



OPEN ACCESS

EDITED BY

Alireza Mani,
University College London, United Kingdom

REVIEWED BY

Roham Mazloom,
Alborz University of Medical Sciences, Iran
Golnar Eftekhari,
University of Virginia, United States

*CORRESPONDENCE

Ling Yang,
✉ lyang@suda.edu.cn
Ying Xu,
✉ yingxu@suda.edu.cn

[†]These authors have contributed equally
to this work

RECEIVED 27 November 2024

ACCEPTED 31 March 2025

PUBLISHED 14 April 2025

CITATION

Shen Y, Fang Z, Zhang T, Yu F, Xu Y and Yang L
(2025) Heart rate variability with circadian
rhythm removed achieved high accuracy for
stress assessment across all times throughout
the day.
Front. Physiol. 16:1535331.
doi: 10.3389/fphys.2025.1535331

COPYRIGHT

© 2025 Shen, Fang, Zhang, Yu, Xu and Yang.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Heart rate variability with circadian rhythm removed achieved high accuracy for stress assessment across all times throughout the day

Yafei Shen^{1,2†}, Zihan Fang^{1,2†}, Tao Zhang³, Feng Yu^{1,2}, Ying Xu^{4*}
and Ling Yang^{1,2*}

¹School of Mathematical Sciences, Soochow University, Suzhou, Jiangsu, China, ²Center for Systems Biology, Soochow University, Suzhou, Jiangsu, China, ³Jiangsu Province Engineering Research Center of Molecular Target Therapy and Companion Diagnostics in Oncology, Suzhou Vocational Health College, Suzhou, Jiangsu, China, ⁴Jiangsu Key Laboratory of Neuropsychiatric Diseases and Cambridge-Su Genomic Resource Center, Suzhou Medical College of Soochow University, Suzhou, Jiangsu, China

Background: Assessing real-time stress in individuals to prevent the accumulation of stress is necessary due to the adverse effects of excessive psychological stress on health. Since both stress and circadian rhythms affect the excitability of the nervous system, the influence of circadian rhythms needs to be considered during stress assessment. Most studies train classifiers using physiological data collected during fixed short time periods, overlooking the assessment of stress levels at other times.

Methods: In this work, we propose a method for training a classifier capable of identifying stress and resting states throughout the day, based on 10 short-term heart rate variability (HRV) feature data obtained from morning, noon, and evening. To characterize the circadian rhythms of HRV features, heartbeat interval data were collected and analyzed from 50 volunteers over three consecutive days. The circadian rhythm trends in the HRV features were then removed using the Smoothness Priors Approach (SPA), and XGBoost models were trained to assess stress.

Results: The results show that all HRV features exhibit 12-h and 24-h circadian rhythms, and the circadian rhythm differences across different days for individuals are relatively small. Furthermore, training classifiers on detrended data can improve the overall accuracy of stress assessment across all time periods. Specifically, when combining data from different time periods as the training dataset, the accuracy of the classifier trained on detrended data increases by 13.67%.

Discussion: These findings indicate that using HRV features with circadian rhythm trends removed is an effective method for assessing stress at all times throughout the day.

KEYWORDS

stress assessment, heart rate variability, circadian rhythm, feature extraction, XGBoost

1 Introduction

Psychological stress refers to a mental state in which an individual feels pressured, tense, or uncomfortable in response to external stressors. This state can result from various challenges, threats, or changes that require the individual to adapt and respond (Folkman, 2013; Selye, 1958). While moderate psychological stress can be beneficial by enhancing motivation, increasing alertness, and improving focus, excessive psychological stress may have detrimental effects. It can contribute to cardiovascular diseases such as hypertension and heart disease, disrupt normal immune function, and cause hormonal imbalances (Cohen et al., 2007; Al-Shargie et al., 2017). Therefore, continuous monitoring of stress and timely intervention are key to managing both physical and mental health.

In recent years, the widespread use of smart wearable devices has facilitated the collection and analysis of physiological signals. An increasing number of researchers are focusing on leveraging these physiological signals to objectively detect psychological stress, enabling timely alerts regarding users' stress levels (Marques et al., 2010). Currently, the primary physiological signals utilized for stress detection include electromyographic signals, electroencephalographic signals, electrocardiographic signals, and electrodermal activity signals (Healey and Picard, 2005; Singh et al., 2013; Peng et al., 2012; Al-Shargie et al., 2016). Among these, heart rate variability (HRV), derived from ECG signals, has gained increasing utilization in the assessment of psychological stress in recent years, owing to its ease of measurement.

Existing research indicates that stress disrupts the homeostasis of sympathetic and parasympathetic nervous system activity, thereby altering the oscillations of the cardiac cycle (Pereira et al., 2017; Petrowski et al., 2010; Camm et al., 1996a; Mukherjee et al., 2011; Luque-Casado et al., 2016; Beauchaine and Thayer, 2015). HRV, calculated from the intervals between two consecutive R-wave peaks (RR intervals) on an ECG, can effectively reflect these oscillations (Camm et al., 1996a). HRV features encompass a set of statistical metrics that provide insights into heart activity across time domain, frequency domain, and non-linear. During periods of psychological stress, significant alterations in HRV features are observed (Acerbi et al., 2016; Hynynen et al., 2011; Cinaz et al., 2013; Taelman et al., 2011; Tharion et al., 2009; Blechert et al., 2006; McDuff et al., 2016; Bernardi et al., 2000; Lucini et al., 2002; Schubert et al., 2009; Melillo et al., 2011). HRV analysis can be conducted over different time intervals, such as 24 h (referred to as long-term HRV analysis), 5 min (referred to as short-term HRV analysis), or even shorter intervals (Camm et al., 1996b). Research findings indicate that both long-term and short-term HRV features reliably reflect stress-related changes in real-life situations (Castaldo et al., 2015; Hernando et al., 2016; Delaney and Brodie, 2000; Salai et al., 2016; Munla et al., 2015). In the research process, the choice between long-term and short-term HRV largely depends on the research question and the type of stress being studied (Malik et al., 1996). In response to short-term stress, the autonomic nervous system (ANS) is influenced, leading to significant changes in HRV (Castaldo et al., 2015), which subsequently return to normal levels. Given its ability to capture rapid fluctuations in ANS activity, short-term HRV has been shown to be particularly

suitable for studying short-term stress (Pereira et al., 2017; Cipresso et al., 2019).

Over the past few years, to better identify short-term stress and mitigate the health risks associated with stress accumulation, short-term HRV analysis (ranging from 5 to 60 s) has become increasingly utilized. The widespread adoption of wearable sensors in devices such as smartphones and smartwatches has further facilitated the convenience of short-term HRV analysis (Athavale and Krishnan, 2017; Pecchia et al., 2018). Castaldo et al. were the first to propose a rigorous methodology for assessing the effectiveness of ultra-short-term HRV features in detecting psychological stress. They demonstrated that HRV analysis can reliably and accurately detect psychological stress when transitioning from short-term (as a reference) to ultra-short-term intervals (Castaldo et al., 2019). Karthikeyan et al. extracted ultra-short-term (32-s) HRV features, including both time-domain and frequency-domain parameters. They used the Fast Fourier Transform (FFT) to select optimal features for model input, and, based on PNN and KNN algorithms, identified the best smoothing factor and K value, achieving high accuracy in classifying stress and normal states (Karthikeyan et al., 2013). Salahuddin et al. utilized the Stroop Color Word Test (SCWT) as a stress-inducing task and successfully collected data from 60 volunteers in both resting and stress states. Using statistical methods, they analyzed changes in various time-domain and frequency-domain features of ultra-short-term HRV under different conditions, providing a theoretical foundation for the effective monitoring of short-term stress (Salahuddin et al., 2007).

In addition to stress, the influence of circadian rhythms on sympathetic and parasympathetic nervous system activity has been increasingly reported in recent years (Niwa et al., 2011; Okada et al., 2013). Naturally, HRV and other physiological indicators have also been observed to follow circadian rhythms (Lee et al., 2023; Stangherlin et al., 2020). However, many existing studies either overlook the effects of these rhythms or mitigate them by selecting a fixed short time window for experimental assessments of stress based on HRV. For instance, Hemakom et al. developed a machine learning model to classify various stress levels using ECG and EEG data, with data collection limited to the hours between 10 a.m. and 12 p.m. to control for circadian influences (Hemakom et al., 2024). Similarly, da Estrela et al. studied low high-frequency HRV to determine whether they can predict stress-related sleep disturbances. To account for diurnal variations in HRV, all testing sessions were conducted in the morning (da Estrela et al., 2020). Despite these efforts, classifiers trained on physiological data collected during fixed short time periods often perform poorly when applied to stress assessment across other times of the day. Hayano et al. demonstrated that, due to the presence of both circadian and ultradian rhythms, applying short-term HRV analysis methods to long-term HRV data collected in free-living conditions can lead to inaccurate conclusions (Hayano and Yuda, 2021). Therefore, it is essential to mitigate the influence of circadian rhythms when evaluating stress based on HRV.

The analysis of the above literature motivated us to explore whether HRV data, with circadian rhythm trends removed, could be effectively used to train a classifier capable of accurately assessing stress levels across all time periods. In this study, we

employed detrended short-term HRV features to evaluate stress. To capture circadian rhythm trends, we collected RR interval data from 50 volunteers continuously over 3 days, during which stress was induced using the SCWT task (Karthikeyan et al., 2014; Al-Shargie et al., 2022; Stroop, 1992). We then applied a smoothing prior method to eliminate the circadian rhythm trend from the HRV feature data. Subsequently, we trained XGBoost classifiers using both raw and detrended HRV feature data collected during the SCWT test, confirming the effectiveness of the detrended HRV features for assessing stress throughout the day. The influence of age and participants' familiarity with the experiment on the final classification accuracy was also investigated. We proposed this approach primarily to enhance the accuracy of stress recognition, making it more applicable in real-life scenarios.

2 Materials and methods

Briefly, the specific process we established for developing a classifier to assess stress is as follows:

1. Data collection: RR interval data from 50 volunteers were collected continuously over 3 days using smartwatches, with stress induced through the SCWT task at 8:30 a.m., 2:00 p.m., and 10:30 p.m. each day.
2. Data preprocessing: Confidence ellipses were used to remove outliers from the collected RR interval data.
3. HRV feature extraction: HRV feature data for each individual was calculated throughout the day based on the denoised RR interval data. Subsequently, circadian rhythm trends in the HRV feature data were obtained and removed.
4. XGBoost classifier training: The detrended HRV feature data from the SCWT test sessions were integrated into a training set, and an XGBoost model was trained as a classifier to assess stress.

Figure 1 provides an overview of the experimental workflow. Below are the specific details of each experimental step.

2.1 Participants

During the data collection phase, fifty healthy volunteers (22 males and 28 females, aged 23–50) were recruited based on pre-screening questionnaires confirming good overall health, absence of psychological and cardiovascular diseases, and no color blindness or color weakness. These participants, comprising students and teachers, contributed heartbeat interval (RR interval) data for subsequent HRV analysis (as shown in Figure 2A). To comprehensively capture circadian rhythm information, RR interval data were collected continuously over a period of 3 days from each participant using a smartwatch. If a participant interrupted the data collection process on any given day, they were asked to reschedule and complete the data collection on a different day. The experimental procedure posed minimal impact on participants and did not interfere with their daily activities. Ethical approval for the study was obtained from the Ethics Committee of Soochow University (Approval No. SUDA20230828H01).

2.2 Experiment setup and task sequence

To collect data under stress conditions, stress was induced using the SCWT task (Karthikeyan et al., 2014; Al-Shargie et al., 2022; Stroop, 1992). To examine the effect of circadian rhythms on HRV, each participant completed the SCWT task 9 times over the 3 days, specifically at 8:30 a.m., 2:00 p.m., and 10:30 p.m. daily. Each SCWT session consisted of three phases (as shown in Figure 2B). Step 1 (resting phase): Participants were instructed to relax for 5 min prior to the task, either by closing their eyes or listening to soft music. Step 2 (stress phase): Participants then performed the SCWT on a computer, where a word describing a color was displayed with the font color differing from the word's meaning. Below the word were five colored blocks. The participants were required to click the block that matched the color described by the word. To ensure pressure conditions during the test, participants are required to complete at least 140 questions within 2 min, with no more than two errors. If these criteria are not met, an immediate retest at the same time point is administered. The total response time must not exceed 5 min; otherwise, the test session is considered invalid. Step 3 (resting phase): Following the test, participants were given time to relax by sitting quietly to alleviate any residual stress.

