$$

$$F1 = \frac{TP}{TP + \frac{1}{2}FP + \frac{1}{2}FN} \tag{31}$$

where TP, TN, FP, and FN were true positive, true negative, false positive, and false negative, respectively.

For a comprehensive comparison, we introduced three widely-used deep learning techniques, i.e., Deep Neural Network (DNN), CNN, and GCN, as the generators, obtaining 3 baseline methods denoted as DNN-GAN, CNN-GAN, and GCN-GAN, respectively. Additionally, ablation experiments were conducted, wherein the AC, MC, and DC layers were replaced by plain DNNs, leading to three variants of SM-GAN: SM-GAN (no AC), SM-GAN (no MC), and SM-GAN (no DC). Furthermore, we explored different network construction methods. In the original SM-GAN, the edges in the multi-view networks were constructed based on Pearson's correlation. In this section, we compared this method with alternatives, specifically Kendall and Spearman correlation coefficients.

Table II showed the performance of the above methods across different tasks, and Table III compared the network construction methods. Fig. 2 displayed the corresponding receiver operator

TABLE I
BASIC INFORMATION OF THE SUBJECTS

Groups	EMCI	LMCI	AD	sEMCI	pEMCI	sLMCI	pLMCI
Number	197	203	233	33	33	48	48
Age (Mean±s.d)	73.4±5.4	74.9±7.9	71.5±5.9	74.4±4.7	73.6±5.7	73.9±7.3	74.2±6.9
Gender (Male/Female)	108/89	108/95	151/82	18/15	19/14	25/23	27/21

TABLE II
PERFORMANCE COMPARISON IN IDENTIFICATION AND PREDICTION TASKS

Methods	Multi-label classification task (EMCI/LMCI/AD)				EMCI-to-LMCI prediction task (pEMCI/sEMCI)				LMCI-to-AD prediction task (pLMCI/sLMCI)			
	ACC (%)	PRE (%)	REC (%)	F1 (%)	ACC (%)	REC (%)	SPE (%)	BAC (%)	ACC (%)	REC (%)	SPE (%)	BAC (%)
DNN-GAN	50.78	53.72	90.28	67.36	65.15	70.27	58.62	64.45	68.75	65.45	73.17	69.31
CNN-GAN	55.47	57.72	93.42	71.36	68.18	69.23	66.67	67.95	70.83	71.15	70.45	70.80
GCN-GAN	60.94	61.90	97.50	75.73	71.21	70.59	71.88	71.23	75.00	71.43	78.72	75.08
SM-GAN (no AC)	66.41	68.55	95.51	79.81	75.76	73.68	78.57	76.13	77.08	73.47	80.86	77.16
SM-GAN (no MC)	68.75	70.40	96.70	81.48	74.24	72.73	75.76	74.24	78.13	72.00	84.78	78.39
SM-GAN (no DC)	71.09	72.22	97.85	83.11	77.27	81.25	73.53	77.39	80.21	80.00	80.39	80.20
SM-GAN	**74.22**	**74.80**	**98.96**	**85.20**	**81.82**	**77.78**	**86.67**	**82.22**	**84.48**	**81.82**	**86.54**	**84.18**

TABLE III
PERFORMANCE COMPARISON OF DIFFERENT NETWORK CONSTRUCTING METHODS

Methods	Multi-label classification task (EMCI/LMCI/AD)				EMCI-to-LMCI prediction task (pEMCI/sEMCI)				LMCI-to-AD prediction task (pLMCI/sLMCI)			
	ACC (%)	PRE (%)	REC (%)	F1 (%)	ACC (%)	SEN (%)	SPE (%)	BAC (%)	ACC (%)	SEN (%)	SPE (%)	BAC (%)
Kendall	72.66	78.81	90.29	84.16	75.76	76.67	75.00	75.83	80.21	78.26	82.00	80.13
Spearman	71.09	77.78	89.22	83.11	78.79	80.00	77.42	78.71	81.25	81.40	81.13	81.26
Pearson	**74.22**	**74.80**	**98.96**	**85.20**	**81.82**	**77.78**	**86.67**	**82.22**	**84.48**	**81.82**	**86.54**	**84.18**

characteristic (ROC) plots, illustrating comparisons between different deep learning architectures (a) and network construction methods (b).

Notably, SM-GAN outperforms other approaches in all tested tasks. Specifically, the multi-label classification accuracy of SM-GAN was 78.13% and the evolution prediction accuracies had reached 86.36% (EMCI-to-LMCI prediction) and 88.54% (LMCI-to-AD prediction). Also, it can be observed that the networks constructed by Pearson's correlation analysis were superior to the others.

C. Results of Model Finalization

In this paper, the model was finalized in two aspects. On the one hand, three hyperparameters in SM-GAN were manually optimized: (1) edge weight threshold $\theta \in [0, 1]$ affecting the edge density of the networks. (2) weight update ratio α in each AC operation. (3) weight update ratio β in each DC operation. On the other hand, adversarial training was applied to optimize the parameters in all trainable matrices.

The performance of SM-GAN under different parameters was examined by ACC and the results were shown in Fig. 3. Based on the results, the optimal parameter combination of $\alpha = 0.5$, $\beta = 0.5$, and $\theta = 0.4$ is determined. In Fig. 4(a) and (b), the changes in loss and the root mean squared error (RMSE) between the real and the generated macro-network in the training process are depicted for all 7 comparative methods. The RMSE is calculated by:

$$RMSE\left(\boldsymbol{f}_{Real}, \boldsymbol{f}_{Gen}\right) = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{f}_{Real}^{(i)} - \boldsymbol{f}_{Gen}^{(i)}\right)^2} \tag{32}$$

where $\boldsymbol{f}_{Real}$ and $\boldsymbol{f}_{Gen}$ respectively represents the vector obtained by flattening the weight matrix of the real and generated

