WILEY | Hindawi

Research Article

Cognitive Load Identification of Pilots Based on Physiological-Psychological Characteristics in Complex Environments

Haibo Wang [ID], Naiqi Jiang [ID], Ting Pan [ID], Haiqing Si [ID], Yao Li [ID], and Wenjing Zou [ID]

School of Civil Aviation and Flight, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

Correspondence should be addressed to Haiqing Si; sihaiqing@126.com

Cognitive load is generated by pilots in the process of information cognition about aircraft control, and it is closely related to flight safety. Cognitive load is the physiological and psychological need that a pilot produces when completing a mission. Therefore, it is meaningful to study the dynamic identification of the cognitive load of the pilot under the complex human-aircraft-environment interaction. In this paper, the airfield traffic pattern flight simulation experiment was designed and used to obtain the ECG physiological and NASA-TLX psychological data. The wavelet transform preprocessing and mathematical statistics analysis were applied on them, respectively. Furthermore, the Pearson correlation analysis method is used to select the characteristic indicators of psycho-physiological data after preprocessing. Based on the psycho-physiological characteristic indicators, the pilot's cognitive load identification model is constructed by combining RNN and LSTM. The results of this study are more accurate compared with the cognitive load identification models established by other methods such as RNN neural network and support vector machine. This research is able to provide a useful reference for preventing and reduction of human error caused by the cognitive load during flight missions. It will be potential to realize intelligent control of aircraft cockpit, improving the flight control behavior and maintaining flight safety.

1. Introduction

With the rapid development of intelligent aircraft cockpits, the complexity and integration of aircraft airborne systems have gradually increased. At the same time, the human-aircraft-environment interaction process becomes more and more complicated and also the difficulty of flight mission increases accordingly. To complete the take-off, leveling, landing, and taxiing processes, pilots have to undertake many tasks in order to accurately control the aircraft. Most parts of the mare main tasks directly related to maneuvering the aircraft, such as information perception, judgment, decision-making, and execution. Secondary tasks mainly include early warning, information interface detection, and voice calls. Due to the limited information processing capacity of pilots, multisource information from different tasks can easily cause "information overload" in the brain. It can

result in the increased cognitive load of pilots, which adversely affects their operating behavior and further poses a potential threat to flight safety. Therefore, accurately identifying the cognitive load of pilots is of great significance for realizing intelligent control of the aircraft cockpit, improving the control behavior of flight crew and maintaining flight safety.

Many researchers have studied the cognitive load of pilots since the early twentieth century [1], [2]. Scholars were interested in the requirements of rapidly cultivating pilots and reducing the impact of overweight cognitive load on missions. Pilots' cognitive load was identified from the perspective of psychological selection and identification using motion coordination response detection, paper pen, or intelligence detection [3–5]. Although these studies were forward-looking, most of them had a certain subjective tendency to interfere. Due to the tremendous advances in the

