Postprint: Mental Fatigue Detection Using Wearable ECG Based on HRV Analysis

Authors: Huang Shitong, Zhang Weiqiang, Zhang Pengzhu

Date: 2018-07-09T00:00:00+00:00

Abstract

Mental fatigue represents a critical contributing factor to numerous chronic diseases, including cardiovascular disease, diabetes, and cancer, yet it remains difficult to quantify, assess, and measure. This study proposes an engineering-feasible scheme for detecting fatigue levels in knowledge workers via smart wearable devices. To detect mental fatigue levels, the Mann-Whitney U test was utilized to evaluate the statistical significance of various HRV metrics in discriminating mental fatigue states, while random forest was employed for feature selection to ascertain the importance of HRV metrics. The findings revealed that the most important HRV metrics were NN.mean, PNN50, VLF, LF, and TP. Finally, four machine learning algorithms—SVM, Naïve Bayes, KNN, and logistic regression—were implemented to classify fatigue states. Experimental results demonstrated that the KNN classifier was the most effective, achieving a cross-validation accuracy of 75.5% and an AUC of 0.74.

Full Text

Detection of Mental Fatigue with Wearable ECG Devices Based on HRV Analysis

Shitong Huang, Weiqiang Zhang, Pengzhu Zhang
Antai College of Economics & Management, Shanghai Jiao Tong University, Shanghai 200030, China

Abstract

Mental fatigue is a key cause of many chronic diseases such as cardiovascular disease, diabetes, and cancer, yet it remains elusive and difficult to measure or detect. This research proposes a feasible, accurate, and cost-effective method for detecting fatigue levels in mental workers via smart wearable devices. To detect mental fatigue levels, this paper first extracts HRV features, then uses the Mann-Whitney U Test to evaluate the statistical significance of HRV features between normal and fatigue states. Random Forest is employed for feature selection to determine the importance of HRV features. The most important HRV features are NN.mean, PNN50, VLF, LF, and TP. This paper then applies four machine learning algorithms to the selected features to predict mental fatigue states. Experiments demonstrate the effectiveness of the KNN classifier, achieving 75.5% accuracy in cross-validation and an AUC of 0.74.

Key words: mental fatigue detection; HRV; Mann-Whitney U test; random forest; machine learning

0 Introduction

Mental fatigue is a subjective feeling of tiredness, with most definitions focusing on mental functional performance [1-3]. Grandjean described mental fatigue as a low-alertness state that impairs brain function [1], while Staals characterized it as a central nervous system state where cellular capacity is insufficient to maintain original alertness [2]. In this paper, mental fatigue is defined as a temporary reduction in maximum cognitive capacity of the brain caused by prolonged cognitive activity, externally manifested as drowsiness, lethargy, or weakened attention. Medical research demonstrates that excessive fatigue has become a primary cause of sudden fatal diseases, with studies indicating strong relationships between overwork-induced mental fatigue and both stroke and karoshi (death from overwork) [4]. High-intensity work increases the probability of cardiovascular disease [5]. Beyond physiological health damage, mental fatigue also significantly impacts memory, judgment, decision-making ability, and emotional management [2], with long-term excessive work leading to stress and tension, which further results in higher accident rates, absenteeism, and lower productivity [6,7].

Currently, over one-quarter of mental workers worldwide face the risk of overwork. Overwork-induced mental fatigue causes symptoms including reduced sleep quality, stress, and anxiety, indirectly leading to chronic disease risks such as cardiovascular and cerebrovascular diseases, diabetes, and cancer. Therefore, using smart wearable devices with physiological indicators to detect fatigue levels in mental workers and prevent further deterioration is highly meaningful.

Despite the importance of mental fatigue research for mental workers' physical and mental health, mental fatigue is difficult to measure in practice. Current mental fatigue quantification methods fall into four categories: subjective scales, objective experiments, observation methods, and physiological measurement. Subjective scale methods require subjects to assess their mental fatigue level through questionnaires, using scales such as the Stanford Sleepiness Scale (SSS), Chalder Fatigue Scale (CFS), and Fatigue Severity Scale (FSS). While simple and efficient, subjective scales suffer from relatively low objective accuracy due to subjects' varying medical knowledge, though they can serve as references for subsequent methods.