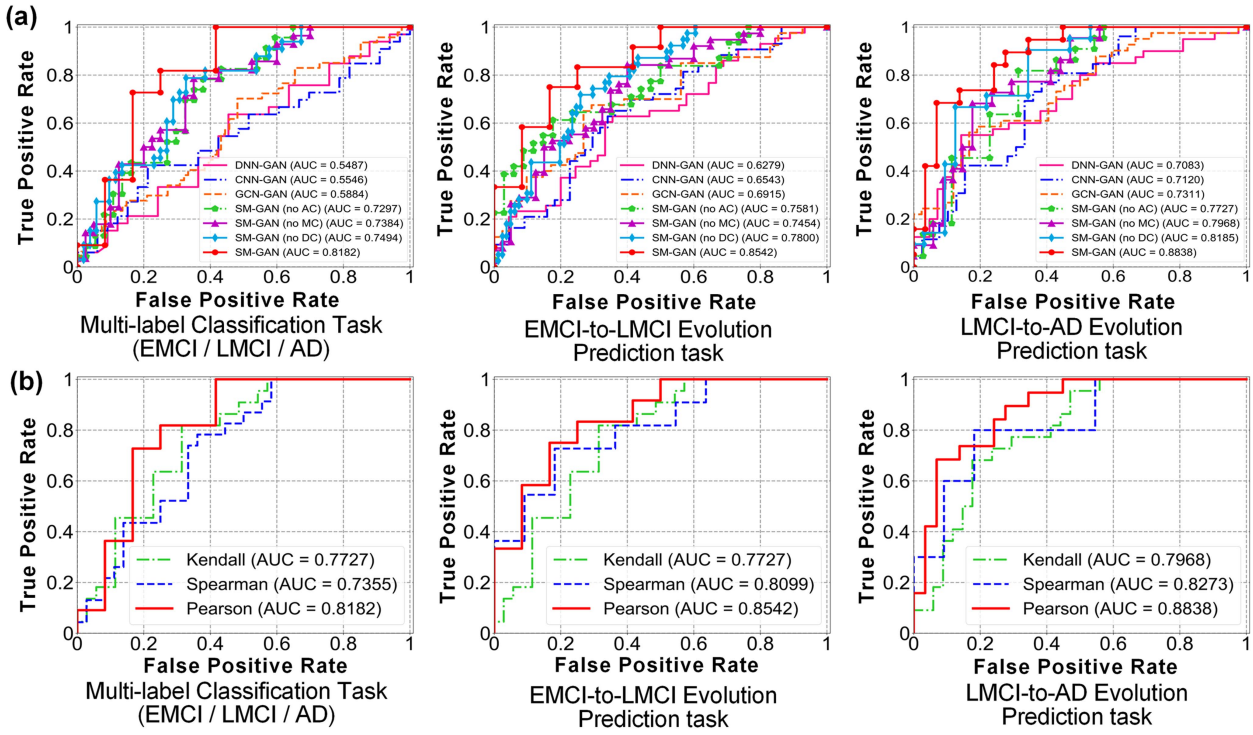


Fig. 2. Comparisons of method performance are depicted through the ROC curves. (a) The comparisons of the deep learning approaches. (b) The comparisons of different network constructing method.

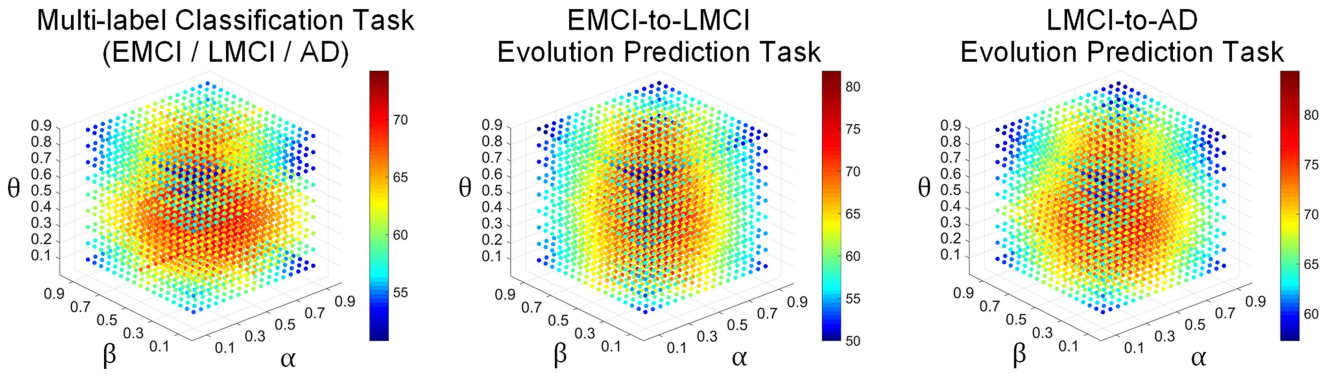


Fig. 3. Effect of parameters on the performance.

network, $\boldsymbol{f}_{Real}^{(i)}$ and $\boldsymbol{f}_{Gen}^{(i)}$ represent the i-th element of the vectors, m is the vector length. In the multi-label classification task, RMSE is obtained from the comparison of real/generated networks of the same category. In the evolution prediction task, the macro-networks generated by SM-GAN was based on the early data but compared with the macro-network of the late data.

As illustrated in Fig. 4, SM-GAN exhibited the earliset convergence and achieved the lowest loss and RMSE values after the training, indicating the superiority of the SM-GAN method.

D. Structural Information Mapping Performance

Two sub-experiments were conducted to compare the multi-view information mapping ability of SM-GAN. First, Fig. 5 visualized the average generated macro-networks with the significant functional connections between brain regions. To show the effectiveness of information mapping, we further calculated the residual matrices and mean absolute error (MAE) between the generated and real networks. For the multi-label classification, as shown in Fig. 5(a), the generated macro-networks were compared with the real macro-networks within the same category. The comparison in evolution prediction is depicted in Fig. 5(b). Using the micro-view data of pEMCI and pLMCI, the corresponding generated macro-networks are compared with the real macro-networks of LMCI and AD subjects. As shown in the results, the macro-networks generated through structural information mapping had the best similarity to the ground truth and the smallest MAE values among all comparative methods.

