Heterogeneous Graph Attention Networks for Depression Identification by Campus Cyber-Activity Patterns

Minqiang Yang, *Member, IEEE*, Zhuoheng Li, Yujie Gao, Chen He, Fuzhan Huang, and Wenbo Chen

***Abstract*—As one of the most prevalent mental disorders, depression is associated with a high rate of self-harm and suicide, particularly among college students. It is urgently needed to discover prospective cases of depression disorder among college students, enabling timely intervention to reduce its impact on their academic performance and daily lives. This study investigates a method for identifying groups that may have early depressive tendencies through their Internet usage on campus networks. This article proposes a heterogeneous graph attention network (H-GAT) model that incorporates an attention mechanism based on ablation experiments in heterogeneous graphs to analyze the patterns and correlations within the surfing behavior data of students. This model makes full use of the interaction relationships between heterogeneous nodes in the graph to capture the affective tendencies reflected in the cyber-activity patterns. The proposed H-GAT model exhibits excellent performance, with nearly 80% accuracy and recall. Our work offers a potential approach to detect depression on college campuses using nonintrusive methods, which could ultimately contribute to early warnings for both individuals experiencing depression and higher education institutions.**

***Index Terms*—Affective computing, cyber-activity patterns, depression identification, heterogeneous graph attention network (H-GAT).**

I. Introduction

DEPRESSION is a prevalent mental disorder distinguished by persistent symptoms such as depressive mood, depressive interest, and reduced energy levels [1]. The widespread influence of depression on individuals' quality of life and social functioning underscores its significance in the field of public health [2].

In recent years, there has been a notable surge in the incidence of depression among college students [3]. Depression research on college campuses has become a prominent focus across disciplines, including psychology, medicine, and computer science [4], [5], [6], [7], [8], [9], [10]. Current research on depression among college students relies primarily on questionnaires [11], biomarkers [12], as well as the analysis of social media and mobile device usage logs [13], and specific smartphone application [14]. Nevertheless, these studies often grapple with limitations due to difficulties in data access, poor data quality, privacy and legal constraints, and subjective evaluation methods. Presently, investigations into campus depression primarily utilize conventional statistical or machine learning methods, which yield low accuracy and lack the capability for early detection and prompt intervention.

Current research in depression or emotion recognition through machine learning predominantly focuses on natural language processing (NLP) [15], eye movement [16], facial expression [17], [18], electroencephalogram (EEG) [19], [20], [21], and audio [22]. Advancements in deep learning have facilitated its widespread application in identifying and predicting depression [23]. Diverse techniques based on deep learning autonomously extract and comprehend the relationships between various features, enabling more precise, efficient, and reliable depression identification and prediction [24], [25]. The research mentioned above requires accumulating vast datasets to train and validate models in strict compliance with data privacy and ethical regulations. However, challenges such as potential data biases and limitations in interpretability might impact the generalizability.

NLP is an active research topic where scholars commonly gather textual data from diverse sources such as social media platforms, online forums, clinical interview records, and everyday conversations to identify depression [26], [27]. Common features include sentiment features and syntactic features, such as emotional vocabulary and polarity [28], topic modeling and word embeddings [29], and sentence complexity and word diversity [30]. The majority of NLP-based depression identification research focuses on deep learning, which encompasses convolutional neural networks (CNNs), recurrent neural

Manuscript received 9 August 2023; revised 24 November 2023; accepted 6 December 2023. This work was supported in part by the Natural Science Foundation of Gansu Province, China, under Grant 22JR5RA488; in part by the Fundamental Research Funds for the Central Universities under Grant lzujbky-2023-16; in part by the National Key Research and Development Program of China under Grant 2019YFA0706200; in part by the National Natural Science Foundation of China under Grants 62227807 and 62202212; in part by the SIT2030-Major Projects under Grant 2021ZD0200800; and in part by the Supercomputing Center of Lanzhou University. *(Corresponding author: Wenbo Chen.)*

Minqiang Yang, Zhuoheng Li, Yujie Gao, Chen He, and Fuzhan Huang are with the School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China (e-mail: yangmq@lzu.edu.cn).

Wenbo Chen is with the School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China (e-mail: chenwb@lzu.edu.cn).

Digital Object Identifier 10.1109/TCSS.2023.3343689

networks (RNNs), long short-term memory networks (LSTMs), and the latest pretrained language models such as bidirectional encoder representations from transformers (BERTs) [31], [32] and generative pretrained transformer (GPT) [33]. Although significant advancements have been achieved through these methodologies, they continue to encounter various challenges such as the effectiveness of model generalizability and ethical considerations. In recent years, there has been a growing trend in integrating research on depression identification using NLP with various disciplines such as computational psychology, multimodal analysis, and online behavior analysis aiming to improve the precision and applicability of depression identification methods [34], [35], [36].

Depression identification through multimodal behavior data always collects information on facial expressions, eye movements, audio, gait, and other paralinguistic aspects for multimodal fusion analysis [37], [38], [39]. Girard et al. extract facial landmarks, facial action unit (AU), and other features to investigate the relationship between depression symptoms and facial expression [40]. Nilsonne et al. focus on parameters such as fundamental frequency, speech rate, and acoustic features to reflect the emotional state and speaking style of the individuals [41]. Kessous et al. fuse and analyze multimodal information such as facial expressions and voice, which improve the effectiveness of depression identification [42]. The most common fusion methods include fusion at the feature level, fusion at the decision level, and multitask learning. Research on depression identification based on multimodal behavior data is closely related to other fields such as psychology, computer vision, and speech processing. This interdisciplinary approach contributes to a more comprehensive understanding of the physiological and psychological characteristics of depression and provides theoretical support for identification methods.

The Internet utilization-based depression assessment has become a focal point along with the booming of mobile internet, such as internet addiction [43], [44] and social network [45], [46]. Analyzing patterns in online communication, including social media posts [47] and forum discussions [48], has shown promise in identifying linguistic markers associated with depressive states. This approach not only allows for continuous monitoring but also offers the possibility of early intervention through targeted support and resources. However, challenges, including data privacy and ethical considerations, need resolution for the continued development and validation of Internet-based depression assessments.

Our research begins with weak-privacy online behavioral data, specifically the categories of websites visited over a period of time, to build a depression assessment model. It aims to detect depression using minimally weak-privacy data in a nondisturbed way, rather than delving into the detailed textual content generated by subjects on these websites. The data on internet browsing behavior encompass multiple dimensions of behavioral characteristics, including daily routines and entertainment preferences.

Prospective signals of depressive psychology can be identified with the robust feature extraction capabilities of deep learning models. The mental health status can be inferred based on the pathogenesis of depression, allowing for the prompt identification of individuals with abnormal symptoms. In this article, the network behavioral data are represented using graph structures that demonstrate the topological relationships among them. Graph convolutional networks (GCNs) are used in processing graph-structured data because of their superior performance [49]. At the same time, using attentional mechanisms can increase the interpretability of neural network structures and learn dynamic interactions between modalities more flexibly [50]. The graph attention network (GAT) [51] has both of these advantages.

We propose a heterogeneous graph attention network (H-GAT) model based on the GAT, which is designed specifically for heterogeneous graphs. The model uses attention mechanisms to compute the weights between heterogeneous nodes, thereby capturing complex network data patterns related to depression. The significance of this research resides in its ability to detect early-stage depression disorders among students, providing a foundation for timely intervention by relevant departments. Under favorable conditions, our method might be able to achieve automated processing and real-time prediction of a large volume of campus behavior data and establish efficient guidance, supervision, and information feed-back mechanisms. In summary, our contributions are listed as follows.

