Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 18 No. 20 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijim.v18i20.51887

PAPER

Deciphering Patterns in Student Emotional Fluctuations: A Big Data Approach in Educational Psychology

Yiyi Zhao(✉)

Department of Art and Design, Shijiazhuang College of Applied Technology, Shijiazhuang, China

fancy998866@163.com

ABSTRACT

With the rapid advancement of big data and information technology, the analysis of student emotional fluctuations has emerged as a new frontier in the field of educational psychology. This study aims to explore the patterns of student emotional fluctuations through big data analysis techniques and to use these patterns to predict students' psychological tendencies, thereby providing educators with real-time, accurate teaching references. The importance of student emotions in the learning process and the potential of big data technology in perceiving and analyzing student emotions are elucidated in the background section. The current state of study discusses the limitations of traditional methods in analyzing student emotions, such as small sample sizes, short data collection time spans, and the lack of timeliness and accuracy in analysis. A student emotional fluctuation identification model based on big data is first established, capable of integrating multi-source data and effectively capturing subtle changes in student emotions. Furthermore, a psychological tendency mixed-frequency prediction model is constructed, utilizing the mixed data sampling (MIDAS) mixed-frequency model, aimed at achieving accurate predictions of trends in student emotional fluctuations. The development and validation of these two models demonstrate the application value of big data analysis in the field of educational psychology, supporting personalized learning and promoting the effectiveness of student emotional management and teaching interventions.

KEYWORDS

educational psychology, big data analysis, student emotional fluctuations, emotion identification model, psychological tendency prediction, personalized teaching

1 INTRODUCTION

In the domain of educational psychology, fluctuations in student emotions are widely regarded as crucial factors affecting learning outcomes and academic achievement [1, 2]. Recent advancements in information technology have ushered

Zhao, Y. (2024). Deciphering Patterns in Student Emotional Fluctuations: A Big Data Approach in Educational Psychology. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(20), pp. 99–114. https://doi.org/10.3991/ijim.v18i20.51887

Article submitted 2024-06-26. Revision uploaded 2024-08-19. Final acceptance 2024-08-24.

in new possibilities for capturing and analyzing student emotions through big data analysis. Data from online learning platforms, social media, and school management systems offer educators unprecedented opportunities to monitor students' emotional states in real-time, understand their dynamic changes, and optimize teaching strategies accordingly [3–5]. Despite these advancements, identifying and analyzing patterns of student emotional fluctuations from the vast amounts of educational big data remains an academic challenge yet to be addressed [6, 7].

The current study has demonstrated a close connection between students' emotional fluctuations and various learning-related variables, such as cognitive load, learning motivation, and strategies [8]. Both positive and negative fluctuations in student emotions significantly impact their learning processes and outcomes [9–11]. Thus, a deeper investigation into these emotional fluctuation patterns is essential for designing effective teaching interventions and enhancing student learning experiences [12, 13]. However, traditional research methods often rely on self-report questionnaires or observations, which are not only time-consuming but also subject to subjectivity, making it difficult to capture the subtleties and true states of emotions.

Although some progress has been made in monitoring and analyzing student emotional fluctuations, significant gaps remain. Most existing studies are limited by small sample sizes and short data collection periods, hindering the generalizability of findings across broader student populations and educational contexts [14–17]. Furthermore, current methods face limitations in capturing the granularity and timeliness of emotional fluctuations, failing to meet the needs for real-time feedback and prediction [18, 19]. Hence, there is an urgent need for a more effective and efficient method for analyzing student emotional fluctuations.

The core content of this thesis focuses on two main aspects. The first is the construction of a student emotional fluctuation identification model using big data analysis techniques. This model, by analyzing emotional indicators from multi-source data, can accurately identify subtle changes in student emotions and reveal the inherent patterns of these fluctuations. Secondly, a psychological tendency mixed-frequency prediction model based on student emotions is established. Utilizing the mixed data sampling (MIDAS) model, this aims to predict future trends in student emotions, thus providing a scientific basis for educational interventions. Through the development and application of these models, this thesis aims to enhance the precision and practicality of study into student emotions within the field of educational psychology, offering robust data support for personalized teaching and emotional guidance with significant theoretical and practical value.

2 CONSTRUCTION OF A STUDENT EMOTIONAL FLUCTUATION IDENTIFICATION MODEL BASED ON BIG DATA ANALYSIS

Against the backdrop of increasing academic pressure, employment competition, and societal expectations faced by university students, psychological issues related to learning are particularly pronounced. Through the construction of a student emotional fluctuation identification model based on big data analysis, not only can the emotional states of students be captured and analyzed in real-time and in detail, but the complex relationship between emotional fluctuations and learning behaviors can also be revealed. Consequently, this provides universities with precise interventions for mental health, promotes personalized adjustments to student learning strategies, and optimizes the teaching environment. This has significant theoretical and practical value, aiding in enhancing students' learning efficacy and overall well-being. It also offers a scientific basis for decision-making for higher education administrators and psychological counseling professionals.

Given the complex spatiotemporal characteristics of emotional fluctuations among university students, this section introduces a multi-scale hemispheric asymmetry model based on convolutional neural networks (CNN), requiring input data that is high-dimensional and reflective of the characteristics of student emotional fluctuations. The primary form of input data is structured time series data, capturing the dynamic changes in students' emotional states through the recording and analysis of online learning behaviors, social media interactions, and physiological responses. After preprocessing, these data are transformed into numerical time series that the model can process, meeting the CNN's requirements for data. Figure 1 demonstrates the two-dimensional matrix representation of the model input data.