In the SCWT test, the incongruent relationship between the word meaning and its font color could easily mislead participants, making it challenging to select the correct color block. This required participants to maintain high levels of concentration to complete the word-color matching tasks quickly and accurately, effectively inducing psychological stress. Furthermore, resting phases were incorporated before and after the SCWT test to minimize the influence of confounding factors, such as physical exertion, on the experimental outcomes.

2.3 Data acquisition

To successfully obtain the collected data, RR interval data from all volunteers were continuously recorded over a 3-day period using a smartwatch (Huawei Watch GT 2, green light, reflection pattern). The sensors used in this experimental equipment could measure ECG to capture RR interval data, with precision down to the millisecond. Besides, collecting ECG data using a smartwatch is very convenient and has minimal impact on participants' daily lives. Around midnight each day, the collected data from the previous day was uploaded to the database via Bluetooth. Finally, we downloaded all the experimental data from the database for subsequent analysis.

2.4 Pre-processing

After the data collection was completed, we accounted for the potential impact of daily activities on the collection of RR interval data—such as intense wrist movements during activities like running, which can cause the smartwatch to produce unstable measurements, leading to missed or extra heartbeat values and, in turn, excessively large or small RR intervals—we converted the collected one-dimensional RR interval data into two-dimensional data. Specifically, the first dimension represents the previous RR interval value, RR_{t-1} , and the second dimension

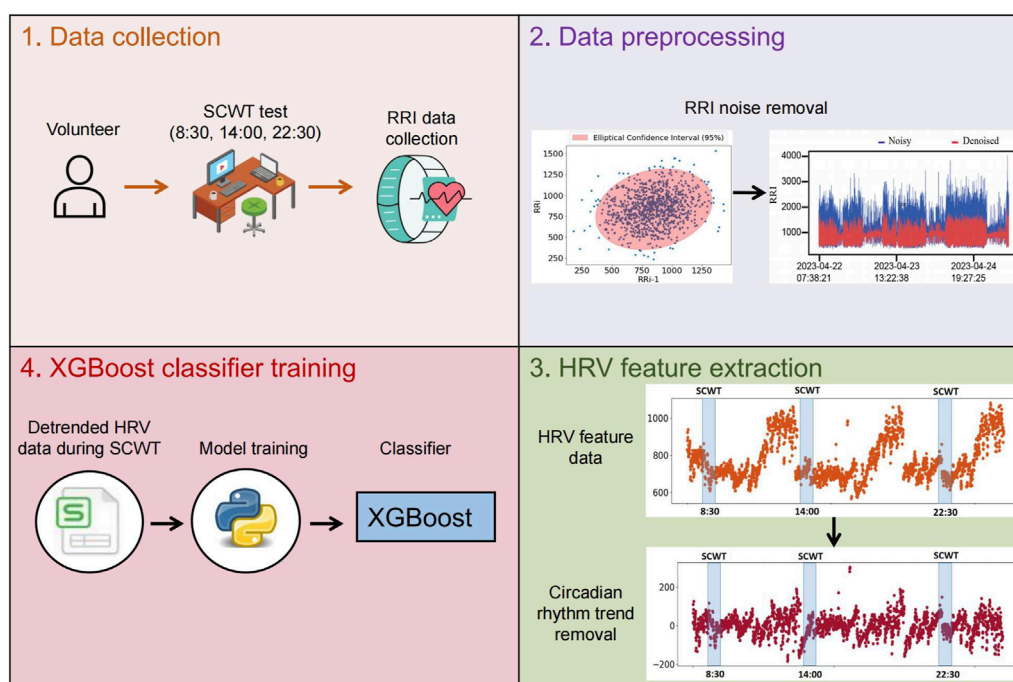


FIGURE 1

The workflow for obtaining the classifier that identify stress across all times throughout the day. The XGBoost model is used as a classifier and trained on detrended HRV feature data from the SCWT experiment.

represents the subsequent adjacent RR interval value, RR_{i+1} , where $i = 1, 2, \dots, N$. The final two-dimensional dataset obtained was $\{(RR_1, RR_2), (RR_2, RR_3), \dots, (RR_{N-1}, RR_N)\}$, where N is the total number of RR intervals. Since the distribution of these two-dimensional data points was approximately elliptical, we applied confidence ellipses to remove noise from the data.

Generally, the mathematical expression for an ellipse is:

$$\left(\frac{x - R_{1p}}{a}\right)^2 + \left(\frac{y - R_{2p}}{b}\right)^2 = 1,$$

where (R_{1p}, R_{2p}) is the center of the ellipse. The covariance matrix of the transformed two-dimensional RR interval data is:

$$C = \begin{bmatrix} \text{COV}(R_1, R_1) & \text{COV}(R_1, R_2) \\ \text{COV}(R_2, R_1) & \text{COV}(R_2, R_2) \end{bmatrix},$$

where $\text{COV}(R_i, R_j)$ represents the covariance between the variables R_i and R_j . Then, the elliptical equation that measures the error of each data point can be derived.

$$\frac{(R_1 - R_{1p})^2}{\lambda_1} + \frac{(R_2 - R_{2p})^2}{\lambda_2} = s, \quad (1)$$

λ_1 and λ_2 are the largest and smallest eigenvalues of the covariance matrix C , respectively, and s is the scale of the ellipse. Additionally, based on the chi-square distribution, we can obtain $P(s < 5.991) = 0.95$. Therefore, we set the value of s to be 5.991 to obtain a 95% confidence ellipse. According to Equation 1, the long axis of the confidence ellipse is length $2\sqrt{5.991\lambda_1}$, and the short axis is length $2\sqrt{5.991\lambda_2}$. If we denote the eigenvector corresponding to the largest

eigenvalue λ_1 of the covariance matrix as v_1 , then the angle between the major axis of the ellipse and the positive direction of the x-axis is:

$$\alpha = \arctan \frac{v_1(y)}{v_1(x)}.$$

$v_1(y)$ and $v_1(x)$ respectively represent the magnitudes of the projections of v_1 onto the positive y-axis and the positive x-axis.

Finally, based on the center of the ellipse (R_{1p}, R_{2p}) , the eigenvalues λ_1 and λ_2 of the data covariance matrix, the scale of the ellipse s , and the angle α between the major axis and the x-axis, we can establish a confidence ellipse to filter out noise from the data.

We randomly selected 50 data points from the collected RR interval dataset to illustrate the results of noise removal using confidence ellipses. As shown in Figure 3, the two-dimensional RR interval data within the confidence ellipse indicate that the differences between adjacent RR intervals are minimal. In contrast, data points outside the confidence ellipse exhibit significant differences between adjacent RR intervals and are therefore considered noise to be removed.

In this work, we employed a sliding window confidence ellipse method to remove noise from each volunteer's RR interval data. The window length was set to 50 data points, with an overlap of 25 data points between consecutive windows. Specifically, the first 50 data points of each volunteer's RR interval data were used to form the initial window. After applying the confidence ellipse to remove noise, the window was shifted back by 25 points to create the second window, where the confidence ellipse was established again for noise removal. This procedure was repeated until the window encompassed all data points. Figure 4 illustrates the results of the noise removal process applied to one volunteer's RR interval data.

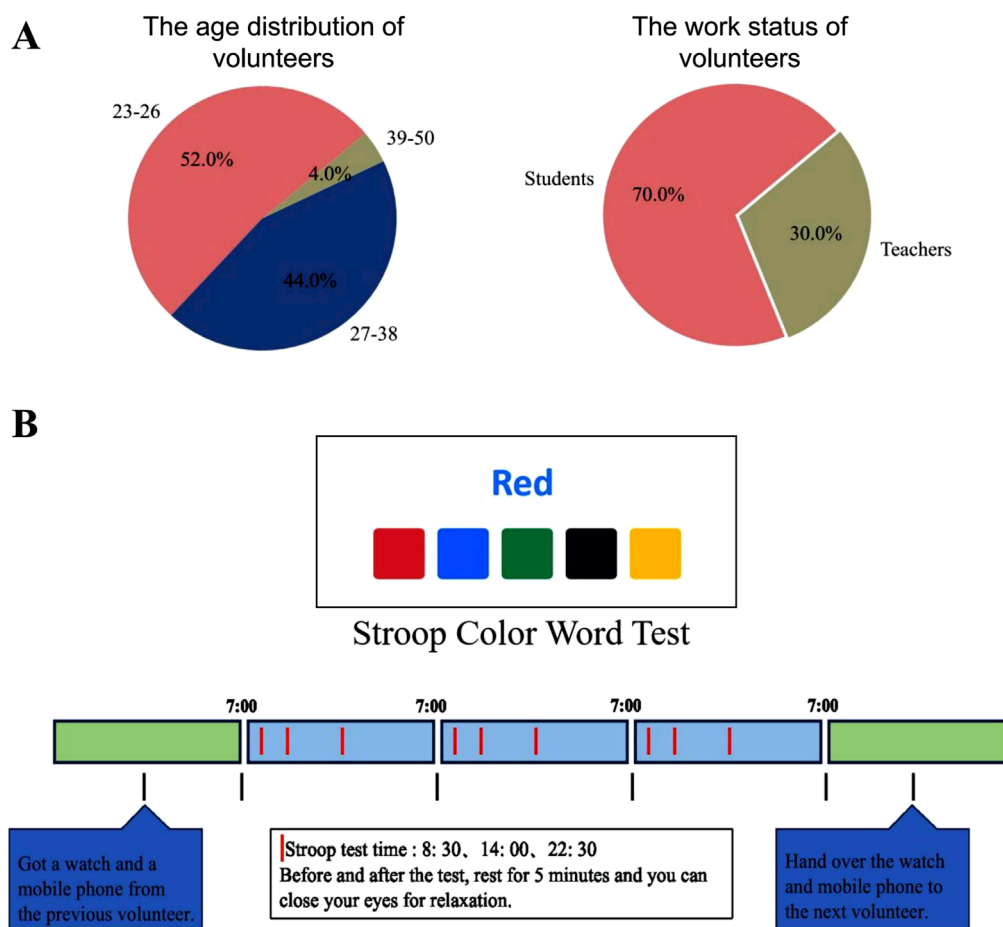


FIGURE 2

Details of data collection. (A) The age distribution and work status of volunteers participating in the SCWT experiment. (B) The process of the SCWT task. The RR interval data for each volunteer was collected over three consecutive days (marked by blue blocks). A Stroop Color Word Test was conducted at 8:30 a.m., 2:00 p.m., and 10:30 p.m. each day to induce stress (indicated by red vertical lines within the blue blocks).

The blue line represents the distribution of the original RR intervals, which exhibits significant fluctuations, while the red line indicates the distribution of the RR intervals following noise removal, showing a much smoother profile. Additionally, we quantified the dispersion of the RR interval data in the dataset by calculating the coefficient of variation (CV), using the following formula:

$$CV = \frac{\sigma}{\mu},$$

where σ is the standard deviation and μ is the mean of the RR interval data. Figure 5 displays the CV of the original RR interval data across all volunteers (average CV = 0.28) and the CV of the RR interval data following noise removal (average CV = 0.2). These results indicate that the sliding window confidence ellipse method effectively reduced abnormal fluctuations in the RR intervals.

2.5 Feature extraction

To evaluate HRV based on the collected RR interval data, we introduced five statistical metrics to extract the time-domain features of the RR interval data. The mean (*meanNN*) was used to

reflect the average level of the RR intervals. The calculation formula is as follows:

$$meanNN = \bar{RR} = \sum_{i=1}^N \frac{RR_i}{N},$$

where N is the number of RR intervals within a 2-min period, $\bar{RR}$ is the average value of the RR intervals within a 2-min period and RR_i represents the i^{th} RR interval.

