



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642

<https://www.inderscience.com/ijict>

Student psychological stress classification model based on sensor data fusion and neural network

Jun Qian, Yuanyuan Ding, Yi Qin

DOI: [10.1504/IJICT.2026.10077190](https://doi.org/10.1504/IJICT.2026.10077190)

Article History:

Received:	10 October 2025
Last revised:	25 November 2025
Accepted:	05 December 2025
Published online:	01 April 2026

Student psychological stress classification model based on sensor data fusion and neural network

Jun Qian*

Teaching Quality Monitoring and Evaluation Center
(Teacher Development Center),

Hangzhou City University,

Hangzhou, 310015, China

Email: qianjun.qj@outlook.com

*Corresponding author

Yuanyuan Ding

School of Education Sciences,

Anhui Normal University,

Wuhu, 241000, China

Email: dingyuayua@163.com

Yi Qin

Hefei Municipal Commission of

Discipline Inspection and Supervision,

Hefei, 231600, China

Email: alisaqin1993@163.com

Abstract: With the development of artificial intelligence and sensor fusion, student psychological stress classification is shifting from subjective assessment to intelligent recognition based on multi-source physiological signals. However, existing methods often rely on single-channel signals and insufficient temporal and multimodal modelling. This study proposes a stress classification method based on sensor data fusion and neural networks. Multi-source EDA, PPG, and body temperature signals are synchronously collected, denoised using biased Kalman filtering, and jointly modelled by 1D-CNN and BiLSTM to capture local and global temporal features. Experimental results show that CNN and BiLSTM achieve accuracies of 78.32% and 82.15%, respectively, while the fused model reaches 89.47% with an F1-score of 0.88. Multimodal fusion improves the F1-score to 0.90, exceeding single-modal EDA (0.70) and PPG (0.71). Kalman filtering reduces RMSE by 42.5% and increases entropy by ≥ 2.0 bits, enhancing classification performance and providing technical support for psychological stress management and intervention.

Keywords: sensor data fusion; neural network; classification of student psychological stress; biased Kalman filter; BiLSTM.

Reference to this paper should be made as follows: Qian, J., Ding, Y. and Qin, Y. (2026) ‘Student psychological stress classification model based on sensor data fusion and neural network’, *Int. J. Information and Communication Technology*, Vol. 27, No. 30, pp.24–48.

Biographical notes: Jun Qian obtained his Master’s in Mental Health Education from Anhui Normal University in 2017. He is been responsible for student mental health affairs for six years and has published three research papers, including ‘Research on the integration of ideological and political education with mental health education for college students’. As a primary contributor (major author), he completed the research report for the Zhejiang Women’s Federation project. He has also provided 288 hours of part-time counselling. Additionally, he has taught nearly 200 class hours of the course ‘Mental Health and Development for College Students’.

Yuanyuan Ding is an Associate Professor and Master’s Supervisor at the School of Education Science, Anhui Normal University, specialises in school mental health education. With three decades of dedicated teaching experience, she has developed courses such as ‘School mental health education’ and published several papers, including ‘Research on school mental health education under the concept of positive psychology’.

Yi Qin obtained her Master’s in Mental Health Education from Anhui Normal University in 2017. She currently serves as a Third-Level Supervisor at the Hefei Municipal Commission for Discipline Inspection and Supervision. She previously worked as a psychological counsellor for primary and secondary schools, possessing years of practical experience in adolescent mental health education and psychological counselling. Her research focuses on the application of psychology in practice.

1 Introduction

In the increasingly fierce academic competition, student groups are facing psychological pressure from various aspects such as academics, interpersonal relationships, and family (Jiménez-Mijangos et al., 2023). Therefore, building a scientific and effective mechanism for identifying and classifying psychological stress is of great practical significance for timely intervention and guidance (Razavi et al., 2024). With the growth of artificial intelligence (AI), sensor technology, etc. data-driven methods based on computers have gradually become the mainstream direction for psychological stress detection and classification (Banerjee et al., 2023). By collecting physiological data including heart rate variability, skin conductance response, body temperature, and respiratory rate, and combining them with machine learning (ML) or deep learning (DL) algorithms for modelling, it is expected to achieve real-time recognition and dynamic classification of psychological stress states (Badr et al., 2024). However, existing methods for psychological stress classification (PSC) still have some limitations. Firstly, most methods rely on a single sensor or limited data dimensions, which cannot fully capture multi-source information features in complex psychological states, leading to unstable classification accuracy (Haque et al., 2024). Secondly, conventional models often rely on data quality and feature engineering, lack robustness, and are susceptible to environmental interference and individual differences (Chen et al., 2023). In addition,

some methods ignore the temporal correlation and semantic fusion relationship between different physiological indicators, making it hard to effectively model the dynamic changes in psychological stress (Narzary et al., 2025; Dogan and Akbulut, 2023). Therefore, this study proposes a student PSC method based on sensor data fusion (SDF) and neural network (NN). Firstly, multiple types of sensors are utilised to real-time collect multidimensional physiological signals (PSs), including electrodermal activity (EDA), photoplethysmography (PPG), and body temperature. Subsequently, a deep pressure classification model is constructed by fusing and denoising multi-source data using the biased Kalman filter (BKF) algorithm. This model combines one-dimensional convolutional neural network (1D-CNN) and bidirectional long short-term memory network (BiLSTM) to efficiently and accurately identify students' psychological stress states. The study aims to efficiently integrate multiple sources of physiological and behavioural data, combined with the powerful feature expression ability of DL, to achieve more accurate student PSC and provide solid technical support for mental health management and intervention.

2 Related works

The growth of computer technology, especially the application of AI, big data, and sensor technology, has provided new technological means for monitoring and classifying students' psychological stress (Mentis et al., 2024). By real-time collection and intelligent analysis of students' PSs, behavioural data, and emotional changes, computer-aided psychological stress detection methods can achieve more objective, accurate, and dynamic stress recognition (Masri et al., 2023). Therefore, more scholars have begun to pay attention to students' psychological pressure. Zhu et al. designed a stress state recognition method built on wrist EDA signals in wearable devices to address the issue of inevitable stress in modern life and its significant impact on physical and mental health. This method could effectively predict and monitor individual stress states, thereby improving the timeliness and scientificity of mental health interventions (Zhu et al., 2023). Women have significant differences in stress response at different stages of their menstrual cycle, and gender factors may affect the accuracy of ML models for stress detection and classification. Given this, Hemakom et al. proposed a gender discrimination and multi-level pressure classification model based on electrocardiogram and electroencephalogram signals. This model could achieve more accurate identification and classification of individual stress states, improving the adaptability and scientificity of the model in personalised and clinical applications (Hemakom et al., 2023). In current social media, people express emotions through text, but existing stress and depression detection methods have shortcomings in feature extraction and model reliability. Ilias et al. proposed introducing additional language information and label smoothing mechanism into the transformer model. This mechanism could achieve more efficient and reliable identification and prediction of stress and depression in social media, improving model performance and confidence calibration ability (Ilias et al., 2023). Talaat and El-Balka proposed a new stress monitoring algorithm for the application of the Internet of Things in medical health monitoring. This algorithm utilised wearable sensors for health status monitoring and improved monitoring accuracy through signal processing and ML algorithms (Talaat and El-Balka, 2023).

In addition, Xu et al. proposed an electronic skin system that can be used for non-invasive, multimodal monitoring of physiological and molecular indicators to address the limitations of traditional stress assessment, which mainly relies on subjective questionnaires and is difficult to achieve continuity and objectivity. This system could achieve high-precision quantification and long-term tracking of psychological stress responses, improving the scientific and practical nature of individualised stress management (Xu et al., 2024). Salankar and Qaisar proposed an automated stress classification method based on multi-channel EEG signal decomposition and shape feature analysis to address insufficient accuracy in processing EEG signals using traditional psychological stress recognition methods. This method used empirical and variational mode decomposition to process EEG signals, achieving high-precision recognition of stress and non-stress states, and improving the intelligence and reliability of psychological state detection (Salankar and Qaisar, 2023). Yang et al. proposed an automated chronic stress recognition system based on Twitter social media text to address the limitations of traditional stress measurement methods including questionnaire surveys in identifying individuals' self-reported chronic stress. This system achieved efficient detection and large-scale monitoring of users' self-reported chronic stress status, improving the timeliness and coverage of chronic stress intervention (Yang et al., 2023). Singh et al. proposed a hybrid ensemble method built on DL to address the limitations of existing mental health monitoring systems in stress recognition, such as strong group targeting and reliance on manual feature extraction. This method achieved automatic feature extraction and accurate classification of various stress states from physiological sensor data collected from intelligent wearable devices, improving the universality and automation level of mental health systems (Singh et al., 2023).

In summary, significant progress has been made in utilising advanced computer technology for multimodal data collection and intelligent analysis of students' psychological stress. These studies cover multiple aspects, from PS monitoring of wearable devices, multi-source data fusion of EEG and ECG, to social media text emotion recognition, greatly improving the precision and real-time performance of stress detection. However, there are still shortcomings in the generalisation ability, cross-group adaptability, and deep mining of multimodal data fusion in current research, especially in the construction of a comprehensive stress recognition framework for student groups, which lacks systematicity and robustness. Therefore, the paper proposes a student PSC method built on SDF and NN. Its innovation lies in the fusion of multi-source wearable PSs, achieving high-precision data synchronisation and denoising processing through BKF, and constructing a deep NN structure composed of 1D-CNN and BiLSTM. This study achieves a closed-loop output from stress recognition to personalised intervention recommendations, enhancing the practical effectiveness and intelligent decision-making ability of the model.

3 Methods and materials

This study proposes a student PSC method that integrates multi-sensor PSs with NN to address the problems of strong subjectivity, weak ability to process multi-source data, and poor model generalisation in traditional methods for identifying and classifying student psychological stress. This method is divided into two stages, the first of which is

to construct a psychological assessment model based on SDF. This model collects multidimensional PSs such as EDA, PPG, and body temperature through multiple types of wearable sensors, and uses BKF to achieve temporal synchronisation and denoising fusion of multi-source data, constructing a dynamic perception and quantitative evaluation mechanism. Secondly, based on the fusion of data and NN classification model, a student PSC model based on NN is proposed. It uses 1D-CNN to extract local high-frequency variation features, combines BiLSTM to model long-term dependencies, and achieves feature fusion and fine classification through deep representation learning. This study achieves precise identification of students' psychological stress status, providing technical support for personalised intervention and intelligent mental health management.

