Application of artificial intelligence in the field of psychological evaluation and intervention of college students

Xurui Liao[1], Yao Chen[2], Heng Wang[1], Gaodong Duan[3], Meina Li[4], Junchen Li[5]

[1] *School of Mathematics and Computer Science / Department of Computer Science, Guangdong Ocean University, Zhanjiang, Guangdong*

[2] *School of Management / Department of Financial Management, Guangdong Ocean University, Zhanjiang, Guangdong*

[3] *School of Physical Education, Hunan Agricultural University, Changsha, Hunan*

[4] *School of Literature and Journalism / Department of International Chinese Language Education, Guangdong Ocean University, Zhanjiang, Guangdong*

[5] *School of Education / Department of Physical Education, Jiangxi Institute of Applied Science and Technology, Nanchang, Jiangxi*

Received: 28 Feb 2025
Revised: 18 March 2025
Accepted: 29 March 2025
Published: 10 April 2025

Abstract*:* With the rapid development of artificial intelligence (AI) technology, its application in the field of mental health is increasingly extensive. As an important group in society, college students' mental health problems have attracted much attention. This paper discusses the present situation of artificial intelligence intervention in college students' psychological evaluation, and analyzes the application limitations. Aiming at the speed and accuracy of evaluation, a multi-modal model based on feature fusion and decision fusion is proposed to integrate multi-dimensional data to improve the efficiency and accuracy of evaluation. The results show that the performance of the model is better than that of the traditional method, which pro vides support for the intelligentization of college students' mental health services. The conclusion is that the model provides an effective way to solve the limitation of evaluation.

Keywords: Artificial intelligence; College students; Psychological evaluation; Psychological intervention; Multimodal model

1 INTRODUCTION

In recent years, the mental health problem of college students has received increasing attention. With the increasing academic pressure, employment pressure and social pressure, there are more and more mental health problems, such as anxiety and depression, among college students. Although traditional psychological assessment and intervention methods, such as face-to-face psychological counseling and standardized psychometric tools, can solve these problems to some extent, they have limitations such as strong subjectivity, time-consuming and limited resources. According to the & quot; 2022 Chinese College Students mental health Survey report & quot; the mental health of key universities and undergraduate students is poor [1]. Most colleges and universities have the importance of the evaluation, teachers are weak, the evaluation system is not perfect. It is difficult to capture the students learning life process of dynamic psychological development change, which affect the students ' psychological status of master and judgment, delay counseling and treatment time will even hinder the students&#039 [2]. Therefore, there is an urgent need to find new and more efficient psychological assessment and intervention methods.

The rise of artificial intelligence provides the possibility for this demand. Artificial intelligence technologies, especially natural language processing (NLP), virtual reality (VR), and big data analysis, have shown great potential in the field of psychological assessment and intervention. These technologies can not only process multiple types of data on a large scale, but also realize the automation and intelligence of psychological interventions, so as to improve the efficiency and effectiveness of psychological services [3-6].

With the development of the intersection field of artificial intelligence and psychology, more and more scholars realize the important role of artificial intelligence in improving the accuracy and effectiveness of psychological evaluation interventions for college students. Related research there are many, but due to less available study sample data, mental illness and diagnostic parameters, clinical data application difficulties, most of the current model are data collection and integration difficulties, psychological assessment results are not comprehensive, psychological assessment accuracy is not high, still need to develop researchers more easily collect, comprehensive and accurate psychological monitoring intervention system to improve the effect of psychological monitoring intervention and the user&#039.

The purpose of this study is to explore the application of artificial intelligence in the field of psychological assessment intervention of college students and to develop a multimodal model with the characteristics of faster assessment speed and higher accuracy for its problems in the mental health assessment of college students [7]. Through this study, it is expected to provide new ideas and methods for college students 'mental health services, promote the deep integration of artificial intelligence and psychology, and promote the development of college students' mental health careers.