The average heart rate (*HR*) was used to reflect the average level of heart rate.

$$HR = \frac{\sum_{i=1}^{N_1} \frac{N_1}{RR_i} + \sum_{i=1}^{N_2} \frac{N_2}{RR_i}}{2},$$

N_1 and N_2 represent the number of RR intervals in the first and second minutes, respectively.

The overall standard deviation (*SDNN*) was introduced to reflect the overall variation in HRV:

$$SDNN = \sqrt{\frac{1}{N} \sum_{i=1}^N (RR_i - \bar{RR})^2},$$

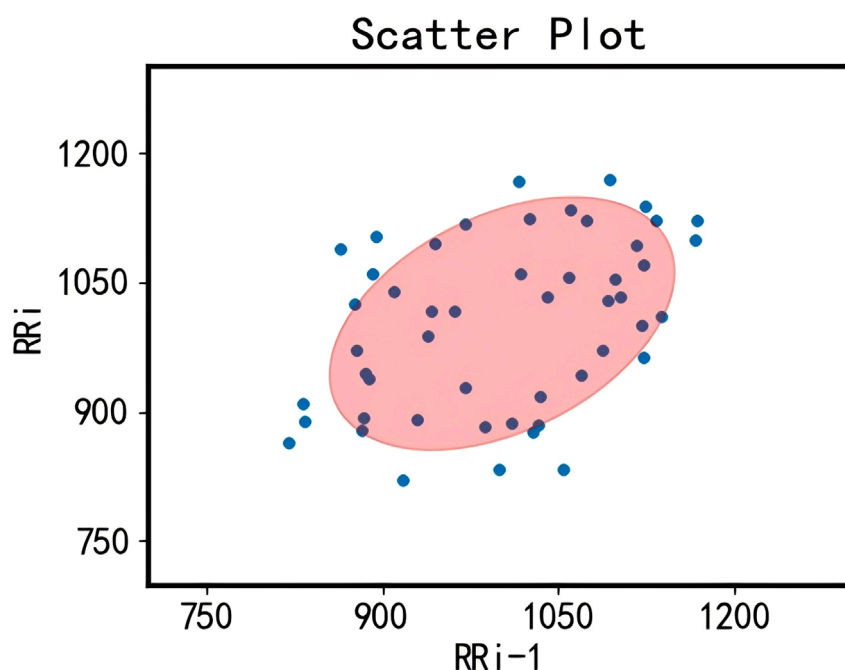


FIGURE 3

Diagram illustrating the removal of outliers in RR intervals using confidence ellipses. The blue points represent the distribution of the transformed two-dimensional RR interval data, where the values of each point's two dimensions correspond to the lengths of two consecutive RR intervals. The red area indicates the 95% confidence ellipse interval.

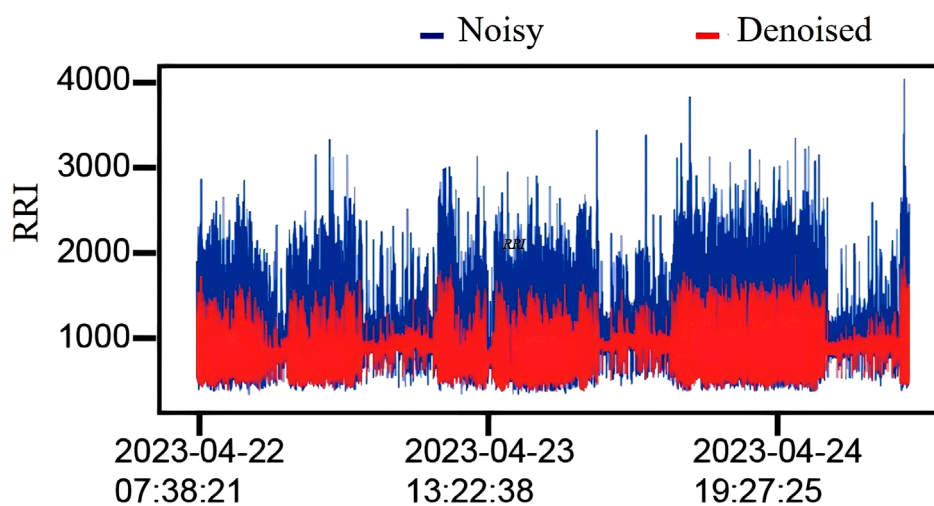


FIGURE 4

Effect of outlier removal. Comparison of 72-h RR interval data (RRI) before and after outlier removal for a volunteer. The blue part represents the data distribution before outlier removal, and the red part represents the data distribution after outlier removal.

and the root mean square of successive differences (RMSSD) was used to estimate the fast component variations in HRV:

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_{i+1} - RR_i)^2}.$$

Additionally, we used the proportion of the number of successive RR interval differences greater than 50 milliseconds to the total

number of RR intervals ($PNN50$) to reflect sudden changes in RR intervals. The calculation formula is as follows:

$$PNN50 = \frac{NN50}{TotalNN},$$

where $NN50$ represents the number of successive RR interval differences greater than 50 milliseconds, and $TotalNN$ represents the total number of RR intervals within a 2-min period.

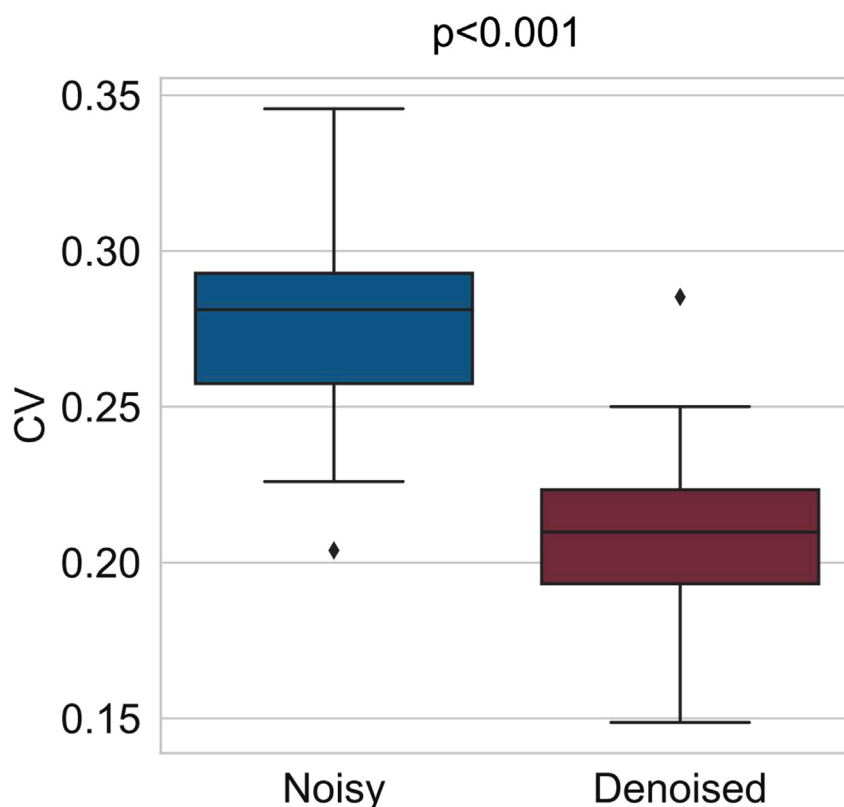


FIGURE 5

Comparison of CV coefficients before and after outlier removal for the RR interval data of all volunteers. After outlier removal, the average CV coefficient decreased from 0.28 to 0.2.

To perform frequency domain analysis of HRV based on the RR interval data, we first estimated the power spectral density ($PSD(f)$) using an autoregressive (AR) model. Subsequently, we integrated the power spectrum over different frequency bands to extract three frequency domain features. The details of the power spectral density estimation are provided in [Supplementary Material](#). Based on the frequency spectral density, we extracted the very low frequency power (vlf).

$$vlf = \int_{0.0033}^{0.04} PSD(f) df,$$

low frequency power (lf)

$$lf = \int_{0.04}^{0.15} PSD(f) df,$$

and high frequency power (hf)

$$hf = \int_{0.15}^{0.5} PSD(f) df.$$

To further analyze HRV from a nonlinear perspective, we extracted two parameters based on the Poincare plot of the two-dimensional RR interval data ([Brennan et al., 2001](#)). The first parameter is the length of the minor axis ($ST1$) of the scatter plot distribution area within a specified time period, which is associated with vagal nerve activity. The second parameter is the length of the major axis ($ST2$) of the scatter plot distribution area within the same

time period, which is indicative of changes in sympathetic nervous system activity. Their calculation formulas are as follows:

$$ST1 = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_i + RR_{i-1} - 2\bar{RR})^2},$$

$$ST2 = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_i - RR_{i-1})^2},$$

2.6 Detrending of data

Since both stress and circadian rhythms influence the activity of the sympathetic and parasympathetic nervous systems, which in turn alter heart rate oscillations, HRV, as a reflection of these oscillations, is naturally related to both stress and circadian rhythms. To better observe the contribution of stress to HRV changes, we considered removing the circadian rhythm trends from each volunteer's HRV feature data to improve the accuracy of stress detection. The Smoothness Priors Approach (SPA) is a nonlinear detrending technique for signals ([Karjalainen, 1997](#)). This method assumes that the original data signal, i.e., the time series X , is composed of two parts:

$$X = X_s + X_t,$$

where X_s is the stationary component and X_t is the nonlinear periodic trend component, which can be expressed as:

$$X_t = H\theta + v,$$

where $H \in \mathbb{R}^{N \times M}$ is the observation matrix, $\theta \in \mathbb{R}^M$ is the regression parameter, v is the observation error. We estimated the parameter θ such that $\hat{X}_t = H\hat{\theta}$ to estimate the trend component in the original data. The estimating of θ was commonly done using the method of least squares. The SPA introduced a differential term $\|D_d(H\theta)\|$ in the process of finding the optimal solution, minimizing it to ensure that θ filtered out the trend component of the data.

$$\theta_\lambda = \arg \min_{\theta} \{\|H\theta - X\|^2 + \lambda^2 \|D_d(H\theta)\|^2\}, \quad (2)$$

λ represents the regularization parameter, and D_d is the matrix of the discretized d^{th} order differential operator. If the sequence X had N extreme points, represented as a column vector $R = [R_1, R_2, \dots, R_N]^T \in \mathbb{R}^N$, then the first-order trend of R was given by $R_1 = [R_2 - R_1, R_3 - R_2, \dots, R_N - R_{N-1}] \in \mathbb{R}^{N-1}$, and the second-order trend of R could be expressed as $R_2 = [R_3 - R_2 - (R_2 - R_1), R_4 - R_3 - (R_3 - R_2), \dots, R_N - R_{N-1} - (R_{N-1} - R_{N-2})] \in \mathbb{R}^{N-2}$, by analogy, we could obtain the discrete representation of the trend component R of any order. Then, D_d was calculated as follows:

$$D_d = \begin{bmatrix} \frac{d(R_d)_1}{dR_1} & \frac{d(R_d)_1}{dR_2} & \dots & \frac{d(R_d)_1}{dR_N} \\ \frac{d(R_d)_2}{dR_1} & \frac{d(R_d)_2}{dR_2} & \dots & \frac{d(R_d)_2}{dR_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d(R_d)_{N-d}}{dR_1} & \frac{d(R_d)_{N-d}}{dR_2} & \dots & \frac{d(R_d)_{N-d}}{dR_N} \end{bmatrix}.$$

Furthermore, we could obtain the solution to Equation 2 as follows:

$$\hat{\theta}_\lambda = (H^T H + \lambda^2 H^T D_d^T D_d H)^{-1} H^T X,$$

$$\hat{X}_t = H\hat{\theta}_\lambda,$$

$\hat{X}_t$ is the estimated value of the trend component. For simplicity, let H be the identity matrix I and the order of D_d be 2. Ultimately, we could obtain the part of the original sequence without the trend component:

$$\hat{X}_s = X - H\hat{\theta}_\lambda = [I - (I + \lambda^2 D_2^T D_2)^{-1}] X = LX.$$

The role of L is akin to a high-pass filter, which filters out the low-frequency components of the sequence. The parameter λ determines the frequency of the filtered signal, i.e., the frequency response. To adequately filter out the rhythmic fluctuations in HRV features, we chose the parameter $\lambda = 10000$, corresponding to a cutoff frequency of 0.0012, which was approximately a 24-h period.