Objective experimental quantification methods design cognitive tasks to evaluate subjects' brain functional performance, measuring reaction time, memory, and decision-making through tasks like the Psychomotor Vigilance Task (PVT) [8], Multiple Sleep Latency Test (MSLT), and Maintenance of Wakefulness Test (MWT) [9]. Observation methods are rarely used for mental fatigue quantification, assessing fatigue through observable physiological states such as blinking and yawning. All three approaches are intrusive, as users must stop current work to complete questionnaires or cognitive tasks, making them unsuitable for monitoring mental fatigue in daily life.

Physiological indicator measurement can detect human fatigue states while maintaining daily work activities. Recent research has extensively studied fatigue detection using physiological indicators. Electroencephalography (EEG) is widely used to measure driver fatigue. An Australian research team developed an EEG-based driver anti-fatigue system to monitor mental fatigue [10,11], using ANOVA analysis to find significant differences between fatigue and alertness states in different frequency bands: delta (σ) (\$$\pm 0 - 4Hz), theta(\theta)(4 - 8Hz), and alpha($ \$) (8-13 Hz). Zhang et al. detected β waves and $(\alpha +)/\beta$ under binaural beat stimulation, finding application prospects for improving attention and alertness to combat mental fatigue [12]. Zhou et al. found significant differences in $(\alpha +)/\beta$ and β/α between fatigue and normal states through entertainment program viewing experiments [13]. Machine learning algorithms have also been widely applied in fatigue driving research. Hu et al. extracted 40 spectral features from EEG signals and trained an SVM model, achieving 86% accuracy in fatigue driving recognition [14]. Hu et al. improved EEG-based driver fatigue detection accuracy to 94% (22 experimental samples) using a Gradient Boosting Decision Tree (GBDT) model [15]. However, EEG-based fatigue detection devices typically have multiple channels and electrodes, making them unsuitable for daily use, particularly in office or home settings. Therefore, a portable wearable device for monitoring mental fatigue is urgently needed.

With recent health information technology development and popularization of smart wearable devices like smart bracelets, real-time and remote health monitoring and management have become possible. Numerous intelligent sensors for continuously acquiring physiological parameters have emerged, such as portable ECG devices and heart rate and blood pressure sensors with Bluetooth wireless transmission. ECG (electrocardiogram) is a promising method for real-time mental fatigue monitoring. Compared to complex EEG systems, ECG signals can be obtained more conveniently. Early studies showed correlations between the autonomic nervous system (ANS) and cardiac rhythm [16], enabling ECG signals to measure mental fatigue states.

Domestic and international research exploring the relationship between ECG and fatigue has primarily focused on visual fatigue and exercise fatigue. Yang et al. used ECG to study visual fatigue from watching 3D/2D movies, finding SDNN increased with fatigue while NN.mean decreased [17]. Song et al. studied HRV changes in athletes before and after fatigue tests, finding time-domain indicators decreased while frequency-domain indicators significantly increased after fatigue tests, validating HRV as a reliable method for detecting exercise fatigue [18]. Wang Jun combined subjective fatigue scales with HRV to detect exercise fatigue, finding that after continuous multiple exercise fatigue sessions, NN.mean, SDNN, RMSSD, and PNN50 values decreased, with HRV showing good correlation with blood biochemical indicators [19]. HRV has proven useful for detecting visual and exercise fatigue, yet while visual and exercise fatigue can be observed from surface characteristics, mental fatigue from complex cognitive work is rarely directly observable and measurable.

This study explores how to detect mental fatigue using portable ECG devices. While single-lead portable ECG devices inevitably have reduced signal accuracy compared to traditional multi-lead instruments, this research aims to propose an engineering feasibility solution based on portable ECG devices, exploring the feasibility of using portable ECG devices for fatigue detection in health management applications rather than for medical diagnosis. To address this research question, this paper designed and conducted an experiment to test the feasibility of wearable ECG devices for mental fatigue measurement.