1) We propose a novel H-GAT that introduces attention mechanisms in heterogeneous graphs for depression detection. The ablation experiments and classification performance demonstrate that the proposed H-GAT model can effectively detect depression.
2) We innovatively utilize weak-privacy internet usage data to screen potential subjects with depressive tendencies. The detection process is inconspicuous, real-time, and objective, rendering it appropriate for extensive applications on a large scale.
3) We find some website categories with high contributions to campus depression, this would facilitate the daily management of related supervisors, which also promotes further cross-disciplinary research.

II. Experimental Data

The research data used in this article are derived from the cyber-activity of undergraduate students on the campus network. A total of 4490 undergraduate subjects participated, and all participants signed the electronic informed consent form. Because the participants in this study are all university students, the terms "subject" and "student" have the same meaning in this article. All the subjects are required to complete the scale of self-rating depression scale (SDS), the scores of which are used as labels of machine learning modeling. The answers, scores, and response time for each question are recorded for further data validation.

A behavior auditing device deployed in the backbone network is used to collect the surfing log with a low level of privacy, i.e., the website, the browse time, instantaneous throughput, terminal type, and flux direction of each access. The

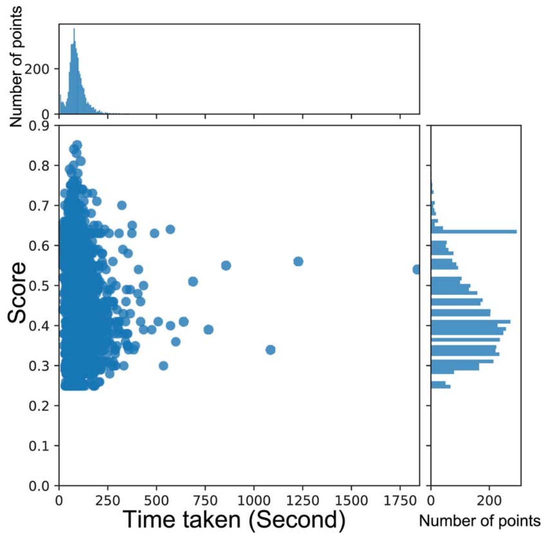


Fig. 1. Distribution of original labels.

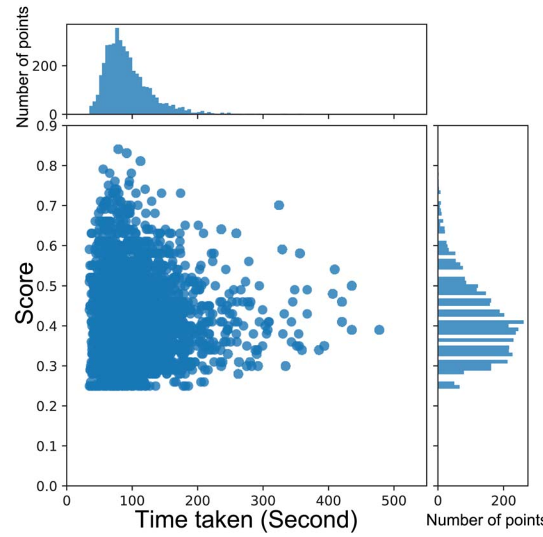


Fig. 2. Distribution of filtered labels.

built-in knowledge base of the behavior audit device identifies well-known application names through URLs and categorizes them into various types, such as web browsers, P2P downloads, office applications, shopping apps, network storage, steam, remote login, forums, instant messaging, microblogs, social networks, and email. The data collection lasts for 73 days. In this article, "cyber-activity" specifically refers to the above-mentioned internet behaviors with weak privacy conducted within the campus network.

Due to the subjective nature of self-assessment scales, a considerable portion of the scale data is not entirely reliable. The distribution of scale data of all subjects is shown in Fig. 1. The x-axis in the graph represents the time required for participants to complete the questionnaire, while the y-axis corresponds to the normalized SDS scores. The upper and right sides of the graph respectively illustrate the distribution of questionnaire scores over time and scores. It is worth noting that the proportion of the depression group in the original distribution is significantly higher than the incidence rate of depression [52]. It means that the initial label data are unreliable and additional screening is necessary. Through detailed observation and analysis of the questionnaire parameters, we found that some participants filled in the questionnaire carelessly or randomly, resulting in an excessively high false positive rate.

Taking advantage of the unique positive and negative scoring mechanisms of the SDS questionnaire, we can detect instances in which positive and negative scoring conflicts occur and filter students who make significant contradictory choices, ensuring the reliability of the label data. The distribution after filtering is shown in Fig. 2, and the constitution and representation of the axis and constitution are consistent with Fig. 1. The count and proportion of scale data before and after filtering are shown in Table I. It shows that after applying the aforementioned screening criteria, the number of depression labels is significantly reduced. In addition, samples with response times of less than 30 s and greater than 3600 s are completely excluded,

TABLE I
SDS Label Data Before and After Filtering

	Group	Count	Proportion
Raw Data	Depression	556	11.15%
	Healthy Control	3934	78.89%
Filtered	Depression	161	3.99%
	Healthy Control	3457	84.67%

as response times exceeding 10 min are considered clinically irrelevant. This indirectly verifies the effectiveness of the filtering method. A stratified undersampling is adapted to ensure balanced samples. The k-fold cross-validation technique is used to generate training and testing datasets for model training and evaluation.

III. Method

A straightforward concept underlying our proposed H-GAT model for the cyber-activity data in this study involves amalgamating traditional statistical methods with the shared pathogenesis of depression and behavioral phenotypes. This fusion aims to formulate an algorithm that assesses the frequency of subjects' visits to various websites and computes fundamental features, including average monthly frequency, average monthly access time, and so on. Even the classic machine learning algorithms like k-nearest neighbor (KNN) can achieve basic classification tasks effectively, graph structures might provide a more accurate representation of cyber-activity due to the complexity and specific structure. As the data of students contain two categories of nodes, students and websites, it is possible to deduce three types of relationships, i.e., edge $< \textit{student, student} >$, edge $< \textit{student, website} >$, and edge $< \textit{website, website} >$. Heterogeneous graphs can more effectively represent this structured data than homogeneous graphs.

Utilizing the well-established hierarchical attention networks (HANs) is a straightforward method for efficiently learning the network behavior characteristics of subjects. The two-level attention mechanism of HAN, i.e., node-level and semantic-level, is ideally adapted to resolve the complex relationships and dependencies between nodes in a heterogeneous graph, making it appropriate for processing graph-structured data with multiple categories of nodes and edges in this study. Nevertheless, exploring data analysis indicates that while HAN performs admirably on the training set, its performance on the test set is subpar. Inevitably, overfitting arises when employing HAN to the dataset in this investigation.

This study proposes the H-GAT model, which is based on the GAT and is designed specifically for heterogeneous graphs. The model uses attention mechanisms to compute the weights between heterogeneous nodes, thereby capturing complex cyber-activity patterns. It attempts to aggregate the features of students using known edge $<$ *student, website* $>$ and edge $<$ *website, website* $>$ features, and then performs binary classification for depression identification. The primary characteristic of the proposed H-GAT model, as opposed to the HAN structure, is the use of attention mechanisms to manage graph data with various categories of nodes and edges. The model uses graph attention layers to analyze relationships between *student-to-student*, *website-to-website*, and *student-to-website* for feature extraction without relying excessively on hierarchical attention mechanisms. It permits us to effectively depict the data using the attention mechanism within the framework of a heterogeneous graph, preventing overfitting and ensuring a better adaptation to the characteristics of cyber-activity.

Here, we define the input layer of the constructed H-GAT based on the analysis conducted. Let the input feature matrix be denoted as

$$X_s \in R^{(N_s \times F_s)} \tag{1}$$

and

$$X_w \in R^{(N_w \times F_w)} \tag{2}$$

where N_s represents the number of student nodes, F_s represents the initial number of features for each student node, N_w represents the number of website nodes, and F_w represents the initial number of features for each website node.