2.7 State classification

Using the detrended HRV feature data obtained during the SCWT task as the training set, the next step involves

establishing a classifier to assess stress. Extreme Gradient Boosting (XGBoost) is a classification model known for its superior accuracy compared to traditional classifiers and its ability to handle diverse types of data (Chen and Guestrin, 2016). In this study, we utilized XGBoost as a classifier to estimate stress states based on the extracted HRV features. The training process for the XGBoost model is outlined as follows:

First, we input the training dataset $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, $x_i = (x_{i1}, x_{i2}, \dots, x_{im}) \in X$, $y_i \in Y$ into XGBoost. x_{ij} represents the j^{th} feature of the i^{th} sample, and y_i represents the label of the i^{th} sample, $i = 1, 2, \dots, n$. Then, we can build the classification model.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F}$$

$$\mathcal{F} = \{f(x) = \omega_q(x)\} (q: \mathbb{R}^K \rightarrow T, \omega \in \mathbb{R}^t).$$

where K represents the number of trees, and f is a function from the function space $\mathcal{F}$. We use the squared error as the loss function, specifically expressed as

$$\text{Obj} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k),$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2,$$

where l is the loss function, ω is the penalty term, T is the number of leaf nodes, and γ and λ are regularization parameters used to control the number of leaf nodes and the output scores of the leaf nodes, respectively. Since the prediction results of the trees generated in each iteration during the training process are fitted to the residuals of the prediction results from the previous iteration, the model at iteration t can be expressed as

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i).$$

Besides, the corresponding objective function can be expressed as

$$\text{Obj}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t)}) + \sum_{k=1}^t \Omega(f_k)$$

$$= \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \sum_{k=1}^{t-1} \Omega(f_k) + \Omega(f_t),$$

if we let $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$, $h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$, then the objective function can be approximated as

$$\text{Obj}^{(t)} = \sum_{i=1}^n \left(l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \sum_{k=1}^{t-1} \Omega(f_k) + \Omega(f_t).$$

Since the training for the first $t-1$ iterations has been completed, $l(y_i, \hat{y}_i^{(t-1)})$ and $\sum_{k=1}^{t-1} \Omega(f_k)$ are constants, and the objective function can be further simplified as

$$\text{Obj}^{(t)} = \sum_{i=1}^n \left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \Omega(f_t)$$

$$= \sum_{i=1}^n \left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2$$

$$= \sum_{j=1}^T \left(\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right) + \gamma T,$$

where ω_j represents the value of the j^{th} leaf node. Let $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$, then G_j and H_j represent the sum of first derivative and second derivative of the loss function of all samples falling on the j^{th} leaf node, respectively. Then the objective function can be further simplified as

$$Obj^{(t)} = \sum_{j=1}^T \left(G_j \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2 \right) + \gamma T. \quad (3)$$

Therefore, the optimal value of ω_j is $\omega_j = -\frac{G_j}{H_j + \lambda}$, and substituting in Equation 3 gives

$$Obj^{(t)} = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T.$$

In general, when building a regression tree, the best partition point of the tree is selected based on the following gain:

$$\begin{aligned} \text{Gain} &= Obj_{L+R} - (Obj_L + Obj_R) \\ &= \left(-\frac{1}{2} \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} + \gamma T \right) \\ &\quad - \left(-\frac{1}{2} \left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} \right) + \gamma(T+1) \right) \\ &= \frac{1}{2} \left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right) - \gamma. \end{aligned} \quad (4)$$

3 Results

3.1 HRV feature analysis

Based on the RR interval data collected from 50 volunteers, we calculated the values of 10 HRV features for each participant during both the rest and stress phases across three SCWT task stages each day. Our findings indicated that the distribution of all features is influenced not only by stress but also by circadian rhythms. For instance, Figure 6 illustrates the distribution of *meanNN* and *HR* for the 50 volunteers during the rest and stress phases at different times of the day. Both *meanNN* and *HR* exhibited variations in their median values between stress and rest states over time, with similar trends in these changes. Furthermore, significant differences were observed in the distributions of *meanNN* and *HR* between the rest and stress states at the same time periods ($p < 0.05$). Notably, around 8:30 and 22:30, the distributions of *meanNN* and *HR* during the stress state overlapped with those during the rest state observed at approximately 14:00.

This finding indicates that *meanNN* and *HR* are sensitive indicators of stress. Additionally, when assessing stress based on these HRV features, it is essential to further mitigate the influence of circadian rhythms. Moreover, the effect of circadian rhythms on the data distribution was also evident for the remaining eight features (Supplementary Figure S1).

3.2 Circadian rhythm analysis

Given that circadian rhythms influence all HRV features, we conducted a detailed analysis of the specific periods associated

with each feature, verifying the presence of circadian rhythms through sinusoidal fitting. Using the 3-day RR interval data from all volunteers, we extracted the values of 10 HRV features every 2 min. Subsequently, we applied FFT to determine the periods and amplitudes of all features (Cooley and Tukey, 1965). We then calculated the mean values of the 10 features for each period across all volunteers, allowing us to perform interval estimation of the overall mean. The amplitude of each feature exhibited significant peaks around 360 min and 720 min, with the 95% interval estimates aligning closely with the trends in the mean values (as shown in Figure 7). These results indicate that each feature demonstrated periodicity, consistently present within the population. Since feature values were extracted at 2-min intervals, the actual corresponding periods corresponded to 12 h and 24 h. This suggests substantial 24-h fluctuations in each feature, alongside potential 12-h fluctuations during daytime hours.

Furthermore, to quantify the differences in circadian rhythms both between individuals and within the same individual across different days, we employed a second-order sine function to fit the time series of each feature for all volunteers collectively.

$$y = A_0 + A_1 \sin(\omega x) + B_1 \cos(\omega x) + A_2 \sin(2\omega x) + B_2 \cos(2\omega x), \quad (5)$$

where A_0 is a constant, A_1 , B_1 , A_2 and B_2 are the amplitudes, and ω is the frequency. These parameters described the diurnal dynamic variations of HRV features, which were fitted using MATLAB according to Equation 5 to obtain the results. As illustrated in Figure 8, the fitting results for *meanNN*, *HR*, *PNN50* and *ST1* correspond closely with the variation trends of the original data over time. The x-axis denotes time, with x values in the intervals 1–501, 735–1,242, and 1,472–1,978 corresponding to the time range of 7:30 a.m. to 12:00 a.m., while the remaining intervals represent the hours from 12:00 a.m. to 7:30 a.m. These features exhibit significant diurnal rhythms. For instance, *meanNN* initially decreases, then stabilizes, and finally increases throughout the day, exhibiting an opposite trend at night. Notably, we also observed that the feature values at night were either significantly higher or significantly lower than those during the day. This results in substantial differences between the distributions of rest and stress state data collected during the day compared to those collected at night. If HRV feature data from different times are input into the same classifier, it may not accurately assess the stress state. For the remaining six features, we similarly applied a second-order sine function for fitting and also observed clear rhythmicity.

Additionally, we fitted the 10 feature data for each volunteer on each day separately, obtaining the corresponding fitted parameter values for each individual on each day (A_1 , B_1 , A_2 and B_2). To compare the variations in HRV within an individual over 3 days and across all the participants, we then calculated the mean values of each parameter across all days for the 50 volunteers (*mean*), as well as the mean standard deviation of each parameter for each individual over the 3 days (std_1), and the overall standard deviation of each parameter across all days (std_2) (as shown in Table 1). By comparing std_1 and std_2 , we found that the differences in rhythmic parameters within the same volunteer across different days were relatively small.

These results indicate that the 10 HRV features exhibit a distinct circadian rhythm trend, and the specific rhythm-related parameters obtained can be utilized for the subsequent removal of these

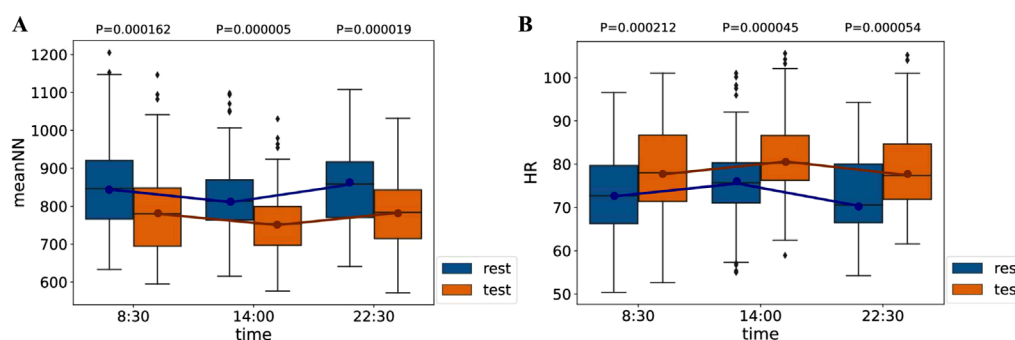


FIGURE 6 Distribution of *meanNN* and *HR* during the SCWT tasks at three time points each day: **(A)** *meanNN*, **(B)** *HR*. The blue box plots represent the data distribution for the resting state, with blue dots indicating the median. The orange box plots represent the data distribution for the stress state, with orange stars indicating the median.