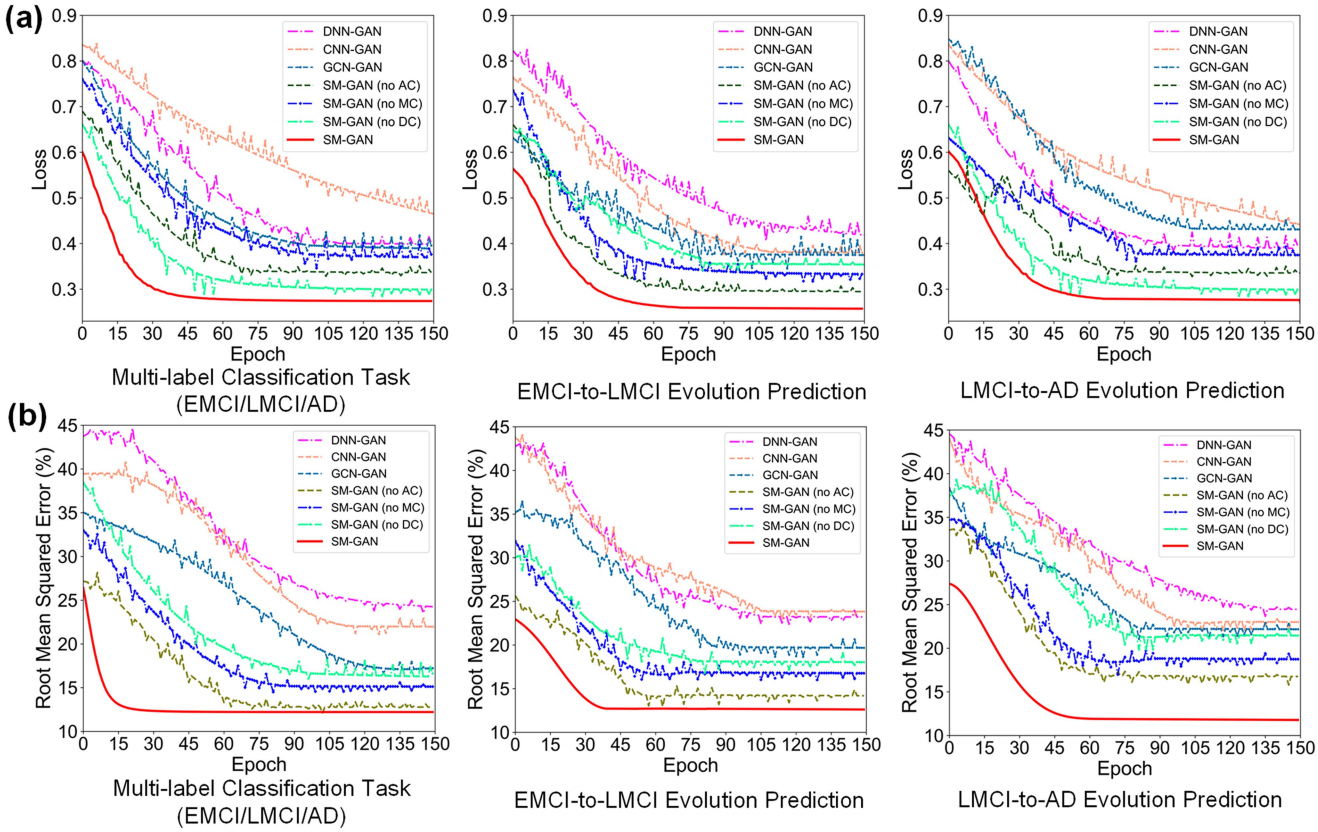


Fig. 4. Training of SM-GAN. (a) The training performance illustrated by loss change. SM-GAN has the smallest fluctuations, fastest speed, and lowest loss value after convergence. (b) The training performance comparisons are illustrated by the root mean squared error (RWSE) describing the difference between the generated and real networks. Similarly, the results indicate the superiority of SM-GAN.

SM-GAN accurately captured the nonlinear information mapping pattern from the micro-view to the macro-view data.

Second, this paper use four metrics [51] to measure the topological properties of real and generated macro-networks. Denoting the node number as N, the metrics were descriced as follows:

1. *Edge number (EN):* Number of the connecting edges in the networks, which can affect the network density.
2. *Global efficiency (GE):* This metric describes the transmission capability of network information:

$$GE = \frac{1}{N\left(N-1\right)} \sum_{i \neq j} d_{ij}^{-1} \tag{33}$$

where d_{ij} represents the length of shortest path from node i to node j.

3. *Characteristic path length (CPL):* This metric refers to the average distance between one vertex to another, representing the information transferring ability. The definition equation is:

$$CPL = \frac{1}{N\left(N-1\right)} \sum_{i \neq j} d_{ij} \tag{34}$$

4. *Clustering coefficient (CC):* The CC value of a network is calculated by averaging of the CC values of the vertices in the network. This metric evaluates the ability of local information transmission, which is defined as:

$$CC = \frac{1}{N} \sum_{i} \frac{t_i}{k_i\left(k_i-1\right)/2} \tag{35}$$

where k_i denotes the neighoring vertex number of vertex i and t_i represents the number of the existing edges between all neighboring vertex of i.

We assessed the similarity between the generated network and the real network using the selected metrics, thereby validating the information mapping capability of SM-GAN. Fig. 6 showed the comparing results. For denotations, I-SM-EMCI, I-SM-LMCI, and I-SM-AD denote the generated macro-networks based on the micro-view data of EMCI, LMCI, and AD samples; P-SM-LMCI and P-SM-AD denote the generated macro-networks based on the micro-view data of EMCI and LMCI samples, respectively.

Table IV further showed the p-value results of t-tests to indicate the significance of the metric difference. As shown, the real macro-networks in different classes had significant differences from each other, indicating that the selected metrics could effectively distinguish samples. In addition, no significant differences were observed between the real/generated macro-networks within the same category, indicating that SM-GAN can effectively perform cross-view information mapping.