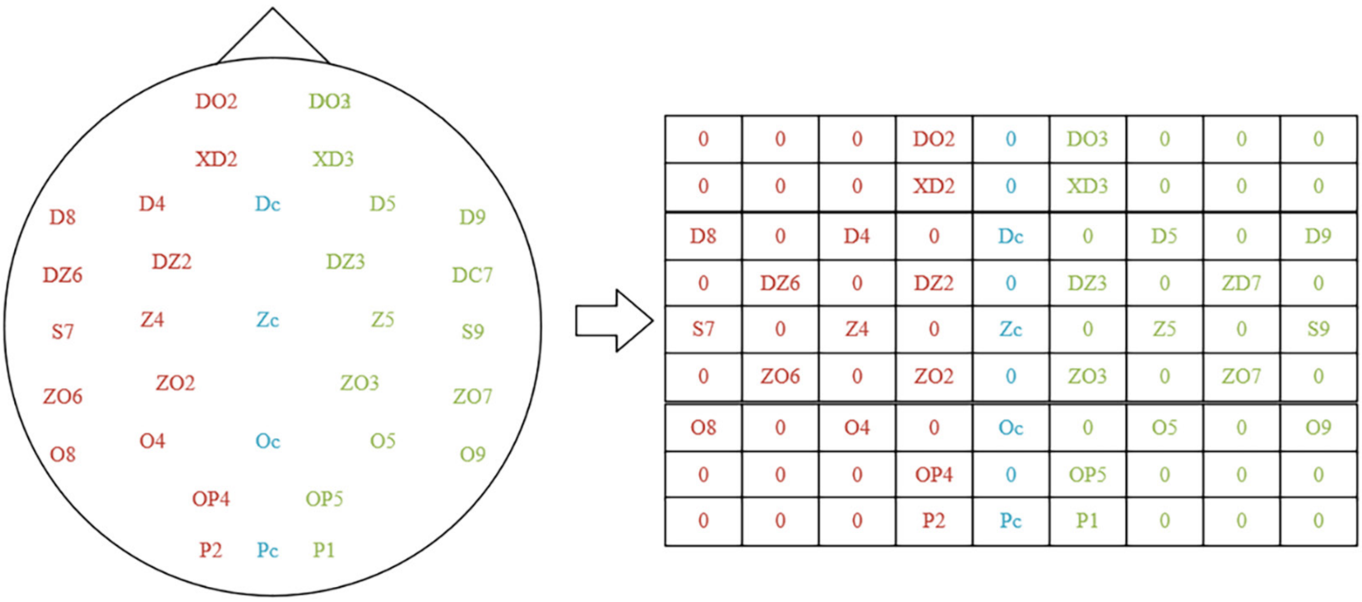


Fig. 1. Two-dimensional matrix representation of the model input data

The model constructed is capable of capturing the spatial distribution features of student emotional states through a spatial feature extraction block, such as the manifestation of emotions in different learning scenarios. Simultaneously, the temporal feature extraction module learns the characteristics of input signals across multiple time scales through various branches, reflecting students' short-term emotional fluctuations and long-term emotional trends, thereby capturing the temporal dynamics and periodicity of emotional changes. Such a design aims to conduct a comprehensive analysis of the spatiotemporal data of emotional fluctuations, taking into account the complexity, variability, and asymmetry of emotions, where the manifestations and impacts of positive and negative emotions may differ. The feature classification module is able to transform these complex features into information useful for educational interventions, providing a new perspective for educational psychology to deeply understand the patterns of student emotional fluctuations.

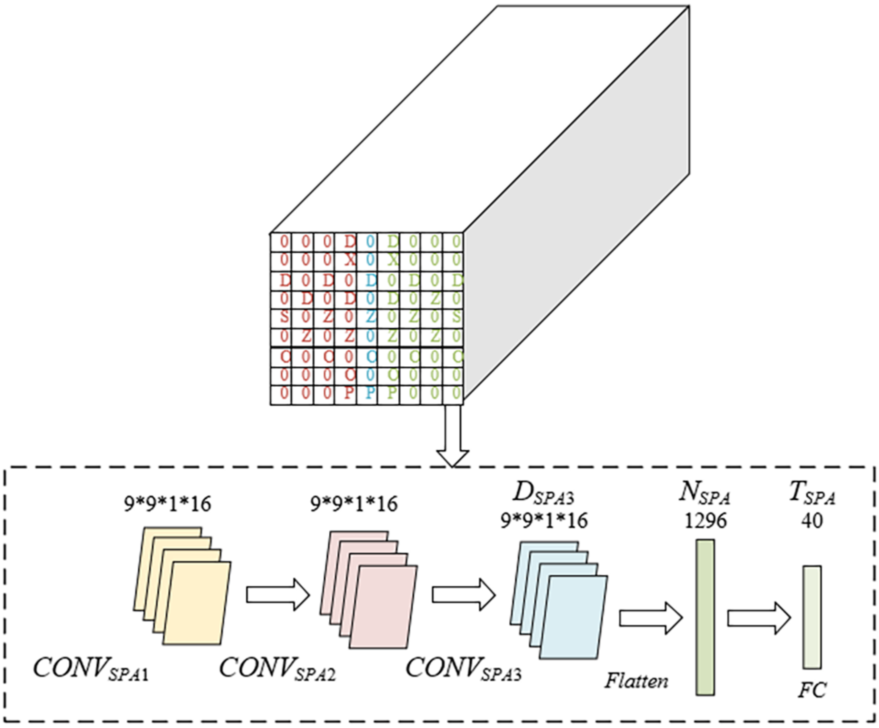


Fig. 2. Structure of the spatial feature extraction module

The spatial feature extraction block, a core component of the multi-scale hemispheric asymmetry model, is designed to capture the spatial characteristics of student emotions, i.e., the expressions of emotions in different learning environments. Figure 2 presents a schematic diagram of the spatial feature extraction module structure. This module consists of three convolutional operations followed by a subsequent fully connected layer. Convolutional operations extract useful feature mappings from raw data, which contain spatial information related to emotions, such as facial expressions and body movements. Convolutional layers, from shallow to deep, abstract these pieces of information layer by layer, capturing emotional expression features from the subtle to the holistic. The fully connected layer further maps these high-dimensional features to a lower-dimensional space, retaining key spatial information in the process and providing the model with a comprehensive representation of spatial features.