3.1 Student psychological stress assessment model based on SDF

The existing psychological stress assessment methods have problems such as relying on subjective questionnaires for data collection, poor adaptability, and insufficient ability to process multimodal PSs, making it difficult to achieve real-time, objective, and high-precision evaluation of students' psychological states. This study proposes a student psychological stress assessment model based on SDF, which utilises wearable sensors to non invasively and continuously monitor multidimensional physiological parameters such as EDA, PPG, and body temperature. Combined with the BKF algorithm, it performs time alignment and noise suppression on heterogeneous data, achieving personalised stress assessment and dynamic state recognition for students. Firstly, multi-source PSs are collected from students, and the raw data are synchronously uploaded to edge processing nodes or servers through various types of sensor devices worn on the wrist and fingers, ensuring temporal consistency and data integrity. The multi-source PS acquisition and synchronisation mechanism is shown in Figure 1.

The system simultaneously collects three types of physiological signals: EDA (4 Hz), PPG (64 Hz), and skin temperature (4 Hz). EDA and temperature data are obtained from a wristband sensor (such as the Empatica E4), while PPG and respiratory signals are recorded by a chest strap device (such as the RespiBAN Professional, 700 Hz). To address the issues of inconsistent sampling frequencies and timestamp drift from multiple data sources, the system employs strategies such as uniform resampling (to 64 Hz), timestamp calibration, transmission delay compensation, and sliding window synchronisation to achieve millisecond-level time alignment of cross-modal signals.

In Figure 1, the system integrates EDA, PPG, and temperature sensors through wearable devices to collect multidimensional physiological data in real-time. After each sensor independently collects data, it is transmitted to wearable devices through Bluetooth/Wi-Fi, and a timestamp synchronisation mechanism is used to ensure the temporal consistency of multi-source signals. Edge nodes (such as smartphones) perform pre-processing such as signal denoising, time alignment, and error correction, and finally upload synchronised data to the server for stress state analysis. During the acquisition of multi-source physiological data, different sensors exhibit variations in sampling frequency, transmission delay, and timestamp accuracy. To ensure the temporal consistency of EDA, PPG, and body temperature signals, this study introduces an engineered synchronisation mechanism based on dynamic time warping, incorporating a unified sampling frequency and delay compensation. All physiological signals are aligned to a common time base according to the highest sampling frequency, and linear

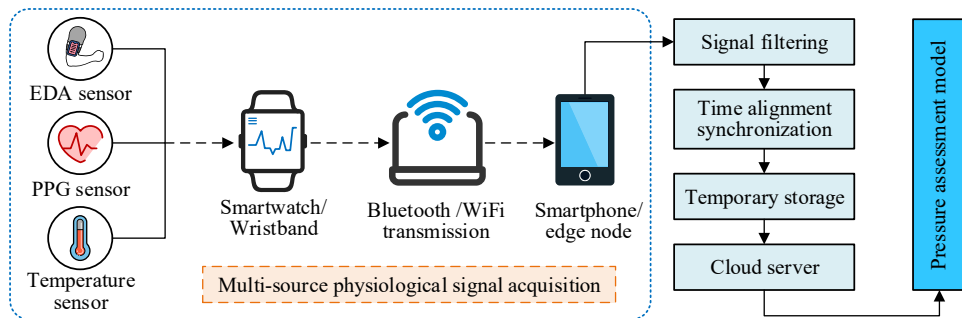
interpolation and resampling strategies are applied to low-frequency signals, ultimately synchronising the three types of signals to a unified time step. Secondly, the transmission delay between sensors is calculated through timestamp differencing, and the timestamps are smoothed and corrected before synchronisation to eliminate Bluetooth/Wi-Fi transmission offsets. Finally, a sliding window weighting strategy is used to compensate for short-term alignment errors across modalities, ensuring high consistency even with slight phase mismatches of 0.1–0.5 seconds. The pre-processing of sensor signals is shown in equation (1) (Jiao et al., 2023).

$$y(t) = \int_{-\infty}^{\infty} x(\tau) \cdot h(t - \tau) d\tau \quad (1)$$

In equation (1), $x(\tau)$ is the original signal (such as EDA or PPG). $h(t - \tau)$ is the impulse response of the filter (such as the time-domain form of Butterworth bandpass filter). $y(t)$ is the filtered signal. Equation (1) eliminates baseline drift (low frequency) and high-frequency noise by frequency domain truncation (such as retaining 0.05–5 Hz), with parameters determined by the cut-off frequency and filter order. Next, its synchronisation mechanism is shown in equation (2) (Zhang et al., 2023a).

$$D(i, j) = \|x_i - y_j\| + \min \begin{cases} D(i-1, j) \\ D(i, j-1) \\ D(i-1, j-1) \end{cases} \quad (2)$$

Figure 1 Multi-source PS acquisition and synchronisation mechanism module (see online version for colours)



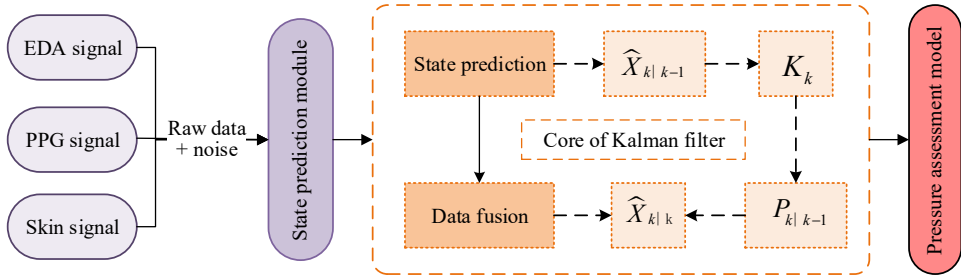
In equation (2), x_i and y_j are the i^{th} and j^{th} sampling points of two signals to be aligned (such as EDA and PPG). $\|x_i - y_j\|$ is the point-to-point distance (such as Euclidean distance). $D(i, j)$ records the minimum cumulative distance from the starting point to (i, j) . Equation (2) uses dynamic programming to find the optimal path, solve the time offset and rate difference between signals, and achieve alignment of non-uniform sampling signals.

Because physiological signals such as EDA and PPG are susceptible to motion artefacts, poor sensor contact, and environmental disturbances, this study introduced a multi-level signal quality control mechanism during the data acquisition and pre-processing stages. First, a real-time artefact detection algorithm monitored for sudden changes in signal amplitude, abnormal conductivity jumps, and saturation segments, automatically marking them as candidate abnormal segments. Second, a combination of

rolling median filtering and Butterworth filtering was used to suppress high-frequency noise and baseline drift, and cross-correlation detection with an offset window was employed to identify phase shifts caused by slight contact problems. Local interpolation repair was performed on clearly abnormal outlier points (e.g., spikes exceeding the mean $\pm 3\sigma$) to prevent artefacts from being passed to the classification model. Finally, low-quality segments were filtered out based on a signal quality index (SQI), and only samples meeting the SQI threshold were used for the fusion model and neural network training.

Due to differences in sampling frequency, delay characteristics, and measurement errors among multiple types of sensors, traditional fusion methods are difficult to balance real-time performance and accuracy. Therefore, this study introduces the BKF algorithm to construct a fusion mechanism, as shown in Figure 2.

Figure 2 BKF data fusion mechanism (see online version for colours)



The mechanism in Figure 2 aims to perform temporal alignment, noise suppression, and dynamic fusion on multi-source PSs (such as EDA, PPG, and body temperature) to improve data consistency and model input quality. This mechanism first predicts the observation values of each sensor at the next moment based on the state space model, and introduces a system bias term to model the signal offset. Subsequently, by observing the update process, the residuals between the actual measured values and the predicted values are compared, and weight optimisation and state updates are achieved using the modified Kalman gain to dynamically adjust the fusion output. The entire process can achieve smooth processing and time-domain alignment of different modal signals, while preserving the small variation features of various signals, providing highly consistent and robust feature inputs for subsequent NN models, and improving the stability and accuracy of pressure state classification. The state prediction equation is given by equation (3) (Singh et al., 2024).

$$\hat{X}_{k|k-1} = A\hat{X}_{k-1|k-1} + Bu_{k-1} + b_k \quad (3)$$

In equation (3), $\hat{X}_{k|k-1}$ denotes the prior state estimate (predicted value) at time K . A means the state transition matrix (describing the dynamic model of the system). Bu_{k-1} represents control input (0 if none). b_k is the sensor bias term (a key improvement used to compensate for inherent biases in signals such as EDA/PPG). This equation predicts the current state by combining the state estimation ($\hat{X}_{k-1|k-1}$) from the previous moment with the system model, and explicitly corrects the sensor systematic error through b_k . The error covariance prediction is shown in equation (4) (Rashid et al., 2023).