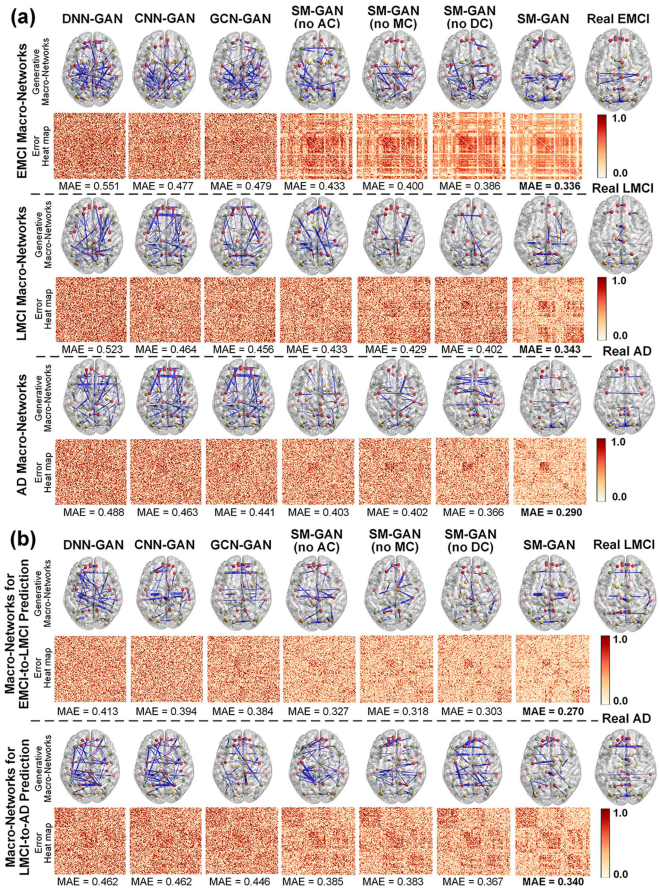


Fig. 5. Visualization of the generated macro-networks and residual matrices. (a) Results for 3-label classification task. (b) Results for risk prediction task. The generated networks are based on the early data (EMCI & LMCI) while the real networks are constructed from the late data (LMCI & AD). From the experimental results, functional connections exist between every pair of brain regions. To better illustrate the characteristic association patterns corresponding to different stages of disease development, the edge weights below the threshold θ are reset to 0.

TABLE IV
T-TEST RESULTS OF NETWORK TOPOLOGICAL PROPERTIES

Methods	EN	GE	CPL	CC
EMCI / LMCI	$3.61*10^{-4}$	$3.61*10^{-4}$	$3.61*10^{-4}$	$3.61*10^{-4}$
EMCI / AD	$5.06*10^{-3}$	$5.06*10^{-3}$	$5.06*10^{-3}$	$5.06*10^{-3}$
LMCI / AD	$3.61*10^{-3}$	$3.61*10^{-3}$	$3.61*10^{-3}$	$3.61*10^{-3}$
EMCI / I-SM-EMCI	$2.99*10^{-1}$	$2.99*10^{-1}$	$2.99*10^{-1}$	$2.99*10^{-1}$
LMCI / I-SM-LMCI	$5.30*10^{-2}$	$5.30*10^{-2}$	$5.30*10^{-2}$	$5.30*10^{-2}$
AD / I-SM-AD	$9.20*10^{-2}$	$9.20*10^{-2}$	$9.20*10^{-2}$	$9.20*10^{-2}$
EMCI / P-SM-LMCI	$5.24*10^{-3}$	$5.24*10^{-3}$	$5.24*10^{-3}$	$5.24*10^{-3}$
LMCI / P-SM-LMCI	$1.09*10^{-3}$	$1.09*10^{-3}$	$1.09*10^{-3}$	$1.09*10^{-3}$
LMCI / P-SM-AD	$2.72*10^{-1}$	$2.72*10^{-1}$	$2.72*10^{-1}$	$2.72*10^{-1}$
AD / P-SM-AD	$6.15*10^{-1}$	$6.15*10^{-1}$	$6.15*10^{-1}$	$6.15*10^{-1}$

EN: edge number; GE: global efficiency; CPL: characteristic path length; CC: clustering coefficient.

E. Effectiveness of the Designed Convolution Operations

The effectiveness of AC, MC, and DC operations was validated by two experiments. For the classification task, we defined $W_{ac,mc,dc}$ as the macro-network obtained by the original SM-GAN, $W_{ac,mc}$ as the macro-network obtained by the SM-GAN variant with DC layers replaced by DNN, W_{ac} as that obtained by the variant with DC and MC layers replaced by DNN, and W were obtained through the SM-GAN with all three types of convolution layers replaced by DNN. Fig. 7 presented the t-SNE results to visualize the samples in the test set. As observed, the t-SNE results of the originally proposed SM-GAN ($W_{ac,mc,dc}$) owned the best clustering effect, indicating the proposed convolution operations improved the ability of sample identification.

To visualize the effectiveness of the proposed convolution in the evolution prediction task, we compared the similarity between the network generated based on early data and the real network of late data. The comparison was conducted between SM-GAN and 3 variants. Specifically, the weight matrix of the network output by the generator in the prediction task was denoted as prediction matrix P, with the name of the late stage as the superscript and the convolutional layers contained in the generator as the subscript. For example, the prediction matrix generated by the original SM-GAN (containing all convolutional layers) in the EMCI-to-LMCI prediction task is denoted as $P^{LMCI}_{AC,MC,DC}$. Also, we calculated the corresponding residual matrices and MAE values between prediction matrices and real networks. All results were displayed in Fig. 8. It can be concluded that the original SM-GAN can generate the best prediction results, having the lowest MAE compared with the real data, which demonstrated the effectiveness of the proposed convolution operations.

F. Comparison With Commonly-used Methods

To provide a more clear indication of the effectiveness of the proposed approach, SM-GAN was compared with several state-of-art studies for AD prodromal stage identification [52], [53], [54], [55], [56] and disease risk prediction, [20], [27], [56], [57], [58] as shown in Table V. We noted that most multi-label classification studies were generally based on normal control (NC), MCI, and AD. However, the similarities between prodromal stages of disease made the classification in this paper more challenging. In this context, SM-GAN still performed the best in both sample identification and evolution prediction tasks.

G. Limitations and Future Works

SM-GAN has achieved promosing performance for the AD progression modeling in this work, demonstrating the effectivness and feasibility for mining the influence pattern from micro-views to macro-views from the perspective of structural information mapping. In this work, SM-GAN focused on the view-level information mapping, while the mapping pattern between specific features remains unknown (e.g., the patterns about how a specific gene influences a particular brain region). We believe that exploring this specific and detailed multi-view information mapping model is a promising direction with great

