# Modeling Alzheimers' Disease Progression from Multi-task and Self-supervised Learning Perspective with Brain Networks

Wei Liang[1,2], Kai Zhang[1,2], Peng Cao[1,2,3(✉)], Pengfei Zhao[4], Xiaoli Liu[5], Jinzhu Yang[1,2,3(✉)], and Osmar R. Zaiane[6]

[1] Computer Science and Engineering, Northeastern University, Shenyang, China
caopeng@mail.neu.edu.cn
yangjinzhu@cse.neu.edu.cn
[2] Key Laboratory of Intelligent Computing in Medical Image of Ministry of Education, Northeastern University, Shenyang, China
[3] National Frontiers Science Center for Industrial Intelligence and Systems Optimization, Shenyang 110819, China
[4] The Affiliated Brain Hospital of Nanjing Medical University, Nanjing, China
[5] DAMO Academy, Alibaba Group, Hangzhou, China
[6] Alberta Machine Intelligence Institute, University of Alberta, Edmonton, Alberta, Canada

**Abstract.** Alzheimer's disease (AD) is a common irreversible neurodegenerative disease among elderlies. Establishing relationships between brain networks and cognitive scores plays a vital role in identifying the progression of AD. However, most of the previous works focus on a single time point, without modeling the disease progression with longitudinal brain networks data. Besides, the longitudinal data is insufficient for sufficiently modeling the predictive models. To address these issues, we propose a **S**elf-supervised **M**ulti-Task learning **P**rogression model SMP-Net for modeling the relationship between longitudinal brain networks and cognitive scores. Specifically, the proposed model is trained in a self-supervised way by designing a masked graph auto-encoder and a temporal contrastive learning that simultaneously learn the structural and evolutional features from the longitudinal brain networks. Furthermore, we propose a temporal multi-task learning paradigm to model the relationship among multiple cognitive scores prediction tasks. Experiments on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset show the effectiveness of our method and achieve consistent improvements over state-of-the-art methods in terms of Mean Absolute Error (MAE), Pearson Correlation Coefficient (PCC) and Concordance Correlation Coefficient (CCC). Our code is available at https ://github.com/IntelliDAL/Graph/tree/main/SMP-Net.

**Keywords:** Self-supervised learning · Multi-task learning · Cognitive scores · Brain networks · Longitudinal prediction.

## 1 Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease, which affects the quality of life as it causes memory loss, difficulty in thinking and learning [12, 19, 21, 23]. Establishing relationships between brain networks and cognitive scores plays a vital role in identifying the early stage of AD [15, 17]. Though there has been substantial progress in AD diagnostics with brain networks [1, 3, 5, 14], most of the current studies focus on a single time point, without exploring longitudinal modeling for disease progression with brain networks. Some learning-based methods are proposed for the longitudinal prediction of AD progression with multi-modal data but generally fail in utilizing brain networks due to the large heterogeneity of brain networks between individuals as well as developmental stages [2, 6, 8, 20].

The cognitive scores prediction with longitudinal brain networks via deep learning models faces many challenges as follows: (*i*) The available longitudinal brain networks are scarce due to few volunteers or subject dropout [10]. Predicting cognitive scores with limited data is extremely challenging for the deep learning model training. (*ii*) Longitudinal brain networks provide rich structure information and disease progression characteristics, accounting for poor generalization for the pure supervised learning due to the insufficient supervision. (*iii*) The relationship between the brain networks and cognitive scores at multiple time points is varied, hindering the accurate prediction performance at multiple time points with a single task model.

To cope with the above challenges, we propose a self-supervised multi-task learning paradigm for AD progression modeling with longitudinal brain networks. The proposed paradigm consists of a self-supervised spatio-temporal representation learning module for exploiting the spatio-temporal characteristics of longitudinal brain networks and a temporal multi-task module for modeling the relationship among cognitive scores prediction tasks at multiple time points. In summary, our contributions are threefold:

1) To the best of our knowledge, our work is the first attempt to predict cognitive scores with longitudinal brain networks through a self-supervised multi-task paradigm.

