Clinical Neuroscience Research, 2026, Volume 4, Issue 2
http://ojs.bbwpublisher.com/index.php/CNR
Online ISSN: 2981-8133
Print ISSN: 3083-4899

Depression Prediction Based on Graph Neural Networks

Rui Hua*

School of Information, Yunnan Normal University, Kunming 650500, Yunnan, China

***Corresponding author:** Rui Hua, 785291309@qq.com

Abstract: Depression, as a highly prevalent mental disorder, is traditionally diagnosed using subjective scale-based assessments, which suffer from limitations such as insufficient objectivity and delayed early warning. With the widespread adoption of electronic health records and the rapid advancement of artificial intelligence technologies, data-driven approaches for depression prediction have become a major research focus. This paper systematically reviews the evolution of depression prediction methods, spanning from traditional assessment scales to machine learning and further to deep learning techniques. The paper places particular emphasis on the current applications and unique advantages of graph neural networks (GNNs) in this domain. Existing studies indicate that GNNs can effectively model complex relationships among patient features, thereby improving predictive performance while enhancing model interpretability and offering a novel technical pathway for early detection of depression. However, several challenges remain, including the lack of standardized graph construction methods, issues related to data privacy and quality, and insufficient model generalizability. Future research is expected to focus on constructing multi-center graph datasets, developing interpretable GNN models, and promoting their practical application in public health screening, ultimately advancing mental health services toward a "data-driven and proactive prevention" paradigm.

Keywords: Depression prediction; Graph neural networks; Deep learning; Graph attention networks

---

1. Introduction

According to statistics from the World Health Organization (WHO), approximately 1 billion people worldwide are affected by mental disorders. During the COVID-19 pandemic, the prevalence of depression increased by 27.6% [1]. Depression has become one of the leading causes of disability in China, ranking second among all causes [2]. In recent years, a series of policies have been introduced to promote mental health initiatives, indicating that the early identification and intervention of depression have become critical public health priorities [3].

Depression Spectrum Disorder (DSD) is a class of neurological disorders characterized by significant heterogeneity [4]. Its core symptoms include persistent low mood, diminished interest, and reduced energy, often accompanied by cognitive and physiological dysfunctions [5]. Current clinical diagnosis primarily relies on clinician interviews and standardized scales (e.g., PHQ-9, HAMD). However, these approaches suffer from limitations such as strong subjectivity, reporting bias from patients, and difficulty in early detection. Delayed diagnosis may result in missed opportunities for optimal intervention, underscoring the importance of early identification. Automated prediction methods hold the potential to improve screening efficiency, optimize the allocation of medical resources, and support research into disease mechanisms.

Advances in machine learning and deep learning have driven the transition of depression prediction from subjective assessment to data-driven approaches [6]. In recent years, these methods have achieved promising results in early detection tasks. Graph Neural Networks (GNNs) are capable of learning representations from graph-structured data by jointly modeling node attributes and relational information to perform classification tasks. In the context of depression prediction, leveraging GNNs to capture inter-patient relational features can enhance the discriminative power of predictive models.

2. Related work

In the early stages of depression prediction research, psychological assessment scales were the primary tools, such as the Hamilton Depression Rating Scale (HAMD), Beck Depression Inventory (BDI), and Center for Epidemiologic Studies Depression Scale (CES-D). However, these scales exhibit notable limitations: different instruments cover inconsistent symptom domains, assessments for the same individual may vary substantially across scales, and they are generally characterized by latency and subjectivity, making them inadequate for early warning purposes.

With the advancement of data mining techniques, machine learning methods have been increasingly applied to depression prediction. Models such as Support Vector Machines (SVM), decision trees, and random forests have achieved promising results. For instance, random forest models have reached an AUC of 0.943 in postpartum depression prediction, while ensemble learning approaches have reported accuracies ranging from approximately 88.6% to 89.7% [7–8]. A related review indicates that the AUC of such models spans from 0.6030 to 0.9976 [6]. Nevertheless, traditional machine learning methods remain limited in handling complex clinical data and multimodal information.

In recent years, deep learning has been widely adopted due to its capability for automatic feature extraction. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are commonly employed to capture spatial and temporal features, respectively. Numerous studies have demonstrated strong performance on electroencephalogram (EEG) and multimodal datasets. For example, an improved Temporal Convolutional Network (TCN) model achieved an accuracy of 91.3% on the DAIC-WOZ dataset, while a CNN-LSTM-based model (DDNet) attained an accuracy of 96.21% on social media data [9–10]. In addition, Graph Neural Network (GNN)-based methods have been introduced; by constructing brain networks and integrating ensemble strategies, these approaches achieved an accuracy of 77% on the MODMA dataset [11].