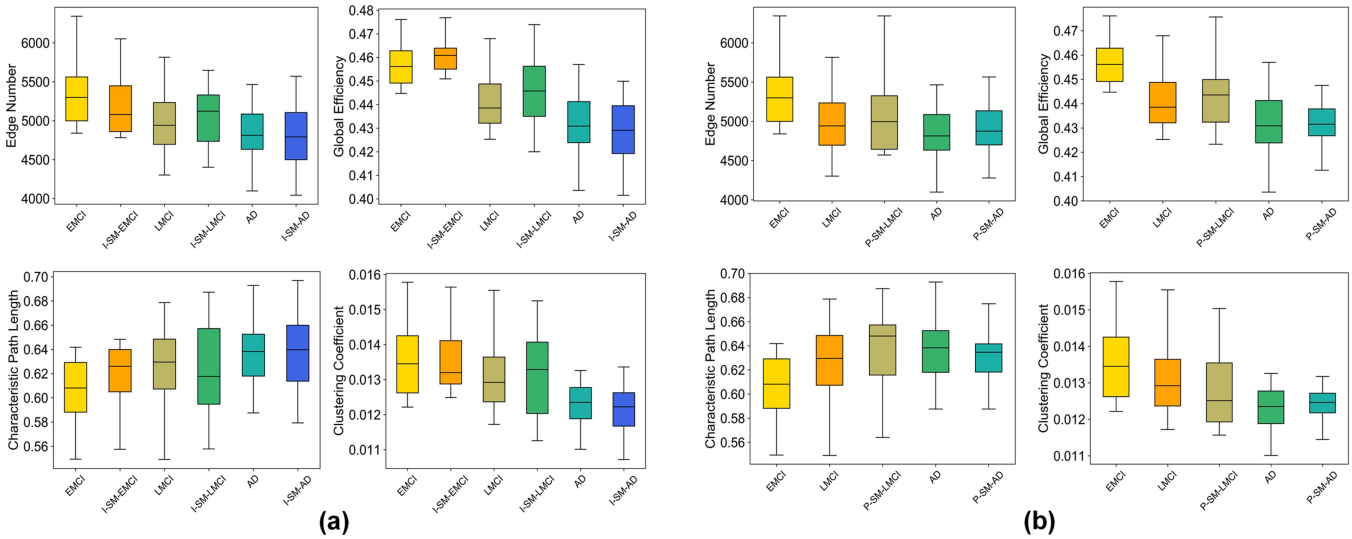


Fig. 6. Box plots comparing the generated and real macro-networks. (a) Comparison for multi-label classification tasks. (b) Comparison for evolution prediction tasks.

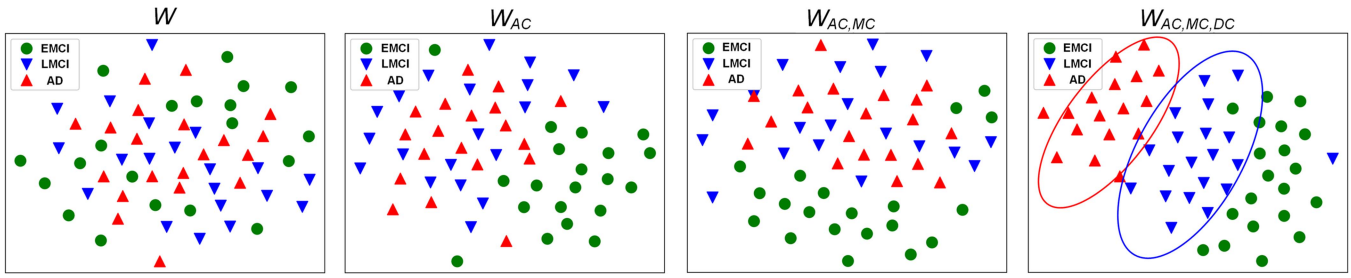


Fig. 7. The t-SNE results of the subjects in three categories under three variants of SM-GAN and the original SM-GAN.

TABLE V
PERFORMANCE COMPARISON WITH THE STATE-OF-ART APPROACHES

Tasks	Methods	Datasets	Subjects	ACC (%)
Multi-label Classification	Wang et al. (2019)	fMRI	154NC/310MCI/90AD	71.8
	Fan et al.(2022)	MRI	100NC/100MCI/100AD	56.1
	Tong et al. (2017)	MRI/PET/CSF	204NC/374MCI/80AD	60.2
	Zhu et al. (2018)	MRI/PET	101NC/118MCI/93AD	72.6
	Zhou et al (2019)	MRI/PET/SNP	204NC/362MCI/171AD	59.8
	Proposed	fMRI/SNP	197EMCI/203LMCI/233AD	74.2
Evolution Prediction	Ashtari-Majlan et al. (2022)	MRI	164pMCI/ 100sMCI	79.9
	Pan et al. (2021)	MRI/PET	40pMCI/ 33sMCI	80.0
	Zhou et al (2019)	MRI/PET/SNP	205pMCI/157pMCI	74.3
	Caod et al. (2022)	MRI/SNP	56pMCI/36sMCI	62.1
	Shang et al (2023)	fMRI/SNP	128pMCI/128sMCI	74.3
	Proposed	fMRI/SNP	197EMCI/203LMCI/33sEMCI/33pEMCI	81.8
			203LMCI/233AD/48sMCI/48pMCI	84.5

PET: Positron emission computed tomography; CSF: Cerebro-Spinal Fluid.