field of artificial intelligence and computer technology, it has been a higher demand for the study of psychological status of pilots such as cognitive load. Many scholars have focused on studies of cognitive load by physiological indicators of pilots obtained by objective experiments. Noel et al. [6] used artificial neural networks to classify the cognitive characteristics combination of pilots at the different flight times. These characteristics include the electrocardiogram (ECG), the electroencephalography (EEG), and eye movements. Gray et al.[7] analyzed the correlation between the pilot's heart rate value, the visual scan area change value, and the aircraft's landing deviation distance. This was due to negative emotions such as anxiety caused by pilots' psychological overload. It showed that the anxiety caused by their psychological load has a significant impact on the flight performance. The problem of different cognitive loads generated by various types of pilots of the same task was studied by Mansikka et al.[8]. Based on the pilot's heart rate data obtained from flight simulation experiments, the researcher used the Shapiro–Wilk test and repeated measures variance to determine the flight performance of the pilot under the different cognitive loads. It was found by Andrew [9] for the rule that heart rate parameter values are most sensitive to changes in cognitive load. Specifically, the psychological burden of the pilot increases as the difficulty of the task increases and the low-frequency component of the heart rate variability decreases. Jonathan et al. [10] used visual behavior to study changes in the pilot's attention when performing landing operations in an anxiety situation. Tjolleng et al. [11] established an artificial neural network based on the cognitive load classification model according to the time-domain index of the pilot's ECG signal and the frequency-domain index. The time-domain indicators mainly include average interval, standard interval of the RR interval, and root-mean-squared error of the RR interval. The frequency-domain indicators include low- and high-frequency power and the ratio of them. This study effectively grades the cognitive load of the pilot, which provides a good theoretical basis for flight training. Wang [12] and others have discussed the increase in the cognitive load of the pilot caused by the large amount of information on the human-machine display interface of the aircraft cockpit. Based on the eye movement index, some researchers established a visual information source matrix to realize the quantitative research of cognitive load. Among them, eye movement indicators mainly include the number of gaze, gaze time, blink time, and pupil diameter. The physiological characteristics associated with brain activity were studied by Gianluca [13] and others, during the transition from normal flight to high cognitive load to fatigue operated by the pilot. Scannella et al. [14] tried to accurately discriminate among pilots' workload levels across real flight phases at the individual level. The averaged values of heart rate (HR), heart rate variablilty (HRV), fixation duration, saccadic rate, and visual entropy from the real flights were used to classify the flight phase of airfield traffic pattern. Antonio [15] introduced physiological signals such as heart rate (HR) and heart rate variability (HRV) to test the psychological state of pilot in order to reduce psychological errors caused by the

cognitive load of pilots. As the difficulty of the mission increases, pilots' cognitive load dose and also physiological indicators such as HR and HRV change accordingly. Wanyan et al. [16] provided electro acupuncture's blink frequency to assess the cognitive load of pilots during flight simulation missions, which showed that the frequency of blinking decreased as the psychological load increased. These findings provide a new evidence for an influence of mental workload on the cognitive function during the flight.

Many researchers have presented a variety of methods to study the cognitive load of pilots from the perspectives of pilots' physiology and psychology. They demonstrate that pilots' physiological characteristics have a significant impact on cognitive load from the different levels. These studies are promising and encouraging a deeper exploration of pilots' physiological and psychological characteristics to inspect the relation between them and cognitive workload in the flight. The existing pilot's cognitive load identification model reflects the intelligence of the pilot to some extent. However, due to the intrinsic analysis of the dynamic and linkage of physiological and psychological features, it could not reflect the actual performance of the flight control process when considering effects of the different cognitive loads. In view of this, this paper firstly provides the physiological and psychological dynamic data of pilots under the airfield traffic pattern through some flight-simulated experiments. The physiological data are mainly ECG characteristics of the pilot. Psychological data are measured by NASA-TLX scale on the task difficulty during the airfield traffic pattern to quantify the cognitive load. Furthermore, based on the preprocessing results of physiological and psychological data, recurrent neural network (RNN) and long short-term memory (LSTM) are used to establish a cognitive load identification model for flying cadets in the complex environments. The present study aims at achieving real-time dynamic, noninvasive, and low-cost identification of pilot's cognitive load based on their physiological-psychological characteristics in complex flight environments.

2. Methods

2.1. Cognitive Load Identification Model Based on LSTM-RNN

2.1.1. RNN Network Structure. RNN can be used to mine semantic information and time-series information in data. A typical cyclic neural network model mainly includes an input layer, a hidden one, and an output one [17], as displayed in Figure 1. I_t is an input of the network at time t. H_t is a hidden state of the network at time t and is also the memory unit of the network. m_t is an output at time t. Also, S_t is a final output of the model U, W, and V represent a weight matrix for the input layer, the hidden layer, and the output layer, respectively. The same weight matrix parameters (U, W, and V) are shared in a recurrent neural network at the different moments in each network layer.