Assuming the convolution operation is denoted by $CONV_{SPA}$, the following is posited:

$$CONV_{SPA}(\cdot) = \delta(CONV_1(\cdot)) \tag{1}$$

Data features obtained after the convolution operation are denoted by D_{SP}. Then D_{SP} is unfolded into a vector and input into a fully connected layer, with the output normalized features represented by T_{SP}. Assuming the weight matrix is denoted by Q and the bias by n, the processing procedure is illustrated by the following expressions:

$$N_{SP} = FLATTEN(D_{SP}) \tag{2}$$

$$N_{dz} = BN(Q \cdot N_{SP} + y) \tag{3}$$

$$= \left[N_1, N_2, \ldots, N_{40} \right] \in E^{40} \tag{4}$$

$$\bar{N}_m = \frac{\exp(N_u)}{\sum_{j=1}^{40} \exp(N_u)}, u = 1, 2, \ldots, 40 \tag{5}$$

$$T_{SP} = \left[\bar{N}_1, \bar{N}_2, \ldots, \bar{N}_{40} \right] \in E^{40} \tag{6}$$

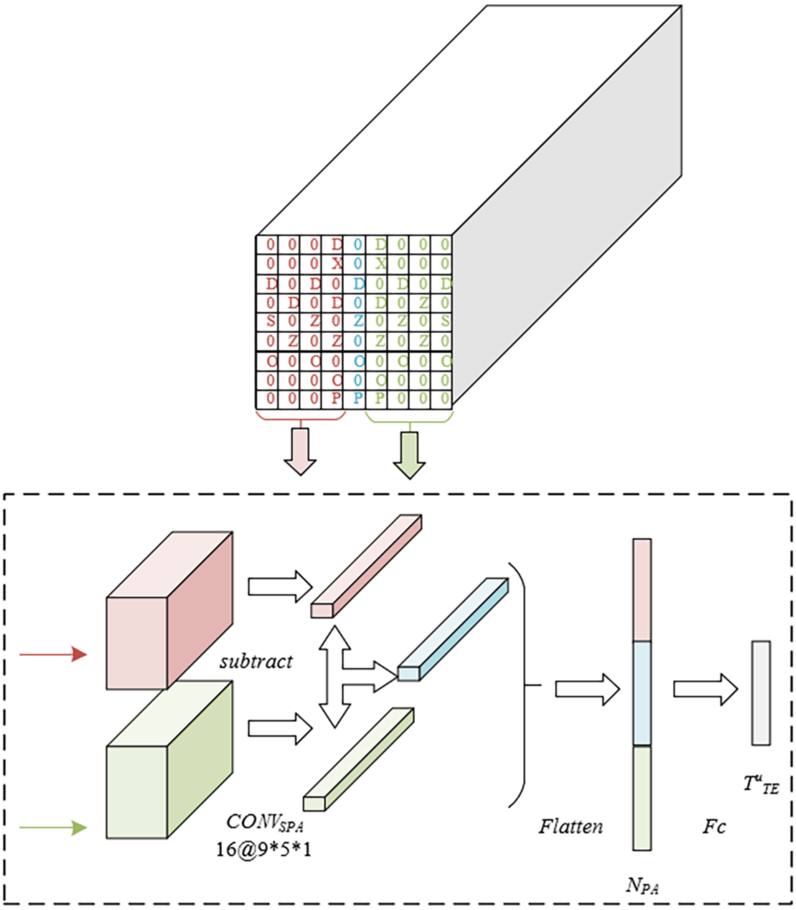


Fig. 3. Structure of the temporal feature extraction module

The temporal feature extraction module of the model is characterized by its multi-branch design, with each branch operating on input signals of different time scales. A schematic diagram of the temporal feature extraction module structure is provided in Figure 3. This module comprises *J* branches, each including a convolution operation $CONV^{j}_{TE}$, a flattening layer, and a feature normalization layer. Convolution kernels of varying sizes are employed across different branches to accommodate input signals of various temporal lengths, thereby capturing the emotional dynamics across different time scales. Branches tailored to short time scales are capable of capturing rapidly changing emotional details, such as instantaneous reactions in a classroom setting, whereas branches designed for longer time scales can detect overall trends and periodic fluctuations in student emotions over time. The flattening layer transforms multidimensional features into one-dimensional vectors, facilitating processing by the feature normalization layer and ensuring standardization of features for subsequent processing—a key to understanding the correlations and differences in student emotional fluctuations across different time scales. Specifically, data capturing the emotional states of students is divided into two parts in the spatial dimension. The first part, denoted by F_m, and the second part, denoted by F_e, represent the emotional state features of students in learning and non-learning scenarios, respectively. F_m and F_e are initially input into the same $CONV_{TE}^{j}$. Assuming the j-th branch is denoted by j and the SELU activation function by δ, the expression is as follows:

$$CONV_{TE}^{j}(\cdot) = \delta(CONV_{j}(\cdot)) \tag{7}$$

Two sets of features obtained using the $CONV_{TE}$ block are denoted by D_m and D_E, with differential features D_x calculated based on D_m and D_E as follows:

$$D_x = D_m - D_e \tag{8}$$

Subsequently, the three sets of features D_m, D_e, and D_x are unfolded into vectors N_m, N_e, and N_x and concatenated into vector N_{CAT}. To achieve a unified representation of multi-scale temporal features, N_{CAT} is input into the feature normalization layer, with the feature of the u-th time scale denoted by T_{TE}^{u}. The expression for the features of all J branches, i.e., the concatenation formula for asymmetric features across multiple time scales, is given as follows:

$$T_{TE} = \left[T_{TE}^{1} \,\|\, T_{TE}^{2} \,\| \cdots \|\, T_{TE}^{J} \right] \tag{9}$$