Next, we define the graph attention convolutional layer. First, calculate the attention scores

$$e_{uv} = \text{LeakyReLU}\left(W_\alpha \times \text{CONCAT}\left(h_v, h_u\right)\right) \tag{3}$$

where u and v represent two adjacent nodes, h_u and h_v represent the features of nodes u and v, $W_\alpha \in R^{\left(2F'\right)}$ is the weight matrix to be learned, and F' represents the dimension of the reduced features.

Next, normalize the attention weights

$$\alpha_{uv} = \exp\left(e_{uv}\right) / \Sigma\left(k \in N\left(v\right)\right) \exp\left(e_{vk}\right) \tag{4}$$

where α_{uv} represents the normalized attention weight between nodes u and v, and $N(v)$ denotes the set of neighboring nodes of node v.

Finally, update node features

$$h'_v = \text{ReLU}\left(W_1 \times \Sigma\left(u \in N\left(v\right)\right) \alpha_{uv} \times h_u\right) \tag{5}$$

where h'_v represents the updated features of node v, and $W_1 \in R^{\left(F' \times F'\right)}$ is the weight matrix to be learned.

After the graph attention convolutional layer, a pooling layer is defined in this article. Global pooling is used as follows:

$$g_s = \left(1/\left|S\right|\right) \times \Sigma\left(h_s\right) \tag{6}$$

$$g_w = \left(1/\left|W\right|\right) \times \Sigma\left(h_w\right) \tag{7}$$

where g_s and g_w represent the features of the student and website nodes after global pooling, respectively, S and W denote the sets of student nodes and website nodes, and $|S|$ and $|W|$ indicate the number of student nodes and website nodes, respectively.

Finally, the concatenated features of the student and website nodes are passed through a fully connected layer for nonlinear transformation

$$g = \text{CONCAT}\left(g_s, g_w\right) \tag{8}$$

$$f = \text{ReLU}\left(W_3 \times g\right). \tag{9}$$

In the equation provided, $g \in R^{\left(2F'\right)}$ represents the concatenated global features, $f \in R^{F''}$ represents the output of the fully connected layer, where F'' represents the output dimension of the fully connected layer. $W_3 \in R^{\left(F'' \times 2F'\right)}$ is the weight matrix to be learned.

The Softmax function is used as the activation function in the output layer

$$y = \text{Softmax}\left(W_4 \times f\right). \tag{10}$$

In the given equation, $y \in [0, 1]$ represents the predicted result, and $W_4 \in R^{\left(1 \times F''\right)}$ is the weight matrix to be learned.

Based on the above, the overall network architecture of H-GAT, an improved version derived from the GAT network model, can be represented as the following equation

$$\begin{aligned} y_i = \text{Softmax}_i\Big(\Big(W_4 \times \text{ReLU}\Big(W_3 \times \text{CONCAT}\Big(\Big(\left(1/\left|S\right|\right) \\ \times \Sigma\Big(\text{ReLU}\Big(W_1 \times \Sigma\Big(\alpha_{uv} \times h_v\Big) + b_1\Big)\Big)\Big), \Big(\left(1/\left|W\right|\right) \\ \times \Sigma\Big(\text{ReLU}\Big(W_1 \times \Sigma\Big(\alpha_{uv} \times h_v\Big) + b_1\Big)\Big)\Big)\Big) + b_3\Big)\Big) + b_4\Big) \end{aligned} \tag{11}$$

where

$$\begin{aligned} \alpha_{uv} = \text{Softmax}_u\Big(\text{LeakyReLU}\Big(a^T \\ \times \text{CONCAT}\left(W_2 \times h_u, W_2 \times h_v\right) + b_2\Big)\Big). \end{aligned} \tag{12}$$

In this model, y_i represents the ith output result, i.e., the probability that a student belongs to the ith class, S represents the set of student nodes, and W represents the set of website nodes. α_{uv} denotes the attention weight between nodes u and v, while h_v represents the feature vector of node v.

W_1, W_3, and W_4 are the weight matrices for the graph attention convolutional layer, fully connected layer, and output layer, respectively. W_2 is the weight matrix used to compute the attention weights. b_1, b_3, and b_4 represent the bias terms for the graph attention convolutional layer, fully connected layer, and output layer, respectively. b_2 is the bias term used in the calculation of attention weights.

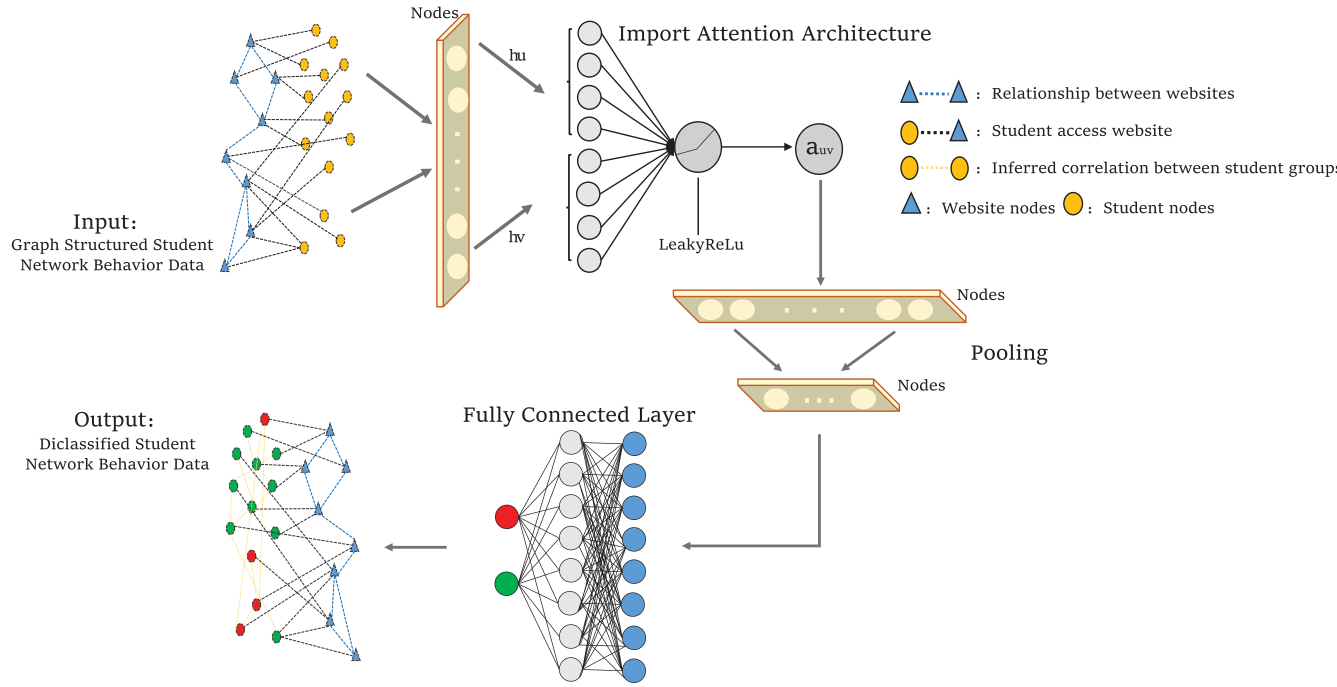
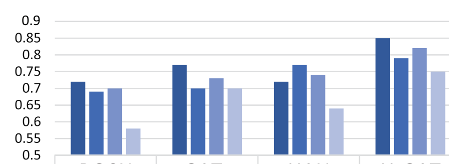


Fig. 3. Architecture diagram of network behavior recognition model based on H-GAT.