2) We design a self-supervised spatio-temporal representation learning module (SSTR), involving masked graph auto-encoder and temporal contrastive learning are jointly pre-trained to capture the structural and evolutional features of longitudinal brain networks simultaneously. The SSTR module can lead to more robust high-level representations for longitudinal brain networks.

3) We assume that inherent correlations exist among the prediction tasks at multiple future time points. Consequently, we propose a temporal multi-task learning paradigm to assist multiple time points cognitive scores prediction, which enhances the model generalization by exploiting the commonalities and differences among different prediction tasks when limited data is available.

# 2 Method

## 2.1 Problem Formalization

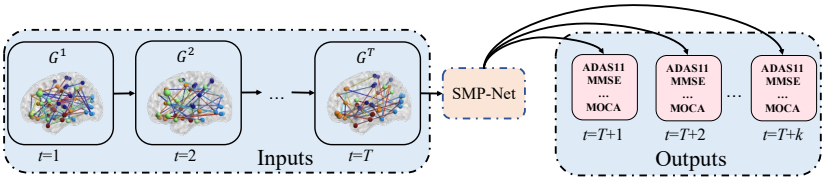


Fig. 1: Illustration of our task. The inputs of the model are $T$ history brain networks, and the outputs are the predicted cognitive scores at the multiple future time points ($T$+1,$T$+2,...,$T$+$k$).

The input to the proposed model is a set of $N$ subjects, each of which has $T$ longitudinal brain networks. Let $X_i = [G_i^1, ..., G_i^t, ..., G_i^T] \in \mathbb{R}^{T*M*M}$ represent the input longitudinal brain networks, where $M$ denotes the number of brain regions based on a specific brain parcellation. $G_i^t = (V_i^t, A_i^t)$ is the brain network of subject $i$ at time $t$; $V_i^t$ and $A_i^t$ are ROIs (nodes) and Pearson correlations between ROIs (edges), respectively. The model outputs are the predicted cognitive scores at $k$ time points $Y_i = [Y_i^{T+1}, ..., Y_i^{T+k}]$ for subject $i$, where $Y_i^{T+k} = [Y_i^{T+k,1}, ..., Y_i^{T+k,p}]$ and $Y_i^{T+k,p}$ is the $p$-$th$ cognitive score of subject $i$ at time $T + k$. As illustrated in Fig. 1, our aim is to build a model $f$ to predict cognitive scores at time $[T + 1, ..., T + k]$ with brain networks at time $[1, ..., T]$, which is formulated as: $\{Y^{T+1}, Y^{T+2}, ..., Y^{T+k}\} = f(G^1, G^2, ..., G^T)$.

## 2.2 Overview

The overview of the proposed SMP-Net is shown in Fig. 2. As shown in Fig. 2, the proposed SMP-Net consists of two modules: Self-supervised spatio-temporal representation learning module (SSTR) for exploiting the spatio-temporal characteristics of longitudinal brain network data itself and a temporal multi-task learning module for modeling the relationship among cognitive scores prediction tasks at multiple time points. SSTR involves a masked graph auto-encoder and a temporal contrastive learning, both of which are jointly pre-trained to learn the structural and evolutional brain networks representation.