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q \quad (4)$$

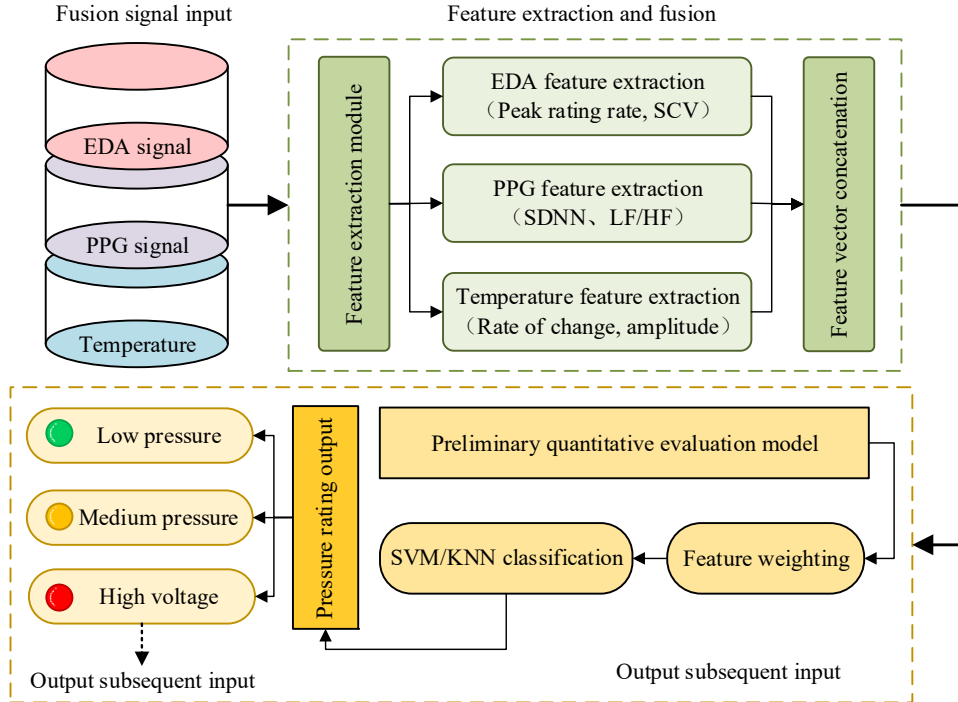
In equation (4), $P_{k|k+1}$ is the prior error covariance (uncertainty of state prediction) at time k . $P_{k-1|k-1}$ is the posterior error covariance of the previous moment. Q is the covariance of process noise (modelling system model error). Equation (4) quantifies the estimation error in the prediction stage by propagating the uncertainty of the previous moment through state transition and adding process noise, providing an uncertainty benchmark for subsequent fusion updates. In the fusion update stage, the calculation of biased Kalman gain K_k is given by equation (5) (Siam et al., 2023).

$$K_k = P_{k|k-1}H^T (HP_{k|k-1}H^T + R + \Delta R)^{-1} \quad (5)$$

In equation (5), $P_{k|k+1}$ is the prior error covariance (prediction uncertainty). H denotes the observation matrix (mapping from state to observation). R is the covariance of observed noise, and ΔR is the bias compensation term (correcting for additional uncertainty introduced by sensor system bias). K_k determines the weight of the observed values. Equation (5) balances the reliability of predicted and observed values during state updates by adjusting the gain K_k . ΔR explicitly compensates for the impact of sensor inherent bias on fusion results. The status update (fusion of multi-source data) is shown in equation (6) (Richer et al., 2024).

$$K_k = P_{k|k-1}H^T (HP_{k|k-1}H^T + R + \Delta R)^{-1} \quad (6)$$

Figure 3 Construction of psychological stress characteristics and preliminary quantitative assessment model (see online version for colours)



In equation (6), $\hat{X}_{k|k}$ is the posterior state estimate (fusion result) at time k . Z_k means the actual observation value (multi-source sensor data). Equation (6) corrects the predicted value through the residual term $(Z_k - H\hat{X}_{k|k-1} - b_k)$. The explicit subtraction of b_k eliminates sensor system bias and achieves high-precision fusion of multi-source data. After integrating multiple PSs, this study constructs a student psychological stress assessment model, as shown in Figure 3.

The model in Figure 3 aims to reflect the characteristic system of students' psychological stress state and achieve preliminary quantitative evaluation of their stress level. A series of static and dynamic statistical features are extracted from signals including EDA, PPG, and body temperature. These features are segmented and locally analysed through a sliding window mechanism to enhance the model's capacity to capture short-term stress responses. On the basis of feature construction, a lightweight ML model is introduced for preliminary stress classification, dividing students' psychological states into low, medium, or high pressure levels. At the same time, the model can integrate weight information of different modal features, establish a weighted scoring mechanism, and improve the ability to distinguish complex states. This stage provides structured and highly expressive input feature vectors for subsequent NN classification models, and can also achieve fast and interpretable prediction of pressure states, which helps to achieve hierarchical pressure recognition and feedback control of the system.

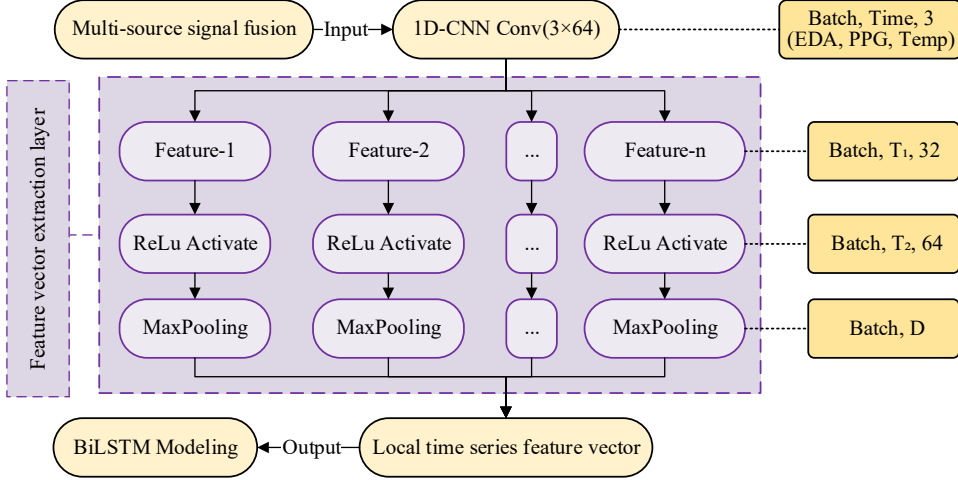
3.2 *Student PSC model based on NN*

This study utilises wearable multimodal physiological sensors to collect multi-source data such as EDA, PPG, and body temperature, and proposes a student psychological stress assessment model based on SDF to achieve dynamic perception and high robustness fusion processing of individual physiological states. This study further proposes a student PSC model that combines 1D-CNN and BiLSTM structures to integrate local variation features with global temporal dependencies. It constructs a feature extraction and temporal modelling system that progresses from shallow to deep, achieving refined classification of psychological stress states. Due to the fact that psychological stress usually manifests as subtle and instantaneous physiological responses such as heart rate fluctuations, skin conductance mutations, or slight temperature changes in a short period of time, traditional statistical features are difficult to fully reflect these high-frequency local characteristics. Therefore, this study constructs a feature extraction module based on 1D-CNN, performs local window convolution operation on time series signals, and extracts short-term variation patterns at different scales, as shown in Figure 4.

In Figure 4, this module aims to effectively capture key temporal patterns of changes from the fused multi-source PSs. Each convolution kernel slides in different channels, automatically learns local temporal dependencies, enhances non-linear expression ability with activation functions such as ReLU, and compresses feature dimensions through pooling operations (such as max pooling) to improve computational efficiency and noise resistance. Finally, the local temporal features extracted from multiple channels are concatenated into a unified feature representation, providing highly expressive and highly temporal preserving feature inputs for subsequent temporal modelling and classification networks. The 1D convolution operation (core feature extraction) is shown in equation (7) (Hafeez and Shakil, 2024).

$$FeatureMap_i(t) = ReLU \left(\sum_{j=0}^{K=1} W_i(j) \cdot X(t+j) + b_i \right) \quad (7)$$

Figure 4 Feature extraction module based on 1D-CNN (see online version for colours)



In equation (7), $X(t+j)$ is the value of the input signal (such as the fused EDA/PPG) at time point $t+j$. W_i is the weight of the i^{th} convolution kernel at position j (trainable parameter). K is the convolution kernel size (e.g., three time points). $ReLU$ is the activation function. Equation (7) extracts local temporal patterns (such as spikes/dips) by sliding the window (step = 1), and outputs the value of the i^{th} feature map at time t . The maximum pooling operation is shown in equation (8) (Jain et al., 2024).

$$Pooled_n(t) = \max_{m \in [t, t+S-1]} FeatureMap_n(m) \quad (8)$$

In equation (8), $FeatureMap_n(m)$ means the value of the n^{th} feature map at time position m , S is the size of the pooling window, and the max operation extracts the maximum value within the window as the output $Pooled_n(t)$. This operation compresses the length of the feature map to $\frac{1}{S}$ (such as 30 second signal $\rightarrow$ 15 seconds) by preserving the local

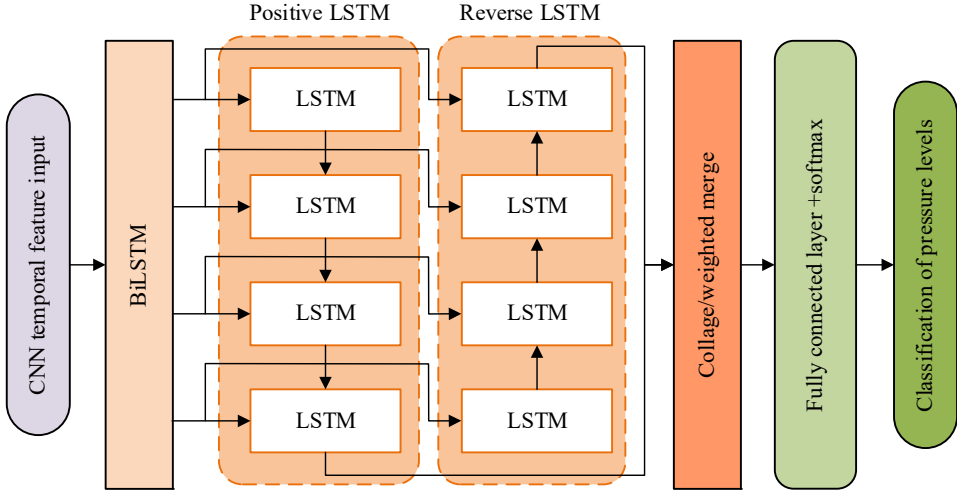
most significant features (such as peaks), reducing computational complexity while enhancing robustness to small temporal fluctuations. This study uses BiLSTM to capture the forward and backward dependencies in time series data, to more comprehensively model the dynamic evolution process of students' psychological stress, as shown in Figure 5.