	RGCN	GATs	HAN	H-GAT
Precision	0.72	0.77	0.72	**0.85**
Recall	0.69	0.7	0.77	**0.79**
F1-score	0.7	0.73	0.74	**0.82**
AUC	0.58	0.7	0.64	**0.75**

Fig. 4. Ablation experiment result.

The aforementioned expression illustrates the calculation process of the H-GAT model, which incorporates activation functions and bias elements. We enhance the model's nonlinear expressive capacity by incorporating rectified linear unit (ReLU) activation functions and bias terms in the graph attention convolutional layer and the fully connected layer. In the output layer, continuous values are transformed into a probability distribution for each class using the Softmax function. The architecture of the proposed H-GAT model is shown in Fig. 3.

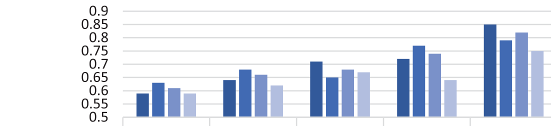
IV. Result and Analyzation

All experiments and code are implemented based on TensorFlow 2.0 framework. We trained our model with a batch size of 32 for a total of 200 epochs. We use the LeakyReLU function in the graph attention convolutional layer and fully connected layer, and use the Softmax function to transform the probability distribution of each category. During the backpropagation phase, Adam optimizer is utilized with an initial learning rate set to 0.001. We maintained the same metrics to evaluate the classification, including precision, recall, and f1 score. For evaluating the performance of the H-GAT model, we conducted an ablation study by removing the attention layer, i.e., recurrent graph convolutional networks (RGCNs), using isomorphic nodes and self-loops to represent features, i.e., GAT, and incorporating semantic-level attention, i.e., HAN. The ablation experiment results are summarized in Fig. 4.

In terms of accuracy and recall, the proposed H-GAT model consistently outperforms the heterogeneous graph RGCN model, which does not include attention mechanisms, as shown in the table of Fig. 4. This demonstrates the substantial benefit of integrating attention mechanisms with heterogeneous graphs to represent complex graph structures in student cyber-activity. In addition, the H-GAT model is significantly more efficient than the GAT model, which incorporates attention mechanisms within a homogeneous graph. This indicates that heterogeneous graphs are adaptable to cyber-activity pattern mining.

The H-GAT model proposed in this article assigns different attention weights to each node within the same hierarchy, thereby better understanding and modeling the complex relationships between nodes. On the other hand, HAN focuses mainly on hierarchical attention mechanisms [53], which model heterogeneous graphs by hierarchically processing multiple types of relationships between nodes. However, our student cyber-activity only includes two types of heterogeneous nodes: students and websites, and no apparent hierarchical structure is observed. This leads to overfitting when using HAN, resulting in less prominent performance on the test set.

We anticipate that various types of edges will acquire different attention mechanisms, designating distinct attention weights



	KNN	RNN	GCN	HAN	H-GAT
Precision	0.59	0.64	0.71	0.72	**0.85**
Recall	0.63	0.68	0.65	0.77	**0.79**
F1-score	0.61	0.66	0.68	0.74	**0.82**
AUC	0.59	0.62	0.67	0.64	**0.75**

Fig. 5. Performance comparison of H-GAT model with other models.

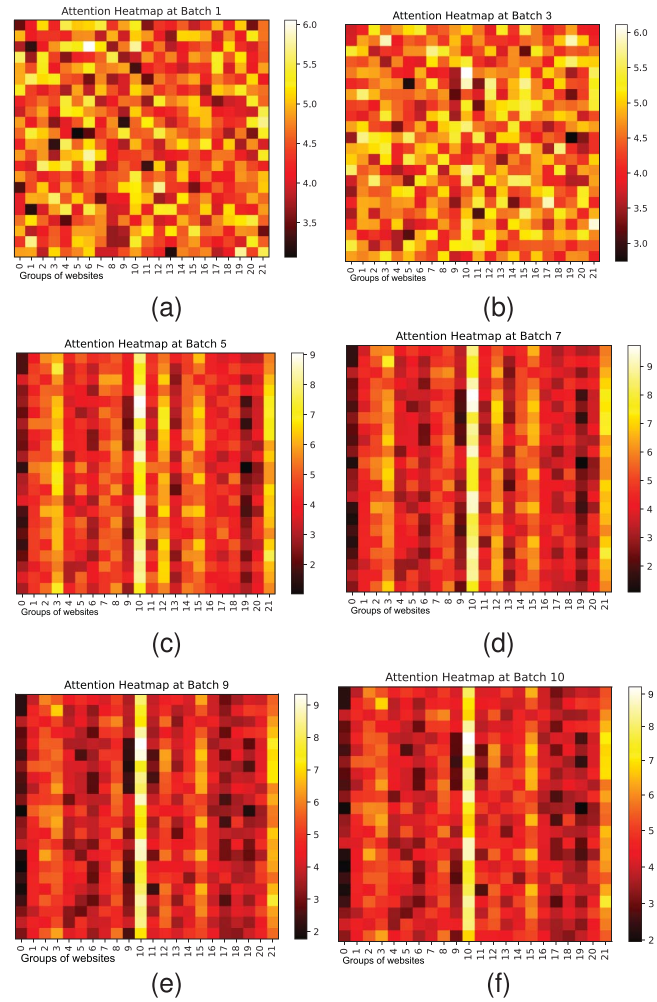


Fig. 6. Attention heatmap at (a) batch 1; (b) batch 3; (c) batch 5; (d) batch 7; (e) batch 9; and (f) batch 10.

to each edge type, Without excessively concentrating on the complexity of semantic levels and meta-paths. This makes the proposed H-GAT architecture potentially more advantageous than the extant HAN architecture when dealing with intricate relationships within the same hierarchy. Fig. 5 presents a performance comparison between the proposed H-GAT model and commonly used classical models on the dataset used in this study. In terms of both precision and recall, it is evident that the proposed H-GAT model outperforms other models. Based on the student dataset utilized for this study, it predicts depression with high accuracy.

Based on the H-GAT model proposed in this study, the heatmaps illustrate variations in attention at different phases of model training, which are shown in Fig. 6. As the number of websites in a collection increases, the disparities in allocated attention become more pronounced. We can attempt to improve the interpretability of the deep learning procedure by introducing attention mechanisms.

These attention heatmaps can help us understand the model's emphasized focal points in the interaction between student nodes and website nodes. This aids in elucidating the model's decision-making process and optimizing its behavior and performance. The color in the attention heatmap indicates the level of focus allocated by the model to various students and websites. When a position in the heatmap is brighter, it indicates that the model attributes a greater attention weight to the relationship between the student and the website. This indicates that in the model's decision-making process, the relationship between the student and the website plays a crucial role in determining the output. Specifically, areas receiving considerable focus indicate the following entities.

1) Key websites: The model may place a greater emphasis on certain websites because they contain more crucial characteristics for determining the depression status of students. These important websites may contain information about the mental health of students, depressive symptoms, or other significant indicators.
2) Important students: The model may pay more attention to certain students, perhaps because they play a greater role in determining the depressive status. These prominent students may exhibit unique characteristics, behavioral patterns, or other significant factors associated with depression.
3) Individual differences: differing regions of the attention heatmap may reflect differing patterns and frequencies of website visits among students. Higher attention weights may indicate that certain students frequently visit specific websites or that their visit patterns differ from those of other students.

The categories of "Tianya Forum," "Weibo," "Aliwangwang Image Transmission," "Steam," "iCloud Drive," "Online Banking," and "Other Online Games" attract more attention in the heatmaps. These websites significantly influence the decision-making process of the model. Through the observation of this series of heatmaps, it is found that their results are consistent with the subsequent visual analysis results in this article.