The design focus of the feature classification module is to combine spatial and temporal features extracted by the first two modules and to perform effective classification. In this module, T_{SP} and T_{TE} are concatenated into a larger feature vector and displayed through feature visualization to more intuitively understand the multidimensional features of student emotions. Then, the possibility of overfitting is reduced through a dropout layer, thereby improving the model's generalizability on unknown data. Finally, the feature vector processed by the dropout layer is input into a fully connected layer, which acts as a classifier, making determinations of emotional states based on the integrated feature vector, i.e., F belongs to the probabilities of various classes ($O(z \mid F)$, $z = 0,1$):

$$b_{PR} = \text{argmax}_{z}\, O(z \mid F), z = 0,1 \tag{10}$$

3 CONSTRUCTION OF A PSYCHOLOGICAL TENDENCY MIXED-FREQUENCY PREDICTION MODEL BASED ON STUDENT EMOTIONS

With the advancement of big data technology and the push towards educational informatization, universities are now able to access multi-source data related to student emotions, including but not limited to social media activities, records of online learning behaviors, and physiological indicators from smart wearable devices. The MIDAS model effectively integrates and analyzes data of different time frequencies, such as high-frequency data on daily emotional changes and low-frequency evaluation data such as semester grades. This approach allows researchers to capture subtle emotional fluctuations and their complex relationships with academic performance, more accurately reflecting the dynamic interaction between students' psychological states and learning outcomes. Therefore, the MIDAS model has been chosen for the construction of a psychological tendency prediction model based on student emotions, presenting clear scientific advantages. Assuming the dependent variable of relatively low frequency is represented by b_s, and the independent variable of relatively high frequency is denoted by $a_s^{(l)}$, with the frequency ratio of the dependent and independent variables represented by l, and the lag operator for high-frequency variables denoted by $M1/l$, satisfying $Mj/la_s^{(l)} = a_{s-j/l}^{(l)}$. Polynomial weight coefficients are represented by $Y(M^{1/l};\varphi)$. The basic MIDAS model expression is provided as follows:

$$b_s = \alpha_0 + \alpha_1 Y\left(M^{\frac{1}{l}};\varphi\right)a_s^{(l)} + \gamma_s, s = 1,2,\ldots S \tag{11}$$

In educational psychology research, a single variable often fails to fully explain the complexity of student emotional fluctuations, as students' emotions and psychological states are influenced by a myriad of factors, such as personal characteristics, educational environment, social interactions, and life events. Consequently, the proposed psychological tendency MIDAS prediction model based on student emotions is optimized and developed into a multivariate MIDAS model to enhance the model's explanatory power and prediction accuracy. By integrating multiple related dependent variables, such as academic performance, engagement, and psychological health indicators, the multivariate MIDAS model can more comprehensively capture the multiple factors affecting student emotional fluctuations and analyze how these factors jointly impact students' psychological tendencies, thereby providing a more precise basis for educational interventions. Assuming low-frequency variables of the same frequency as b_s are denoted by c_{us}, and high-frequency variables of the same frequency are represented by $a_{ks}^{(l)}$, with the weight function denoted by $Y(M^{1/l};\varphi k)$. Its expression is as follows:

$$b_s + \alpha_0 + \sum\nolimits_{u=1}^{o} \alpha_u c_{us} + \sum\nolimits_{k=1}^{w} Y\left(M^{\frac{1}{l}};\varphi_k\right)a_{ks}^{(l)} + \gamma_s \tag{12}$$

In constructing the psychological tendency MIDAS prediction model based on student emotions to overcome the imbalance in quality and structure of input big data signals and effectively capture the relative importance of different data sources and frequency data, this paper introduces a weight function. Considering high-frequency data such as daily emotional tracking may have more noise and volatility, while low-frequency data such as end-of-semester psychological assessments or academic grades are more stable but update slowly; the weight function provides a mechanism for adjusting the influence of these data in prediction. By assigning greater weight to high-quality signals, reducing the impact of noisy signals, and differentiating data based on timeliness and prediction objectives, an effective balance is achieved.

Initially, the step method is employed to set the weight function in this study. This approach allows for the treatment of data weights within different time periods as piecewise constants, permitting researchers to assign varying weights to data based on specific stages of university students' learning psychology. Given that the emotional and psychological states of university students may fluctuate significantly at different points in the semester, the step method simplifies the modeling of seasonal or cyclical changes, thereby enabling the model to better adapt to and reflect the specific patterns of changes in student psychological states over time. This method is particularly useful when time series data exhibit clear breakpoints or phase changes, such as the cyclical changes in student emotions under a semester system. Specifically, assuming the weight function is a piecewise function, the Step method predicts parameters φ, with $\varphi_1 > \varphi_2 > \varphi_3 > 0$ and $\varphi_1 + \varphi_2 + \varphi_3 = 1$, resulting in the function expression as follows:

$$z(j;\varphi)\begin{cases}\varphi_1 0 \le u < j_1 \\ \varphi_2 j_1 \le u < j_2 \\ \varphi j_2 \le u < J\end{cases} \tag{13}$$

Given that the psychological states of university students are not instantaneous reactions but the result of the accumulation of past experiences and behaviors, the Almon lag polynomial is further utilized to set the weight function in this paper. This method provides a means to smoothly handle the lag effects in time series data, effectively estimating the impact of different lag periods on the current emotional state. This is crucial for capturing and understanding the dynamic process of student emotional fluctuations, especially in long-term psychological tracking studies, revealing potential long-term psychological trends. Assuming the weight function coefficients conform to a polynomial equation, the expression is as follows:

$$z(j;\varphi) = \frac{\varphi_0 + \varphi_1 j + \varphi_2 j^2 + \ldots + \varphi_o j^o}{\sum_{j=1}^{J}(\varphi_0 + \varphi_1 j + \varphi_2 j^2 + \ldots + \varphi_o j^o)} \tag{14}$$

Compared to static models, dynamic models are capable of more accurately capturing and modeling the complex changes in student emotional states over time. This enables the earlier identification of subtle signals that may lead to psychological issues, thus providing strong support for the timely implementation of school psychological health intervention measures. Consequently, this study further proposes the construction of a P-order Almon-MIDAS psychological tendency prediction model for students. By integrating historical emotional fluctuation data with current psychological states and employing a dynamic lag structure, this model predicts potential changes in student psychological tendencies, thereby enhancing the early warning capability for psychological risk mutations.