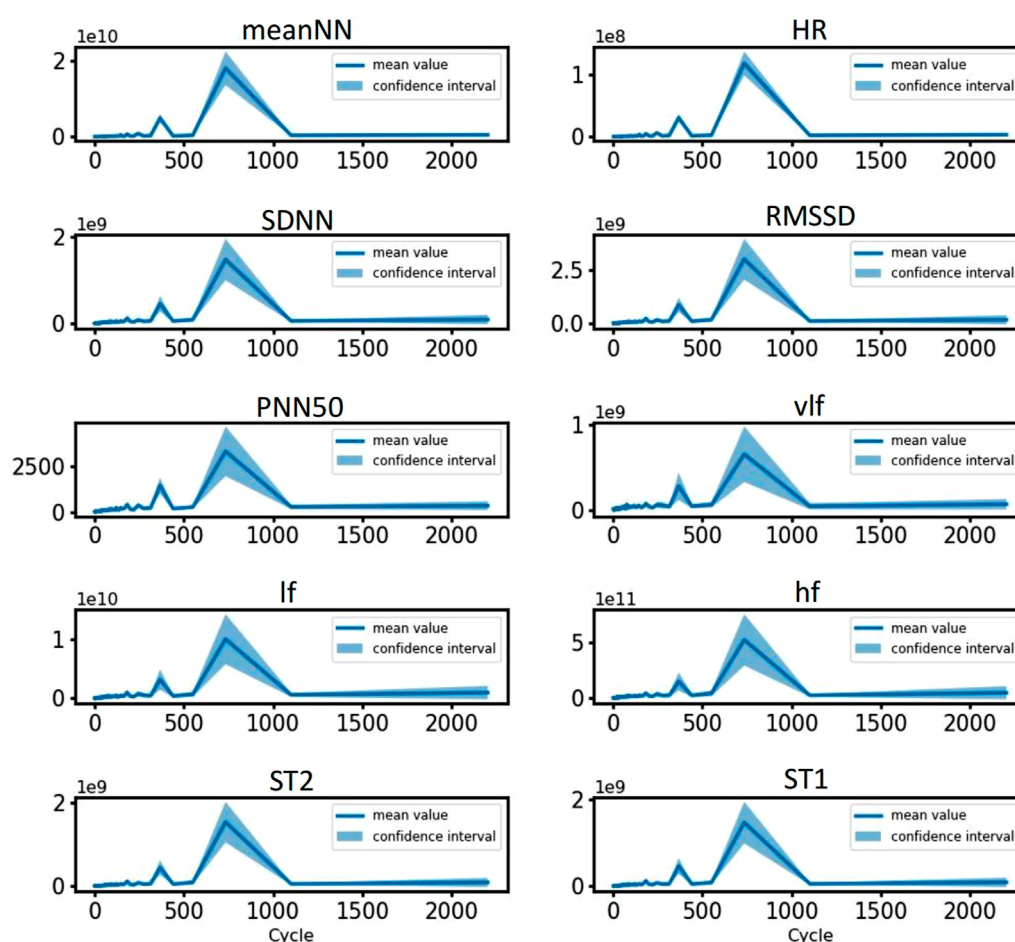
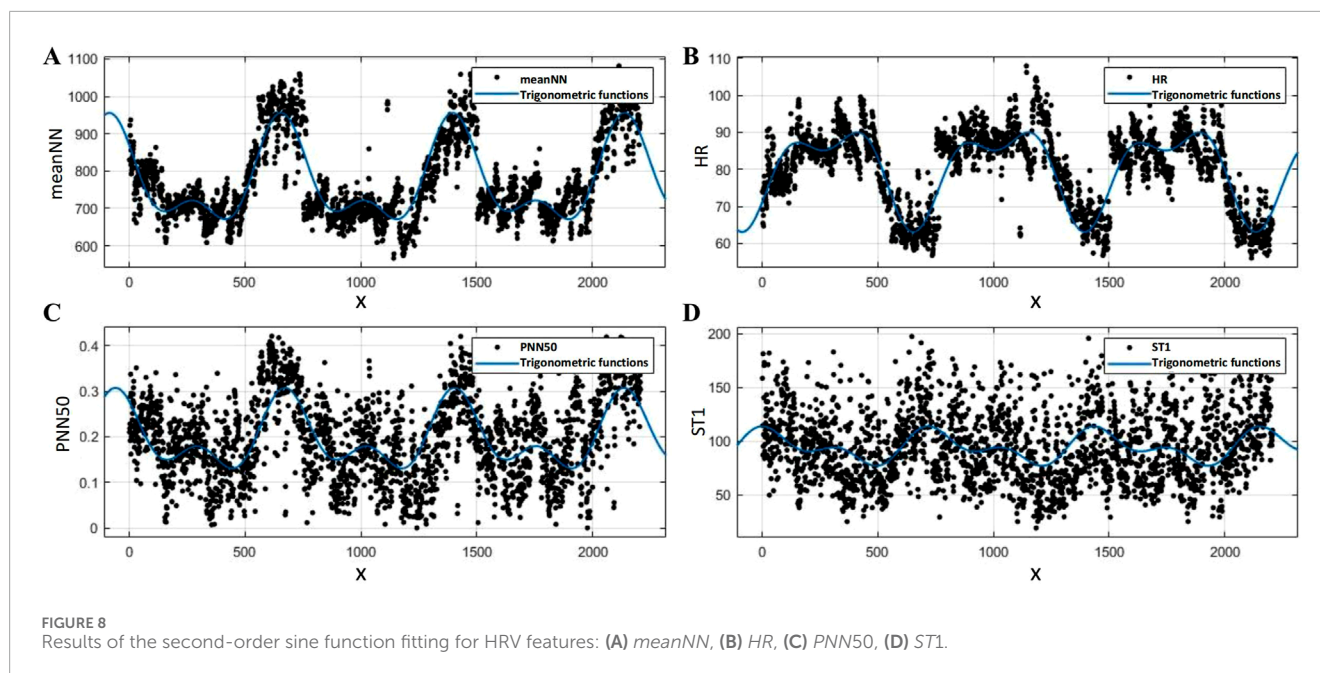


FIGURE 7 Period spectrum plots of HRV features for all volunteers. The blue line represents the average amplitude, and the shaded area corresponds to the interval estimate.

trends. Moreover, the rhythmic patterns of HRV parameters exhibit minimal fluctuations within individuals over three consecutive days, reflecting relative intra-individual stability. Consequently, data from the previous day can be utilized as a reference for detrending

circadian rhythm trends when assessing stress throughout the day. However, considerable inter-individual variability is observed, suggesting that the data collected from all volunteers encompass a certain degree of diversity.



3.3 Detrended HRV feature analysis

Considering the influence of circadian rhythms on HRV features in stress assessment, we further investigated the effects of removing these trends from the HRV feature data. The distribution of the 10 HRV feature values for all volunteers during the rest and stress phases of the SCWT experiment is illustrated in Figure 9. Each panel presents two graphs: the left graph displays the value distribution of each feature during the rest and stress phases before detrending, while the right graph shows the distribution after detrending. Notably, distinct differences in the distribution of each feature between the rest and stress phases are observed both before and after the detrending process ($p < 0.05$). To further investigate, we examine whether detrending enhances these differences, making the separation between the two states more pronounced.

First, we quantified the differences in value distributions of each feature between groups and within groups during the rest and stress phases using the following equations:

$$SSA = n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2,$$

$$SSE = \sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2 + \sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2,$$

where n_1 and n_2 represent the sample number of the feature data during the rest and stress phases, respectively, $\bar{x}_1$ and $\bar{x}_2$ represent the means of the feature data during the rest and stress phases, respectively, and $\bar{x}$ is the overall mean of the features. Additionally, the degrees of freedom for SSA and SSE are 1 and $n_1 + n_2 - 2$, respectively. Then, we can calculate the F-statistic for each feature using the following formula:

$$F = \frac{SSA}{SSE/(n_1 + n_2 - 2)}.$$

The larger the F value, the greater the difference in data distribution between the rest and stress phases for the corresponding feature.

The calculated F -statistic results for each feature before and after detrending are presented in Table 2. All features exhibited significantly increased F values following detrending, indicating that the distribution differences between the rest and stress phases became more pronounced.

In summary, the influence of circadian rhythms leads to significant overlap in the HRV feature distributions for stress and rest states throughout the day. By removing the circadian rhythm trend, the differences in HRV features between stress and rest states are enhanced.

3.4 Stress assessment using XGBoost

To investigate the impact of detrended HRV feature data obtained during the SCWT task on stress assessment throughout the entire time period, we trained XGBoost models as classifiers. We determined the optimal hyperparameters of the model through a grid search approach. Specifically, we adjusted each parameter individually within a predefined range and utilized these parameters to train the model. In each training session, 20% of the training dataset was reserved as the test dataset, with the F1 score serving as the evaluation criterion for model performance. Ultimately, we identified the parameter settings that yielded the highest accuracy on the test dataset as the optimal values for each hyperparameter.

Additionally, XGBoost, as a gradient boosting tree model, has the capability to output feature importance. Specifically, for each node of each decision tree, XGBoost calculates the gain resulting from splitting the node, which reflects the extent to which the split enhances model performance (Equation 4). The importance score for each feature is then obtained by weighting and summing the results across all decision trees, followed by averaging these scores (as shown in Figure 10). To enhance the model's accuracy, we ultimately selected the top seven important features for classification.

TABLE 1 Parameter results of the second-order sinusoidal fitting for HRV feature data of all volunteers. For each fitted parameter, the first row represents the mean value of the parameters obtained by fitting 3 days of data together for each volunteer, resulting in the mean value for 50 volunteers. The second row represents the mean of the standard deviations of the fitting parameters for each volunteer over 3 days. The third row represents the standard deviation of the 50 parameters obtained by fitting 3 days of data together for each volunteer.

Parameter	Statistic	<i>meanNN</i>	<i>HR</i>	<i>SDNN</i>	<i>RMSSD</i>	<i>PNN50</i>
<i>A</i> ₁	<i>mean</i>	77.60	−7.20	−18.72	−32.94	−0.08
	<i>std</i> ₁	11.08	0.99	4.77	10.02	0.09
	<i>std</i> ₂	41.98	4.09	19.67	28.69	0.22
<i>B</i> ₁	<i>mean</i>	−61.04	4.68	11.99	16.53	−0.13
	<i>std</i> ₁	10.80	1.12	3.378	13.45	0.18
	<i>std</i> ₂	35.18	2.28	21.08	22.43	0.47
<i>A</i> ₂	<i>mean</i>	2.28	−0.43	−1.21	−4.33	−0.03
	<i>std</i> ₁	1.16	1.23	0.69	2.33	0.05
	<i>std</i> ₂	27.12	2.04	13.74	18.56	0.12
<i>B</i> ₂	<i>mean</i>	−37.29	3.40	11.96	13.34	0.06
	<i>std</i> ₁	11.17	1.55	3.25	3.54	0.07
	<i>std</i> ₂	22.41	2.09	17.60	24.69	0.16
Parameter	Statistic	<i>vlf</i>	<i>lf</i>	<i>hf</i>	<i>ST2</i>	<i>ST1</i>
<i>A</i> ₁	<i>mean</i>	−9.44	−82.62	−551.8	−23.39	−18.42
	<i>std</i> ₁	2.28	19.51	132.1	3.51	1.00
	<i>std</i> ₂	16.13	97.15	341.34	20.39	22.85
<i>B</i> ₁	<i>mean</i>	10.08	52.62	317.2	11.74	14.91
	<i>std</i> ₁	2.35	22.01	162.9	1.08	9.21
	<i>std</i> ₂	22.33	70.34	458.0	15.95	19.51
<i>A</i> ₂	<i>mean</i>	−1.18	−20.15	−102.4	−3.09	0.95
	<i>std</i> ₁	0.43	8.14	23.0	1.31	0.56
	<i>std</i> ₂	10.24	55.80	151.5	13.19	14.42
<i>B</i> ₂	<i>mean</i>	10.08	34.07	204.2	9.49	11.53
	<i>std</i> ₁	2.82	13.75	63.9	9.20	2.92
	<i>std</i> ₂	18.32	69.02	195.9	17.56	20.11

The classification accuracy of the XGBoost model trained on various HRV feature datasets is depicted in [Figure 11](#). [Figures 11A–C](#) illustrate the accuracy of classifiers trained on HRV feature data obtained from the SCWT experiments conducted in the morning, afternoon, and evening, respectively, using both raw data and detrended data for stress assessment across different time periods. For the raw dataset, classifiers trained on data from a specific SCWT task exhibit high accuracy in assessing stress during that time period. However, these classifiers show significantly lower accuracy in assessing stress for other SCWT tasks compared to those trained

on detrended data. Furthermore, classifiers trained on detrended data from each SCWT task demonstrate higher overall accuracy in assessing stress across all SCWT tasks compared to those trained on raw data (morning: +1.62%, afternoon: +3.05%, evening: +5.22%) and exhibit less variation in accuracy across different time periods. These results indicate that the presence of circadian rhythms adversely affects stress detection throughout the day, while detrending significantly enhances classification accuracy. Furthermore, we trained the XGBoost model using integrated HRV feature data from all SCWT tasks (as shown in [Figure 12](#)). To