In Figure 5, BiLSTM processes data through two independent LSTM units, one forward LSTM gradually processes data from the beginning to the end of the sequence, and the other reverse LSTM processes data from the end of the sequence forward. Through this bidirectional structure, the network can simultaneously consider contextual information, enhancing its understanding of short-term and long-term dependencies in time series. At the output, the results of forward and backward LSTM are concatenated or weighted to form a unified feature representation. Ultimately, the features generated by

BiLSTM are used for subsequent PSC, achieving accurate prediction of students' stress state. The forward and backward LSTM calculations are shown in equation (9) (Kang et al., 2023).

$$\begin{cases} \bar{h}_t = LSTM(x_t, \bar{h}_{t-1}, \bar{c}_{t-1}) \\ \tilde{h}_t = LSTM(x_t, \tilde{h}_{t+1}, \tilde{c}_{t+1}) \end{cases} \quad (9)$$

Figure 5 Temporal modelling based on BiLSTM (see online version for colours)



In equation (9), x_t is the input feature of time step t (extracted from 1D-CNN), h is the hidden state, and c is the cell state. Forward $LSTM(\bar{h}_t)$ processes the sequence from $t = 1$ to T , and backward $LSTM(\tilde{h}_t)$ processes it from $t = T$ to 1, jointly modelling global temporal dependencies. The forward and backward LSTM outputs are merged as shown in equation (10) (Zhang et al., 2023b).

$$h_t = \alpha \bar{h}_t + (1 - \alpha) \tilde{h}_t \quad (10)$$

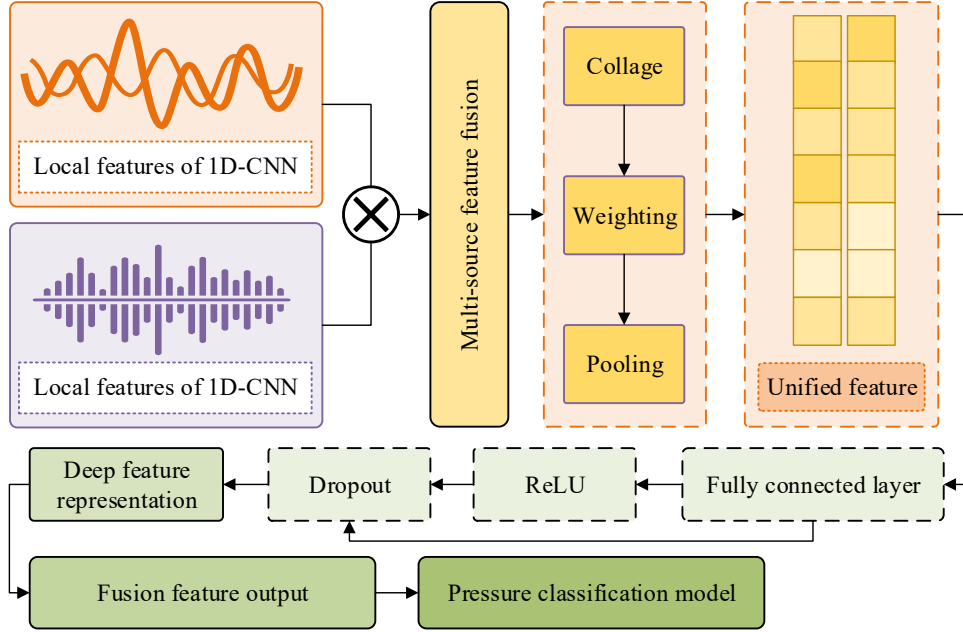
In equation (10), $\bar{h}_t$ and $\tilde{h}_t$ are the hidden states of forward/backward LSTM at t . α is the learnable weighting coefficient ($0 \leq \alpha \leq 1$). The concatenation operation preserves bidirectional complete information, while weighted merging adaptively adjusts the forward and backward contributions, ultimately forming a unified feature representation that includes global temporal dependencies. Multi-source features are fused to propose a feature fusion and deep representation learning mechanism, as shown in Figure 6.

In Figure 6, by integrating multi-source features from different networks (1D-CNN and BiLSTM), potential patterns of student psychological stress are further explored and classification performance is improved. Firstly, the local temporal features extracted by 1D-CNN are fused with the temporal dependent features modelled by BiLSTM. Common methods include feature concatenation, weighted averaging, or max pooling to form a unified high-dimensional feature vector. Next, deep neural networks such as multilayer perceptrons are used to perform deep representation learning on the fused features, learning more abstract feature representations. These deep features can capture complex

nonlinear relationships and subtle changes in psychological stress states. This mechanism optimises NN parameters through backpropagation, continuously extracts the expressive power of features, and ultimately outputs high-quality feature vectors for PSC. The feature weighted fusion extracted by 1D-CNN and BiLSTM is shown in equation (11) (Prakash et al., 2024; Pan et al., 2024).

$$h_{fused} = g(W_c h_{CNN} + W_b h_{BiLSTM} + b) \quad (11)$$

Figure 6 Feature fusion and deep representation learning mechanism (see online version for colours)



In equation (11), h_{CNN} (dimension $T \times D_1$) and h_{BiLSTM} (dimension $T \times D_2$) are the local features extracted by 1D-CNN and the temporal dependent features extracted by BiLSTM. T is the time step and D means the feature dimension. W_c and W_b are trainable weight matrices, b denotes the bias term, and g is the activation function (such as ReLU) or normalisation operation. Equation (11) adaptively adjusts the contributions of two types of features through a weight matrix, supports concatenation or weighted summation, and preserves the temporal structure while maintaining a unified dimension (with the time step T unchanged). Finally, the deep representation learning through fully connected layers is shown in equation (12) (Luo et al., 2023).

$$h_{deep} = \eta \cdot (W_2 ReLU(W_1 h_{fused} + b_1) + b_2) \quad (12)$$

In equation (12), h_{fused} is the fused feature, W_1 and W_2 are trainable weight matrices, b_1 and b_2 are bias terms, and η is the Sigmoid activation function (used for classification tasks). Equation (12) uses a two-layer fully connected network (including ReLU nonlinear activation) to perform nonlinear transformation and dimensionality reduction on the fused features, ultimately outputting deep abstract features h_{deep} while maintaining

temporal structural integrity. Dropout layers can be selectively added to prevent overfitting.

To enable the model to meet the requirements of real-time psychological stress monitoring scenarios, the research optimised the inference stage through lightweighting and pipelining after network design. First, depthwise separable convolution was introduced into the convolutional layers to reduce the number of parameters and computational cost, and a parameter sharing strategy was adopted in the fully connected layers to reduce redundant weights. Second, structured pruning combined with eight-bit quantisation technology was used to compress the network weights, reducing the overall model size by approximately 30%–35%, while improving inference speed with less than 1% change in accuracy. Furthermore, considering the temporal characteristics of continuous physiological signals, a sliding window incremental inference mechanism was employed, updating features only for the newly added time slices, avoiding overall redundant calculations. Finally, the classification module and the intervention strategy recommendation module operate asynchronously and in parallel within the system architecture, allowing stress identification and strategy generation to be scheduled independently in the background, further reducing end-to-end response latency. As shown in Figure 7.

Figure 7 A classification and intervention model of student psychological stress based on NN (see online version for colours)

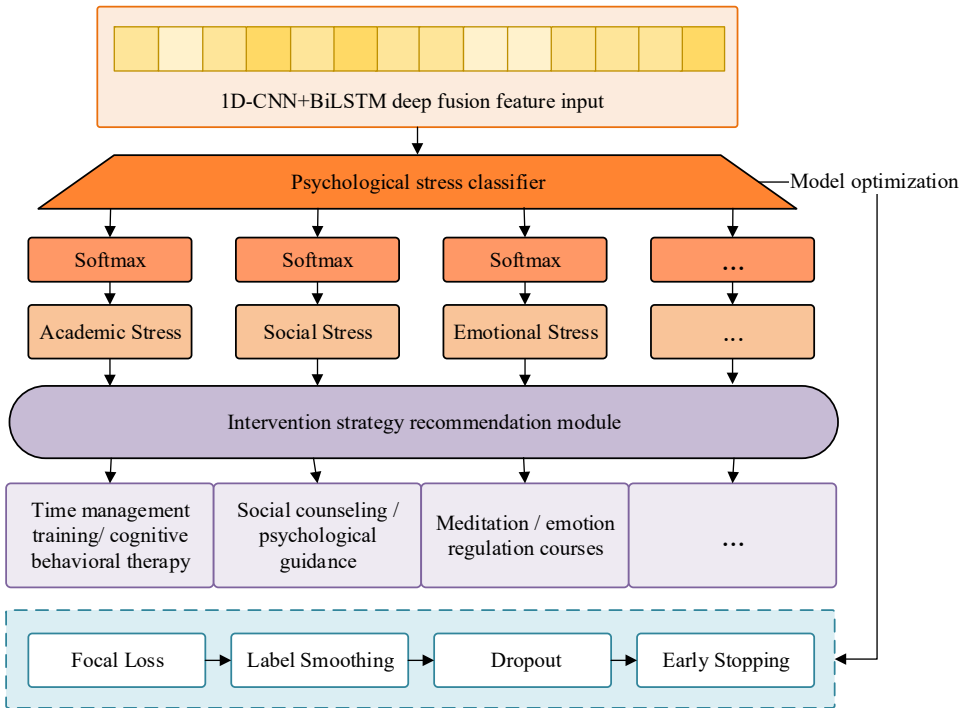


Figure 7 shows the overall architecture of the student PSC and intervention strategy model, forming a closed-loop process from multi-source physiological data input to the final intervention strategy output. The model first uses a 1D-CNN to capture local

short-term physiological change patterns, and then models global temporal dependencies using a BiLSTM. After feature compression and deep abstraction, the data enters the classification module, which is used to identify the student's stress type (e.g., academic stress, social stress, emotional stress, etc.). Based on this, the system further links to the intervention strategy recommendation submodule. Unlike traditional static rule engines, the recommendation module adopts a hybrid mechanism of 'multi-indicator weighted scoring + category constraint rules'. On the one hand, it uses the stress type probability from the classification output, the current physiological fluctuation amplitude, historical stress trends, and model confidence to construct an intervention score S , as shown in equation (13).

$$S = w_1 P_{class} + w_2 V_{volatility} + w_3 T_{history} + w_4 C_{confidence} \quad (13)$$

As shown in equation (13), w_1 – w_4 is obtained through tuning on the validation set and is used to quantify the degree of match between the intervention plan and the individual's state. On the other hand, the system incorporates hierarchical intervention rules commonly used in clinical psychology as constraints, so that different stress categories are prioritised to match specific types of intervention strategies (e.g., academic stress is biased towards time management training, emotional stress towards emotion regulation training, and social stress towards social skills counselling). Based on the above mechanism, to improve the adaptability of the intervention plan, this study further introduces a personalised intervention strategy enhancement mechanism. This mechanism adds individual behavioural and habitual characteristics (such as learning pace, social sensitivity, self-regulation tendency, and circadian rhythm preference) to the original intervention score, forming a personalised correction term, as shown in equation (14).

$$S_{personalised} = S + \lambda_1 H_{habit} + \lambda_2 B_{behaviour} + \lambda_3 P_{preference} \quad (14)$$

As shown in equation (14), λ_1 – λ_3 is learned through the validation set and is used to adjust the degree of matching between the intervention strategy and the individual characteristics of the student. This mechanism ensures that different students still receive differentiated support strategies under the same type of stress, elevating the recommendation module from 'category-level decision-making' to 'individual-level decision-making'. Finally, the system jointly decides and outputs the optimal intervention plan based on the principle of maximising the personalised weighted score and category rules, achieving a closed-loop intelligent decision-making process from stress identification to personalised support strategy generation.