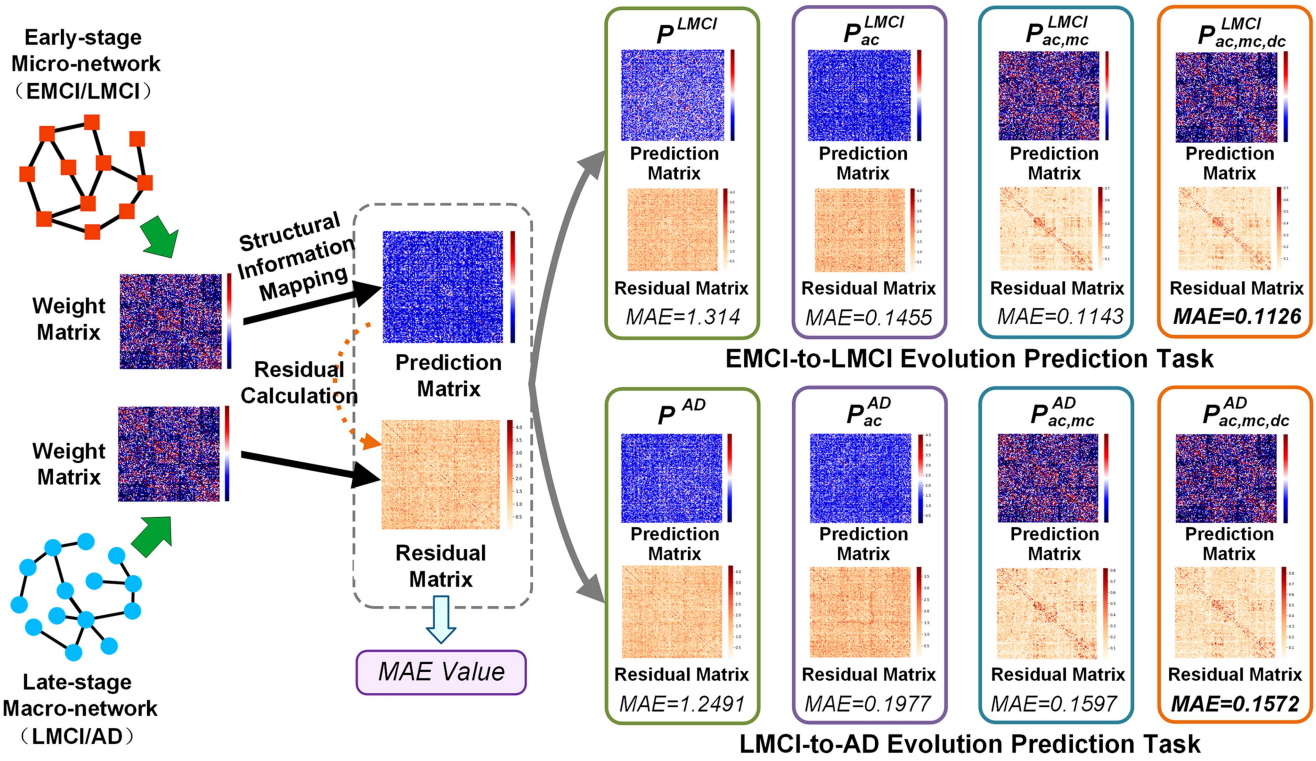


Fig. 8. Comparation of the generated/real macro-networks in the evolution prediction tasks.

practical value. We are now working on several related works aiming to portray a more detailed and specific multi-view information mapping model.

The experiment setting of this work utilizes public-domain biomedical data, which were collected and preprocessed under protocols with better quality controll and curation than in clinical practice. We acknowledge that the current data size and diversity might be insufficient to fully validate the generalization performance of SM-GAN. We are actively exploring the feasibility of applying the SM-GAN model to larger-scale datasets collected in a clinical practice setting. By training on a larger, more diversified dataset, we can develop a more stable and generalizble multi-view data mining framework.

V. Conclusion

This paper proposes a novel multi-view learning method to explore the role of micro-views on macro-views from the perspective of view information mapping. In this paper, the main contributions involve a structural information mapping model and a GAN-based deep learning framework called SM-GAN. The generator includes an information mapping module based on the regular "encoder-decoder" structure to generate a macro-network based on the micro-network. We conduct rich experimental validation on ADNI imaging genetic data, and demonstrate the effectiveness of SM-GAN on the tasks of multi-label classification and data evolution prediction.

References

[1] X. Liu et al., "Late fusion incomplete multi-view clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 10, pp. 2410–2423, Oct. 2019.
[2] X. Yan, S. Hu, Y. Mao, Y. Ye, and H. Yu, "Deep multi-view learning methods: A review," *Neurocomputing*, vol. 448, pp. 106–129, 2021.
[3] K. Karthik and S. S. Kamath, "A deep neural network model for content-based medical image retrieval with multi-view classification," *Vis. Comput.*, vol. 37, no. 7, pp. 1837–1850, 2021.
[4] J. Huang, W. Yan, G. Li, T. Li, and S. Liu, "Learning disentangled representation for multi-view 3D object recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 2, pp. 646–659, Feb. 2022.
[5] K. Sun, J. Zhang, J. Liu, R. Yu, and Z. Song, "DRCNN: Dynamic routing convolutional neural network for multi-view 3D object recognition," *IEEE Trans. Image Process.*, vol. 30, pp. 868–877, 2020.
[6] A. Paul et al., "Generalized zero-shot chest X-ray diagnosis through trait-guided multi-view semantic embedding with self-training," *IEEE Trans. Med. Imag.*, vol. 40, no. 10, pp. 2642–2655, Oct. 2021.
[7] N. Mushtaq et al., "Brain tumor segmentation using multi-view attention based ensemble network," *Comput., Mater. Continua*, vol. 72, no. 3, pp. 5793–5806, 2022.
[8] J. Tang, H. He, S. Fu, Y. Tian, G. Kou, and S. Xu, "Robust multi-view learning with the bounded LINEX loss," *Neurocomputing*, vol. 518, pp. 384–400, 2023.
[9] X. Ning et al., "A review of research on co-training," *Concurrency Computation: Pract. Experience*, vol. 35, 2021, Art. no. e6276.
[10] P. Yildirim Taser, "A novel multi-view ordinal classification approach for software bug prediction," *Expert Syst.*, vol. 39, no. 7, 2022, Art. no. e13044.
[11] O. Graa and I. Rekik, "Multi-view learning-based data proliferator for boosting classification using highly imbalanced classes," *J. Neurosci. Methods*, vol. 327, 2019, Art. no. 108344.
[12] Q. Ye, P. Huang, Z. Zhang, Y. Zheng, L. Fu, and W. Yang, "Multiview learning with robust double-sided twin SVM," *IEEE Trans. Cybern.*, vol. 52, no. 12, pp. 12745–12758, Dec. 2022.
[13] Y. Hao, X.-Y. Jing, R. Chen, and W. Liu, "Learning enhanced specific representations for multi-view feature learning," *Knowl.-Based Syst.*, vol. 272, 2023, Art. no. 110590.

[14] A. Khan and P. Maji, "Multi-manifold optimization for multi-view subspace clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 3895–3907, Aug. 2022.

[15] M. Heilbron, K. Armeni, J.-M. Schoffelen, P. Hagoort, and F. P. De Lange, "A hierarchy of linguistic predictions during natural language comprehension," *Proc. Nat. Acad. Sci. USA*, vol. 119, no. 32, 2022, Art. no. e2201968119.

[16] X. Chen, C.-C. Zhu, and J. Yin, "Ensemble of decision tree reveals potential miRNA-disease associations," *PLoS Comput. Biol.*, vol. 15, no. 7, 2019, Art. no. e1007209.

[17] Y. Zhao, J. Wang, J. Chen, X. Zhang, M. Guo, and G. Yu, "A literature review of gene function prediction by modeling gene ontology," *Front. Genet.*, vol. 11, 2020, Art. no. 400.

[18] R. Su, H. Wu, B. Xu, X. Liu, and L. Wei, "Developing a multi-dose computational model for drug-induced hepatotoxicity prediction based on toxicogenomics data," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 16, no. 4, pp. 1231–1239, 2018.