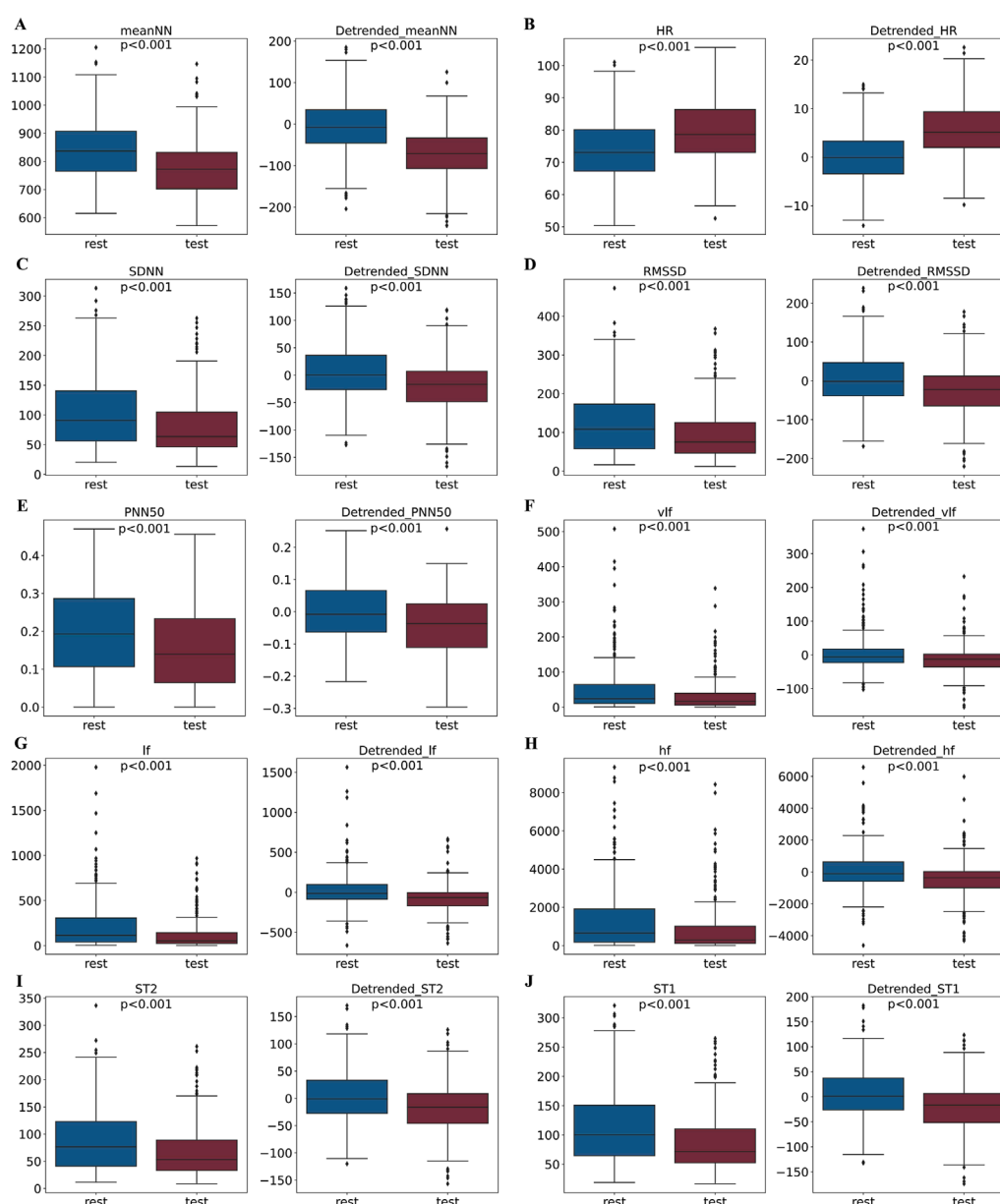


FIGURE 9

Comparison of feature distributions between rest and stress states before and after removing circadian rhythm trends. For each module, the left panel shows the feature distribution before trend removal, and the right panel shows the distribution after trend removal.

ensure a consistent sample size in the training dataset, we randomly selected one-third of the data from each SCWT task and combined them to form the training dataset.

For the non-detrended data, the classifier trained on the integrated dataset from all SCWT tasks exhibited overall lower accuracy in assessing stress compared to classifiers trained on data from each individual SCWT task. This finding indicates that simply integrating data from different time periods into a single training dataset does not enhance the overall accuracy of stress assessment across all periods, highlighting the significant impact of circadian rhythms.

In contrast, for the detrended data, the classifier trained on integrated data from different SCWT tasks maintained a high classification accuracy. As illustrated in Figure 12, the accuracy of

the classifier reaches 74.82% for overall stress assessment across all SCWT tasks, reflecting a notable improvement of 13.67% compared to the classifier trained on the raw data. This substantial enhancement highlights the critical role of detrending in improving stress recognition performance and underscores its necessity for achieving more reliable and accurate analyses.

Notably, the detrended data used for training the stress classification model was collected from all volunteers. In general, HRV varies among individuals, with differences potentially influenced by factors such as age. For instance, the results presented in Table 1 demonstrate the diversity of circadian rhythm trends in HRV features across individuals. To further investigate this, we examined the impact of inter-individual HRV differences on the classification accuracy of the XGBoost model trained on detrended data.

TABLE 2 The F-statistics for the distribution of each HRV feature in resting and stress states. The first column shows the results for HRV feature data without removing circadian rhythm trends, while the second column shows the results for HRV feature data after removing circadian rhythm trends.

Features	Original F-statistic	Detrended F-statistic
<i>meanNN</i>	54.410	151.796
<i>HR</i>	47.975	152.578
<i>SDNN</i>	23.514	33.829
<i>RMSSD</i>	19.352	26.978
<i>PNN50</i>	19.465	30.536
<i>vlf</i>	15.438	20.984
<i>lf</i>	23.298	30.256
<i>hf</i>	16.811	22.750
<i>ST2</i>	19.391	27.030
<i>ST1</i>	25.722	37.384

Since nearly half of the recruited volunteers were 26 years old or younger (as shown in [Figure 2A](#)), we further investigated the impact of HRV differences across age groups on the classification accuracy. We categorized all participants into two groups: Group 1 (age ≤ 26) and Group 2 (age > 26). In [Figure 13A](#), “Model1” and “Model2” represent the overall classification accuracy of stress assessment across all time periods for all participants, obtained from classifiers trained on detrended data from volunteers in Group 1 and Group 2, respectively. Due to the relatively concentrated age distribution of our recruited volunteers and the removal of circadian rhythm trends, which partially reduces HRV feature differences across age groups, the overall accuracy of stress assessment using classifiers trained on detrended data from these two groups remains comparable. However, it is important to note that when there is a larger age disparity among participants, the impact of age-related HRV differences on stress assessment accuracy should be carefully considered.

Moreover, during the data collection phase, as volunteers become more familiar with the experimental procedure, their stress levels and patterns may change. To investigate the impact of this familiarity on the accuracy of the final classifier, we conducted further analyses. In general, volunteers’ familiarity with the experiment gradually increases from the first to the third day. To examine this effect, we trained three classifiers separately using detrended SCWT task data from the first, second, and third days, respectively. We then evaluated the accuracy of these classifiers in assessing stress levels for all volunteers on each individual day, as well as their overall accuracy across the 3 days. As shown in [Figure 13B](#), classifiers trained on detrended data from a specific day achieved consistently high accuracy in stress assessment for other days, and the overall accuracy of the three classifiers remained relatively similar. These findings suggest that changes in stress levels and patterns due to increased familiarity with the experiment have a minimal impact on the accuracy of the final classifier in our study.

4 Discussion

This study analyzed the rhythmicity of HRV features influenced by the autonomic nervous system and examined its impact on stress detection throughout the day. First, outlier removal was conducted on the RR interval data collected from 50 volunteers over three consecutive days using a confidence ellipse with a sliding window. Next, 10 HRV features were extracted. The rhythmic periods of these HRV features were calculated using the FFT algorithm, and the circadian rhythm trends were removed using the SPA method. Finally, XGBoost classifiers were trained on the detrended HRV feature data from various SCWT tasks to assess psychological states.

Specifically, during the preprocessing step, a confidence ellipse with a sliding window was applied to the collected RR interval data to effectively eliminate outlier data points. The exclusion of these outliers reduced the average CV of the RR interval data from 0.28 to 0.20, thereby enhancing data quality and reliability. Then, in the analysis of circadian rhythms in HRV features, we employed FFT to compute the periods and amplitudes of HRV features, revealing distinct 12-h and 24-h periodic components. Additionally, a second-order sine function was utilized to model the diurnal fluctuations of these features. Subsequently, to eliminate circadian rhythmic trends in HRV features, we applied the SPA method to filter out these trends in order to more clearly observe the contribution of stress to HRV variations. After detrending, the distribution differences of each HRV feature between resting and testing states were more distinctly evident. Finally, XGBoost classifiers were trained using the data collected from the SCWT task to evaluate stress. The classifier’s accuracy improved by 13.67% when trained on the detrended data.

When analyzing the circadian rhythmic trends in HRV features, we found that the day-to-day variability within the same individual was smaller than the variability between individuals. This suggests that circadian rhythm in individual experimental data can be removed by using data from the previous day as a reference. However, these HRV indicators are generally influenced by a variety of internal and external factors. In this study, the participants were primarily young students and teachers with relatively regular daily routines. Given the sample size of 50, further stratification based on factors such as menstrual cycle and menopausal status in females, individual intelligence levels, personality traits, occupations, and day-night work patterns would result in subgroups too small to yield reliable conclusions. Nevertheless, the potential impact of these factors should not be overlooked. Studies have found that children and older adults exhibit longer reaction times and higher error rates when performing the Stroop task ([Ménétré and Laganaro, 2023](#)). In a small-scale study, it was found that men were consistently slower than women across trial blocks by approximately 46 milliseconds, although their error rates did not differ significantly ([Mekarski et al., 1996](#)). In cognitive function tests, patients with Alzheimer’s disease (AD) performed significantly worse than the normal control group on the Mini-Mental State Examination (MMSE) scores and various sub-scores of the Stroop test. Additionally, certain indices showed a significant correlation with glucose metabolism in specific regions of the prefrontal cortex ([Yun et al., 2011](#)). Studies on sleep deprivation have found that after a night of sleep loss, participants exhibit significantly prolonged reaction times in the Stroop task ([Cain et al., 2011](#)). The potential roles of these factors in stress detection based on daily HRV features will be an important focus of future research.

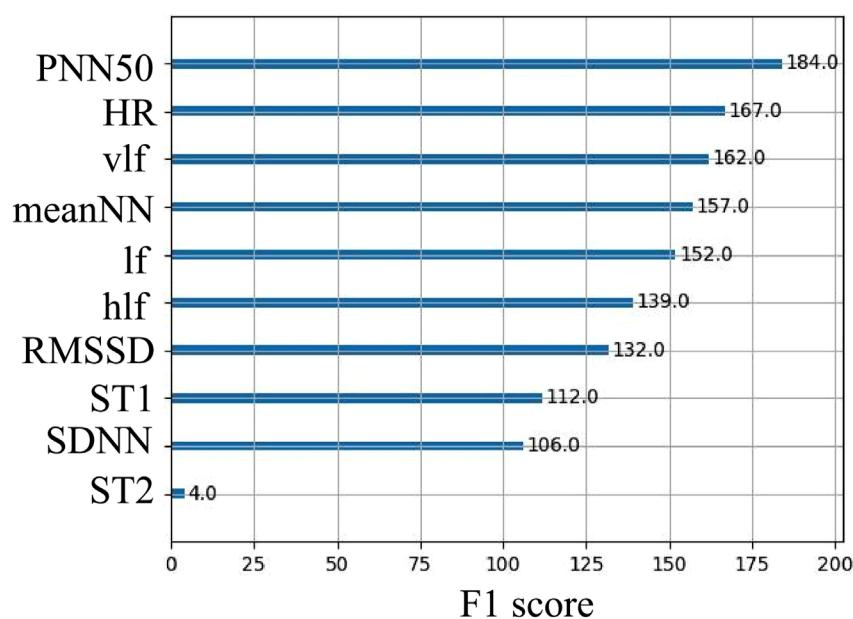


FIGURE 10
Ranking of the importance of the 10 HRV features obtained using XGBoost.

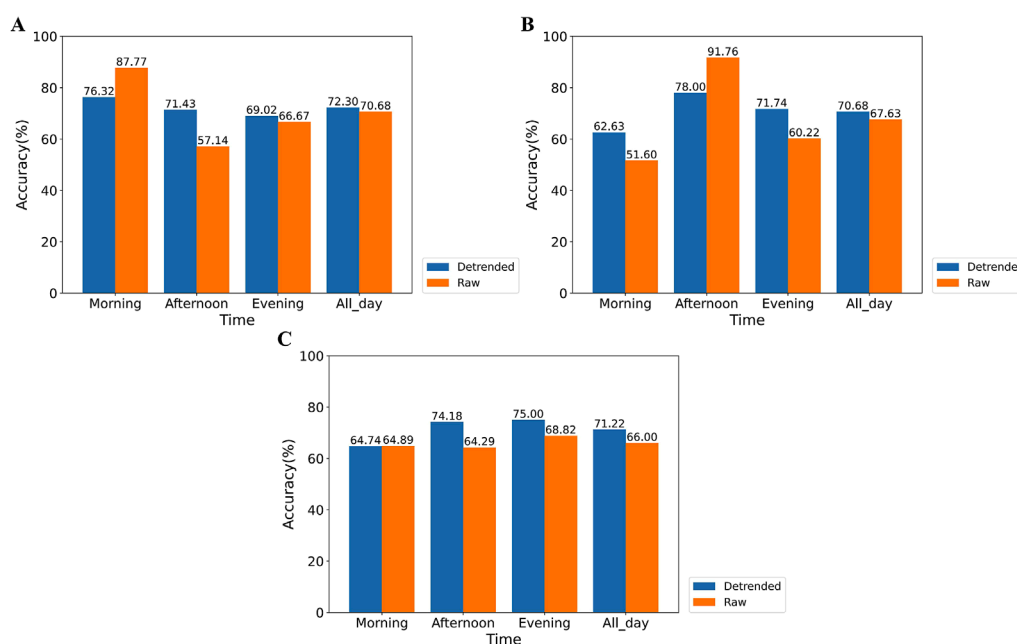


FIGURE 11
Comparison of the accuracy of stress assessment for different time periods using XGBoost classifiers trained on SCWT experimental data for each time period. This includes classifiers trained using only morning (A), afternoon (B), or evening (C) experimental data.