2.1.2. LSTM Network Structure. Based on the traditional RNN, input gates, output gates, and forgotten gates are added into LSTM, as shown in Figure 2. Among them, σ
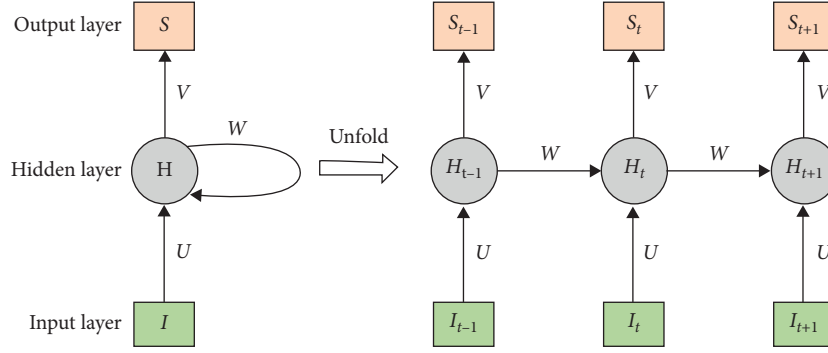
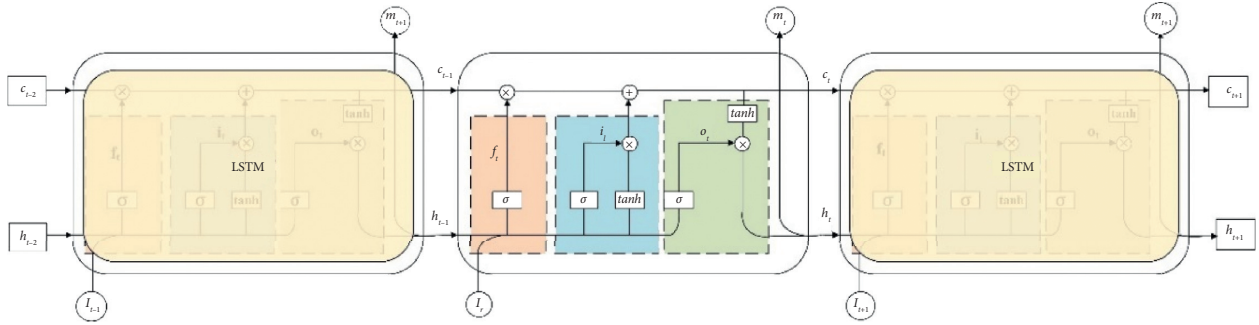FIGURE 1: Unfolded drawing of the cyclic neural network model.

FIGURE 2: Schematic of the LSTM structure.

represents the Sigmoid function, and its role is to control which parts of the connected quantity can continue to pass. In this paper, the cell state at time t is assigned to c_t, W is a corresponding weight matrix, and b_f, b_c, and b_o represent a corresponding offset vector, respectively [18].

(1) *Forgotten gate*. The function of the forgetting gate is to multiply the retained information by c_{t-1} after partially discarding the hidden layer state H_{t-1} at the previous moment and the input variable I_t at the moment. The formula is exhibited as follows:

$$f_t = \sigma\left(W_f \cdot [H_{t-1}, I_t] + b_f\right). \qquad (1)$$

(2) *Input gate*. The input gate is used to judge which parts of the information output by the forgetting gate are stored in the new c_t, and the formulas are shown as follows:

$$i_t = \sigma\left(W_i \cdot [H_{t-1}, I_t] + b_i\right), \qquad (2)$$

$$c_t = f_t * c_{t-1} + i_t * \tan h\left(W_c \cdot [H_{t-1}, I_t] + b_c\right). \qquad (3)$$

(3) *Output gate*. The output gate is used to determine the final output value m_t, which is obtained by multiplying c_t and o_t (the hidden layer state H_{t-1} at the previous moment and the reserved portion of the input variable I_t at the moment) by the tanh function. This is a node value output by the LSTM unit at this time. The formula is displayed as follows:

$$o_t = \sigma\left(W_o [H_{t-1}, I_t] + b_o\right), \qquad (4)$$

$$m_t = o_t * \tan h\left(c_t\right). \qquad (5)$$

2.2. Data Acquisition

2.2.1. Experimental Participants. The sample size of the experiment was 100 male flying cadets who were in good health had a visual acuity of 1.0 or more. They were familiar with flight procedures and do not have any chronic drug history. All of them aged 22.0 ± 2.1. They had 250–300 hours of flight experience.