The binary response variable, indicating whether a psychological tendency mutation occurs for the *u*-th student in year '*s*,' is denoted by $B_{u,s}$, where $B_{u,s} = 1$ signifies a psychological tendency mutation for the *u*-th student at the end of year *k*. The constant intercept is represented by α_0, the weight coefficients for explanatory variables by α_s, and the annual explanatory variable for the *u*-th student in year *s–j* by $A_{u,s-j}$. The weight coefficient polynomial for high-frequency variables with a lag order of *w* is denoted by $Y(M^{1/l};\varphi_w)$, where $c^{(l)}_{w,s}$ represents the high-frequency explanatory variable with a lag order of w and the random disturbance by γ_s. This can be expressed as follows:

$$B_{u,s} = \alpha_0 + \sum_{j=1}^{o} \alpha_j a_{u,s-j} + \sum_{w=1}^{W} Y\left(M^{\frac{1}{l}}; \varphi_w \right) c_{u,s}^{(l)} + \gamma_s \tag{15}$$

Specifically, this paper opts to construct the student psychological tendency prediction model using a 3rd-order Almon-MIDAS function, achieving the capture of the dynamism and continuity in emotional fluctuations while ensuring the model's flexibility and fit. A 3rd-order polynomial provides sufficient parameters to accommodate potential nonlinear relationships, assuming the Almon weight function is represented by $Y(M;\varphi)$, with the specific expression as follows:

$$B_{u,s} = \alpha_0 + \alpha_s a_{u,s-1} + \sum_{w=1}^{3} Y\left(M^{\frac{1}{12}}; \varphi_w \right) c_{w,s}^{(12)} + \gamma_s \tag{16}$$

By setting the piecewise function to take two values, 0 or 1, for $B_{u,k}$, the following is obtained:

$$B_{u,k} = \begin{cases} 0, B_{(u,k)} < 0.5 \\ 1, B_{(u,k)} \geq 0.5 \end{cases} \tag{17}$$

4 EXPERIMENTAL RESULTS AND ANALYSIS

Table 1 presents the accuracies of various models for recognizing emotional fluctuations among students on the training set, encompassing four dimensions of emotion: valence, arousal, dominance, and likability. It is observed from the table data that the model proposed in this study exhibits exceptional performance in the dimensions of valence, dominance, and likability, with accuracies of 98.87%, 98.58%, and 98.36%, respectively, and very small standard deviations (0.45 and 0.42). This indicates not only a high level of accuracy but also a very stable performance across different training sets. Although the accuracy in arousal is slightly lower (91.45%), its standard deviation remains at a low level (0.42%), indicating high stability in this dimension as well. Compared to other models, such as multimodal residual long short-term memory (MMResLSTM), which achieved the highest accuracy in arousal at 98.36%, the model proposed in this study still performs best in valence and dominance. Additionally, the performance of the proposed model in likability (98.36%) is comparable to that of the pyramid convolutional recurrent neural network (PCRNN), but with a smaller standard deviation, implying greater stability of the proposed model. PCRNN and MMResLSTM also show excellent performance in valence, arousal, and dominance, but they lack standard deviation data for at least one dimension, suggesting their stability is not as high as that of the proposed model. Other models such as residual network (ResNet), visual geometry group network (VGGNet), extreme version of inception (Xception), densely connected convolutional networks (DenseNet), EfficientNet, SqueezeNet, and capsule network generally exhibit lower accuracies in valence and arousal, and some models lack standard deviation data for certain dimensions, preventing a full assessment of their stability.

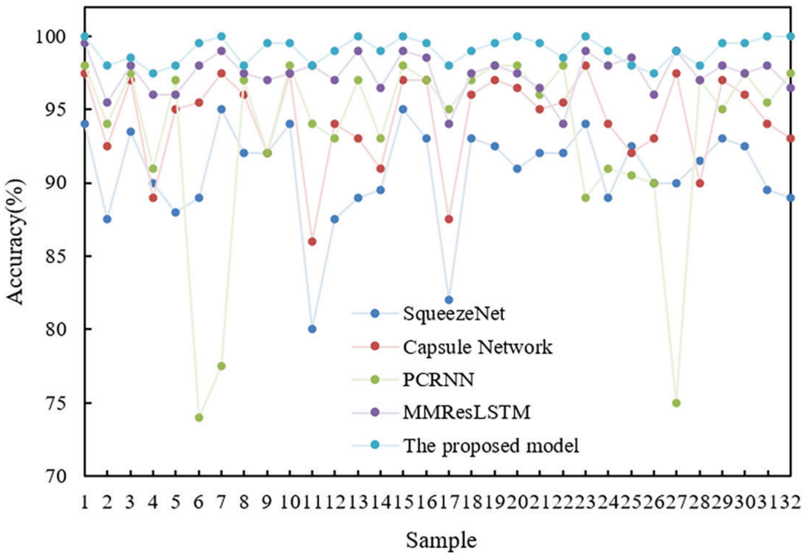
Table 1. Accuracies (mean ± standard deviation) of various models for recognizing emotional fluctuations among students on the training set

Method	Valence (%)	Arousal (%)	Dominance (%)	Likability (%)
ResNet	84.52 ± 3.45	83.26 ± 2.15	91.25 ± 4.23	/
VGGNet	85.62 ± /	83.69 ± /	/	/
Xception	91.25 ± 3.14	91.25 ± 2.56	/	/
DenseNet	91.25 ± 3.17	92.58 ± 2.89	/	/
EfficientNet	92.36 ± /	92.36 ± /	/	/
SqueezeNet	92.45 ± 1.45	91.25 ± 2.14	/	/
Capsule Network	93.68 ± 2.57	93.68 ± 2.75	/	/
PCRNN	93.68 ± /	94.58 ± /	93.69 ± /	94.25 ± /
MMResLSTM	95.36 ± 2.58	98.36 ± 2.14	/	/
The model proposed in this study	98.87 ± 0.45	91.45 ± 0.42	98.58 ± 0.42	98.36 ± 0.45