4 Results

4.1 Validity verification of SDF mechanism

To verify the reliability and effectiveness of the student PSC model based on SDF and NN, this study conducts experimental performance verification and analysis. The experimental parameters are shown in Table 1.

Table 1 shows the experimental environment and key parameters. Based on this, two publicly available datasets, WESAD and AMIGOS, were selected as the main data

sources. The WESAD dataset includes 17 participants (eight males, nine females), aged 24–35 years. The AMIGOS dataset includes 40 participants (22 males, 18 females), aged 21–40 years. Both datasets cover the multimodal physiological features required for stress induction experiments. WESAD used RespiBAN Professional (chest device, 700 Hz) and Empatica E4 (wrist device, EDA 4 Hz, PPG 64 Hz, body temperature 4 Hz) for data acquisition. In AMIGOS, the EDA sampling rate was 32 Hz, PPG was 64 Hz, and body temperature was 4 Hz. To ensure consistency of multi-source data in the same model, all physiological signals were uniformly resampled to PPG 64 Hz, EDA 32 Hz, and body temperature 4 Hz, and processed with timestamp alignment. The final samples were divided into training, validation, and test sets in a 7:2:1 ratio, maintaining a balanced proportion of different genders and age groups to improve the model’s generalisation ability in cross-individual scenarios.

Table 1 Experimental environment and key parameters

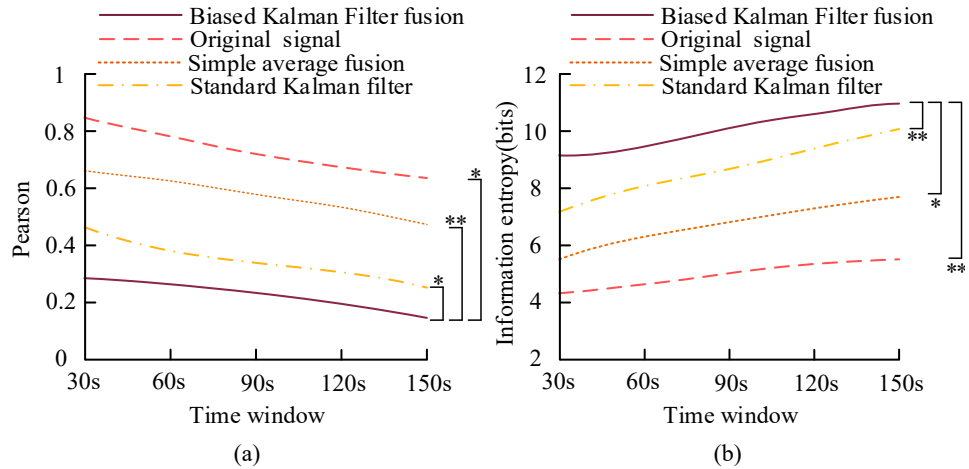
<i>Experimental environment</i>		
Runtime environment	Operating system	Windows 11
	Hardware platform	Intel Core i7-12700
	GPU	NVIDIA RTX 3060 (6 GB)
	RAM	16 GB DDR4
Software and frameworks	Python version	Python 3.9
	Deep learning framework	PyTorch 1.12
	Major packages	NumPy, SciPy, Scikit-learn, Matplotlib
<i>Key configuration</i>		
Dataset configuration	Datasets used	WESAD, AMIGOS
	Data split	70% train / 20% Val / 10% test
	Unified sampling rate	EDA: 32 Hz, PPG: 64 Hz, temp: 4 Hz
Model key parameters	CNN kernel size	3×1
	CNN channels	32 / 64
	BiLSTM hidden size	128
	Optimiser	Adam (lr = $1e-3$)
	Batch size	64
	Epochs	80

In terms of variable control, the independent variables are the input data types (single-mode, dual-mode, multi-mode), and the control methods include the original signal, unbiased Kalman filtering, weighted moving average (WMA), RNN, GRU, Transformer, etc. The variables are classification accuracy, F1-score, RMSE, alignment error, Pearson correlation coefficient, intervention matching degree, etc. The experimental process is divided into four steps. Firstly, the signals from multi-source sensors are collected and pre-processed, and biased Kalman filtering is used for timing alignment and noise suppression. Secondly, a preliminary quantitative assessment model for psychological stress is constructed to distinguish between low pressure, medium pressure and high pressure levels. Then, local feature extraction is carried out using 1D-CNN, sequence dependency is modelled by BiLSTM, and pressure state classification is

achieved through deep feature fusion. Finally, input the classification results into the intervention recommendation module and output the personalised regulation strategies.

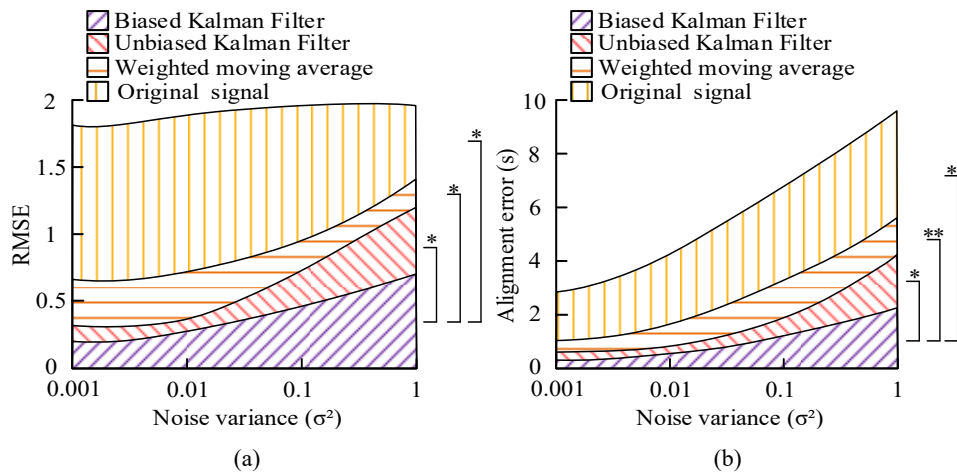
The experiment first compares the feature quality before and after SDF, selecting the original signal (no fusion), simple average fusion (no filtering), and standard Kalman filter (SKF) as controls. The experimental results are shown in Figure 8.

Figure 8 Comparison of multimodal feature fusion effects, (a) comparison of feature redundancy (b) comparison of information entropy (see online version for colours)



Notes: * indicates $p < 0.05$, ** indicates $p < 0.01$. Significance testing is based on independent sample t-tests from three independent experiments, and all markings indicate a significant difference compared to the control model.

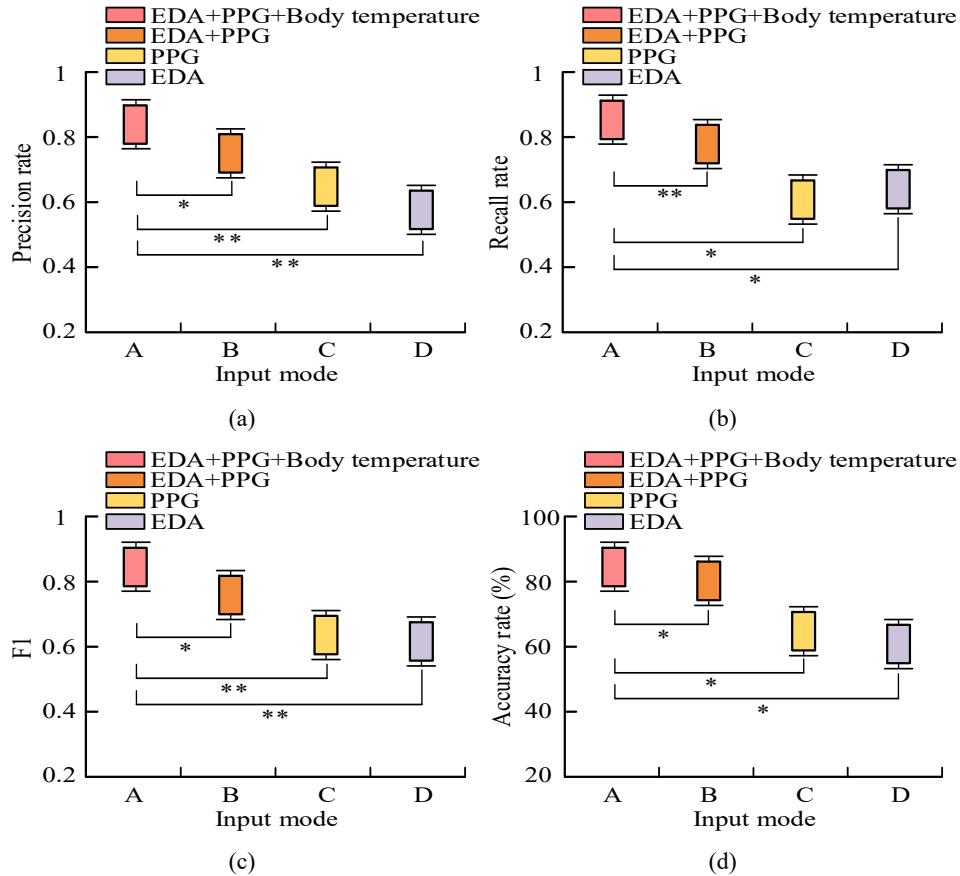
Figure 9 Comparison of the influence of deviation modelling on filtering performance, (a) comparison of filtered signals (b) alignment error comparison (see online version for colours)



Notes: * indicates $p < 0.05$, ** indicates $p < 0.01$. Significance testing is based on independent sample t-tests from three independent experiments, and all markings indicate a significant difference compared to the control model.