The overall graph of the model after training is shown in Fig. 7. The blue node clusters represent different types of websites, and the green and red nodes aggregated inside them represent healthy and depression students, respectively. In this figure, due to the large number of nodes involved, most of the edges overlap and form a complex black mesh region. To understand the model more clearly, we intercept some of the student nodes and website nodes to explore them below.

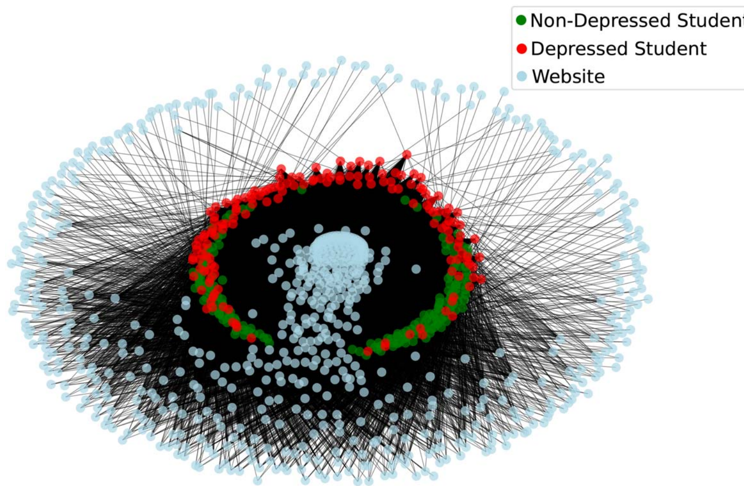


Fig. 7. Model morphological structure. The figure shows the general morphological structure of the model as a whole in two dimensions, where the red and green dots represent the two groups of students, and the blue dots connected to them represent the different sites and correlations.

It is worth pointing out that although there are a total of 680 types of websites in the dataset, many of them have very few views, and some of them have only been viewed by one or two students, this type of website is not uncommon in the dataset. To effectively extract the websites that have a high impact on identifying whether a student is depressed or not from a large amount of website data, we introduced a multiple attention mechanism during group training. By segmenting the input sequence of student-website relationships and calculating attention weights for each attention head, the outputs of multiple attention heads are finally combined by weighted summation to form the final attention representation.

We assign the frequency information of the websites visited by students to the website feature matrix and perform a dimension-increasing operation in the feature matrix function so that it can be mapped to the features corresponding to all 680 websites. Subsequently, attention computation is conducted on the two input high-dimensional matrices designated as the $<$ *website, website* $>$ attention matrix, and the website-features matrix. This calculates the average attention value between each website and other websites and returns the updated website features. This process enables each attention head to learn different attention weight distributions in parallel. The learned attention weights are then utilized to perform a weighted summation and combination of the input student-website relationship sequences, facilitating the integration of information between various feature representations and enhancing noise resistance. Consequently, the H-GAT model acquires a more global and comprehensive representation of the sequences, as well as a richer and more extensive representational capacity. This benefits the effective capture of diverse feature representations between student nodes and website nodes. Meanwhile, since each attention head is independent of the others, the multihead attention mechanism allows parallel computation among multiple attention heads without serial processing, accelerating the training and inference process of this modeling. We divided the website data into seven groups for input, and Fig. 8 shows the change process of attention weight allocation in the first group of data. The numbers of the top five websites with the largest attention weights at that moment in each graph are shown in the picture.

TABLE II
WEBSITE NUMBER WITH LARGEST ATTENTIONAL WEIGHT

Step	Website Number				
Step 1	0	1	2	3	4
Step 2	0	39	16	9	21
Step 3	0	3	16	21	39
Step 4	0	3	16	39	66
Step 5	0	3	16	21	39
Step 6	39	21	16	3	0
Step 7	39	3	16	66	0
Step 8	39	6	11	16	0
Step 9	39	66	16	3	0
Step 10	39	66	16	3	0

In Fig. 8, the more vivid color shows the attention allocation at that moment, in which colors such as red and orange are the highest, and the weight decreases sequentially from yellow, black to white, which represents a corresponding decrease in the allocated attention. For example, in the late iteration, almost all the self-loops are categorized as white. At the initial moment of learning, we sequentially select the first five website nodes, i.e., five websites numbered "0, 1, 2, 3, 4," respectively, as the initial object of attention distribution, and it can be seen that the weights are gradually shifted as learning proceeds. At the beginning of the training, the attention distribution is more random and uniform. As the learning process proceeds, the more influential nodes are aggregated to the inner center with the allocation of attention, while the unimportant website nodes are excluded to the outer circle, forming a radial graph. This process is very intuitively reflected in Fig. 8, where lines of various colors are richly distributed in a random radial pattern at the very beginning, and as nodes with high attentional weights are gradually aggregated to the interior of the image, the vivid colors in the graph representing the high weights are also reduced. In the above process, the number of websites with the largest

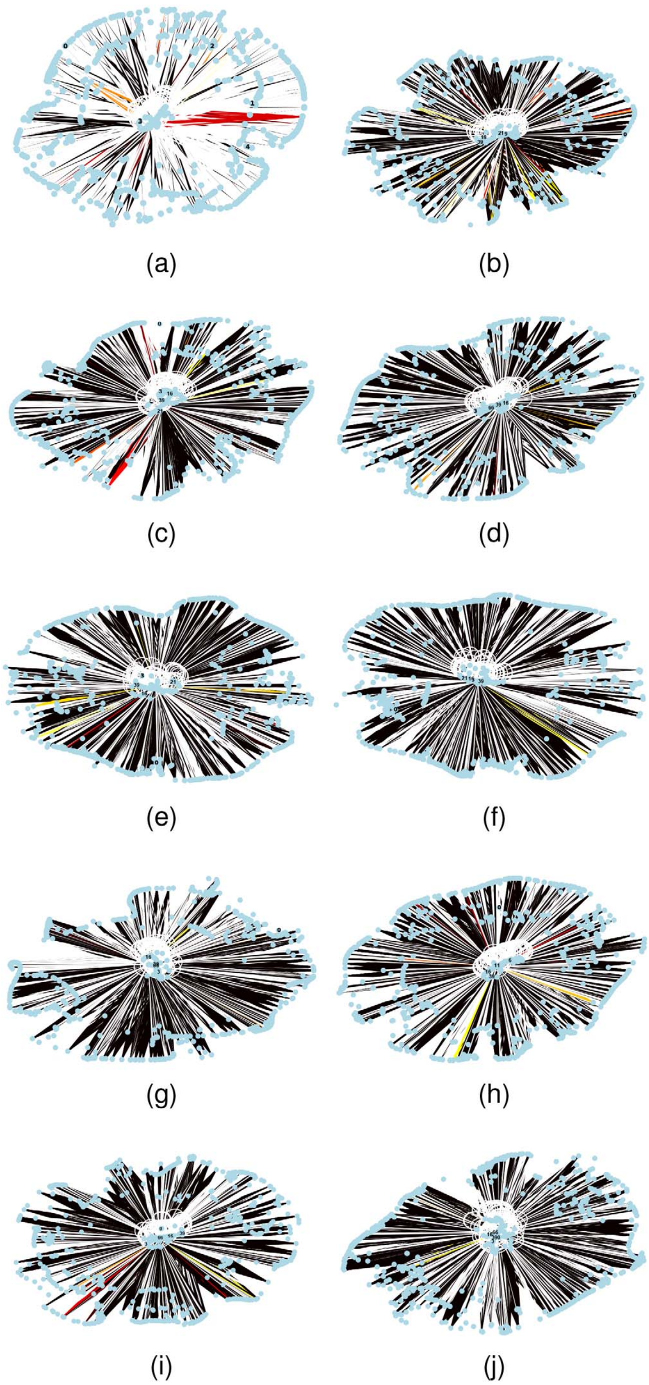


Fig. 8. Website weights at (a) step 1; (b) step 2; (c) step 3; (d) step 4; (e) step 5; (f) step 6; (g) step 7; (h) step 8; (i) step 9; and (j) step 10.

attentional weight at each moment and the corresponding website name are shown in Tables II and III. It can be seen from the table that, with the increase in the number of training times, the attentional weight gradually stabilizes on certain types of websites, which are similar to the above-mentioned "Tianya Forum," "Weibo," "Aliwangwang Image Transmission," "Steam," "iCloud Drive," "Online Banking," and "Other Online Games," which also confirms the effectiveness of the model.