3. Dataset and data preprocessing

3.1. Dataset

In depression prediction studies based on the DBIDE dataset, preprocessed data can be accessed via publicly available Amazon S3 storage. The derived data for each subject and different processing pipelines are stored as independent files, enabling on-demand download.

To construct a population graph that integrates both imaging and non-imaging information, this study utilizes rs-fMRI, sMRI, and phenotypic data from the DBIDE dataset. The sample selection is consistent with that of Parisot et al., comprising a total of 871 subjects, including 403 patients with DSD and 468 healthy controls [8].

For fair comparison, the fMRI data are preprocessed using the same C-PAC pipeline [9], which includes temporal and motion correction, normalization, band-pass filtering (0.01–0.1 Hz), registration, and the construction of functional connectivity matrices ((S_1, ⋯, S_N)). The sMRI data are processed using the CIVET pipeline, including denoising, registration, and segmentation. Finally, phenotypic information such as age, sex, full-scale IQ (FIQ), acquisition site, and diagnosis is incorporated jointly for depression prediction modeling.

This study employs two distinct graph construction strategies—P-Edge and PF-Edge—to build the population graph edge structures required for depression research. These two methods differ in construction logic, data sources, and the aspects of topological characteristics they emphasize, making them suitable for modeling subject relationships from different perspectives.

3.2. Data preprocessing

The P-Edge method defines graph node connectivity primarily based on subjects' clinical phenotypic data. It selects key phenotypic attributes, such as age and sex, as the basis for graph construction and determines inter-node relationships through differentiated rules. For continuous phenotypic variables, such as age and full-scale IQ (FIQ), a connection is established if the numerical difference is below a predefined threshold. For categorical variables, such as sex, nodes are connected if the values are identical. Edge label matrices are generated by aggregating the association results of individual phenotypic attributes, and a binary adjacency matrix is subsequently derived through logical rules to complete the population graph construction. This method relies on basic demographic and clinical indicators, offering simplicity, interpretability, and low computational complexity. However, it yields limited variability in edge weights and has a relatively weak capacity to capture deeper associations among subjects.

In contrast, the PF-Edge method measures inter-node correlations (i.e., edge weights) by integrating both phenotypic similarities and the Pearson correlation coefficients between fMRI data of different subjects. By jointly considering phenotypic and imaging features, this approach computes comprehensive inter-subject correlations based on both clinical attribute relationships and functional imaging similarities, thereby defining edge weights and connectivity. This method leverages the complementary strengths of phenotypic clinical attributes and brain network characteristics derived from functional imaging data. It not only preserves the foundational constraints imposed by demographic and clinical features but also incorporates deeper similarity information from functional neuroimaging. Compared with the single P-Edge approach, PF-Edge provides a more comprehensive and multidimensional characterization of inter-subject topological relationships, thereby offering a higher-quality graph structure for graph neural network modeling.

4. Design and implementation of a graph neural network-based depression prediction model

4.1. Implementation of model components

The proposed model consists of four modules: a graph construction module, a feature processing module, a Graph Attention Network (GAT) module, and a classification output module. The implementation details of each module are described as follows.

4.2. Graph structure construction module

All subjects are treated as graph nodes, and brain functional connectivity features are used as node feature vectors. An undirected weighted graph is constructed based on similarity computation. The Pearson correlation distance between subject features is calculated:

$$dist = pdist\left(\text{features,metric=correlation}\right) \tag{1}$$

A similarity adjacency matrix is then generated using a Gaussian kernel function:

$$adj = \exp\left(-\frac{\mathrm{dist}^2}{2\sigma^2}\right) \tag{2}$$

where represents the mean distance. A threshold (h_x = 0.8) is applied to filter valid edges, and the resulting graph structure is represented as edge_index for use in PyTorch Geometric. Functional connectivity matrices from the DBIDE dataset are used as the original fMRI features. Feature selection is performed using analysis of variance (ANOVA), retaining the top 2000 most significant features. These features are subsequently standardized to ensure training stability.

4.3. Graph Attention Network (GAT) Module

A two-layer Graph Attention Network (GAT) architecture is adopted, with the following core design:

(1) First Layer: The input node features are projected into a hidden representation space with a dimensionality of (hid_c = 12). In this stage, attention weights are learned to capture the relative importance of neighboring nodes.

(2) Second Layer: The learned hidden representations are further transformed to produce the final output for binary classification.