In Figure 8, within a 150 second time window, the Pearson correlation coefficient of the BKF fusion algorithm decreases to 0.17 (0.14 lower than SKF's 0.31 and 0.55 lower than the original signal's 0.72). The information entropy has been increased to 10.82 bits (1.87 bits higher than SKF's 8.95 bits and 5.48 bits higher than the original signal's 5.34 bits). All methods improve with the extension of the window, but the BKF fusion algorithm still maintains optimal performance (redundancy of 0.28, information entropy of 9.32 bits) in short windows (30 s), confirming its ability to effectively integrate the advantages of multi-source signals and suppress inherent sensor biases. Compared with other models, the test results of the proposed model show significant differences ($p < 0.05$). The experiment conducts an analysis of the bias modelling effect in Kalman fusion, selecting the original signal, unbiased Kalman filter (UKF), and WMA as controls, as shown in Figure 9.

Figure 10 The influence of multimodal signals on the classification of psychological stress, (a) precision comparison (b) recall rate comparison (c) F1-score comparison (d) accuracy rate comparison (see online version for colours)



Notes: * indicates $p < 0.05$, ** indicates $p < 0.01$. Significance testing is based on independent sample t-tests from three independent experiments, and all markings indicate a significant difference compared to the control model.

In Figure 9, when the noise variance σ^2 is 0.1, the RMSE of BKF is 0.42, which is lower than UKF's 0.73 and WMA's 0.89. The alignment error is 0.95s, which has a significant advantage over UKF's 1.82s and WMA's 2.97s. As the noise variance increases to $\sigma^2 = 1.0$, the RMSE of BKF only rises to 0.76, much lower than UKF's 1.29. Its alignment error of 2.03 seconds is also significantly smaller than UKF's 4.15 seconds and WMA's 5.62 seconds. Data show that BKF effectively improves the stability of the system in noisy environments by explicitly modelling sensor bias. Compared with other models, the test results of the proposed model show significant differences ($p < 0.05$). Next, experiments are conducted to compare the performance of multimodal signal classification, selecting single-modal EDA, single-modal PPG only, and dual-modal EDA+PPG as controls. The results are shown in Figure 10.

In Figure 10, the F1-score of the multimodal combination (EDA+PPG+body temperature) reaches 0.90, which is better than the bimodal combination (0.82) and the unimodal combination (EDA 0.70, PPG 0.71). Its accuracy reaches 92.4%, which is also higher than bimodal (84.6%) and unimodal (PPG-only 73.8, EDA-only 71.3). In terms of indicator balance, the accuracy (0.89) and recall (0.91) of multimodal analysis remain optimal, with the smallest standard deviation (F1-score ± 0.01). Single-modal presentation is significantly biased (EDA has a recall rate of 0.72 but an accuracy rate of 0.68, while PPG is the opposite). The data confirm that multimodal signals enhance the stability of the model while improving classification performance through feature complementarity. Compared with other models, the test results of the proposed model show significant differences ($p < 0.05$).

4.2 Performance validation of NN classification model

After verifying the effectiveness of the SDF mechanism, performance validation is conducted on the 1D-CNN and BiLSTMNN classification models. The experiment first compares the performance of feature extraction methods, selecting traditional time-domain+frequency-domain extraction, multilayer perceptron (MLP), and recurrent neural network (RNN) as controls. The experimental parameters and environment are the same as above, and the results are shown in Figure 11.

In Figure 11, under 1024 time point samples, 1D-CNN only takes 35.47 ms to complete feature extraction, which is less than one tenth of the time required by RNN (342.56 ms). At the same time, it achieves a classification accuracy of 90.24%, which is significantly improved compared to RNN (79.85%). As the sample size increased from 128 to 1,024, the accuracy of 1D-CNN remains steadily increasing (82.37% to 90.24%), while traditional methods only improves from 68.42% to 72.31%. In all test dimensions, the accuracy of 1D-CNN is superior to other methods, and the processing time is always the shortest (only 5.16 ms is required for 128 dimensions). The data confirm that 1D-CNN can efficiently capture temporal features. Compared with other models, the test results of the proposed model show significant differences ($p < 0.05$). The experiment compares the performance of time series modelling methods, selecting unidirectional LSTM, GRU, and transformer encoder structures as controls, as shown in Figure 12.

In Figure 12, BiLSTM achieves a time-dependent score of 9.73 in a 100-second long sequence task, higher than transformer encoder (8.68), GRU (8.17), and unidirectional LSTM (7.51). The gradient change amplitude is only 0.98, much lower than that of unidirectional LSTM (4.02) and GRU (3.42). As the sequence length increases from 20

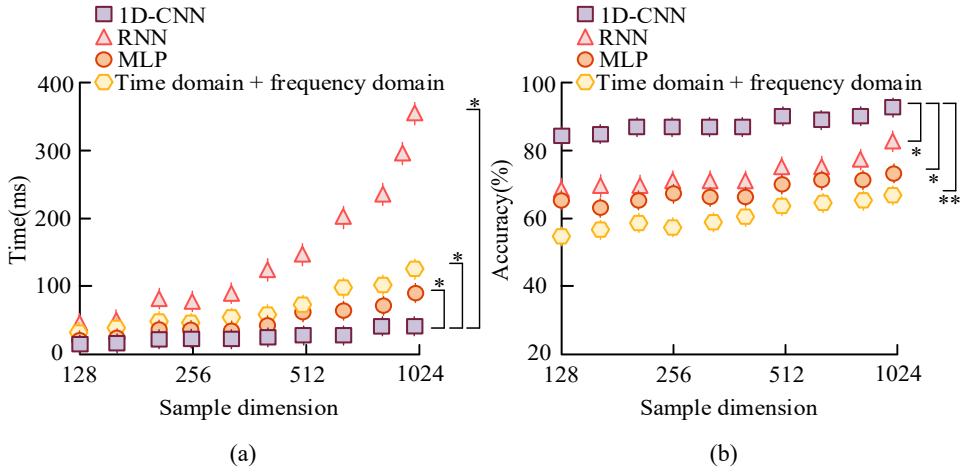
seconds to 100 seconds, the modelling score of BiLSTM continues to improve (7.94 to 9.73), while the gradient fluctuation decreases (1.62 to 0.98), demonstrating a unique stability advantage. This confirms that BiLSTM is more suitable for processing temporal features of long-term psychological stress. Compared with other models, the test results of the proposed model show significant differences ($p < 0.05$). The experiment conducts a comparison of deep feature fusion strategies, selecting CNN+classifier, BiLSTM+classifier, and MLP as controls, as listed in Table 2.

- 1 The t-SNE separation score is used to measure the separability of different pressure categories in the t-SNE dimensionality reduction space. A higher value indicates a greater distance between classes and less overlap between classes. Its calculation method is shown in equation (15).

$$t_s = \frac{1}{K(K-1)} \sum_{i \neq j} (d_{ij}^{between} - d_{ij}^{within}) \quad (15)$$

In equation (15), $d_{ij}^{between}$ represents the distance between class centroids, and d_{ij}^{within} represents the average distance within a class; both are dimensionless.

Figure 11 Performance comparison of feature extraction methods, (a) comparison of feature extraction time (b) comparison of classification accuracy (see online version for colours)

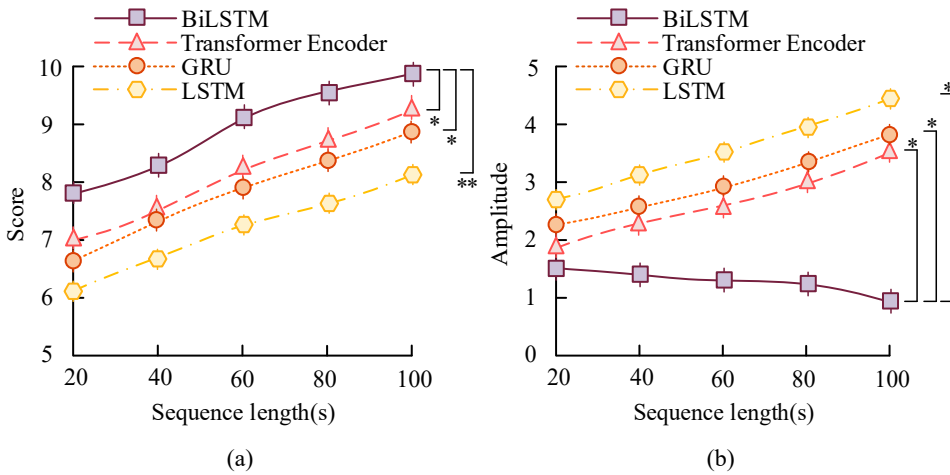


Notes: * indicates $p < 0.05$, ** indicates $p < 0.01$. Significance testing is based on independent sample t-tests from three independent experiments. All markings indicate a significant difference compared to the control model. Error bars in the figure represent the standard deviation.

- 2 Floating point operations (2FLOPs) represents the number of floating-point calculations required for one forward inference of the model, and is typically used to evaluate model complexity. This study uses the 'total computation per inference' as the metric for FLOPs statistics, with the unit being millions of floating-point operations (MFLOPs).