Fig. 9 shows the distribution of all network nodes connected to the node corresponding to a particular student under the multiple attention mechanism after training. The red nodes near the inner circle are the nodes of the websites that are calculated to require more attention allocation, while the websites with relatively low allocation of attention weights are quadratically dispersed to the outer side, which we consider to have less impact on recognizing depression.

Fig. 10(a) shows the distribution of the top five websites and their associated students selected for ranking influence weights, where websites are blue nodes, depression students are red nodes, and healthy control are green nodes. Fig. 10(b) shows the students associated with five general website nodes selected. It can be seen that the student nodes connected to the top five

TABLE III
WEBSITE NAME MAP

Number	Application
0	League of Legends
1	Mobile Tiexue Forum
2	IT home
3	AliWangwang Image Transfer
4	Gameloft
6	NTP
9	Dajie Web
16	Steam
21	Xiaomi Store
39	Zhaopin
66	Software download

TABLE IV
P-VALUES OF PERMANOVA

Website	P-Value
Tianya Forum	0.042
Weibo	0.157
Aliwangwang Image Transmission	0.354
Steam	0.013
iCloud Drive	0.047
Online Banking	0.011
Other Online Games	0.601
Average of all over 600 websites	0.582

Fig. 9. Network nodes distribution. In this figure, we chose one subject node and show the architecture graph of all the websites associated with that node. It can be seen that under the influence of the attention mechanism, the stronger the association, the closer the website is to the center.

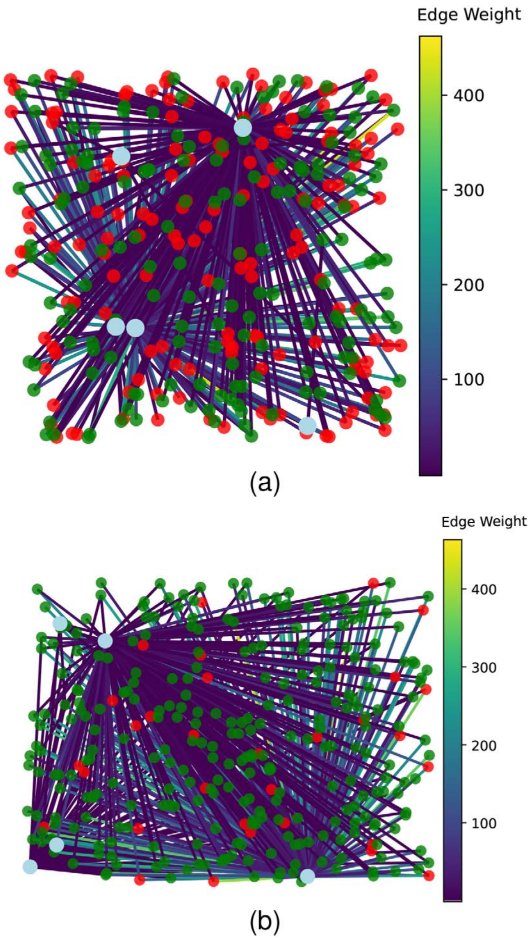


Fig. 10. Selected websites and connected students. In (a), we select five highly weighted website nodes and print the architecture graph of all student nodes associated with them, while (b) shows the architecture graph of the normal five website nodes. The difference between the student population associated with the nodes screened by the attention mechanism and the normal nodes can be clearly seen.

weighted websites that we have filtered out do show significant differences. In this graph, the weights influence the color of the edges, the darker the color of the edge, the deeper the influence of the website on the decision-making of the corresponding student node, and the greater the contribution of the connection between the student nodes of that website to the overall model.

Considering that the dimensionality of the model is too high and the presentation of all the data is poor, in Fig. 11. The whole graph of all the data is hard to draw. We randomly select ten students, including five depression students and five healthy students represented in red and green, respectively, and draw a partial image of all the websites they are connected to as a way to help us understand the model in a general way. In Fig. 11, the red or green nodes with hash coding are the elected student nodes, and the outer blue nodes are all the website nodes connected by the ten student nodes. The darker the color of the edge connected to the student node, the closer the website node is to the selected student node, and the greater the attention weight assigned to this edge by the model learning in the model construction process. Similarly, the closer the website node is to the center, the more important the website node is. The seven most prominent website nodes in the big ring are "Tianya Forum," "Weibo," "Aliwangwang Image Transmission," "Steam," "iCloud Drive," "Online Banking," and "Other Online Games." These categories possess great repetition rate with the websites that contribute more to the model as given in Figs. 6 and 8.

In our study, websites such as "Tianya Forum," "Weibo," "Aliwangwang Image Transmission," and "Steam" were found to have a significant impact on the updating of the attention node graph. The empirical evidence substantiates the assertion that these sites contributed better to the sample group of college students. But this does not mean that the group with depressive tendencies loves visiting these sites more, it just means that these sites play a greater decision-making weight under our sample. Additionally, while our data provide insights into specific online environments, it is incumbent upon us to consider the potential for unexplored digital domains that may exhibit relevance. It is also prudent to recognize that there is inherent variability in digital engagement patterns across different national contexts and sociodemographic groups. The conclusions drawn from our dataset are primarily reflective of the behaviors of university students within the context of Chinese higher education institutions.

To explore whether there were significant differences in the behavior of depressed and normal groups when accessing different websites, we conducted hypothesis testing on the selected websites with higher weights and normal websites in the decision-making process. Due to the fact that the network behavior data do not follow a conventional normal distribution, we chose the nonparametric method permutational multivariate analysis of variance (PERMANOVA) for statistical significance testing to compare the differences between the two groups in accessing different websites. The null hypothesis (H0) is formulated as: there is no significant difference in the behavior of visiting the website between the depressed group and the normal group. Consequently, the alternative hypothesis (H1) posits that there is a significant difference in the behavior of visiting the website between the depressed group and the normal group. By constructing a permutation distribution, the p-values of different websites are shown in Table IV.

From Table IV, it can be seen that the p-values of Tianya Forum, Steam, iCloud Drive, and Online Banking are less than the predetermined significance level (0.05), indicating that there are significant differences in the visiting behavior of different groups on these websites, thus accepting the alternative hypothesis. In the remaining websites, such as Weibo, although its p-value is closer to the significance level of 0.05 than normal websites, it is still not enough to reject the null hypothesis. We cannot consider that there is sufficient evidence to suggest significant differences in the visiting behavior of the depressed population compared to the normal population on these websites.