## 2.3 Self-supervised Spatio-Temporal Representation Learning

Although brain networks provide rich structure information, the pure supervised learning scheme limits the representation capacity of the models due to insufficient supervision. To solve this problem, we introduce the self-supervised spatio-temporal representation learning module, SSTR. The procedure of SSTR involves two stages as follow:

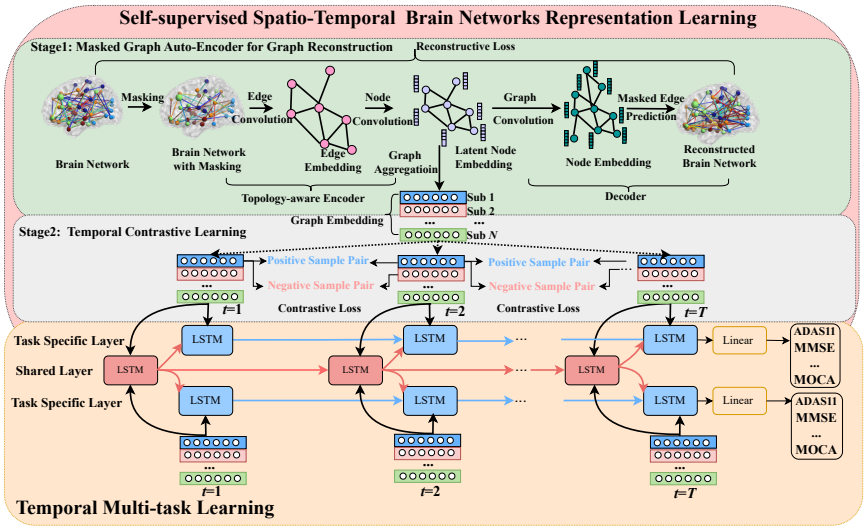


Fig. 2: Illustration of the proposed SMP-Net framework.

**Stage 1 Masked Graph Auto-Encoder for Graph Reconstruction** In stage 1, a masked graph auto-encoder, containing a topology-aware encoder and a decoder, is designed to exploit the crucial structural information in brain networks. To sufficiently exploit the graph structure, we randomly mask some nodes and the associated edges. The unmasked nodes and edges fed into the topology-aware encode to learn the latent representations. Let $H_u$ indicate the feature map in the encoding stage. We define the adjacent matrix of the unmasked nodes as $A_u$, which is taken as the input of the topology-aware encoder, that is $H_u^{(0)} = A_u$. The topology-aware encoder consists of three parts: **1)** The edge convolution with multiple cross-shaped filters for capturing the locality in the graph according to $H_u^{(l)} = EC(H_u^{(l-1)}) = \sum_{i=0}^{M}\sum_{j=0}^{M} H_{u_{(i,\cdot)}}^{(l-1)} \boldsymbol{w}_r + H_{u_{(\cdot,j)}}^{(l-1)} \boldsymbol{w}_c$, where $\boldsymbol{w}_r \in \mathbb{R}^{1\times M}$ and $\boldsymbol{w}_c \in \mathbb{R}^{M\times 1}$ are convolution kernels. **2)** The node convolution for learning the latent node embedding. It is defined as: $H_n^{(l)} = NC(H_u^{(l-1)}) = \sum_{i=1}^{M} H_{u_{(i,\cdot)}}^{(l-1)} \boldsymbol{w}_n^{l-1}$, where $\boldsymbol{w}_n$ is the learned filter vector, $H_n^{(l)} \in \mathbb{R}^{M\times D_n}$ is the latent unmasked node embedding and $D_n$ is the channels in $NC$. **3)** The graph aggregation for achieving the global graph embedding through: $H_g^{(l)} = GA(H_n^{(l-1)}) = \sum_{i=1}^{M} H_{n_{(i,\cdot)}}^{(l-1)} \boldsymbol{w}_g$, where $\boldsymbol{w}_g$ is the learned filter vector, $H_g \in \mathbb{R}^{M\times D_g}$ is the graph embedding and $D_g$ is the dimensionality in $GA$.