In Table 2, the accuracy of CNN+BiLSTM reaches 89.47%, higher than that of BiLSTM+classifier (82.15) and CNN+classifier (78.32). The separation degree of t-SNE is 8.62, which is at least 1.77 higher than other methods, indicating a significant advantage in feature separability. In terms of model efficiency, the FLOPs of CNN+BiLSTM are 5.23×10^8 , which is higher than MLP (0.95×10^8), but the F1-score (0.88) is significantly higher than MLP (0.72). Compared with other models, the test results of the proposed model show significant differences ($p < 0.05$). The data confirms that CNN+BiLSTM achieves optimal classification performance while ensuring the model's expressive ability through spatiotemporal feature fusion. Next, the experiment compared the effectiveness of pressure classification and intervention strategy recommendations, using rule-based expert system, CBT static matching model, and SVM+hand-crafted features as controls. The experimental results are shown in Table 3.

Figure 12 Performance comparison of time series modelling methods, (a) time-dependent score comparison (b) comparison of the magnitude of gradient changes (see online version for colours)



Notes: * indicates $p < 0.05$, ** indicates $p < 0.01$. Significance testing is based on independent sample t-tests from three independent experiments. All markings indicate a significant difference compared to the control model. Error bars in the figure represent the standard deviation.

Table 2 Performance of deep feature fusion strategies

Evaluation index	CNN+classifier	BiLSTM+classifier	MLP	CNN+BiLSTM
Accuracy rate (%)	78.32 ± 1.25	82.15 ± 1.08	74.63 ± 1.47	$89.47 \pm 0.85^*$
t-SNE separation degree	5.72 ± 0.38	6.85 ± 0.42	4.93 ± 0.45	$8.62 \pm 0.33^*$
F1-scores	0.76 ± 0.03	0.81 ± 0.02	0.72 ± 0.04	$0.88 \pm 0.01^*$
FLOPs ($\times 10^8$)	1.82 ± 0.12	3.57 ± 0.21	0.95 ± 0.08	$5.23 \pm 0.34^*$
Training time (min)	18.45 ± 1.32	34.27 ± 2.15	12.83 ± 1.05	$42.16 \pm 2.87^*$

Notes: * indicates that the indicator is statistically superior to other models ($p < 0.05$).

As shown in Table 3, the intervention matching degree of the rule-based expert system was only 6.12 ± 0.63 , and the classification accuracy was $71.23\% \pm 2.05$. The matching degree and accuracy of the CBT static matching model increased to 6.85 ± 0.58 and $74.68\% \pm 1.98$, respectively. Furthermore, the SVM+artificial feature method performed slightly better, with an intervention matching degree of 7.21 ± 0.52 and a classification accuracy of $76.84\% \pm 1.89$, but the decision coverage rate was still only $78.62\% \pm 3.51$. In contrast, the CNN+BiLSTM recommendation model proposed in this study significantly outperformed all other methods in all metrics, achieving an intervention matching degree of 8.92 ± 0.42 , a classification accuracy of $89.47\% \pm 1.08$, a decision coverage rate of $95.36\% \pm 2.14$, and a user satisfaction score of 8.63 ± 0.55 , while maintaining a response time of 89.76 ± 5.23 ms. Overall, the proposed method outperformed the three control methods in terms of intervention accuracy, classification ability, and system coverage, with all differences reaching statistical significance ($p < 0.05$). Finally, to verify the contribution of each key component (Kalman filter module, multimodal input, CNN/BiLSTM structure fusion) to the final performance, an ablation experiment was designed, and the results are shown in Table 4.

Table 3 Performance comparison of recommended methods for psychological intervention strategies

<i>Evaluation index</i>	<i>Rule-based expert system</i>	<i>CBT static matching model</i>	<i>SVM+hand-crafted features</i>	<i>CNN+BiLSTM</i>
Intervention matching degree	6.12 ± 0.63	6.85 ± 0.58	7.21 ± 0.52	$8.92 \pm 0.42^*$
Classification accuracy (%)	71.23 ± 2.05	74.68 ± 1.98	76.84 ± 1.89	$89.47 \pm 1.08^*$
Decision coverage (%)	82.47 ± 3.18	86.35 ± 2.74	78.62 ± 3.51	$95.36 \pm 2.14^*$
Customer satisfaction	6.02 ± 0.71	6.58 ± 0.66	6.95 ± 0.78	$8.63 \pm 0.55^*$
Response time (ms)	45.67 ± 3.42	58.32 ± 4.26	68.23 ± 4.15	$89.76 \pm 5.23^*$

Notes: * indicates that the indicator is statistically superior to other models ($p < 0.05$).

Table 4 Experimental results of component ablation studies

<i>Component removal items</i>	<i>Accuracy (%)</i>	<i>F1-score</i>	<i>RMSE</i>	<i>Alignment error</i>
No Kalman filtering	76.84 ± 1.93	0.78 ± 0.03	1.82 ± 0.11	3.47 ± 0.18
Unbiased Kalman filter	82.15 ± 1.08	0.81 ± 0.02	0.73 ± 0.05	1.82 ± 0.12
EDA+PPG	84.60 ± 1.15	0.82 ± 0.02	0.58 ± 0.04	1.54 ± 0.10
EDA-only	71.30 ± 2.05	0.70 ± 0.04	2.00 ± 0.13	6.83 ± 0.29
PPG-only	73.80 ± 1.89	0.71 ± 0.03	1.95 ± 0.12	4.12 ± 0.21
CNN-only	78.32 ± 1.25	0.76 ± 0.03	1.29 ± 0.09	2.97 ± 0.17
BiLSTM-only	82.15 ± 1.08	0.81 ± 0.02	0.89 ± 0.06	1.82 ± 0.12
Complete model	$89.47 \pm 0.85^*$	$0.88 \pm 0.01^*$	$0.42 \pm 0.03^*$	$0.95 \pm 0.08^*$

Notes: * indicates that the indicator is statistically superior to other models ($p < 0.05$).

As shown in Table 4, the accuracy without Kalman filtering is only 76.84%, and the F1-score is 0.78, while the RMSE and alignment error increase to 1.82 and 3.47s, respectively. The accuracy of the single-modal EDA-only/PPG-only models is only

71.30%/73.80%, and the F1-score is only 0.70–0.71. Even the dual-modal EDA+PPG model only improves to 84.60%. In terms of structure, the accuracy of the CNN-only and BiLSTM-only models are 78.32% and 82.15%, respectively, still significantly lower. In contrast, the complete model achieves an accuracy of 89.47% and an F1-score of 0.88, with an RMSE of only 0.42 and an alignment error of 0.95s, demonstrating the best overall performance. The results show that removing any key component significantly weakens the classification ability and signal quality, and all the above differences are statistically significant ($p < 0.05$), further proving the necessity and irreplaceability of each module for the model's performance.

5 Discussion and interpretation

The proposed student PSC method that integrated multi-source PSs with NN has achieved collaborative optimisation in multimodal data processing, temporal modelling, and intervention recommendation, effectively improving classification accuracy and system adaptability. Compared with Zhu et al.'s single channel pressure recognition method based on wrist EDA signals, this study achieved performance with an F1-score of 0.90 and a classification accuracy of 92.4% through multimodal fusion of EDA, PPG, and body temperature, surpassing the limitations of its single channel detection in accuracy and robustness (Zhu et al., 2023). Compared with the gender discrimination multi-level classification model based on EEG and ECG constructed by Hemakom et al. CNN+BiLSTM had more advantages in sequence modelling and feature fusion. Its time-dependent score in long sequence tasks was as high as 9.73, which better captured the dynamic evolution characteristics of psychological stress than its traditional LSTM strategy (Hemakom et al., 2023). In addition, compared to Salankar et al.'s automatic classification model that used empirical mode decomposition and variational mode decomposition to process EEG signals, this study adopted an end-to-end DL strategy to avoid complex signal pre-processing processes. Its training efficiency was higher, with a classification accuracy of 89.47% and an intervention matching degree of 8.92, making it more practical and valuable for engineering deployment (Salankar and Qaisar, 2023).

Given the high interpretability requirements of deep neural networks in mental health applications, this study further introduced Shapley additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME) to conduct explanatory analysis of the classification model's decision-making basis. SHAP calculates the marginal contribution of each physiological feature to the classification result, revealing the key physiological change patterns that the model focuses on under different stress levels, such as short-term EDA peak fluctuations, PPG rhythmic changes, and slight decreases in body temperature. LIME, on the other hand, explains individual predictions through local linear approximation, demonstrating the model's reasoning at the individual level. The results show that the temporal attention points of the neural network are highly consistent with traditional psychophysiological stress responses, further enhancing the transparency of model decisions and user trust, and supporting the interpretability requirements in mental health applications.

Although the multi-source physiological signal fusion model proposed in this study achieved good performance in identifying instantaneous psychological stress in students, it still has certain limitations. First, in actual deployment scenarios, the continuous

collection of sensitive physiological data such as EDA, PPG, and body temperature by wearable devices may involve privacy leakage risks. The research data all come from public datasets, but in future applications in real campus environments, it is necessary to strictly follow the principle of data minimisation, anonymisation processing specifications, and obtain informed consent from students before data collection, establishing a data management process that meets ethical review requirements to ensure the acceptability and compliance of the system. Secondly, the current model mainly focuses on classifying ‘instantaneous stress’ and does not incorporate time series prediction capabilities for the stress evolution process, thus failing to capture the continuous change trend of daily stress on an hourly-to-daily time scale. In the future, it can be further combined with sequence prediction models (such as temporal CNN, Seq2Seq, transformer, state-space models, etc.) to build a continuous stress index prediction framework, enabling the system to predict the risk of increased stress in advance, extending from ‘state identification’ to ‘trend prediction’ and ‘early intervention’.

6 Conclusions and recommendations

A student PSC method that integrated sensor data and NN was proposed to address the shortcomings of existing PSC methods in weak multi-source data fusion ability, low classification accuracy, and extensive intervention recommendation. The BKF mechanism was introduced to achieve temporal alignment and noise suppression of multimodal PSs. A NN model integrating 1D-CNN and BiLSTM was constructed to enhance local feature perception and temporal modelling capabilities, improve classification accuracy, and enhance the effectiveness of intervention recommendations. In the experiment, in terms of SDF, the redundancy of BKF algorithm was only 0.28 and the information entropy reached 9.32 bits, which was significantly better than SKF and the original signal processing method. In the comparison of feature fusion strategies, the combination of CNN+BiLSTM outperformed CNN or BiLSTM alone in classification accuracy (89.47%), t-SNE separation (8.62), and F1-score (0.88). The final classification and intervention recommendation module had a matching degree of 8.92, a decision coverage rate of 95.36%, and a response time of 89.76 ms, which was better than SVM and expert rule systems. The research method has achieved collaborative optimisation in PS fusion, spatiotemporal feature modelling, and closed-loop classification intervention, improving the accuracy and real-time recognition of students’ psychological stress, and providing intelligent solutions and technical support for personalised mental health intervention.