[19] C. Lian, M. Liu, J. Zhang, and D. Shen, "Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 880–893, Apr. 2020.

[20] Y. Pan, M. Liu, Y. Xia, and D. Shen, "Disease-image-specific learning for diagnosis-oriented neuroimage synthesis with incomplete multi-modality data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 6839–6853, Oct. 2022.

[21] M. Kim, R. Wu, X. Yao, A. J. Saykin, J. H. Moore, and L. Shen, "Identifying genetic markers enriched by brain imaging endophenotypes in Alzheimer's disease," *BMC Med. Genomic.*, vol. 15, no. 2, pp. 1–12, 2022.

[22] L. Du et al., "A novel SCCA approach via truncated ℓ 1-norm and truncated group lasso for brain imaging genetics," *Bioinformatics*, vol. 34, no. 2, pp. 278–285, 2018.

[23] L. Du et al., "Associating multi-modal brain imaging phenotypes and genetic risk factors via a dirty multi-task learning method," *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3416–3428, Nov. 2020.

[24] V. Pasquadibisceglie, A. Appice, G. Castellano, and D. Malerba, "A multi-view deep learning approach for predictive business process monitoring," *IEEE Trans. Serv. Comput.*, vol. 15, no. 4, pp. 2382–2395, Jul./Aug. 2022.

[25] H. Gao, Y. Liu, and S. Ji, "Topology-aware graph pooling networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4512–4518, Dec. 2021.

[26] L. Cai, J. Li, J. Wang, and S. Ji, "Line graph neural networks for link prediction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 5103–5113, Sep. 2022.

[27] J. Shang et al., "GCCN: Graph capsule convolutional network for progressive mild cognitive impairment prediction and pathogenesis identification based on imaging genetic data," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 6, pp. 2968–2979, Jun. 2023.

[28] X. Wang, D. Peng, P. Hu, and Y. Sang, "Adversarial correlated autoencoder for unsupervised multi-view representation learning," *Knowl.-Based Syst.*, vol. 168, pp. 109–120, 2019.

[29] X. Song et al., "Graph convolution network with similarity awareness and adaptive calibration for disease-induced deterioration prediction," *Med. Image Anal.*, vol. 69, 2021, Art. no. 101947.

[30] A. Rafi, T. M. Madni, U. I. Janjua, M. J. Ali, and M. N. Abid, "Multi-level dilated convolutional neural network for brain tumour segmentation and multi-view-based radiomics for overall survival prediction," *Int. J. Imag. Syst. Technol.*, vol. 31, no. 3, pp. 1519–1535, 2021.

[31] P. Trirat, S. Yoon, and J.-G. Lee, "MG-TAR: Multi-view graph convolutional networks for traffic accident risk prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 3779–3794, Apr. 2023.

[32] J. Zhao, X. Xie, X. Xu, and S. Sun, "Multi-view learning overview: Recent progress and new challenges," *Inf. Fusion*, vol. 38, pp. 43–54, 2017.

[33] W. Cai, H. Zhou, and L. Xu, "A multi-view co-training clustering algorithm based on global and local structure preserving," *IEEE Access*, vol. 9, pp. 29293–29302, 2021.

[34] X. Liu et al., "Multiple kernel *k*-means with incomplete kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 5, pp. 1191–1204, May 2020.

[35] L. Du et al., "Detecting genetic associations with brain imaging phenotypes in Alzheimer's disease via a novel structured SCCA approach," *Med. image Anal.*, vol. 61, 2020, Art. no. 101656.

[36] D. Kumar and P. Maji, "Discriminative deep canonical correlation analysis for multi-view data," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 2, 2023, doi: 10.1109/TNNLS.2023.3277633.

[37] D. M. Vigneault, W. Xie, C. Y. Ho, D. A. Bluemke, and J. A. Noble, "Ω-net (omega-net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks," *Med. Image Anal.*, vol. 48, pp. 95–106, 2018.

[38] G. Qu et al., "Ensemble manifold regularized multi-modal graph convolutional network for cognitive ability prediction," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 12, pp. 3564–3573, Dec. 2021.

[39] A. Semwal and N. D. Londhe, "MVFNet: A multi-view fusion network for pain intensity assessment in unconstrained environment," *Biomed. Signal Process. Control*, vol. 67, 2021, Art. no. 102537.

[40] D. W. Shu, W. Jang, H. Yoo, H.-C. Shin, and J. Kwon, "Deep-plane sweep generative adversarial network for consistent multi-view depth estimation," *Mach. Vis. Appl.*, vol. 33, pp. 1–10, 2022.

[41] T. Sasagawa, S. Kawai, and H. Nobuhara, "Recommendation system based on generative adversarial network with graph convolutional layers," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 25, no. 4, pp. 389–396, 2021.

[42] D. Valsesia, G. Fracastoro, and E. Magli, "Learning localized representations of point clouds with graph-convolutional generative adversarial networks," *IEEE Trans. Multimedia*, vol. 23, pp. 402–414, 2020.

[43] R. Mariappan and V. Rajan, "Deep collective matrix factorization for augmented multi-view learning," *Mach. Learn.*, vol. 108, pp. 1395–1420, 2019.

[44] S. Affeldt, L. Labiod, and M. Nadif, "Spectral clustering via ensemble deep autoencoder learning (SC-EDAE)," *Pattern Recognit.*, vol. 108, 2020, Art. no. 107522.

[45] H.-R. Kim, T. Lee, J. K. Choi, and Y. Jeong, and A. S. D. N. Initiative, "Polymorphism in the MAGI2 Gene Modifies the Effect of Amyloid β on Neurodegeneration," *Alzheimer Dis. Assoc. Disord.*, vol. 35, no. 2, pp. 114–120, 2021.

[46] B. W. Kunkle et al., "APOE-stratified genome-wide association analysis identifies novel Alzheimer disease candidate risk loci for African Americans," *Alzheimer's Dement.*, vol. 17, 2021, Art. no. e056383.

[47] M. Blume et al., "Dab1 contributes differently to the morphogenesis of the hippocampal subdivisions," *Develop. Growth Differentiation*, vol. 59, no. 8, pp. 657–673, 2017.

[48] J. Fang et al., "Fast and accurate detection of complex imaging genetics associations based on greedy projected distance correlation," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 860–870, Apr. 2018.

[49] P. Peng et al., "Group sparse joint non-negative matrix factorization on orthogonal subspace for multi-modal imaging genetics data analysis," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 19, no. 1, pp. 479–490, 2020.