The decoder takes the masked nodes and the latent unmasked node embeddings as inputs, and then produces predictions for the masked nodes and edges by graph convolution operations and the masked edge prediction. The graph convolution is defined as: $H^{(l+1)} = \sigma(\widetilde{A} H_n^{(l)} W^{(l)}) + b^{(l)}$, where $\widetilde{A} \in \mathbb{R}^{M\times M}$ is the binary adjacency matrix, $W$ denotes trainable weight, $H \in \mathbb{R}^{M\times D_n'}$ is

the node embedding and $D'_n$ is the hidden layer size of graph convolutional layers. The masked edge prediction is defined as: $\hat{A} = H^{(l+1)}(H^{(l+1)})^T$. The reconstruction loss between the prediction graphs and corresponding targets is $L_{rec} = \sum_{i=1}^{N}\sum_{t=1}^{T} ||A_i^t - \hat{A}_i^t||_2^2$, where $\hat{A}_i^t$ is the reconstructed brain networks of subject $i$ at time $t$.

**Stage 2 Temporal Contrastive Learning** The longitudinal brain networks of a subject acquired at multiple visits characterize gradual disease progression of the brain over time, which manifests a temporal progression trajectory when projected to the latent space. We assume that brain networks features at two consecutive time points from the same subject are similar, while dissimilar from different subjects. Based on this assumption, we introduce a temporal contrastive loss by enforcing an across-sample relationship in the learning process. Specifically, $H_{g(i)}^t$ is the brain network features of subject $i$ at time $t$, $H_{g(i)}^t$ and $H_{g(j)}^{t+1}$ are considered as the positive sample pair if $i = j$, otherwise they are considered as the negative sample pair. The temporal contrastive framework aims to enlarge the similarity between positive sample pair, and reduce it between the negative sample pair. The similarity calculation function $s$ can be any distance function, and here we utilize cosine similarity. The loss for temporal contrastive learning can be represented as:

$$L_{con} = -log \sum_{i=1}^{N}\sum_{t=1}^{T-1} \frac{exp(s(H_{g(i)}^t, H_{g(i)}^{t+1})/\tau)}{\sum_{j=1,j\neq i}^{N} exp(s(H_{g(i)}^t, H_{g(j)}^{t+1})/\tau)}, \tag{1}$$

where $\tau$ is a temperature factor that controls the model's discrimination against negative sample pair and $exp(.)$ is an exponential function.

### 2.4 Temporal Multi-task learning

Existing studies have demonstrated the effectiveness of multi-task learning for the extraction of a robust feature representation [9][24]. In this regard, to further exploit the correlation among the prediction tasks at multiple future time points, we design a temporal multi-task learning paradigm. Specifically, the temporal multi-task learning module consists of a shared network and multiple task-specific networks, all of which are designed with a Long Short-Term Memory (LSTM) architecture [4]. The shared network is trained for modeling the shared information $h_s^t$ among cognitive scores prediction tasks at multiple time points. The $q$-th task-specific network aims to capture the task-specific information $h_q^t$ from the shared network and the brain networks features at time $t$. The temporal multi-task learning module can be seen as an end-to-end architecture with the shared and task-specific parameters of $W_s, W_q$. By learning these parameters jointly, we arrive at a collaborative learning method to jointly improve the performance of the prediction tasks at multiple time points. The shared information $h_s^t$ and task-specific information $h_q^t$ are formulated as $h_s^t = LSTM(Hg^t, h_s^{t-1}, W_s)$ and $h_q^t = LSTM([Hg^t, h_s^t], h_q^{t-1}, W_q)$. The output of the temporal multi-task

learning module is formulated as: $\hat{Y}^t = W_2(W_1 h_q^t + b_1) + b_2$, where $W_1$, $W_2$, $b_1$, $b_2$ are learnable parameters of LSTM. Errors between the actual observations $Y^t$ and predictions $\hat{Y}^t$ are used to update the model parameters through the regression loss as follow:

$$L_{reg} = \sum_{i=1}^{k} \sum_{t=2}^{T} (||Y^t - \hat{Y}^t||_1 + ||Y^{T+i} - \hat{Y}^{T+i}||_1) \tag{2}$$