Additionally, due to the combined influence of stress and circadian rhythms on HRV, we ultimately used the HRV data with the circadian rhythmic trends removed to train the classifier for stress assessment. It is important to note that studies have shown stress can also affect circadian rhythms (Weibel et al., 2002; Tahara et al., 2016), and this influence should be considered when removing circadian rhythm components

from HRV. However, in our data collection, the stress induced by the SCWT test was short in duration and the stressors were controlled, minimizing the impact of stress on circadian rhythms (Thompson et al., 2013).

In the comparison between the XGBoost classifiers trained on detrended data and non-detrended data, when SCWT experimental data from only the morning, afternoon, or evening were used as

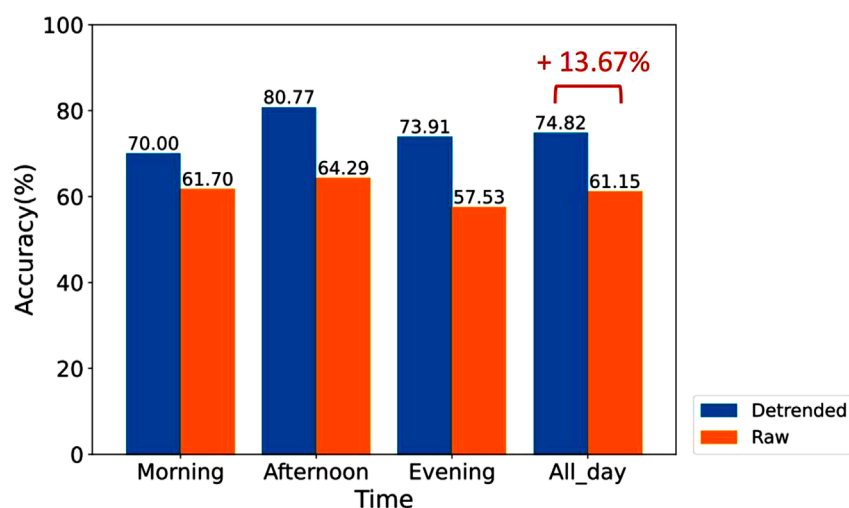


FIGURE 12

Comparison of the accuracy of stress assessment for different time periods using XGBoost classifiers trained on a dataset that randomly samples one-third of the integrated data from different SCWT tasks.

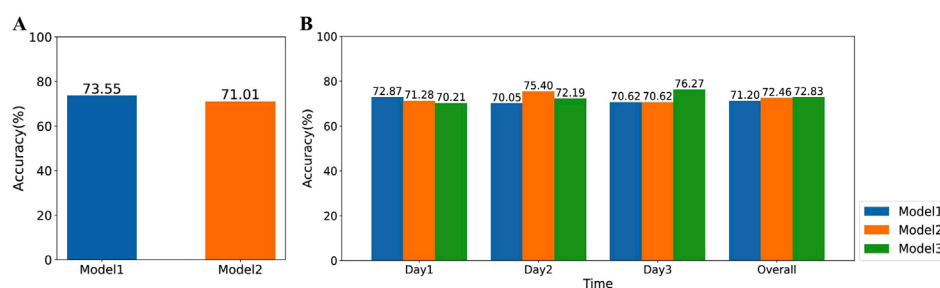


FIGURE 13

The impact of volunteers' age and familiarity with the experiment on the accuracy of stress assessment. (A) Comparison of the accuracy of XGBoost classifiers trained on data from volunteers of different age groups. Model1 and Model2 represent classifiers trained on data from volunteers aged 26 and below and those aged above 26, respectively. (B) Comparison of the accuracy of XGBoost classifiers trained on data from different days. Model1, Model2, and Model3 represent classifiers trained on volunteer data collected on the first, second, and third days, respectively.

the training dataset, classifiers trained on non-detrended data from each SCWT task demonstrated high accuracy, but only for stress assessment within the corresponding time period. This finding suggests that previous studies, which collected data during fixed time periods, may have partially reduced the influence of circadian rhythms (Hemakom et al., 2024; da Estrela et al., 2020). However, stress classification models developed in this way are less reliable for detecting stress at other times of the day. In contrast, classifiers trained on detrended HRV data from each SCWT task exhibited consistently high accuracy in stress classification across all SCWT tasks and achieved higher overall accuracy in stress assessment throughout the day (morning: +1.62%, afternoon: +3.05%, evening: +5.22%). These results indicate that classifiers trained on detrended HRV feature data provide a more robust and effective assessment of stress across different time periods.

Furthermore, when the data from different time periods of the SCWT task were integrated as the training set, classifiers trained on the non-detrended data exhibited low accuracy in stress assessment for each SCWT task, confirming the negative impact

of circadian rhythms on stress evaluation. In contrast, classifiers trained on the detrended data demonstrated significantly higher accuracy in stress assessment across all SCWT tasks compared to those trained on non-detrended data, achieving an overall accuracy improvement of 13.67%. The significant improvement in accuracy suggests that fluctuations in HRV features caused by intrinsic circadian rhythms may mask the changes induced by short-term stress, particularly when the stress levels are not extreme. Directly analyzing non-detrended data may overlook many stress-related problems, potentially leading to severe adverse consequences. Additionally, the ability of the combined detrended data from different time periods to produce highly accurate classifiers allows for more flexible data collection times, enhancing the convenience of the experimental process. Moreover, existing studies have also evaluated stress using physiological indicators beyond HRV, such as electroencephalography and functional near-infrared spectroscopy (Al-Shargie et al., 2017; 2016). Since most physiological indicators are influenced by circadian rhythms (Stangherlin et al., 2020), our approach can be widely applied to these stress assessment methods,

thereby improving both the accuracy of stress detection and the convenience of data collection.

Although our findings demonstrate that removing circadian rhythms from HRV data can enhance stress detection accuracy across different time periods, there are still some limitations. Due to constraints in experimental conditions and the available dataset, this study did not directly perform stress recognition across all time periods. Furthermore, the final classifier we obtained can only identify the presence or absence of stress, without providing a specific stress level. We aim to address this gap in future research. In the SCWT task, the need to repeat the experiment following a failure may be influenced by the participant's personality traits, which could affect their experienced stress levels. Additionally, repeated exposure to the experimental procedure may introduce a training effect. These limitations will be thoroughly addressed in future large-scale studies. Moreover, we treated pulse rate variability (PRV) derived from smartwatch data as equivalent to HRV. However, prior studies have shown that PRV and HRV may differ under certain conditions (Mejía-Mejía et al., 2020). To further improve the accuracy of stress detection, these differences should be carefully considered.

5 Conclusion

In this study, HRV features with circadian rhythm trends removed were used to assess stress throughout the day. We found that the extracted HRV features exhibited periodic fluctuations with 12-h and 24-h cycles, causing overlaps between stress-induced HRV feature distributions during one time period and resting states during another, which reduced classification accuracy. Additionally, circadian rhythm variations between different days for individuals were considered negligible. Compared to non-detrended HRV data, XGBoost classifiers trained on detrended data from each SCWT task demonstrated improved overall accuracy in stress assessment across all SCWT tasks (morning: +1.62%, afternoon: +3.05%, evening: +5.22%), with smaller differences in accuracy across different tasks. Moreover, when data from different time periods were combined to train the classifier, the overall accuracy of the classifier trained on non-detrended data for all SCWT tasks significantly decreased. In contrast, the classifier trained on detrended data maintained high overall accuracy, outperforming the non-detrended classifier by 13.67%. These findings indicate that using detrended HRV feature data is an effective method for assessing stress throughout the day and allows for more flexible data collection.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of Soochow University. The studies were conducted in accordance with the local legislation and institutional

requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

YS: Data curation, Formal Analysis, Investigation, Methodology, Visualization, Writing – original draft. ZF: Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft. TZ: Data curation, Investigation, Methodology, Writing – original draft. FY: Data curation, Investigation, Methodology, Writing – original draft. YX: Conceptualization, Project administration, Resources, Supervision, Writing – review and editing. LY: Conceptualization, Project administration, Resources, Supervision, Writing – review and editing, Methodology, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the National Natural Science Foundation of China (12471467), the National Key Research and Development Program of China (2018YFA0801103), the National Natural Science Foundation of China (12071330) to Ling Yang, Interdisciplinary basic Frontier Innovation Program of Suzhou Medical College of Soochow University (YXY2303022) to Ying Xu, the National Natural Science Foundation of China (32100931) to Tao Zhang, and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX23_3225) to Yafei Shen.

Acknowledgments

The authors thank all the volunteers for their efforts in data collection.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of