Declarations

The authors have no relevant financial or non-financial interests to disclose.

References

- Badr, Y., Tariq, U., Al-Shargie, F. and Mir, T. (2024) 'A review on evaluating mental stress by deep learning using EEG signals', *Neural Comput. Appl.*, May, Vol. 36, No. 21, pp.12629–12654, DOI: 10.1007/s00521-024-09809-5.
- Banerjee, J.S., Mahmud, M. and Brown, D. (2023) 'Heart rate variability-based mental stress detection: an explainable machine learning approach', *SN Comput. Sci.*, February, Vol. 4, No. 2, p.176, DOI: 10.1007/s42979-022-01605-z.
- Chen, H., Shi, H., Liu, X. and Wang, Y. (2023) 'SMG: a micro-gesture dataset towards spontaneous body gestures for emotional stress state analysis', *Int. J. Comput. Vis.*, February, Vol. 131, No. 6, pp.1346–1366, DOI: 10.1007/s11263-023-01761-6.
- Dogan, G. and Akbulut, F.P. (2023) 'Multi-modal fusion learning through biosignal, audio, and visual content for detection of mental stress', *Neural Comput. Appl.*, October, Vol. 35, No. 34, pp.24435–24454, DOI: 10.1007/s00521-023-09036-4.
- Hafeez, M.A. and Shakil, S. (2024) 'EEG-based stress identification and classification using deep learning', *Multimed. Tools Appl.*, October, Vol. 83, No. 14, pp.42703–42719, DOI: 10.1007/s11042-023-17111-0.
- Haque, Y., Zawad, R.S., Rony, C.S.A. and Hasan, M.M. (2024) 'State-of-the-art of stress prediction from heart rate variability using artificial intelligence', *Cogn. Comput.*, October, Vol. 16, No. 2, pp.455–481, DOI: 10.1007/s12559-023-10200-0.
- Hemakom, A., Atiwiwat, D. and Israsena, P. (2023) 'ECG and EEG based detection and multilevel classification of stress using machine learning for specified genders: a preliminary study', *PLoS One*, September, Vol. 18, No. 9, pp.e0291070–e0291071, DOI: 10.1371/journal.pone.0291070.
- Ilias, L., Mouzakitis, S. and Askounis, D. (2023) 'Calibration of transformer-based models for identifying stress and depression in social media', *IEEE Trans. Comput. Soc. Syst.*, April, Vol. 11, No. 2, pp.1979–1990, DOI: 10.1109/TCSS.2023.3283009.
- Jain, P., Yedukondalu, J., Chhabra, H. and Gupta, R. (2024) 'EEG-based detection of cognitive load using VMD and LightGBM classifier', *Int. J. Mach. Learn. Cybern.*, April, Vol. 15, No. 9, pp.4193–4210, DOI: 10.1007/s13042-024-02142-2.
- Jiao, Y., Sun, Z., Fu, L. and Chen, W. (2023) 'Physiological responses and stress levels of high-speed rail train drivers under various operating conditions – a simulator study in China', *Int. J. Rail Transp.*, June, Vol. 11, No. 4, pp.449–464, DOI: 10.1080/23248378.2022.2086638.
- Jiménez-Mijangos, L.P., Rodríguez-Arce, J., Martínez-Méndez, R. and García-Ceja, L. (2023) 'Advances and challenges in the detection of academic stress and anxiety in the classroom: a literature review and recommendations', *Educ. Inf. Technol.*, September, Vol. 28, No. 4, pp.3637–3666, DOI: 10.1007/s10639-022-11324-w.
- Kang, S., Choi, W., Park, C.Y. and Kim, J. (2023) 'K-emophone: a mobile and wearable dataset with in-situ emotion, stress, and attention labels', *Sci. Data*, June, Vol. 10, No. 1, pp.351–352, DOI: 10.1038/s41597-023-02248-2.
- Luo, N., Yu, H., You, Z., Li, Y., Zhou, T. and Han, N. (2023) 'Fuzzy logic and neural network-based risk assessment model for import and export enterprises: a review', *J. Data Sci. Intell. Syst.*, June, Vol. 1, No. 1, pp.2–11, DOI: 10.47852/bonviewJDSIS32021078.
- Masri, G., Al-Shargie, F., Tariq, U. and Mir, T. (2023) 'Mental stress assessment in the workplace: a review', *IEEE Trans. Affect. Comput.*, September, Vol. 15, No. 3, pp.958–976, DOI: 10.1109/TAFFC.2023.3312762.
- Mentis, A.F.A., Lee, D. and Roussos, P. (2024) 'Applications of artificial intelligence – machine learning for detection of stress: a critical overview', *Mol. Psychiatry*, April, Vol. 29, No. 6, pp.1882–1894, DOI: 10.1038/s41380-023-02047-6.
- Narzary, D., Sharma, U. and Khanna, A. (2025) 'An automated stress detection model based on dual approach of clinical psychologist prediction and machine learning', *Int. J. Inf. Technol.*, September, Vol. 17, No. 2, pp.755–765, DOI: 10.1007/s41870-024-02213-1.

- Pan, Y., Li, F., Liang, H. and Wang, Q. (2024) 'Effectiveness of mindfulness-based stress reduction on mental health and psychological quality of life among university students: a grade-assessed systematic review', *Evid.-Based Complement. Alternat. Med.*, February, Vol. 2024, No. 1, pp.8872685–8872686, DOI: 10.1155/2024/8872685.
- Prakash, P., Prabu, P., Sakthivel, V. and Kumar, R. (2024) '[RETRACTED] An efficient deep learning based stress monitoring model through wearable devices for health care applications', *Opt. Quantum Electron.*, December, Vol. 56, No. 2, p.200, DOI: 10.1007/s11082-023-05801-w.
- Rashid, N., Mortlock, T. and Al Faruque, M.A. (2023) 'Stress detection using context-aware sensor fusion from wearable devices', *IEEE Internet Things J.*, August, Vol. 10, No. 16, pp.14114–14127, DOI: 10.1109/JIOT.2023.3265768.
- Razavi, M., Ziyadidegan, S., Mahmoudzadeh, A. and Moradi, M.H. (2024) 'Machine learning, deep learning, and data preprocessing techniques for detecting, predicting, and monitoring stress and stress-related mental disorders: scoping review', *JMIR Ment. Health*, October, Vol. 11, No. 1, pp.e53714–e53715, DOI: 10.2196/53714.
- Richer, R., Koch, V. and Abel, L. (2024) 'Machine learning-based detection of acute psychosocial stress from body posture and movements', *Sci. Rep.*, April, Vol. 14, No. 1, pp.8251–8252, DOI: 10.1038/s41598-024-59043-1.
- Salankar, N. and Qaisar, S.M. (2023) 'EEG based stress classification by using difference plots of variational modes and machine learning', *J. Ambient Intell. Humaniz. Comput.*, April, Vol. 14, No. 12, pp.16347–16360, DOI: 10.1007/s12652-022-03856-3.
- Siam, A.I., Gamel, S.A. and Talaat, F.M. (2023) 'Automatic stress detection in car drivers based on non-invasive physiological signals using machine learning techniques', *Neural Comput. Appl.*, March, Vol. 35, No. 17, pp.12891–12904, DOI: 10.1007/s00521-023-08428-w.
- Singh, G., Phukan, O.C. and Gupta, R. (2024) 'Hybrid deep learning model for wearable sensor-based stress recognition for internet of medical things (IoMT) system', *Int. J. Commun. Syst.*, November, Vol. 37, No. 3, pp.5657–5658, DOI: 10.1002/dac.5657.
- Singh, G., Phukan, O.C. and Kumar, R. (2023) 'Stress recognition with multi-modal sensing using bootstrapped ensemble deep learning model', *Expert Syst.*, February, Vol. 40, No. 6, pp.e13239–e13240, DOI: 10.1111/exsy.13239.
- Talaat, F.M. and El-Balka, R.M. (2023) 'Stress monitoring using wearable sensors: IoT techniques in medical field', *Neural Comput. Appl.*, June, Vol. 35, No. 25, pp.18571–18584, DOI: 10.1007/s00521-023-08681-z.
- Xu, C., Song, Y., Sempionatto, J.R. and Gao, W. (2024) 'A physicochemical-sensing electronic skin for stress response monitoring', *Nat. Electron.*, January, Vol. 7, No. 2, pp.168–179, DOI: 10.1038/s41928-023-01116-6.
- Yang, Y., Xie, A., Kim, S. and Lee, J. (2023) 'Automatic detection of twitter users who express chronic stress experiences via supervised machine learning and natural language processing', *Comput. Inform. Nurs.*, September, Vol. 41, No. 9, pp.717–724, DOI: 10.1097/CIN.0000000000000985.
- Zhang, L., Zhang, T., Ren, Z. and Li, X. (2023a) 'Predicting compassion fatigue among psychological hotline counselors using machine learning techniques', *Curr. Psychol.*, April, Vol. 42, No. 5, pp.4169–4180, DOI: 10.1007/s12144-021-01776-7.
- Zhang, X., Wei, X., Zhou, Z. and Liu, Y. (2023b) 'Dynamic alignment and fusion of multimodal physiological patterns for stress recognition', *IEEE Trans. Affect. Comput.*, June, Vol. 15, No. 2, pp.685–696, DOI: 10.1109/TAFFC.2023.3290177.
- Zhu, L., Spachos, P., Ng, P.C. and Plataniotis, K.N. (2023) 'Stress detection through wrist-based electrodermal activity monitoring and machine learning', *IEEE J. Biomed. Health Inform.*, May, Vol. 27, No. 5, pp.2155–2165, DOI: 10.1109/JBHI.2023.3239305.