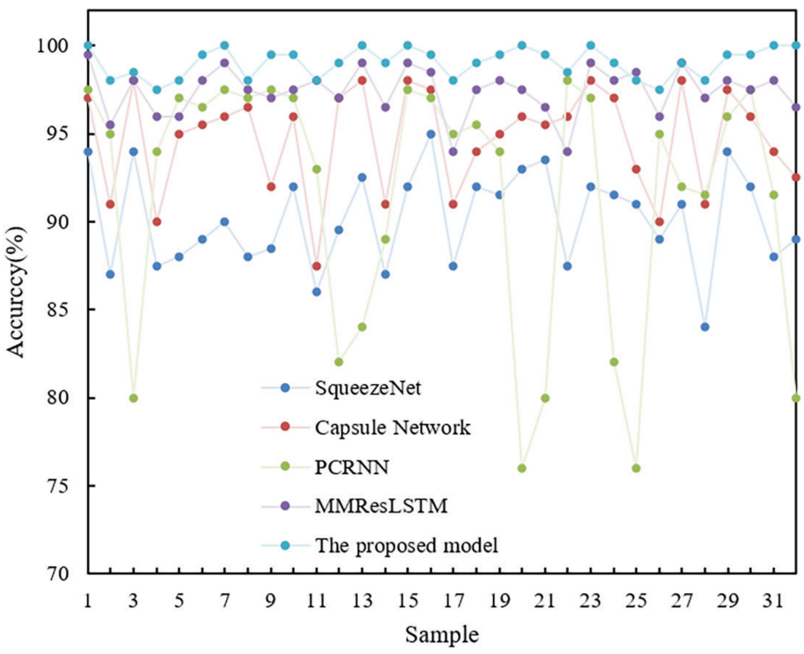
Table 2. Accuracies (mean ± standard deviation) of various models for recognizing emotional fluctuations among students on the test set

Method	Valence (%)	Arousal (%)	Dominance (%)
SqueezeNet	92.45 ± /	91.25 ± /	91.23 ± /
Capsule Network	94.25 ± 2.15	96.32 ± 2.75	/
PCRNN	98.26 ± 0.88	97.85 ± 1.32	99.23 ± /
MMResLSTM	97.24 ± 0.66	98.36 ± 0.98	97.85 ± 0.83
The model proposed in this study	98.99 ± 0.25	98.99 ± 0.21	98.58 ± 0.23

Table 2 provides accuracies of various models for recognizing emotional fluctuations among students on the test set, aiding in the assessment of the models' generalization capability on unknown data. It is observed that the model proposed in this study achieved the highest accuracies in the dimensions of valence, arousal, and dominance, with respective values of 98.99%, 98.99%, and 98.58%, and the lowest standard deviations (0.25, 0.21, and 0.23) among all models. This indicates that the proposed model not only exhibits exceptionally high accuracy on the test set but also demonstrates very stable performance, indicating strong generalization capabilities. Compared to PCRNN and MMResLSTM, two models with relatively good performance, although PCRNN reached an accuracy of 99.23% in dominance, slightly higher than the proposed model, its accuracy and stability in valence and arousal are lower than those of the proposed model. MMResLSTM shows a performance close to that of the proposed model in arousal (98.36%), but its accuracy and stability are still slightly lower in the other two dimensions. Other models, such as SqueezeNet and capsule networks, show significantly lower accuracies across all dimensions compared to the proposed model. Additionally, their standard deviations are either not reported or higher, suggesting these models' performance is less stable than that of the proposed model.



a) Accuracy for the valence dimension



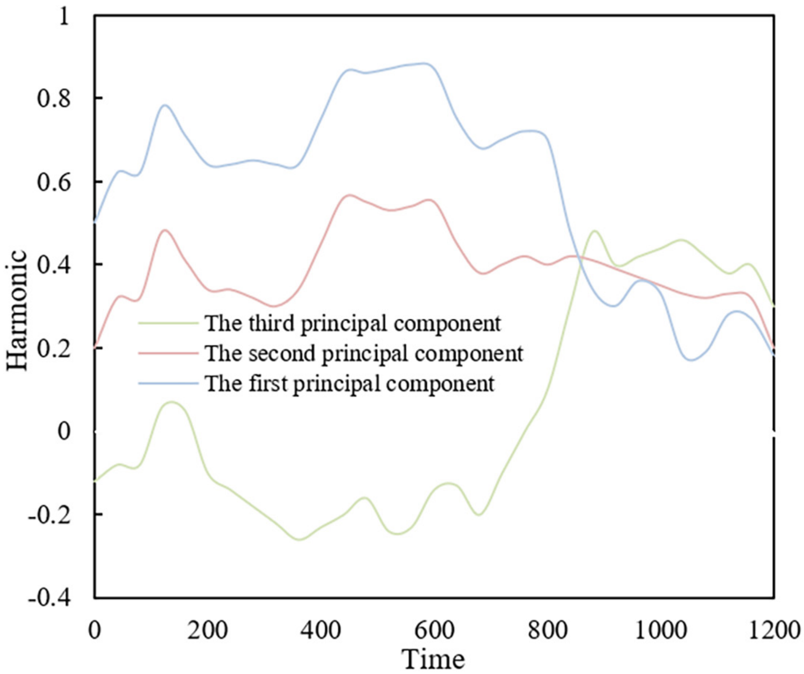
b) Accuracy for the arousal dimension

Fig. 4. Accuracy of the student emotional fluctuation identification model proposed in this study on the test set

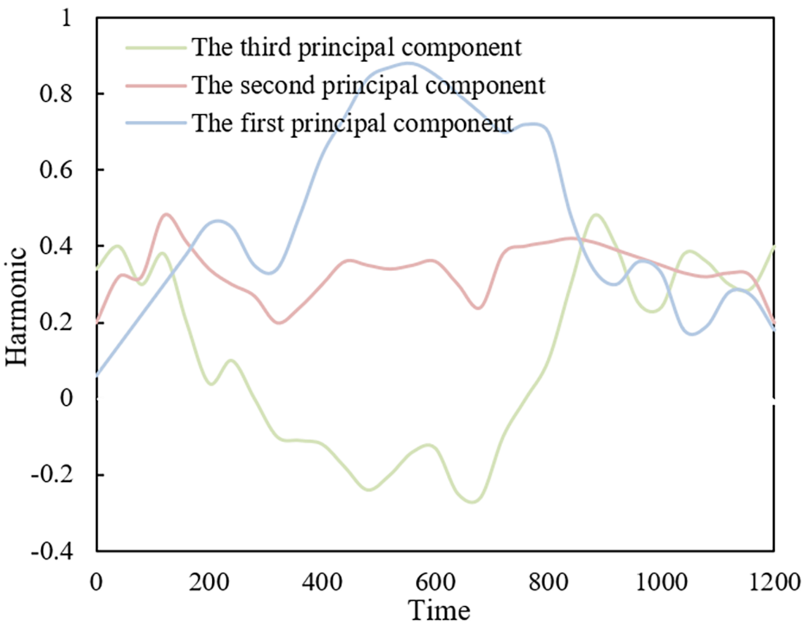
In summary, the performance of the proposed model on the test set highlights its superior generalization ability and stability. The high accuracies and low standard deviations together indicate that the proposed model can provide reliable and consistent emotional fluctuation recognition, both during the training phase and when facing unknown data in the test phase. Thus, these results further validate the effectiveness of the proposed model in recognizing emotional fluctuations, underscoring its potential value in practical applications.