their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Acerbi, G., Rovini, E., Betti, S., Tirri, A., Rónai, J. F., Sirianni, A., et al. (2016). "A wearable system for stress detection through physiological data analysis," in *Italian forum on active and assisted living*.
- Al-Shargie, F., Katmah, R., Tariq, U., Babiloni, F., Al-Mughairbi, F., and Al-Nashash, H. (2022). Stress management using fmirs and binaural beats stimulation. *Biomed. Opt. express* 13 (6), 3552–3575. doi:10.1364/BOE.455097
- Al-Shargie, F., Kiguchi, M., Badruddin, N., Dass, S. C., Hani, A. F. M., and Tang, T. B. (2016). Mental stress assessment using simultaneous measurement of eeg and fmirs. *Biomed. Opt. express* 7 (10), 3882–3898. doi:10.1364/BOE.7.003882
- Al-Shargie, F., Tang, T. B., and Kiguchi, M. (2017). Assessment of mental stress effects on prefrontal cortical activities using canonical correlation analysis: an fmirs-eeg study. *Biomed. Opt. express* 8 (5), 2583–2598. doi:10.1364/BOE.8.002583
- Athavale, Y. R., and Krishnan, S. S. (2017). Biosignal monitoring using wearables: observations and opportunities. *Biomed. Signal Process. Control.* 38, 22–33. doi:10.1016/j.bspc.2017.03.011
- Beauchaine, T. P., and Thayer, J. F. (2015). Heart rate variability as a transdiagnostic biomarker of psychopathology. *Int. J. Psychophysiol. official J. Int. Organ. Psychophysiol.* 98 (2 Pt 2), 338–350. doi:10.1016/j.ijpsycho.2015.08.004
- Bernardi, L., Wdowczyk-Zsulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G., et al. (2000). Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. *J. Am. Coll. Cardiol.* 35 (6), 1462–1469. doi:10.1016/s0735-1097(00)00595-7
- Blechert, J., Lajtman, M., Michael, T., Margraf, J., and Wilhelm, F. H. (2006). Identifying anxiety states using broad sampling and advanced processing of peripheral physiological information. *Biomed. Sci. Instrum.* 42, 136–141.
- Brennan, M., Palaniswami, M. S., and Kamen, P. (2001). Do existing measures of poincare plot geometry reflect nonlinear features of heart rate variability? *IEEE Trans. Biomed. Eng.* 48, 1342–1347. doi:10.1109/10.959330
- Cain, S. W., Silva, E. J., Chang, A. M., Ronda, J. M., and Duffy, J. F. (2011). One night of sleep deprivation affects reaction time, but not interference or facilitation in a stroop task. *Brain Cognition* 76, 37–42. doi:10.1016/j.bandc.2011.03.005
- Camm, A. J., Malik, M., Bigger, J. T., Breithardt, G., Cerutti, S., Cohen, R. J., et al. (1996a). Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology. *Circulation* 93, 1043–1065. doi:10.1161/01.cir.93.5.1043
- Camm, A. J., Malik, M., Bigger, J. T., Breithardt, G., Cerutti, S., Cohen, R. J., et al. (1996b). Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation* 93, 1043–1065. doi:10.1161/01.cir.93.5.1043
- Castaldo, R., Melillo, P., Bracale, M., Caserta, M., Triassi, M., and Pecchia, L. (2015). Acute mental stress assessment via short term hrv analysis in healthy adults: a systematic review with meta-analysis. *Biomed. Signal Process. Control.* 18, 370–377. doi:10.1016/j.bspc.2015.02.012
- Castaldo, R., Montesinos, L., Melillo, P., James, C. J., and Pecchia, L. (2019). Ultra-short term hrv features as surrogates of short term hrv: a case study on mental stress detection in real life. *BMC Med. Inf. Decis. Mak.* 19, 12. doi:10.1186/s12911-019-0742-y
- Chen, T., and Guestrin, C. (2016). Xgboost: a scalable tree boosting system. *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 785–794. doi:10.1145/2939672.2939785
- Cinaz, B., Arnrich, B., Marca, R. L., and Tröster, G. (2013). Monitoring of mental workload levels during an everyday life office-work scenario. *Personal Ubiquitous Comput.* 17, 229–239. doi:10.1007/s00779-011-0466-1
- Cipresso, P., Colombo, D., and Riva, G. (2019). Computational psychometrics using psychophysiological measures for the assessment of acute mental stress. *Sensors* 19, 781. doi:10.3390/s19040781
- Cohen, S., Janicki-Deverts, D., and Miller, G. E. (2007). Psychological stress and disease. *JAMA* 298 (14), 1685–1687. doi:10.1001/jama.298.14.1685
- Cooley, J. W., and Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Math. Comput.* 19, 297–301. doi:10.1090/s0025-5718-1965-0178586-1
- da Estrela, C., McGrath, J. J., Booi, L., and Gouin, J. (2020). Heart rate variability, sleep quality, and depression in the context of chronic stress. *Ann. Behav. Med. a Publ. Soc. Behav. Med.* 55, 155–164. doi:10.1093/abm/kaa039
- Delaney, J. P. A., and Brodie, D. A. (2000). Effects of short-term psychological stress on the time and frequency domains of heart-rate variability. *Percept. Mot. Ski.* 91, 515–524. doi:10.2466/pms.2000.91.2.515
- Folkman, S. (2013). Stress: Appraisal and coping. *Encyclopedia of Behavioral Medicine*. Editor M. D. Gellman, and J. R. Turner (New York, NY: Springer New York), 1913–1915. doi:10.1007/978-1-4419-1005-9_215
- Hayano, J., and Yuda, E. (2021). Assessment of autonomic function by long-term heart rate variability: beyond the classical framework of lf and hf measurements. *J. Physiological Anthropol.* 40, 21. doi:10.1186/s40101-021-00272-y
- Healey, J., and Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intelligent Transp. Syst.* 6, 156–166. doi:10.1109/tits.2005.848368
- Hemakom, A., Atiwiwat, D., and Israsena, P. (2024). Ecg and eeg based machine learning models for the classification of mental workload and stress levels for women in different menstrual phases, men, and mixed sexes. *Biomed. Signal Process. Control.* 95, 106379. doi:10.1016/j.bspc.2024.106379
- Hernando, A., Lázaro, J., Gil, E., Valdés, A. A., Rey, J. M. G., López-Antón, R., et al. (2016). Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment. *IEEE J. Biomed. Health Inf.* 20, 1016–1025. doi:10.1109/JBHI.2016.2553578
- Hynynen, E., Kontinen, N., Kinnunen, U., Kyröläinen, H., and Rusko, H. (2011). The incidence of stress symptoms and heart rate variability during sleep and orthostatic test. *Eur. J. Appl. Physiology* 111, 733–741. doi:10.1007/s00421-010-1698-x
- Karjalainen, P. A. (1997). *Regularization and bayesian methods for evoked potential estimation*. Ph.D. Thesis. Kuopio, Finland: University of Kuopio.
- Karthikeyan, P., Murugappan, M., and Yaacob, S. (2014). Analysis of stroop color word test-based human stress detection using electrocardiography and heart rate variability signals. *Arabian J. Sci. Eng.* 39, 1835–1847. doi:10.1007/s13369-013-0786-8
- Karthikeyan, P., Murugappan, M., and Yaacob, S. B. (2013). Detection of human stress using short-term ecg and hrv signals. *J. Mech. Med. Biol.* 13, 1350038. doi:10.1142/s0219519413500383
- Lee, D., Han, C. H., Kim, H., Uhm, J.-S., Yoon, D., and Park, J. Y. (2023). Changes in the circadian rhythm of high-frequency heart rate variability associated with depression. *J. Korean Med. Sci.* 38, e142. doi:10.3346/jkms.2023.38.e142
- Lucini, D., Norbiato, G., Clerici, M., and Pagani, M. (2002). Hemodynamic and autonomic adjustments to real life stress conditions in humans. *Hypertens. J. Am. Heart Assoc.* 39, 184–188. doi:10.1161/hy0102.100784
- Luque-Casado, A., Perales, J. C., Cárdenas, D., and Sanabria, D. (2016). Heart rate variability and cognitive processing: the autonomic response to task demands. *Biol. Psychol.* 113, 83–90. doi:10.1016/j.biopsycho.2015.11.013
- Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., et al. (1996). Heart rate variability. standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* 17, 354–381. doi:10.1093/oxfordjournals.eurheartj.a014868
- Marques, A. H., Silverman, M. N., and Sternberg, E. M. (2010). Evaluation of stress systems by applying noninvasive methodologies: measurements of neuroimmune biomarkers in the sweat, heart rate variability and salivary cortisol. *Neuroimmunomodulation* 17, 205–208. doi:10.1159/000258725
- McDuff, D. J., Hernández, J., Gontarek, S., and Picard, R. W. (2016). Cogcam: contact-free measurement of cognitive stress during computer tasks with a digital camera. *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst.*, 4000–4004. doi:10.1145/2858036.2858247
- Mejía-Mejía, E., Budidha, K., Abay, T. Y., May, J. M., and Kyriacou, P. A. (2020). Heart rate variability (hrv) and pulse rate variability (prv) for the assessment of autonomic responses. *Front. Physiology* 11, 779. doi:10.3389/fphys.2020.00779
- Mekarski, J. E., Cutmore, T. R. H., and Suboski, W. (1996). Gender differences during processing of the stroop task. *Percept. Mot. Ski.* 83, 563–568. doi:10.2466/pms.1996.83.2.563
- Melillo, P., Bracale, M., and Pecchia, L. (2011). Nonlinear heart rate variability features for real-life stress detection. case study: students under stress due to university examination. *Biomed. Eng. OnLine* 10, 96. doi:10.1186/1475-925X-10-96
- Ménétré, E., and Laganaro, M. (2023). The temporal dynamics of the stroop effect from childhood to young and older adulthood. *PLOS ONE* 18, e0256003. doi:10.1371/journal.pone.0256003

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2025.1535331/full#supplementary-material>

- Mukherjee, S., Yadav, R., Yung, I. O., Zajdel, D. P., and Oken, B. S. (2011). Sensitivity to mental effort and test-retest reliability of heart rate variability measures in healthy seniors. *Clin. Neurophysiol.* 122, 2059–2066. doi:10.1016/j.clinph.2011.02.032
- Munla, N., Khalil, M., Shahin, A., and Mourad, A. (2015). “Driver stress level detection using hrv analysis,” in *2015 international conference on advances in biomedical engineering (ICABME)*, 61–64.
- Niwa, F., Kuriyama, N., Nakagawa, M., and Imanishi, J. (2011). Circadian rhythm of rest activity and autonomic nervous system activity at different stages in Parkinson's disease. *Aut. Neurosci.* 165, 195–200. doi:10.1016/j.autneu.2011.07.010
- Okada, Y., Yoto, T. Y., aki Suzuki, T., Sakuragawa, S., and Sugiura, T. (2013). “Wearable eeg recorder with acceleration sensors for monitoring daily stress: office work simulation study,” in *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 4718–4721.
- Pecchia, L., Castaldo, R., Montesinos, L., and Melillo, P. (2018). Are ultra-short heart rate variability features good surrogates of short-term ones? state-of-the-art review and recommendations. *Healthc. Technol. Lett.* 5, 94–100. doi:10.1049/htl.2017.0090
- Peng, H., Hu, B., Zheng, F., Fan, D., Zhao, W., Chen, X., et al. (2012). A method of identifying chronic stress by eeg. *Personal Ubiquitous Comput.* 17, 1341–1347. doi:10.1007/s00779-012-0593-3
- Pereira, T., Almeida, P. R., Cunha, J. P. S., and Aguiar, A. (2017). Heart rate variability metrics for fine-grained stress level assessment. *Comput. methods programs Biomed.* 148, 71–80. doi:10.1016/j.cmpb.2017.06.018
- Petrowski, K., Herold, U., Joraschky, P., Mück-Weymann, M., and Siepmann, M. (2010). The effects of psychosocial stress on heart rate variability in panic disorder
- Salahuddin, L. B., Cho, J., Jeong, M. G., and Kim, D. (2007). “Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings,” in *2007 29th annual international conference of the IEEE Engineering in Medicine and Biology Society*, 4656–4659.
- Salai, M., Vassányi, I., and Kósa, I. (2016). Stress detection using low cost heart rate sensors. *J. Healthc. Eng.* 2016, 5136705. doi:10.1155/2016/5136705
- Schubert, C., Lambert, M., Nelesen, R. A., Bardwell, W. A., Choi, J. B., and Dimsdale, J. E. (2009). Effects of stress on heart rate complexity—a comparison between short-term and chronic stress. *Biol. Psychol.* 80, 325–332. doi:10.1016/j.biopsycho.2008.11.005
- Selye, H. (1958). The stress of life. *Med. J. Aust.* 1.
- Singh, R. R., Conjeti, S., and Banerjee, R. (2013). A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals. *Biomed. Signal Process. Control.* 8, 740–754. doi:10.1016/j.bspc.2013.06.014
- Stangherlin, A., Watson, J. L., Wong, D. C. S., Barbiero, S., Zeng, A., Seinkmane, E., et al. (2020). Publisher Correction: compensatory ion transport buffers daily protein rhythms to regulate osmotic balance and cellular physiology. *Nat. Commun.* 12, 6988. doi:10.1038/s41467-021-26725-7
- Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *J. Exp. Psychol. General* 18, 643–662. doi:10.1037/h0054651
- Taelman, J., Vandeput, S., Vleminx, E., Spaepen, A., and Huffel, S. V. (2011). Instantaneous changes in heart rate regulation due to mental load in simulated office work. *Eur. J. Appl. Physiology* 111, 1497–1505. doi:10.1007/s00421-010-1776-0
- Tahara, Y., Aoyama, S., and Shibata, S. (2016). The mammalian circadian clock and its entrainment by stress and exercise. *J. Physiological Sci. JPS* 67, 1–10. doi:10.1007/s12576-016-0450-7
- Tharion, E., Parthasarathy, S., and Neelakantan, N. (2009). Short-term heart rate variability measures in students during examinations. *Natl. Med. J. India* 22 (2), 63–66.
- Thompson, R. S., Christianson, J. P., Maslanik, T. M., Maier, S. F., Greenwood, B. N., and Fleshner, M. (2013). Effects of stressor controllability on diurnal physiological rhythms. *Physiology and Behav.* 112, 32–39. doi:10.1016/j.physbeh.2013.02.009
- Weibel, L., Maccari, S., and van Reeth, O. (2002). Circadian clock functioning is linked to acute stress reactivity in rats. *J. Biol. Rhythms* 17, 438–446. doi:10.1177/074873002237138
- Yun, J.-Y., Lee, D. Y., Seo, E. H., Choo, I. H., Park, S. Y., Kim, S. G., et al. (2011). Neural correlates of stroop performance in alzheimer's disease: a fdg-pet study. *Dementia Geriatric Cognitive Disord. EXTRA* 1, 190–201. doi:10.1159/000329517