1 Empirical and Analysis Methods

1.1 Data Preparation

This study used the wearable ECG device "LaPatch" developed by Langlang Technology Company, with an ADS1292R chip from Texas Instruments for ECG signal acquisition, transmitting data to mobile terminals via Bluetooth. Thirty-five healthy college students without heart disease history participated in the experiment, with a male-to-female ratio of 1:1.3 and age range of 23$±4*years.Duringtheexperiment, subjectsworea"LaPatch"ECGdevicetocollectreal−timeECGsignals.Beforetheexperiment, subjectscompletedthe*14−*itemChalderFatigueScale*[20]*tocharacter* minutes. After the test, subjects completed the 14-item Chalder Fatigue Scale again according to their actual feelings. Based on the optimal 3/4 threshold from literature [20], scores above 50 were labeled as "fatigue" (fatigue = 1) and scores below 50 as "normal" (fatigue = 0).

1.2 HRV Index Extraction

ECG signals represent bioelectrical changes generated by the pacemaker, atria, and ventricles during each cardiac cycle. Heart Rate Variability (HRV) is a quantitative indicator reflecting autonomic nervous system activity and quantitatively assessing cardiac sympathetic and vagal tone and their balance [21]. By measuring variability in consecutive R-R intervals, HRV reflects heart rate change processes and patterns to evaluate comprehensive effects on systemic organs, including mental activity. Using HRV indicators to assess human fatigue states can provide a more comprehensive evaluation from a medical theory perspective.

The HRV index system is typically divided into time-domain and frequency-domain indicators. Time-domain indicators are extracted through statistical analysis of RR interval sequences, including: normal heartbeat interval mean (NN.mean); normal heartbeat interval standard deviation (SDNN); short-term average normal heartbeat interval standard deviation; root mean square of successive differences between adjacent normal heartbeat intervals (rMSSD); number of adjacent normal heartbeat intervals differing by more than 50ms (NN50); and PNN50, obtained by dividing NN50 by the total number of normal heartbeat intervals and multiplying by 100. The calculation formulas are shown in Table 1 .

Frequency-domain indicators are obtained by converting original ECG time-domain signals to frequency domain through Fourier transform and calculating power spectral density (PSD) in different frequency bands. Frequency-domain indicators are divided into four bands: total power TP (0-0.4Hz); very low frequency power VLF (0.003-0.04 Hz); low frequency power LF (0.04-0.15Hz); high frequency power HF (0.15-0.4 Hz); and low-to-high frequency power balance ratio LF/HF. Research shows VLF reflects thermoregulation [22], vasodilation, and fluid system regulation [23,24]; LF reflects dual regulation of cardiac sympathetic and parasympathetic activity [25,26]; HF primarily reflects vagal activity regulation [27]; and LF/HF balance ratio reflects sympathetic and parasympathetic system balance.

In this experiment, ECG signal sampling frequency was 250 Hz. The time window for calculating HRV indicators was 5 minutes. This paper calculated ten HRV indicators: NN.mean, SDNN, SDANN, PNN50, rMSSD, TP, HF, LF, VLF, and LF/HF.

1.3 Data Analysis

After removing outliers and missing values caused by device instability and special circumstances, data from 29 subjects were considered qualified for subsequent research. The entire experiment lasted 60-80 minutes (with test time of 54$±$8 minutes). We selected the first 10 minutes of ECG data to calculate HRV time-domain and frequency-domain indicators as pre-experiment physiological indicators corresponding to pre-experiment Chalder Fatigue Scale results, and used the last 10 minutes of ECG data to obtain HRV indicators as post-experiment physiological indicators corresponding to post-experiment Chalder Fatigue Scale results. Ultimately, 58 samples with fatigue labels (29 pre-test + 29 post-test) were used for subsequent data analysis.

1.3.1 Index Collinearity Analysis Considering the impact of multicollinearity on model performance, this paper used Pearson coefficients for redundancy analysis. When the correlation coefficient between two indicators exceeded 0.7, the indicators were considered to have multicollinearity and only one was retained. As shown in Figure 1 [Figure 1: see original paper], SDNN and rMSSD showed obvious collinearity; TP, LF, and HF also showed high collinearity. Therefore, in subsequent modeling, only one of SDNN and rMSSD could be retained, and similarly, only one variable could be retained among TP, LF, and HF due to their collinearity.