Figure 4 illustrates the accuracies of the model proposed in this study for the dimensions of valence and arousal on the test set. It is revealed that the proposed model demonstrated exceedingly high accuracy in valence recognition, with all samples achieving accuracies above 98% and several samples reaching 100% identification accuracy. Compared to the other four models, the performance of the proposed model is observed to be superior, exhibiting not only the highest accuracy but also exceptional stability and reliability overall. Notably, the PCRNN model's performance on certain samples was significantly lower than on others, particularly with accuracies of only 74% and 75% for samples six and 27, indicating the PCRNN model's deficiency in recognizing specific types of emotional fluctuations under certain conditions. While the MMResLSTM model generally performed well, it still fell short of the proposed model. In the dimension of arousal, the proposed model similarly excelled, with all samples showing accuracies above 98% and several samples achieving 100% identification accuracy. Again, the proposed model demonstrated the highest accuracy and best stability compared to other models. The PCRNN model's performance was less satisfactory on certain samples, such as samples three, 20, 21, 24, and 32, with accuracies of 80%, 76%, 80%, 82%, and 80%, respectively, further indicating PCRNN's limitations in processing arousal information.

Based on the data from the perturbation graph of principal components of the psychological tendency function after incorporating student emotions, as provided in Figure 5, the trends of each principal component over time can be observed. The following analysis examines the trends of each principal component and discusses their significance for the effectiveness of the psychological tendency model. The first PCA corresponds to the frequency of emotional fluctuations. In both datasets, the first principal component increases over time, albeit at different rates and points in time. In the first dataset (with a variability percentage of 57.1%), this component begins to decrease after the 1000 time point, while in the second dataset (with a variability percentage of 33.8%), it maintains an upward trend after initially rising. This indicates that emotional fluctuations are not unidirectional over time and can change due to external events or internal psychological dynamics. The second PCA corresponds to the ratio of positive to negative emotions. In both datasets, the trend of the second principal component is relatively stable, but there is more fluctuation in the first dataset before the 1200 time point. This suggests that the ratio of positive to negative emotions undergoes significant changes initially but stabilizes over time. The stabilization of this trend is related to students adapting to their environment or learning how to manage their emotions better. The third PCA corresponds to heart rate variability. In the first dataset, the third principal component shows a clear trend of rising and falling over time, whereas in the second dataset, the fluctuations appear more subdued. These changes in heart rate variability reflect fluctuations in students' stress levels and their physiological responses to stress.



a) Variability percentage of 57.1%



b) Variability percentage of 33.8%

Fig. 5. Perturbation graph of principal components of the psychological tendency function after incorporating student emotions

It can be concluded that both datasets reveal distinct performances of the three main components at different time points, emphasizing the dynamic nature of student emotions and the complexity of psychological tendencies. The trend changes in the first principal component suggest that emotional fluctuations are not linear over time, and the model's ability to capture these nonlinear changes is crucial for understanding emotional dynamics. The fluctuations in the second principal component indicate that the model can reflect the complexity of emotions, which is important in predicting psychological tendencies. The changes in the third principal component highlight those physiological states are an important indicator of emotional states, and the inclusion of this physiological parameter in the model aids in a more comprehensive understanding of students' psychological tendencies.

In summary, the mixed-frequency prediction model based on student emotions captures key changes in emotional and physiological indicators, crucial for predicting students' psychological tendencies. The model's effectiveness is demonstrated by its ability to identify dynamic changes in key factors such as the frequency of emotional fluctuations, the ratio of positive to negative emotions, and heart rate variability, providing a complex yet detailed perspective for predicting psychological tendencies.

5 CONCLUSION

The construction of a student emotional fluctuation identification model has been completed in this study, which involves analyzing emotional indicators from multi-source data to identify subtle changes in student emotions and reveal their inherent patterns. Additionally, the construction of a psychological tendency mixed-frequency prediction model has been realized, employing the MIDAS model to forecast the future trends of student emotions, thereby providing a scientific basis for educational interventions.

Experiments have demonstrated the identification accuracies of different models on training and testing datasets. In the dimensions of valence (the positivity or negativity of emotions) and arousal (the intensity of emotions), the model proposed in this study has shown higher accuracies on the testing set. The prediction results of the model have been provided, showcasing the model's capability to forecast future emotional trends. Principal component graphs of the psychological tendency function after incorporating student emotions have been drawn. Perturbation graphs of the functional principal components have been presented, analyzing the impact of emotions on psychological tendencies.

It can be concluded that the model constructed in the study effectively identifies and predicts student emotional fluctuations, holding significant value for understanding students' emotional states and psychological tendencies. The experimental results indicate the complexity of student emotional fluctuations, with the model capable of capturing the multidimensional changes in emotions. The application of big data analysis techniques has proven their potential and value in the fields of emotion recognition and psychological prediction. Future research may continue to enhance model precision, expand data sources, and explore the relationships between emotions and other behaviors or academic performances.

In summary, this study provides a comprehensive approach for identifying and predicting student emotional fluctuations, offering not only theoretical insights for educators and psychologists but also potential significant impacts in practical applications. Through a deep understanding of the subtle changes in student emotions, better support can be provided for students' learning and psychological health while also offering a basis for educational interventions, promoting personalized teaching, and facilitating student development.