[50] M. Wang, T.-Z. Huang, J. Fang, V. D. Calhoun, and Y.-P. Wang, "Integration of imaging (epi) genomics data for the study of schizophrenia using group sparse joint nonnegative matrix factorization," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 17, no. 5, pp. 1671–1681, Sep./Oct. 2020.

[51] C. J. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens, "Small-world networks and functional connectivity in Alzheimer's disease," *Cereb. Cortex*, vol. 17, no. 1, pp. 92–99, 2007.

[52] M. Wang, C. Lian, D. Yao, D. Zhang, M. Liu, and D. Shen, "Spatial-temporal dependency modeling and network hub detection for functional MRI analysis via convolutional-recurrent network," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 8, pp. 2241–2252, Aug. 2020.

[53] C.-C. Fan et al., "TR-GAN: Multi-session future MRI prediction with temporal recurrent generative adversarial Network," *IEEE Trans. Med. Imag.*, vol. 41, no. 8, pp. 1925–1937, Aug. 2022.

[54] T. Tong, K. Gray, Q. Gao, L. Chen, D. Rueckert, and A. S. D. N. Initiative, "Multi-modal classification of Alzheimer's disease using nonlinear graph fusion," *Pattern Recognit.*, vol. 63, pp. 171–181, 2017.

[55] Y. Zhu, X. Zhu, M. Kim, J. Yan, D. Kaufer, and G. Wu, "Dynamic hypergraph inference framework for computer-assisted diagnosis of neurodegenerative diseases," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 608–616, Feb. 2019.

[56] T. Zhou, M. Liu, K.-H. Thung, and D. Shen, "Latent representation learning for Alzheimer's disease diagnosis with incomplete multi-modality neuroimaging and genetic data," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2411–2422, Oct. 2019.

[57] M. Ashtari-Majlan, A. Seifi, and M. M. Dehshibi, "A multi-stream convolutional neural network for classification of progressive MCI in Alzheimer's disease using structural MRI images," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 8, pp. 3918–3926, Aug. 2022.

[58] L. W. Caod, J. E. Galvinf, and M. F. Bega, "Machine learning based multimodal neuroimaging genomics dementia score for predicting future conversion to Alzheimer's," *J. Alzheimers Dis.*, vol. 87, no. 3, pp. 1345–1365, 2022.

Xia-an Bi (Member, IEEE) received the PhD degree in computer applications from Hunan University, in 2012. He is currently a professor with the College of Information Science and Engineering, Hunan Normal University. His current research interests include machine learning, brain science, and artificial intelligence. He is an author or coauthor of several technical papers including several ESI Highly Cited Papers according to the most recent Clarivate Analytics ESI report and also a very active reviewer for many international journals and conference. He is a member of MICCAI and CCF.

YangJun Huang received the BE degree in geographic information science from the Central South University of Forestry and Technology, Changsha, China, in 2023. He is currently working toward the MS degree with the College of Information Science and Engineering, Hunan Normal University, Changsha, China. The major of him is computer technology. His main research interests include fields include data mining, brain science, and artificial intelligence.

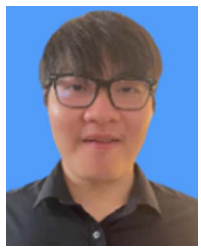
Zicheng Yang received the BE degree in software engineering from the Hubei University of Economics, Wuhan, China, in 2023. He is currently working toward the MS degree with the College of Information Science and Engineering, Hunan Normal University, Changsha, China. The major of him is computer technology. His main research interests include fields include data mining, brain science, and artificial intelligence.

Ke Chen received the BE degree in Internet of Things engineering from Anhui Polytechnic University, Wuhu, China, in 2021. He is currently working toward the MS degree with the College of Information Science and Engineering, Hunan Normal University, Changsha, China. The major of him is computer technology. His main research interests include data mining, brain science, and artificial intelligence.

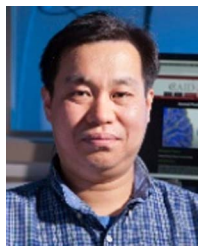
Zhaoxu Xing received the BE degree in computer science and technology from Hunan Normal University, Changsha, China, in 2019. He is currently working toward the the MS degree with the College of Information Science and Engineering, Hunan Normal University, Changsha, China. The major of him is computer science and technology. His main research interests include data mining, brain science, and artificial intelligence.

Luyun Xu received the PhD degree in business administration from Hunan University, in 2018. She is currently an associate professor with Business School, Hunan Normal University. Her research interests include focus on knowledge management, data mining, machine learning and brain science. She has published in the *Journal of Technology Transfer, Technological Analysis & Strategic Management*, *Computational and Mathematical Organization Theory, Briefings in Bioinformatics*.

Xiang Li is an instructor of Investigation with Massachusetts General Hospital and Harvard Medical School, Center for Advanced Medical Computing and Analysis (CAMCA). He joined MGH in 2016 as a research fellow, mentored by Dr. James Thrall, Chairman Emeritus of MGH Department of Radiology. His research interests include developing artificial intelligence solutions for analyzing healthcare data, especially fusion across imaging and non-imaging data, and developing medical informatics systems for smart data management and AI deployment in the clinical workflow.



Zhengliang Liu received the BA and MS degrees in computer science from the University of Wisconsin-Madison and Washington University in St. Louis, in 2018 and 2021, respectively. He is currently working toward the PhD degree with the School of Computing, University of Georgia, Athens, GA. His research interests include biomedical natural language processing, biomedical image analysis and the intersection of machine learning, and radiation oncology.



Tianming Liu (Senior Member, IEEE) serves as the director with Center for Brain-inspired Artificial Intelligence and is a distinguished research professor in computer science with the University of Georgia. In the past two decades, his research has focused on brain imaging, computational neuroscience and brain-inspired artificial intelligence. He has published more than 400 peer reviewed manuscripts on these topics with H-index 54. He received the prestigious NSF CAREER Award and NIH Career Award for this pioneering research works. He is a fellow of American Institute for Medical and Biological Engineering (AIMBE) and was the General Chair of Medical Image Computing and Computer Assisted Interventions (MICCAI) 2019. He serves on an editorial boards of multiple prestigious international journals including *Medical Image Analysis*, *IEEE Transactions on Medical Imaging*, *IEEE Reviews in Biomedical Engineering*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, and *IEEE Journal of Biomedical and Health Informatics*.