1.3.2 Correlation Analysis Between HRV Indicators and Mental Fatigue Considering that HRV indicators are typically non-uniformly distributed, this paper used the Mann-Whitney U test to examine statistical differences in HRV indicators between fatigue and normal states. The Mann-Whitney U test is currently the most widely applied rank sum test method, suitable as an alternative to T-test when normality and homogeneity of variance assumptions are not met. By considering the rank of each measurement value in each sample, it tests whether the means of two populations have significant differences. The p-value from the Mann-Whitney U test serves as a reference for examining the significance of each HRV indicator.

Figure 2 [Figure 2: see original paper] shows boxplots of HRV indicators in fatigue and normal states. The figure indicates that PNN50 is significantly lower in the fatigue state than in the normal state. NN.mean is also a significant indicator, showing obvious decline in the fatigue state. Indicators rMSSD, TP, VLF, and LF show slight increases in the fatigue state compared to the normal state.

Table 2 presents Mann-Whitney U test results. At the 0.1 significance level, significant differences exist between fatigue and normal states for the PNN50 indicator (p=0.02). NN.mean, LF, and VLF approach significance (p_{mean} = 0.112, p_{LF} = 0.166, p_{VLF} = 0.109). Other indicators such as SDNN, TP, HF, and LF/HF did not pass the Mann-Whitney U test. However, the Mann-Whitney U test only detects significance differences from a mean perspective without exploring the relationship between HRV indicators and fatigue.

1.3.3 Random Forest-Based HRV Indicator Evaluation The Random Forest classifier builds multiple independent decision tree models based on randomly sampled information, combining their predictions through voting to determine the class of variables to be classified. Assuming K decision trees are generated with M feature dimensions, bootstrap sampling extracts sub-training sets, and a specified number of features are randomly selected from all features to build multiple decision tree classifiers {h(x, _i), i = 1,2,3 ···K}, where { _i} are independent and identically distributed random vectors.

Each decision tree' s growth depends on an independent and identically distributed random vector; the generalization error of the ensemble depends on individual tree classification capability and correlation among trees. Beyond serving as a classifier, Random Forest can also be used for feature selection. Random Forest-based feature evaluation mainly includes two methods:

1) **Gini Index-Based Importance**: Decision trees typically determine splitting operations based on maximum node purity principles during prun-

ing. Node purity is usually measured by the Gini index. A smaller Gini index indicates higher purity. Based on changes in the Gini index (Equation (1)), each feature's importance can be reflected—the smaller the value, the higher the purity, indicating greater feature importance [28,29].

2) **Classification Accuracy Decrease-Based Importance**: Out-of-bag error (OOB) is commonly used to measure classification accuracy and can also measure feature importance. After obtaining each decision tree' s OOB error (B0), for each feature variable involved in the decision tree operation, keep other feature values unchanged while randomly permuting the OOB data values of the specified feature, then recalculate the decision tree' s OOB error (B1). The mean difference between these two types of OOB errors across all decision trees is the feature importance score VIM. For feature M, the importance score formula is:

$$VIM(M) = \sum (B1_t^M - B0_t)$$

A higher VIM score indicates higher feature importance [30,31].

Figure 3 [Figure 3: see original paper] shows importance scores based on these two Random Forest evaluation methods. In the classification accuracy decrease-based evaluation system, the five most important indicators are VLF, LF, TP, NN.mean, and PNN50, in descending order of importance. In the Gini index-based evaluation system, the five most important indicators are NN.mean, PNN50, rMSSD, VLF, and LF, in descending order. Combining Mann-Whitney U test and Random Forest results, this paper identifies NN.mean, PNN50, TP, LF, and VLF as important indicators for subsequent classification model validation.