The results of the hypothesis testing align with our previous conclusion that the selected websites exhibit high weights in the updating process of the graph neural network, indicating their significant contribution to the decision-making process. However, these results should not be simplistically interpreted to infer a positive or negative correlation with the depressed

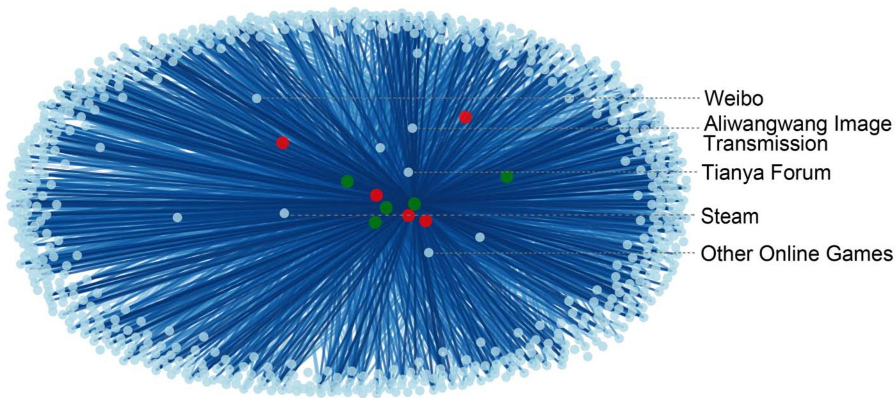


Fig. 11. Selected subjects and connected websites. In this figure, five healthy subjects and five subjects with depressive tendencies were selected and all the website nodes they were associated with were printed. The website nodes closer to the middle section have a greater impact on model discrimination. It can be seen that the sites discussed in the previous section that have a high degree of influence on the model are also more prominently featured in this architecture diagram.

group. Identifying characteristics of the depressed group requires nonlinear decision-making of complex networks and multiple considerations.

Despite the limitations and specificity of our dataset, we can still try to analyze why those kinds of websites get high attention from the perspective of cognitive and behavioral psychology on this basis. First, these websites or platforms attract numerous users, who often share their daily lives, emotions, and thoughts on these platforms. Especially on social media platforms such as “Tianya Forum” and “Weibo,” people are more inclined to share their emotional states, whether they are happy or depressed. This could potentially provide researchers with rich data to analyze and identify potential signs of depression. Similarly, “Steam” is a gaming platform, and the fact that some players may be addicted to gaming to escape from real-life stresses or problems may be a clue to their emotional state. However, it should be noted that online games require a lot of energy, and depressed patients usually show a loss of interest and energy at specific times of the day, etc. Combining the periods when players log on to the game may help analyze the problem. “Aliwangwang Image Transmission” is a web record left by users when they login to Taobao and chat with customer service. Taobao is the largest e-commerce platform in China, and users’ interactions with it are mainly related to shopping. People’s shopping behaviors, evaluations, interactions, or search records may indirectly reflect their mental states and emotions. For example, some people may regulate their emotions through shopping, and frequent impulse shopping or the purchase of a particular item may be related to their emotional state. In addition, the communication records between buyers and sellers on “Aliwangwang” may also provide some clues, such as their language use and frequency of communication. However, as the data on the content of the chat are too private to be accessed, we can only make a general guess based on the time in which the incident took place.

V. Conclusion

This study proposes a H-GAT model, which incorporates attention mechanisms and makes use of the interactions between heterogeneous nodes in a heterogeneous graph to better capture complex patterns in cyber-activity data. On our collected dataset, the H-GAT architecture avoids the problem of overfitting induced by the hierarchical attention mechanism of the conventional HAN architecture. Utilizing attention mechanisms, the model adaptively allocates weights to each node during the feature extraction and training phases, thereby facilitating improved learning of node features and achieving favorable recognition performance.

As a prospective and interdisciplinary discipline in the era of data science, depression recognition research based on digital behavior phenotype has significant practical significance and enormous growth potential. In future research, obtaining more comprehensive network behavior data while addressing privacy concerns is crucial for the development of effective depression assessment models. Long-term monitoring research may be conducted to evaluate the long-term effects of depression recognition and intervention measures. Longitudinal studies can also contribute to research on model interpretability and causal inference, allowing for a greater understanding of the causes and influencing factors of depression.

References

[1] M. T. Shea and R. M. Hirschfeld, “Chronic mood disorder and depressive personality,” *Psychiatr. Clin. North Amer.*, vol. 19, no. 1, pp. 103–120, 1996.
[2] D. L. Evans and D. S. Charney, “Mood disorders and medical illness: A major public health problem,” *Biol. Psychiatry*, vol. 54, no. 3, pp. 177–180, 2003.
[3] R. Beiter et al., “The prevalence and correlates of depression, anxiety, and stress in a sample of college students,” *J. Affect. Disorders*, vol. 173, pp. 90–96, 2015.
[4] J. S. Westefeld and S. R. Furr, “Suicide and depression among college students,” *Prof. Psychol., Res. Pract.*, vol. 18, no. 2, p. 119, 1987.
[5] S. R. Furr, J. S. Westefeld, G. N. McConnell, and J. M. Jenkins, “Suicide and depression among college students: A decade later,” *Prof. Psychol., Res. Pract.*, vol. 32, no. 1, p. 97, 2001.
[6] H. Eastman-Mueller, T. Wilson, A.-K. Jung, A. Kimura, and J. Tarrant, “iRest yoga-nidra on the college campus: Changes in stress, depression, worry, and mindfulness,” *Int. J. Yoga Therapy*, vol. 23, no. 2, pp. 15–24, 2013.
[7] J. L. Buchanan, “Prevention of depression in the college student population: A review of the literature,” *Arch. Psychiatr. Nurs.*, vol. 26, no. 1, pp. 21–42, 2012.
[8] D. Lester, “Depression and suicidal ideation in college students: A preliminary study of campus variables,” *Psychol. Rep.*, vol. 112, no. 1, pp. 106–108, 2013.

[9] R. Wang et al., "Tracking depression dynamics in college students using mobile phone and wearable sensing," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 1, pp. 1–26, 2018.

[10] X. Gong, C. P. Chen, and T. Zhang, "Cross-cultural emotion recognition with EEG and eye movement signals based on multiple stacked broad learning system," *IEEE Trans. Comput. Social Syst.*, 2023, doi: 10.1109/TCSS.2023.3298324.

[11] W. W. Eaton, C. Muntaner, C. Smith, A. Tien, and M. Ybarra, "Center for epidemiologic studies depression scale: Review and revision," in *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment*, NJ, USA: Lawrence Erlbaum Associates, 2004.

[12] S. Suseelan and G. Pinna, "Heterogeneity in major depressive disorder: The need for biomarker-based personalized treatments," in *Advances in Clinical Chemistry*, Elsevier: Amsterdam, Netherlands, vol. 112. 2022.

[13] A. Dhir, Y. Yossatorn, P. Kaur, and S. Chen, "Online social media fatigue and psychological wellbeing—A study of compulsive use, fear of missing out, fatigue, anxiety and depression," *Int. J. Inf. Manage.*, vol. 40, pp. 141–152, 2018.

[14] R. Wang et al., "StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 3–14.

[15] M. Trotzek, S. Koitka, and C. M. Friedrich, "Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 3, pp. 588–601, Mar. 2020.

[16] M. Yang, C. Cai, and B. Hu, "Clustering based on eye tracking data for depression recognition," *IEEE Trans. Cogn. Develop. Syst.*, vol. 15, no. 4, pp. 1754–1764, Dec. 2023.

[17] M. Yang, Y. Ma, Z. Liu, H. Cai, X. Hu, and B. Hu, "Undisturbed mental state assessment in the 5G era: A case study of depression detection based on facial expressions," *IEEE Wireless Commun.*, vol. 28, no. 3, pp. 46–53, Jun. 2021.

[18] B. Hu, Y. Tao, and M. Yang, "Detecting depression based on facial cues elicited by emotional stimuli in video," *Comput. Biol. Med.*, vol. 165, 2023, Art. no. 107457.