The overall loss function $L$ is described as Eq. (3), where $\lambda_{con}$ and $\lambda_{rec}$ are the weights for contrastive loss and reconstruction loss, respectively.

$$L = L_{reg} + \lambda_{con} L_{con} + \lambda_{rec} L_{rec} \tag{3}$$

## 3 Experiments

### 3.1 Dataset and Experimental Settings

In this work, we choose 219 longitudinal resting-state fMRI scans of 73 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [11][7]. AAL template is used to obtain 90 ROIs for every subject [18]. We predict nine cognitive scores at time M24, M36 and M48 with brain networks times of M0, M6 and M12 to evaluate our proposed SMP-Net. The number of samples for the three tasks are 73, 35 and 31, respectively.

During the model training, the Adam optimizer is used with a momentum of 0.9 and a weight decay of 0.01. The learning rate is set to $10^{-3}$. The hidden layer size of LSTM and graph convolutional layers are set to 64 and 48, respectively. The values of hyperparameter $\lambda_c$ and $\lambda_r$ are set to 1. The model is trained with 20 epochs in the self-supervised spatio-temporal representation learning stage and 300 epochs in the temporal multi-task learning stage with a batch size of 16. To avoid over-fitting due to the limited subjects, in all experiments, we repeat the 5-fold cross-validation 10 times with different random seeds. We finally report the average results. Three commonly used metrics are adopted to evaluate all methods, including Mean Absolute Error (MAE), Pearson Correlation Coefficient (PCC) and Concordance Correlation Coefficient (CCC). CCC reflects both the correlation and the absolute error between the true and the predicted cognitive scores. Due to limited space, we report the results in terms of CCC in this paper. The results in terms of MAE and CC are shown in the supplementary material. To ensure a fair comparison, the hyperparameters of comparable methods are optimized to achieve their best performance.

### 3.2 Effectiveness Evaluation

We compare the performance of our SMP-Net with three state-of-the-art (SOTA) sequential graph learning methods: evolveGCN [13], stGCN [22] and DySAT [16]

[7] http://adni.loni.usc.edu/.

Table 1: Experimental results in terms of CCC. The best results are bold and the superscript symbol * indicates that the proposed method significantly outperformed that method with $p$-value=0.01.

| | Cognitive scores | CDRSB | ADAS11 | MMSE | RAVLT imm | RAVLT learn | ADAS13 | RAVLT forget | RAVLT perc forget | MOCA | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M24 | GCN | 0.190* | 0.286* | 0.139* | 0.346* | 0.280* | 0.319* | 0.320* | 0.311* | 0.201* | 0.266 (0.045)* |
| | evolveGCN | 0.351* | 0.415* | 0.141* | 0.487* | 0.399* | 0.427* | 0.283* | 0.433* | 0.272* | 0.356 (0.078)* |
| | stGCN | 0.310* | 0.391* | 0.265* | 0.458* | 0.323* | 0.415* | 0.279* | 0.405* | 0.373* | 0.358 (0.078)* |
| | DySAT | 0.574* | 0.416* | 0.438* | 0.342* | **0.588** | 0.353* | 0.597* | 0.230* | 0.531* | 0.452 (0.051)* |
| | **SMP-Net** | **0.617** | **0.798** | **0.658** | **0.798** | 0.525 | **0.809** | **0.654** | **0.819** | **0.666** | **0.705 (0.038)** |
| M36 | GCN | 0.152* | 0.229* | 0.122* | 0.314* | 0.055* | 0.254* | 0.099* | 0.214* | 0.213* | 0.184 (0.076)* |
| | evolveGCN | 0.350* | 0.450* | 0.114* | 0.563* | 0.508* | 0.450* | 0.341* | 0.431* | 0.219* | 0.381 (0.124)* |
| | stGCN | 0.246* | 0.410* | 0.323* | 0.439* | 0.333* | 0.430* | 0.226* | 0.491* | 0.341* | 0.360 (0.131)* |
| | DySAT | **0.556** | 0.355* | 0.561* | 0.543* | **0.577** | 0.281* | **0.629** | 0.140* | 0.569* | 0.468 (0.048)* |
| | **SMP-Net** | 0.490 | **0.754** | **0.593** | **0.801** | 0.496 | **0.788** | 0.571 | **0.832** | **0.663** | **0.665 (0.060)** |
| M48 | GCN | 0.144* | 0.313* | 0.159* | 0.431* | 0.141* | 0.316* | 0.168* | 0.324* | 0.097* | 0.233 (0.091)* |
| | evolveGCN | 0.471* | 0.342* | 0.103* | 0.484* | 0.397* | 0.368* | 0.512* | 0.542* | 0.108 * | 0.370 (0.084)* |
| | stGCN | 0.276* | 0.300* | 0.296* | 0.424* | 0.392* | 0.321* | 0.450* | 0.557* | 0.299* | 0.368 (0.096)* |
| | DySAT | **0.609** | 0.347* | 0.416* | 0.378* | 0.640* | 0.268* | 0.675* | 0.146* | 0.441* | 0.436 (0.085)* |
| | **SMP-Net** | 0.561 | **0.694** | **0.554** | **0.798** | **0.518** | **0.752** | **0.749** | **0.869** | **0.570** | **0.674 (0.083)** |