To enhance feature representation, a multi-head attention mechanism is employed (**Figure 1**). This mechanism aggregates information from neighboring nodes by assigning attention coefficients, enabling the model to capture diverse structural patterns. The Exponential Linear Unit (ELU) activation function is applied to introduce nonlinearity and improve the model's expressive capacity. In the final layer, a LogSoftmax function is used to generate class probability distributions. The output dimension is set to 2, corresponding to the two categories: depression (DSD) and healthy control (TC).

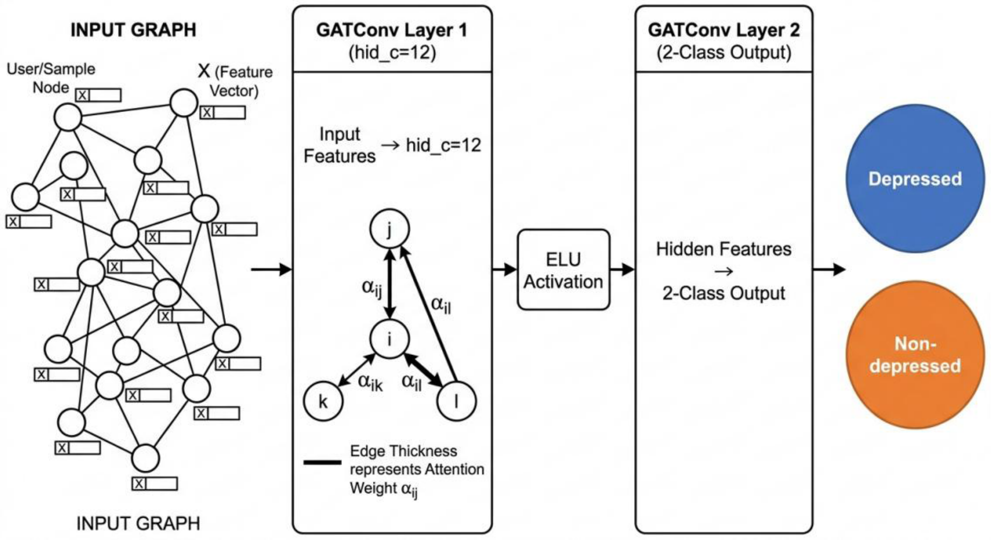

Figure 1. Schematic of attention network structure

5.Experimental results

5.1. Model performance evaluation

The model training process demonstrates stable convergence, achieving a final classification accuracy of 69.32% and an AUC value of 0.6863 on the test set. These results indicate that the proposed depression classification model, based on population graph structures and a Graph Attention Network (GAT), can effectively capture both brain functional connectivity features and topological relationships among subjects, thereby exhibiting a certain capability for assisting in depression identification.

The training loss curve shows a continuous downward trend with gradual convergence, decreasing from an initial value of 0.7935 to 0.0957 at the 200th epoch. This indicates that the model is capable of effectively learning feature representations without exhibiting significant overfitting, and the overall training process is stable and reliable.

5.2. Confusion matrix analysis

The visualization of the confusion matrix is shown in **Figure 2**. In the matrix, rows represent the true labels of samples, while columns represent the predicted labels. Class 0 and Class 1 correspond to healthy controls and depression spectrum disorder (DSD), respectively. The key values in the matrix are as follows: among healthy control samples, 25 cases are correctly classified as Class 0, while 16 cases are misclassified as Class 1; among depression samples, 37 cases are correctly classified as Class 1, while 10 cases are misclassified as Class 0.

From the distribution, the model demonstrates relatively strong recognition capability for depression samples, with the number of true positives significantly exceeding false negatives, indicating that the model effectively captures brain imaging features associated with depression. Meanwhile, the classification accuracy for healthy controls remains at a reasonable level, with a relatively low false positive rate. This is of particular importance in clinical decision-support scenarios, as it helps reduce unnecessary interventions and misdiagnoses.