6 REFERENCES

[1] W. Liu and S. Bi, "Influence mechanism of students' learning enthusiasm based on educational big data," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 6, pp. 96–109, 2023. https://doi.org/10.3991/ijet.v18i06.38717

[2] K. Mou and B. Xu, "Research on emotional analysis platform of college students' public opinion information based on big data," in *Proc. 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC)*, 2022, pp. 2105–2110. https://doi.org/10.1109/ITOEC53115.2022.9734607

[3] Y. Liu and W. Qi, "Application of flipped classroom in the era of big data: What factors influence the effect of teacher-student interaction in oral English teaching," *Wireless Communications and Mobile Computing*, vol. 2021, no. 1, pp. 1–7, 2021. https://doi.org/10.1155/2021/4966974

[4] P. Qin, "A study on the effect of music on college students' depression and anxiety based on big data analysis," *Applied Mathematics and Nonlinear Sciences*, vol. 9, no. 1, pp. 1–16, 2024. https://doi.org/10.2478/amns.2023.2.00183

[5] F. Liu and D. Liu, "Mining the mental health crisis among college students using big data: A data analysis based on TikTok," in *Proc. 2022 3rd International Conference on Education, Knowledge, and Information Management (ICEKIM 2022)*, 2022, pp. 859–862. https://doi.org/10.1109/ICEKIM55072.2022.00188

[6] J. C. Yang and X. M. Du, "Strategies for monitoring and managing online public opinion in universities under the background of big data," *Journal of Computers*, vol. 34, no. 6, pp. 163–171, 2023. https://doi.org/10.53106/199115992023123406013

[7] Y. Tan, "How to effectively infiltrate emotional education in primary school Chinese teaching from perspective of big data," in *Proc. 2020 International Conference on Computers, Information Processing and Advanced Education (CIPAE 2020),* 2020, pp. 138–141. https://doi.org/10.1109/CIPAE51077.2020.00044

[8] Z. Yin and Q. Lin, "Analysis based on the data of 2660 college students during COVID-19 epidemic: Emotional state and influence factors," in *Cyber Security Intelligence and Analytics: International Conference on Cyber Security Intelligence and Analytics (CSIA2021)*, 2021, pp. 31–38. https://doi.org/10.1007/978-3-030-70042-3_5

[9] J. Hu, Y. Liu, Q. Zheng, and F. Liu, "Emotional mechanisms in supervisor-student relationship: Evidence from machine learning and investigation," *Journal of Social Computing*, vol. 4, no. 1, pp. 30–45, 2023. https://doi.org/10.23919/JSC.2023.0005

[10] D. Ore-Betancur, S. Alarcon-Luna, and M. Gonzales-Macavilca, "Exploration of emotional characteristics in 2nd grade elementary students from the district of Independencia, Lima (Peru)," in *Proc. 2023 IEEE 3rd International Conference on Advanced Learning Technologies on Education & Research (ICALTER)*, 2023, pp. 1–4. https://doi.org/10.1109/ICALTER61411.2023.10372882

[11] P. J. M. Valdez, Z. M. Padilla, P. J. Sison, and L. R. De Luna, "Emotional intelligence, self-efficacy and resilience among Filipino nursing university students in the blended learning setup: A mediation analysis using PLS-SEM," in *Proc. 2023 14th International Conference on E-Education, E-Business, E-Management and E-Learning*, 2023, pp. 25–31. https://doi.org/10.1145/3588243.3588263

[12] L. Wang, Z. Ye, and S. Zhu, "Research on college students' emotional experience in online learning," *International Journal of Continuing Engineering Education and Life Long Learning*, vol. 33, no. 6, pp. 523–535, 2023. https://doi.org/10.1504/IJCEELL.2023.134350

[13] F. A. Hernandez, M. Navarro-Estrella, J. C. Rojas, and J. Alvarez, "Emotional tool design: A perspective on how to generate empathy in higher education students," in *2023 IEEE Global Engineering Education Conference (EDUCON)*, 2023, pp. 1–6. https://doi.org/10.1109/EDUCON54358.2023.10125212

[14] E. Treadway, K. Tubbs, M. J. Caserto, M. Lee, and J. E. Swenson, "Assessment of a survey instrument for measuring affective pathways," in *ASEE Annual Conference Proceedings*, 2023, pp. 1–16.

[15] Z. Zhu, S. Qin, L. Yang, A. Dodd, and M. Conti, "Emotion regulation tool design principles for arts and design university students," *Procedia CIRP*, vol. 119, pp. 115–120, 2023. https://doi.org/10.1016/j.procir.2023.03.085

[16] Y. Wu, K. Li, Y. Qiu, and Y. Zhou, "A study on the relationship between online learning and emotional quotient of college students during the period of epidemic prevention and control," in *2021 International Conference on Data Analytics for Business and Industry (ICDABI)*, 2021, pp. 478–485. https://doi.org/10.1109/ICDABI53623.2021.9655787

[17] T. Zhang and S. Yu, "Effects of asynchronous interaction on positive emotional experiences of learners during online learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 18, no. 8, pp. 48–60, 2023. https://doi.org/10.3991/ijet.v18i08.38703

[18] A. Sabaritha, P. Sinduja, and R. Priyadharshini, "Knowledge awareness, and practice on emotional intelligence and academic performance among lecturers," in *2023 International Conference on Business Analytics for Technology and Security (ICBATS)*, 2023, pp. 1–7. https://doi.org/10.1109/ICBATS57792.2023.10111176

[19] Z. Gong, "Psychological resilience characteristics and improvement strategies of 'Internet Plus' engineering normal college students," in *Journal of Physics: Conference Series*, 2020, vol. 1533, pp. 1–8. https://doi.org/10.1088/1742-6596/1533/4/042084

7 AUTHOR

Yiyi Zhao, Graduated from North China Electric Power University in 2013, works at Shijiazhuang College of Applied Technology and is engaged in the research of education and art design (E-mail: fancy998866@163.com; ORCID: https://orcid.org/0009-0005-3170-9682).