as well as a baseline method: GCN [7]. Table 1 summarizes the results of all methods in terms of CCC on the ADNI dataset. As reported in the supplementary material, consistent conclusions are obtained by SMP-Net in terms of MAE and CC. Based on the experimental results, we have the following observations: First, evolveGCN, stGCN and DySAT consistently outperform GCN, indicating that evolveGCN, stGCN and DySAT are able to capture the dynamism underlying a brain networks sequence through a recurrent model, which contributes to improve performance in disease prediction. Second, DySAT shows a higher average CCC than evolveGCN and stGCN. One possible reason is that DySAT utilizes joint structural and temporal self-attention, which enables it to learn more efficient dynamic graph representation compared with evolveGCN and stGCN. Finally, our proposed SMP-Net maintains a stable and competitive performance at all the time points, demonstrating that 1) SMP-Net can learn more expressive representations of brain networks structure by masked graph auto-encoder for graph reconstruction module. 2) SMP-Net suffciently takes advantage of the temporal and subject correlation in disease progression by temporal contrastive learning. 3) The temporal multi-task paradigm of SMP-Net effectively exploits the inherent correlation among multiple prediction tasks at different time points, which facilitate to improve the model performance.

### 3.3 Discussion

**Ablation analysis** To valid the effect of each proposed module, we consider the following variants for evaluation: 1) SMP-Net-c: the temporal contrastive loss is removed; 2) SMP-Net-r: the reconstruction loss is removed; 3) SMP-Net-rc: both temporal contrastive loss and graph reconstruction loss are removed; 4) SMP-Net-m: the temporal multi-task paradigm is ignored. Table 2 summarizes the results of ablation studies in terms of CCC. It is apparent that SMP-Net

Table 2: Average CCC results of ablation studies. The best results are bold and the superscript symbol * indicates that the proposed method significantly outperformed that method with $p$-value=0.01.

| Methods | SMP-Net-rc | SMP-Net-r | SMP-Net-c | SMP-Net-m | **SMP-Net** |
|---|---|---|---|---|---|
| M24 | 0.299* | 0.382* | 0.472* | 0.635* | **0.705** |
| M36 | 0.319* | 0.410* | 0.490 * | 0.590* | **0.665** |
| M48 | 0.313* | 0.434* | 0.468* | 0.544* | **0.674** |