2 MULTIMODAL MODEL BASED ON FEATURE FUSION AND DECISION FUSION

Aiming at the above problems of mental health assessment of college students, this paper proposes a multimodal model based on feature fusion and decision fusion. Modal fusion refers to extracting and combining relevant information between multiple modes to obtain better performance than the unimodal scheme [8]. According to the different stages of modal fusion, modal fusion can be divided into three ways: input level fusion (Early Fusion), intermediate fusion (Intermediate Fusion), and decision level fusion (Late Fusion).

2.1 Characteristic layer fusion

2.1.1 Multi-space self-attention layer

Feature layer fusion is feature fusion after data extraction of different modes. Since feature extraction can be divided into multiple stages, intermediate-level fusion can be more flexible in the selection of fusion timing. This chapter uses the characteristics of layer fusion realized through direct splicing multimodal features, namely through the EEG signal, voice text, eye movement signal, and expression data feature extraction, combined with the cascade of attention and joint attention mechanism multimodal fusion algorithm to get the corresponding feature data splicing multimodal feature layer [9]. EEG signals, speech text, eye movement signals, and expression data were collected during the emotion-induced video. There is a certain process in the induction of emotions. Emotions are hidden in the eye movement features and EEG features, etc. Self-attention highlights the higher-order features that can best express emotions [10]. EEG and speech text is a signal of high time resolution, eye movement signal and expression data is a low time resolution, through high time resolution and low time resolution high order features of the joint attention, to low time resolution of high order feature dimension screening, in order to get stronger emotional representation of higher order features and make the dimension space low resolution signal features and high resolution features have higher complementarity.

How to effectively highlight the feature space related to emotion is a key issue for emotion recognition. A multi-space self-attention mechanism is proposed in a multimodal fusion network based on high and low resolution features, which can effectively highlight the high and low resolution features in the feature space that best expresses emotional states.

This mechanism first maps input features to different feature spaces through nonlinear mapping, which represent different feature types, and different types of features are different expressions of emotions. Feature extraction and transformation into the corresponding feature tensor in different feature spaces. Summarize the characteristic tensor under different space, and from the summary of the attention vector, the attention vector in different feature space, to highlight the strong correlation with emotions under different feature space, and get the corresponding feature space with the attention feature tensor, finally on the different feature space with the attention feature output. The specific implementation of the multi-space self-attention mechanism is shown in Figure 1:

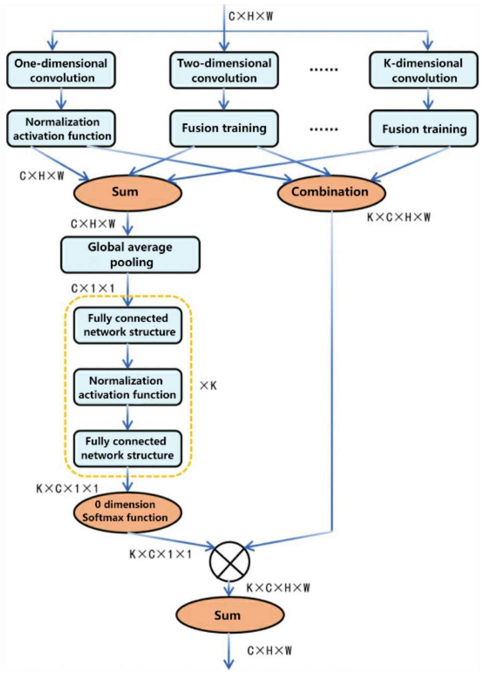

Fig. 1: Specific implementation of the multi-space self-attention mechanism

Taking the eye movement signal self-attention mechanism as an example, it is the input of the eye movement self-attention mechanism, namely the output of the eye movement feature extraction layer. The feature dimension of eye movement is regarded as the channel dimension, assuming that the number of channels of the input feature tensor is and the number of the mapped feature space is K, then the input feature is equally divided into K parts in the channel dimension, and the number of channels of each feature. Next, the convolution operation is performed in parallel to extract the high-order eye movement features on different feature space, and then the component output on the corresponding feature space is obtained. This output is summed according to the corresponding feature dimensions to obtain the eye movement summary features, and the global average pooling is made in the time dimension of the eye movement. The global tie pooling in the time dimension can obtain the time global information on each channel dimension (feature dimension). The temporal global information on the channel dimension is passed through two layers of full connection and the group channel attention vector after Softmax. This attention vector can highlight the important channels while inhibit the unimportant channels, give a high weight value to the important channels, and highlight the characteristics that are highly related to emotions. Each group in this set of attention vectors affects the importance of the output on the corresponding feature space in the channel dimension, and finally sums the weighted output on the group feature space. The whole process can be represented by the following formula:

$$Att_s = Softmax\left(FC\left(ReLU\left(BN\left(FC\left(Avg_{global}(x_{eye}^{sum})\right)\right)\right)\right)\right) \quad (1)$$

$$x'_{eye} = \sum_{i=1}^{K} Att_s^i \cdot x_{eye}^i \quad (2)$$

Where is the splicing of the group attention vector, representing the group attention vector, and representing the global average pooling in the temporal dimension of eye movement. Represents the output of the eye-movement self-attention mechanism.

Taking the self-attention layer of the eye-movement signal as an example, the input of the self-attention mechanism was adjusted with the output of the self-attention mechanism and summed with the features of the linear transformation to obtain the final output of the eye-movement self-attention layer. The process formula is as follows:

$$x_{eye}^{att} = ReLU\left(W_k(x_{eye}) + BN\left(W_v(x'_{eye})\right)\right) \quad (3)$$

Represents the final output of the eye movement self-attention layer, and are two linear transformation matrices.

2.1.2 Joint attentional layer

The higher order EEG features and eye movement features are integrated to obtain joint attention, and the features of low temporal resolution are screened to obtain the eye movement features with stronger emotional representation ability and make the eye movement features and the EEG features in this dimension space have higher complementarity. The framework structure of the joint attention mechanism based on EEG and eye movement is shown in Figure 2:

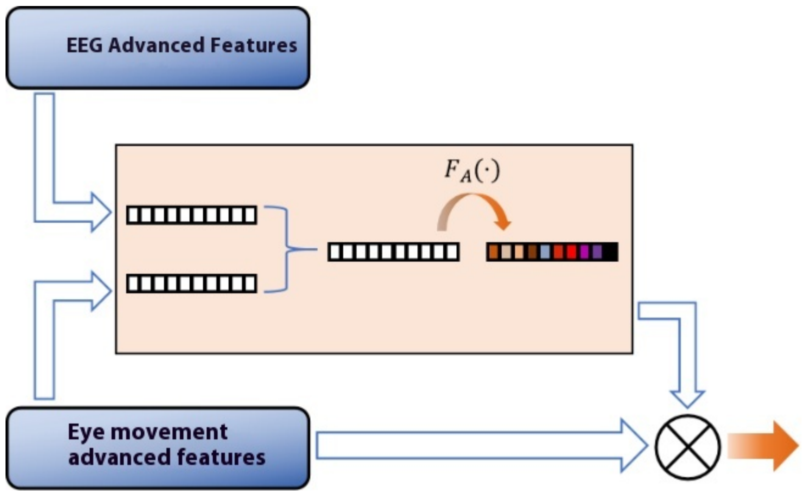

Fig. 2: Frame structure of the joint attention mechanism based on EEG and eye movement

Joint attention mechanism first through linear changes from the attention layer with attention and the EEG features map to the same dimension space joint feature tensor, from the joint features the joint attention vector, the joint attention vector can highlight the highly related to the emotions and can better fusion with EEG advanced features of the important features. Allocation of joint attention vectors to eye movement features enables further screening of eye movement features. Flowchart of the joint attention mechanism based on EEG and eye movement:

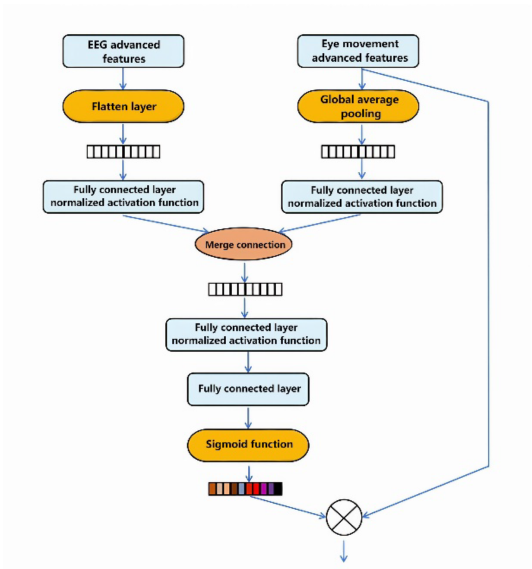

Fig. 3: Flowchart of the joint attention mechanism based on EEG and eye movement

Specifically, in the joint attention mechanism based on EEG and eye movement, and for the joint attention mechanism, that is, the output of the multi-space self-attention layer. And can be understood as eye movement advanced features and EEG advanced features. Considering that there are temporal dimensions on the eye movement features, the global average pooling of the advanced eye movement features in the temporal dimension is conducted to obtain the temporal global information on each channel dimension (feature dimension). Then leveling the EEG high-level features yields a global representation of the working state of different brain regions. The transformed features were individually linearly transformed and mapped to the same dimensional space to obtain the eye-movement attention tensor and the EEG attention tensor. These two kinds of attention tensors are stitched together and passed through the fully connected layer and sigmoid layer to obtain the joint attention vector. This joint attention vector can give different weight values to the eye movement high-level features of different channels, highlighting the important features that are highly correlated with emotion and can be better integrated with the EEG high-level features. Taking the eye-movement features and EEG features as an example, the joint attention vector product influences the importance of the eye-movement features in the channel dimension, obtaining higher-order eye-movement features based on the joint attention of EEG and eye movement. The process formula is as follows:

$$Att_{eye} = ReLU\left(BN\left(FC\left(GAP(x_{eye}^{att})\right)\right)\right) \tag{4}$$

$$Att_{eeg} = ReLU\left(BN\left(FC\left(Flatten(x_{eeg}^{att})\right)\right)\right) \tag{5}$$

$$Att_{joint} = ReLU\left(BN\left(FC([Att_{eye}, Att_{eeg}])\right)\right) \tag{6}$$

$$x_{eye}^{att'} = Att_{joint} \cdot x_{eye}^{att} \tag{7}$$

Based on the flexibility of feature-level fusion in the choice of fusion timing, this chapter proposes a multimodal fusion network based on self-attention and joint attention. Specifically, the multi-space self-attention mechanism in the network maps to different feature spaces to extract the features under different feature spaces and highlight the features that are strongly correlated with emotions. And the joint attention mechanism can, through high-time resolution of EEG features and low-time resolution features, joint attention, and through the attention vector of low-time resolution of characteristic dimension screening, reduce the modal redundancy between the degrees and extract the complementarity, which is stronger, to better integrate the dimension space on multimodal features and further improve the accuracy of emotion recognition.

2.2. Decision-making layer integration

Existing MKGR methods mainly project different modal representations into a unified space, using a multimodal unified representation interlay for prediction. However, in the prediction process, each mode may provide unique information, and the contribution to MKGR is also different. Decision fusion rationally assigns weights for each mode to achieve weighted fusion, capturing complementary information between different modes to generate more reliable prediction results.

Specifically, the entity embedding representation and the relational embedding representation of each modal are obtained through γiγthe encoder and input into the shared decoder to generate each modal prediction result. Then, the prediction results of each mode are weighted averaged to generate the final comprehensive prediction, and the mean square error loss function is used to measure the gap between the predicted and true values. The decision fusion calculation formula is as follows:

$$y = \frac{\gamma_s \gamma_s + \gamma_t \gamma_t + \gamma_v \gamma_v + \gamma_m \gamma_m}{\gamma_s + \gamma_t + \gamma_v + \gamma_m} \tag{8}$$

Where represents the weight parameter for each mode. Combining the alignment loss and the contrast loss in the previous section, the overall training goal of DFKR is to minimize the loss, and the calculation formula is as follows:

$$L_{total} = \gamma_s L_s + \gamma_t L_t + \gamma_v L_v + \gamma_m L_m + \lambda L_{cL} + L_{KA} \quad (9)$$

The mean square error loss function representing the mode is the weight parameter of the above prediction stage and is the weight hyperparameter learned by comparison.