1.4 Classifier Algorithm Theory and Performance Comparison

1.4.1 Classifier Algorithm Theory 1) Support Vector Machine (SVM) Classification Algorithm

The SVM classifier learning concept is to find a separating hyperplane in sample space that divides samples of different classes. The hyperplane can be represented by the linear equation:

$$w^{Tx} + b = 0$$

Assuming the hyperplane (w, b) can correctly classify samples, then for (x_i, y_i) D:

$$y_i(w_i^{Tx} + b) \geq 1, i = 1, 2..., m$$

The training samples closest to the hyperplane that satisfy this equation are called support vectors. The sum of distances from two heterogeneous support vectors to the hyperplane is $\gamma = 2/\|w\|$. To find the hyperplane with maximum margin, maximizing $\|w\|^{-1}$ is equivalent to minimizing $\|w\|^2$, solving the optimization equation:

$$\min \frac{1}{2}\|w\|^2$$
$$\text{s.t. } y_i(w_i^{Tx} + b) \geq 1, i = 1, 2, ..., m$$

2) Naïve Bayes Classifier

The Naïve Bayes classification approach operates under the Bayesian probability framework. With all relevant probabilities known, it selects optimal class labels based on these probabilities and misclassification loss. The decision criterion is based on posterior probability P(c|x) and misclassification loss, forming an optimal classifier that minimizes conditional risk. The Naïve Bayes classifier is based on the "attribute conditional independence assumption" : assuming all attributes are independent given the class. Under this assumption, the class-conditional probability P(c|x) for sample x relative to class label c is:

$$P(c|x) = P(c)P(x|c) = P(c)\prod_{i=1}^{d} P(x_i|c)$$

where d represents feature categories and x_i represents the value of x on the i-th attribute. Since P(x) is identical for all feature categories, the Naïve Bayes optimal classifier that minimizes classification error rate is:

$$h_{nb}(x) = \arg\max_{c\in Y} P(c)\prod_{i=1}^{d} P(x_i|c)$$

The Naïve Bayes classifier training process estimates class prior probability P(c) from training set D, estimates conditional probability P(x_i|c) for each feature according to the formula, then obtains sample x' s class-conditional probability P(c|x) relative to class label c, assigning the sample to the class with maximum conditional probability.

3) K-Nearest Neighbors (KNN) Classifier

The K-Nearest Neighbors (KNN) classifier, given a test sample, finds the k training samples closest to it based on a distance metric function. Typically through "voting," it selects the most frequent class label among these k samples as the prediction result. Given sample x, if its nearest neighbor is z, the error probability of the nearest neighbor classifier—that is, the probability that x and z have different class labels—can be expressed as:

$$P(err) = 1 - \sum_{c\in Y} P(c|x)P(c|z)$$

4) Logistic Regression

Logistic regression is a generalized linear regression for binary classification variables. It introduces a logit function on top of the linear regression model to make the dependent variable y approximate 0 or 1, formally retaining linear regression characteristics while actually seeking a nonlinear function mapping from input to output space. The logistic regression model expression is:

$$y = w^{Tx} + b$$

If we treat y in the equation as posterior probability p(y=1|x), then:

$$\frac{p(y=1|x)}{p(y=0|x)} = w^{Tx} + b$$

From p(y=0|x) + p(y=1|x) = 1, we obtain:

$$p(y=1|x) = \frac{e^{w^T+b}}{1+e^{w^T+b}}$$

$$p(y=0|x) = \frac{1}{1+e^{w^T+b}}$$

Estimating the maximum likelihood function for the logistic regression model yields w and b:

$$l(w,b) = \sum_{i=1}^{m} \ln p(y_i|x_i; w, b)$$

1.4.2 Classifier Performance Comparison This paper compared four machine learning classification algorithms: Support Vector Machine (SVM), Naïve Bayes classifier, KNN, and Logistic Regression. Optimal parameters were found through grid search, and classifier performance was evaluated through cross-validation using cross-validation accuracy and AUC as final performance metrics.

Tables 3-6 show the performance of the four classification models considering different numbers of indicator combinations. Results show KNN algorithm performance is significantly superior to the other three algorithms, achieving 75.5% cross-validation accuracy (K=3) with the HRV indicator combination of NN.mean, TP, and LF. SVM classifier ranks second, achieving 71.6% cross-validation accuracy with the HRV indicator combination of NN.mean and PNN50. Naïve Bayes and Logistic Regression models performed slightly worse, with cross-validation accuracies of 65.6% and 69.9%, respectively.