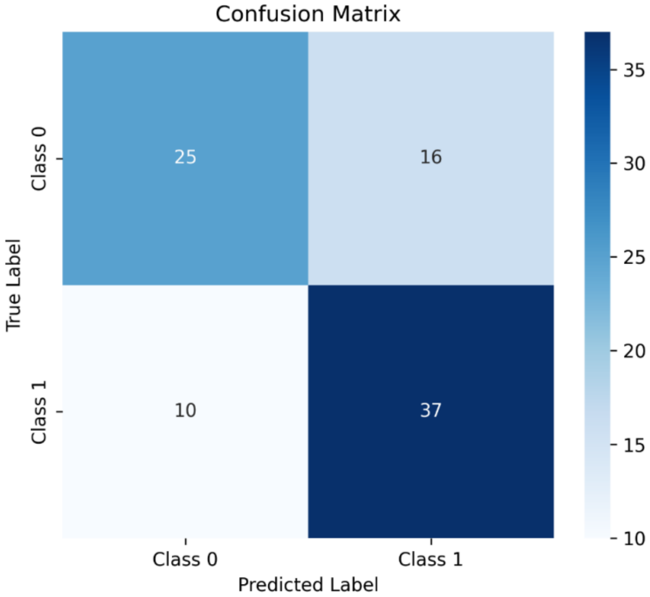

Figure 2. Confusion matrix

6. Conclusion

This study focuses on resting-state functional magnetic resonance imaging (rs-fMRI) data from the DBIDE dataset and addresses the task of automatic identification of depression spectrum disorder by constructing and implementing a classification model based on a Graph Attention Network (GAT).

Multi-center data from DBIDE were preprocessed through cleaning, normalization, and feature selection to extract brain functional connectivity features. To address the limitation of traditional machine learning methods in capturing inter-sample relationships, this study constructs a population graph based on inter-subject feature similarity, transforming tabular data into graph-structured data that incorporates both node features and topological relationships, thereby providing a solid foundation for graph neural network modeling.

The GAT model is implemented using the PyTorch Geometric framework, leveraging the attention mechanism to adaptively learn the importance weights of different subject nodes, enabling deep feature representation and classification of brain imaging data. The model is trained using a stratified cross-validation strategy, optimized with the Adam optimizer and a negative log-likelihood loss function. Over 200 training epochs, the loss converges steadily without significant overfitting.

The model achieves a classification accuracy of 69.32% and an AUC value of 0.6863 on the test set. The confusion matrix analysis indicates that the model performs well in identifying healthy control samples while maintaining a relatively low false positive rate, demonstrating its potential value in clinical decision support. Furthermore, visualization of GAT attention weights confirms that the model can adaptively focus on key inter-subject relationships, thereby enhancing the interpretability of the decision-making process.

Disclosure statement

The author declares no conflict of interest.

References

[1] Anokye R, Acheampong E, Budu-Ainooson A, et al., 2018, Prevalence of Postpartum Depression and Interventions Utilized for its Management. Annals of General Psychiatry, 2018(17): 18.

[2] Chai J, Zeng H, Li A, et al., 2021, Deep Learning in Computer Vision: A Critical Review of Emerging Techniques and Application Scenarios. Machine Learning with Applications, 2021(6): 100134.

[3] Dadi AF, Akalu TY, Baraki AG, et al., 2020, Epidemiology of Postnatal Depression and Its Associated Factors in Africa: A Systematic Review and Meta-Analysis. PloS One, 15(4): e0231940.

[4] Paulson JF, Bazemore SD, 2010, Prenatal and Postpartum Depression in Fathers and Its Association with Maternal Depression: A Meta-Analysis. JAMA: The Journal of the American Medical Association, 303(19): 1961–1969.

[5] Saidi A, Othman SB, Saoud SB, 2020, Hybrid CNN-SVM Classifier for Efficient Depression Detection System. 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET), 229–234.

[6] Iqbal N, Chowdhury AM, Ahsan T, 2018, Enhancing the Performance of Sentiment Analysis by Using Different Feature Combinations. 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), 1–4.

[7] Johansson M, Higgins KL, Nzefa LD, et al., Postpartum Depression and Life Experiences of Mothers with an Immigrant Background Living in the South of Sweden. International Journal of Qualitative Studies on Health and Well-being, 18(1): 2187333.

[8] Karmiani D, Kazi R, Nambisan A, et al., 2019, Comparison of Predictive Algorithms: Backpropagation, SVM, LSTM and Kalman Filter for Stock Market. 2019 Amity International Conference on Artificial Intelligence (AICAI), 228–234.

[9] de Lima OF, Martin-Santos R, 2020, Postpartum Depression: A Clinical and Research Update. Nova Science Publishers, Incorporated.

[10] Medhat W, Hassan A, Korashy H, 2014, Sentiment Analysis Algorithms and Applications: A Survey. Ain Shams Engineering Journal, 5(4): 1093–1113.

[11] Wisner KL, Chambers C, Sit DKY, 2006, Postpartum Depression: A Major Public Health Problem. JAMA: The Journal of the American Medical Association, 296(20): 2586–2593.