3 CONCLUSIONS

The experimental results show that the model has superior performance in the psychological evaluation of college students, which can not only complete the evaluation task quickly and accurately but also provide more accurate data support for the subsequent mental health intervention. This not only provides new ideas and methods for the intelligent development of college students' mental health services, but also provides a useful reference for the application of artificial intelligence in other related fields.

In the future, we will continue to develop an AI-based psychological assessment system that can collect students 'mental health data through questionnaires and online chats and quickly and accurately assess students' mental health status. Analysis of potential mental health problems in students using NLP techniques. Long-term tracking of the factors affecting the mental health of college students, such as academic stress, interpersonal relationships, etc., to take measures in advance to reduce the negative impact. With the continuous progress of artificial intelligence technology and the increasing demand for mental health services, it is believed that the multi-modal model proposed in this study will play a more important role in the field of psychological assessment intervention for college students. At the same time, we also expect that more scholars and experts can join in the research in this field to jointly promote the deep integration and development of artificial intelligence and mental health services.

ACKNOWLEDGEMENTS

This work was supported by the Innovation Training Project of Guangdong Ocean University (Project No. CXXL2024113).

REFERENCES

[1] Fang, Y., Wang, L., & Chen, Z. (2023). A survey report on the mental health stat us of Chinese college students in 2022. *The Social Sciences Academic Press*, 70-99.

[2] Lin, L. (2021). On the dynamic evaluation system of university and college students' mental health based on big data technology. *Journal of Qiqihar University (Philosophy and Social Sciences Edition)*, (11), 181-185.

[3] Pang, H., Wang, C., Li, G., & Dong, Y. (2022). The construction of student mental health monitoring and evaluation system based on artificial intelligence. *Educational measurement and Evaluation*, (03), 31-39.

[4] Deng, X. (2021). Innovative application of artificial intelligence in self-diagnosis and treatment of psychological problems and health education in Colleges and universities. *Journal of Jilin Medical College*, *42*(06), 427-428.

[5] Zhao, T., Chen, G., Gatewongsa, T., & Busababodhin, P. (2025). Forecasting Agricultural Trade Based on TCN-LightGBM Models: A Data-Driven Decision. *Research on World Agricultural Economy*, 207-221.

[6] Jin, Z., Su, R., & Zhao, X. (2024). Psychological assessment method based on the heterograph network. *Journal of Jilin University (Engineering Edition)*, *54*(4), 1078-1085.

[7] MAO, J., Li, C., Zhu, H., & Lv, B. (2020). Design of students' psychological state early warning system based on data mining and intelligent algorithm. *Electronic Design Engineering*.

[8] Wang, Z., Zhang, Z., Su, T., Ding, Z., & Zhao, T. (2024, December). Research on Supply Chain Network Optimisation Based on the CNNs-BiLSTM Model. In *2024 International Conference on Information Technology, Comunication Ecosystem and Management (ITCEM)* (pp. 197-202). IEEE.

[9] Zhao, Z., Bao, Z., Zhang, Z., Deng, J., Cummins, N., Wang, H., ... & Schuller, B. (2019). Automatic assessment of depression from speech via a hierarchical attention transfer network and attention autoencoders. *IEEE Journal of Selected Topics in Signal Processing*, *14*(2), 423-434.

[10] Chen, G., Chutiman, N., Zhao, T., Panta, C., & Busababodhin, P. (2024). Identification of Earthquake Source Attributes Based on DAPSO-BP Combined Model. *Lobachevskii Journal of Mathematics*, *45*(12), 6271-6285.