Table 7 and Figure 4 [Figure 4: see original paper] present AUC values and ROC curves for the four algorithms. Results show KNN achieves the highest AUC value of 0.74, SVM' s AUC is 0.68, while NB and LR have relatively lower AUC values of only 0.64 and 0.65. Overall, we conclude that the KNN classifier demonstrates significant effectiveness in classifying mental fatigue in mental workers. Based on optimal classifier indicator selection results, the most effective indicators for detecting mental fatigue states are PNN50, NN.mean, LF, VLF, and HF. These four indicators prove most effective both in previous Random Forest evaluation scores and optimal classifier indicator selection results.

2 Conclusion

This study provides an engineering feasibility solution for mental fatigue detection in mental workers through portable ECG devices based on HRV analysis. Under a simulated mental work scenario, this research combines questionnaire scales with ECG data collected by the "Langlang Heart" portable ECG device for statistical analysis and fatigue state recognition.

This paper addresses two primary questions: which HRV indicators correlate with mental fatigue states, and how to establish a fatigue state recognition model based on the HRV indicator system. For the first question, Mann-Whitney U tests reveal that PNN50 is significantly effective, while NN.mean, LF, and VLF approach significance. Using Random Forest classifiers, HRV indicator importance was evaluated from both Gini index and classification accuracy perspectives. From the classification accuracy perspective, the five most important indicators are VLF, LF, TP, NN.mean, and PNN50, while from the Gini index perspective, the five most important indicators are NN.mean, PNN50, rMSSD, VLF, and LF. Combining Mann-Whitney U test and Random Forest results, this paper identifies NN.mean, PNN50, TP, LF, and VLF as important indicators for mental fatigue degree.

For the second question, this paper evaluates the performance of four machine learning classifiers. Results show KNN algorithm performs best, achieving 75.5% cross-validation accuracy (K=3) with the HRV indicator combination of NN.mean, TP, and LF. KNN also achieves the highest AUC value of 0.74. Overall, the KNN classifier demonstrates significant effectiveness in classifying mental fatigue in mental workers.

Excessive fatigue causes tremendous physical and mental harm and affects normal human thinking and work, making fatigue state recognition research of great practical significance. This paper proposes an engineering feasibility solution for recognizing human fatigue states based on portable ECG devices. This study explores and analyzes the relationship between ECG signals and mental fatigue, laying a research foundation for further mental fatigue studies, effective identification, prediction, and intervention of human mental workload.

References

[1] Grandjean E. Fitting the task to the man: an ergonomic approach [M]. Taylor & Francis, 1969: 210.

[2] Staal M A. Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework [J]. National Aeronautics & Space Administration, 2004.

[3] Cheng Shan, Ma Jin, Hui Duoduo, et al. The research advancement of objective assessment methods for mental fatigue [J]. Progress in Modern Biomedicine, 2014, 14 (32): 6387-6390.

[4] Ke D. Overwork, stroke, and karoshi-death from overwork [J]. Acta Neurol Taiwan, 2012, 21 (2): 54-59.

[5] Backé E M, Seidler A, Latza U, et al. The role of psychosocial stress at work for the development of cardiovascular diseases: a systematic review [J]. International Archives of Occupational & Environmental Health, 2012, 85 (1): 67-79.

[6] Beer L T D, Pienaar J, Jr S R. Investigating the reversed causality of engagement and burnout in job demands-resources theory [J]. South African Journal of Industrial Psychology, 2013, 39 (1): 9-pages.

[7] Dewa C S, Loong D, Bonato S, et al. How does burnout affect physician productivity? A systematic literature review [J]. Bmc Health Services Research, 2014, 14 (1): 1-10.

[8] Urrila A S, Stenuit P, Huhdankoski O, et al. Psychomotor vigilance task performance during total sleep deprivation in young and postmenopausal women [J]. Behavioural Brain Research, 2007, 180 (1): 42-47.

[9] Sangal R B, Thomas L, Mitler M M. Maintenance of wakefulness test and multiple sleep latency test. Measurement of different abilities in patients with sleep disorders [J]. Chest, 1992, 101 (4): 898-902.

[10] Lal S K, Craig A. A critical review of the psychophysiology of driver fatigue [J]. Biological Psychology, 2001, 55 (3): 173-194.