outperforms all of the variants. Specifically, SMP-Net consistently outperforms SMPT-Net-m 11.0%, 12.7% and 23.8% at time M24, M36 and M48, respectively, indicating the effectiveness of the temporal multi-task paradigm. It also indicates that the multi-task paradigm in SMP-Net is more helpful for the prediction at farther time points. The reason is that prediction tasks at farther time points are more difficult due to the insignificant relationship between the brain networks and the cognitive scores. Temporal multi-task paradigm enforces the long-term prediction to benefit from short-term prediction, making the prediction tasks at farther time points gain more improvements. Moreover, we can observe that models with SSTR perform better than the ones without SSTR. For instance, SMP-Net-m and SMP-Net show superior performance than SMP-Net-r, SMP-Net-c and SMP-Net-rc. This demonstrates that SSTR facilitates the learning of structural and evolutional features in the condition of limited samples and insufficient supervision, thereby leading to more robust high-level representations for downstream tasks.

**Evaluating Robustness** To evaluate the robustness of the SSTR module, we pre-train SMP-Net with fMRI at three time points (M0, M6, M12) and fine-tune it with different downstream tasks of predicting cognitive scores at different time points. As shown in Fig. 3, MSP-Net provides comparatively stable performance on different fine-tuning tasks, demonstrating that features learned with our pre-trained model are robust to the different fine-tuning tasks.

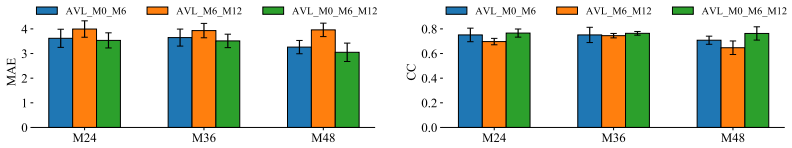


Fig. 3: Average MAE and CC of the pre-trained model fine-tuning on three different downstream tasks. AVL_M0_M6 denotes that fMRI data at M0 and M6 are available.

## 4 Conclusion

This paper proposes an AD progression model SMP-Net from multi-task and self-supervised learning perspective with longitudinal brain networks. In the proposed SMP-Net, self-supervised spatio-temporal representation learning is designed to learn more robust structural and evolutional features from longitudinal brain networks. The temporal multi-task paradigm is designed for boosting the ability of cognitive score prediction at multiple time points. Experimental results on the ADNI dataset with fewer samples demonstrate the advantage of self-supervised spatio-temporal representation learning and temporal multi-task learning.

**Acknowledgements** This research was supported by the National Natural Science Foundation of China (No.62076059), the Science Project of Liaoning Province (2021-MS-105) and the 111 Project (B16009), .

## References

1. Aviles-Rivero, A.I., Runkel, C., Papadakis, N., Kourtzi, Z., Schönlieb, C.B.: Multimodal hypergraph diffusion network with dual prior for alzheimer classification. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part III. pp. 717–727. Springer (2022)
2. Brand, L., Wang, H., Huang, H., Risacher, S., Saykin, A., Shen, L., ADNI: Joint high-order multi-task feature learning to predict the progression of alzheimer's disease. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I. pp. 555–562. Springer (2018)
3. Chen, Z., Liu, Y., Zhang, Y., Li, Q., Initiative, A.D.N., et al.: Orthogonal latent space learning with feature weighting and graph learning for multimodal alzheimer's disease diagnosis. Medical Image Analysis **84**, 102698 (2023)
4. Graves, A., Graves, A.: Long short-term memory. Supervised sequence labelling with recurrent neural networks pp. 37–45 (2012)
5. Huang, Y., Chung, A.C.: Disease prediction with edge-variational graph convolutional networks. Medical Image Analysis **77**, 102375 (2022)
6. Jung, W., Jun, E., Suk, H.I., Initiative, A.D.N., et al.: Deep recurrent model for individualized prediction of alzheimer's disease progression. NeuroImage **237**, 118143 (2021)
7. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
8. Liang, W., Zhang, K., Cao, P., Liu, X., Yang, J., Zaiane, O.: Rethinking modeling alzheimer's disease progression from a multi-task learning perspective with deep recurrent neural network. Computers in Biology and Medicine **138**, 104935 (2021)
9. Liao, W., Xiong, H., Wang, Q., Mo, Y., Li, X., Liu, Y., Chen, Z., Huang, S., Dou, D.: Muscle: Multi-task self-supervised continual learning to pre-train deep models for x-ray images of multiple body parts. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII. pp. 151–161. Springer (2022)