[19] J. Shen et al., "Depression recognition from EEG signals using an adaptive channel fusion method via improved focal loss," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 7, pp. 3234–3245, Jul. 2023.

[20] J. Shen et al., "Exploring the intrinsic features of EEG signals via empirical mode decomposition for depression recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 356–365, 2023.

[21] B. Liu, J. Guo, C. P. Chen, X. Wu, and T. Zhang, "Fine-grained interpretability for EEG emotion recognition: Concat-aided Grad-CAM and systematic brain functional network," *IEEE Trans. Affect. Comput.*, 2023, doi: 10.1109/TAFFC.2023.3288885.

[22] S. Sardari, B. Nakisa, M. N. Rastgoo, and P. Eklund, "Audio based depression detection using convolutional autoencoder," *Expert Syst. Appl.*, vol. 189, 2022, Art. no. 116076.

[23] M. Z. Uddin, K. K. Dysthe, A. Følstad, and P. B. Brandtzaeg, "Deep learning for prediction of depressive symptoms in a large textual dataset," *Neural Comput. Appl.*, vol. 34, no. 1, pp. 721–744, 2022.

[24] M. Li, X. Sun, and M. Wang, "Detecting depression with heterogeneous graph neural network in clinical interview transcript," *IEEE Trans. Comput. Social Syst.*, 2023, doi: 10.1109/TCSS.2023.3263056.

[25] Y. Tao, M. Yang, Y. Wu, K. Lee, A. Kline, and B. Hu, "Depressive semantic awareness from vlog facial and vocal streams via spatio-temporal transformer," *Digit. Commun. Netw.*, early access, Mar. 27, 2023, doi: 10.1016/j.dcan.2023.03.007.

[26] J. Robinson et al., "Social media and suicide prevention: A systematic review," *Early Intervention Psychiatry*, vol. 10, no. 2, pp. 103–121, 2016.

[27] T. Saha, S. M. Reddy, S. Saha, and P. Bhattacharyya, "Mental health disorder identification from motivational conversations," *IEEE Trans. Comput. Social Syst.*, vol. 10, no. 3, pp. 1130–1139, Jun. 2023.

[28] F. Jia and C.-C. Chen, "Emotional characteristics and time series analysis of internet public opinion participants based on emotional feature words," *Int. J. Adv. Robotic Syst.*, vol. 17, no. 1, 2020, Art. no. 1729881420904213.

[29] F. Viegas et al., "CluWords: Exploiting semantic word clustering representation for enhanced topic modeling," in *Proc. 12th ACM Int. Conf. Web Search Data Min.*, 2019, pp. 753–761.

[30] T. Alexopoulou, M. Michel, A. Murakami, and D. Meurers, "Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques," *Lang. Learn.*, vol. 67, no. S1, pp. 180–208, 2017.

[31] X. Wang et al., "Depression risk prediction for Chinese microblogs via deep-learning methods: Content analysis," *JMIR Med. Inform.*, vol. 8, no. 7, 2020, Art. no. e17958.

[32] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[33] X. Han et al., "Pre-trained models: Past, present and future," *AI Open*, vol. 2, pp. 225–250, 2021.

[34] E. A. Ríssola, D. E. Losada, and F. Crestani, "A survey of computational methods for online mental state assessment on social media," *ACM Trans. Comput. Healthcare*, vol. 2, no. 2, pp. 1–31, 2021.

[35] M. Schindler and E. Domahidi, "The computational turn in online mental health research: A systematic review," *New Media Soc.*, vol. 25, no. 10, pp. 2781–2799, 2023.

[36] D. Sik, M. Rakovics, J. Buda, and R. Németh, "The impact of depression forums on illness narratives: A comprehensive NLP analysis of socialization in e-mental health communities," *J. Comput. Social Sci.*, pp. 1–22, 2023.

[37] S. Alghowinem et al., "Multimodal depression detection: Fusion analysis of paralinguistic, head pose and eye gaze behaviors," *IEEE Trans. Affect. Comput.*, vol. 9, no. 4, pp. 478–490, Oct.–Dec. 2018.

[38] M. Yang, Y. Wu, Y. Tao, X. Hu, and B. Hu, "Trial selection tensor canonical correlation analysis (TSTCCA) for depression recognition with facial expression and pupil diameter," *IEEE J. Biomed. Health Inform.*, 2023, doi: 10.1109/JBHI.2023.3322271.

[39] M. Niu, J. Tao, B. Liu, J. Huang, and Z. Lian, "Multimodal spatiotemporal representation for automatic depression level detection," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 294–307, Jan.–Mar. 2023.

[40] J. M. Girard, J. F. Cohn, M. H. Mahoor, S. Mavadati, and D. P. Rosenwald, "Social risk and depression: Evidence from manual and automatic facial expression analysis," in *Proc. 10th IEEE Int. Conf. Workshops Autom. Face Gesture Recognit. (FG)*, Piscataway, NJ, USA: IEEE, 2013, pp. 1–8.

[41] Å. Nilsonne, J. Sundberg, S. Ternström, and A. Askenfelt, "Measuring the rate of change of voice fundamental frequency in fluent speech during mental depression," *J. Acoust. Soc. Amer.*, vol. 83, no. 2, pp. 716–728, 1988.

[42] L. Kessous, G. Castellano, and G. Caridakis, "Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis," *J. Multimodal User Interfaces*, vol. 3, pp. 33–48, 2010.

[43] P. Wang, J. Wang, Y. Yan, Y. Si, X. Zhan, and Y. Tian, "Relationship between loneliness and depression among Chinese junior high school students: The serial mediating roles of internet gaming disorder, social network use, and generalized pathological internet use," *Front. Psychol.*, vol. 11, 2021, Art. no. 529665.

[44] X. Dong et al., "Depression mediates the relationship between childhood trauma and internet addiction in female but not male Chinese adolescents and young adults," *J. Clin. Med.*, vol. 10, no. 21, 2021, Art. no. 5015.

[45] S. Yoon, M. Kleinman, J. Mertz, and M. Brannick, "Is social network site usage related to depression? A meta-analysis of Facebook–depression relations," *J. Affect. Disorders*, vol. 248, pp. 65–72, 2019.

[46] F. Cacheda, D. Fernandez, F. J. Novoa, and V. Carneiro, "Early detection of depression: Social network analysis and random forest techniques," *J. Med. Internet Res.*, vol. 21, no. 6, 2019, Art. no. e12554.

[47] T. Yang et al., "Fine-grained depression analysis based on Chinese micro-blog reviews," *Inf. Process. Manage.*, vol. 58, no. 6, 2021, Art. no. 102681.

[48] J. Feldhege, M. Moessner, and S. Bauer, "Who says what? Content and participation characteristics in an online depression community," *J. Affect. Disorders*, vol. 263, pp. 521–527, 2020.

[49] M. Ye, C. P. Chen, and T. Zhang, "Hierarchical dynamic graph convolutional network with interpretability for EEG-based emotion recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, 2022, doi: 10.1109/TNNLS.2022.3225855.

[50] T. Zhang, S. Li, B. Chen, H. Yuan, and C. P. Chen, "AIA-Net: Adaptive interactive attention network for text–audio emotion recognition," *IEEE Trans. Cybern.*, vol. 53, no. 12, pp. 7659–7671, Dec. 2023.

[51] P. Veličković G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2018, p. 12.

[52] A. K. Ibrahim, S. J. Kelly, C. E. Adams, and C. Glazebrook, "A systematic review of studies of depression prevalence in university students," *J. Psychiatric Res.*, vol. 47, no. 3, pp. 391–400, 2013.

[53] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proc. Conf. North Amer. Chap. Assoc. Comput. Ling.: Human Lang. Technol.*, 2016, pp. 1480–1489.