[11] Lal S K L, Craig A. Electroencephalography activity associated with driver fatigue: Implications for a fatigue countermeasure device [J]. Journal of Psychophysiology, 2001, 15 (3): 183-189.

[12] Zhang Lu, Jiao Xuejun, Gao Xiang, et al. Effects of beta band binaural beats on brain physiological status [J]. Manned Spaceflight, 2016, 22 (2): 254-261.

[13] Zhou Jiaping, Zhang Pengzhu. Study on fatigue identification based on wearable electroencephalogram [J]. Application Research of Computers, 2018, 35 (8): .

[14] Hu Shuyan, Zheng Gangtie, Peters B. Driver fatigue detection from electroencephalogram spectrum after electrooculography artefact removal [J]. Iet Intelligent Transport Systems, 2013, 7 (1): 105-113.

[15] Hu Jianfeng, Min Jianliang. Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model [J]. Cognitive Neurodynamics, 2018, (12): 1-10.

[16] Sherwood L. Human physiology: from cells to systems [M]. Brooks//Cole, Cengage Learning, 2012.

[17] Yang Xinpan, Wang Danli, Hu Haichen, et al. P-31: visual fatigue assessment and modeling based on ECG and EOG caused by 2D and 3D displays [J].

Sid Symposium Digest of Technical Papers, 2016, 47 (1): 1237-1240.

[18] Song Tao, Yuan Chunping, Wu Yuxiang, et al. Research on the application of HRV in the evaluation of exercise fatigue for experimental teaching [J]. Laboratory Science, 2017, 20 (2): 8-11.

[19] Wang Jun. Research on sports fatigue detection method based on the combination of RPE and HRV [D]. Wuhan: Wuhan Sports University, 2015.

[20] Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale [J]. Journal of Psychosomatic Research, 1993, 37 (2): 147-153.

[21] Mulder L J. Measurement and analysis methods of heart rate and respiration for use in applied environments [J]. Biological Psychology, 1992, 34 (2-3): 205-236.

[22] Aoki K, Stephens D P, Johnson J M. Diurnal variation in cutaneous vasodilator and vasoconstrictor systems during heat stress [J]. Am J Physiol Regul Integr Comp Physiol, 2001, 281 (2): R591.

[23] Braga A N, Da S L M, Da S J, et al. Effects of angiotensins on day-night fluctuations and stress-induced changes in blood pressure [J]. Ajp Regulatory Integrative & Comparative Physiology, 2002, 282 (6): R1663-R1668.

[24] Jr P G, Rivkees S A. Ontogeny of humoral heart rate regulation in the embryonic mouse [J]. Ajp Regulatory Integrative & Comparative Physiology, 2001, 281 (281): R401-7.

[25] Lanfranchi P A, Somers V K. Arterial baroreflex function and cardiovascular variability: interactions and implications [J]. American Journal of Physiology Regul Integr Comp Physiol, 2002, 283 (4): R815.

[26] Malpas S C. Neural influences on cardiovascular variability: possibilities and pitfalls [J]. American Journal of Physiology Heart & Circulatory Physiology, 2002, 282 (1): H6.

[27] Rentero N, Cividjian A, Trevaks D, et al. Activity patterns of cardiac vagal motoneurons in rat nucleus ambiguus [J]. Ajp Regulatory Integrative & Comparative Physiology, 2002, 283 (6): R1327-34.

[28] Nunes L O, Biscainho L W P, Lee B, et al. Degradation type classifier for full band speech contaminated with echo, broadband noise, and reverberation [J]. IEEE Trans on Audio Speech & Language Processing, 2011, 19 (8): 2516-2526.

[29] Zhang Hongqiang, Liu Guangyuan, Lai Xiangwei. Application of random forest algorithm in important feature selection from EMG signal [J]. Computer Science, 2013, 40 (1): 200-202.

[30] Belgiu M, Drăguţ L. Random forest in remote sensing: A review of applications and future directions [J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2016, 114: 24-31.

[31] Liu Shu, Jiang Qigang, Mayue, et al. Object-oriented wetland classification based on hybrid feature selection method combining with relief f, multi-objective genetic algorithm and random forest [J]. Progress in Modern Biomedicine, 2017, 48 (1): 119-127.

Note: Figure translations are in progress. See original paper for figures.