10. Logothetis, N.K.: What we can do and what we cannot do with fmri. Nature **453**(7197), 869–878 (2008)
11. Marinescu, R.V., Oxtoby, N.P., Young, A.L., Bron, E.E., Toga, A.W., Weiner, M.W., Barkhof, F., Fox, N.C., Klein, S., Alexander, D.C., et al.: Tadpole challenge: Prediction of longitudinal evolution in alzheimer's disease. arXiv preprint arXiv:1805.03909 (2018)
12. Nguyen, H.D., Clément, M., Mansencal, B., Coupé, P.: Interpretable differential diagnosis for alzheimer's disease and frontotemporal dementia. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part I. pp. 55–65. Springer (2022)
13. Pareja, A., Domeniconi, G., Chen, J., Ma, T., Suzumura, T., Kanezashi, H., Kaler, T., Schardl, T., Leiserson, C.: Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In: Proceedings of the AAAI conference on artificial intelligence. vol. 34, pp. 5363–5370 (2020)
14. Parisot, S., Ktena, S.I., Ferrante, E., Lee, M., Guerrero, R., Glocker, B., Rueckert, D.: Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease. Medical image analysis **48**, 117–130 (2018)
15. Petersen, E., Feragen, A., da Costa Zemsch, M.L., Henriksen, A., Wiese Christensen, O.E., Ganz, M., Initiative, A.D.N.: Feature robustness and sex differences in medical imaging: A case study in mri-based alzheimer's disease detection. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part I. pp. 88–98. Springer (2022)
16. Sankar, A., Wu, Y., Gou, L., Zhang, W., Yang, H.: Dysat: Deep neural representation learning on dynamic graphs via self-attention networks. In: Proceedings of the 13th international conference on web search and data mining. pp. 519–527 (2020)
17. Seyfioğlu, M.S., Liu, Z., Kamath, P., Gangolli, S., Wang, S., Grabowski, T., Shapiro, L.: Brain-aware replacements for supervised contrastive learning in detection of alzheimer's disease. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part I. pp. 461–470. Springer (2022)
18. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. Neuroimage **15**(1), 273–289 (2002)
19. Xiao, T., Zeng, L., Shi, X., Zhu, X., Wu, G.: Dual-graph learning convolutional networks for interpretable alzheimer's disease diagnosis. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII. pp. 406–415. Springer (2022)
20. Xu, L., Wu, H., He, C., Wang, J., Zhang, C., Nie, F., Chen, L.: Multi-modal sequence learning for alzheimer's disease progression prediction with incomplete variable-length longitudinal data. Medical Image Analysis **82**, 102643 (2022)
21. Yang, F., Meng, R., Cho, H., Wu, G., Kim, W.H.: Disentangled sequential graph autoencoder for preclinical alzheimer's disease characterizations from adni study. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24. pp. 362–372. Springer (2021)

22. Yu, B., Yin, H., Zhu, Z.: Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. arXiv preprint arXiv:1709.04875 (2017)
23. Zhang, S., Chen, X., Ren, B., Yang, H., Yu, Z., Zhang, X.Y., Zhou, Y.: 3d global fourier network for alzheimer's disease diagnosis using structural mri. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part I. pp. 34–43. Springer (2022)
24. Zhu, J., Li, Y., Ding, L., Zhou, S.K.: Aggregative self-supervised feature learning from limited medical images. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII. pp. 57–66. Springer (2022)