The experimental process is strictly in accordance with the Helsinki Declaration. All experimental personnel received a complete explanation of the experimental procedures and equipment and signed a written informed consent form to participate in the experiment.

2.2.2. Experimental Apparatus. This experiment includes physiological and psychological tests. The apparatus in physiological experiment mainly includes Cessna 172R simulation, high-level simulation flight platform, BIOPAC MP150 signal acquisition equipment, notebook computer, high-definition camera, and other experimental equipment. Also, the present platform is mainly composed of a joystick, throttle and foot rudder control system, high level simulation instrument system, and virtual scene system based on Prepar 3D v4. The flight simulation experiment is shown in Figure 3.
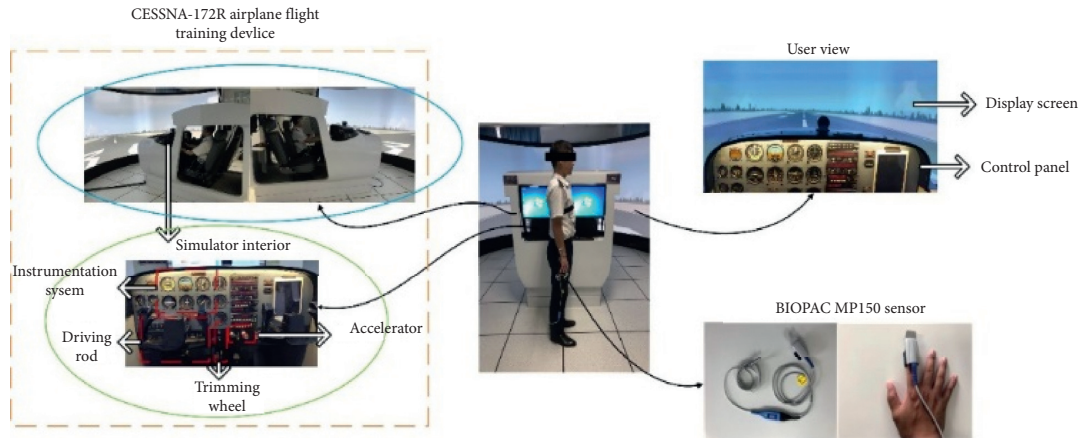
FIGURE 3: Simulator flight experiment diagram.

Psychological experiment is mainly carried out using the NASA Task Load Index (NASA-TLX). The specific description of each dimension of the NASA-TLX scale is given in [19] (for example, the psychological demand dimension is exhibited in Figure 4).

Participants could draw a mark on the scale that matches their perceived workload. The six dimensions of cognitive load paired together could form fifty cognitive load dimension combinations. Participants were asked to tick out the dimensions that contribute more to the cognitive load size in each combination. The weight value of the corresponding dimension to the total cognitive load is determined according to the number of times where each cognitive load dimension is marked (the value is 0–5). In this way, the total load level is calculated. The higher the score, the greater the psychological load of the participants; conversely, it is small.

2.2.3. Experimental Procedure

(1) *Flight scenario*. The experiment was carried out in a simulated environment with high visibility and without wind. The setting of the environment is mainly based on CCAR-61-R4 "Rules for the certification of pilots and ground instructors of civil aircraft" [20], including the flight skill requirement (11) of the 61.127 private pilot license and the flight experience requirement of the driver of the commercial pilot license aircraft category of the 61.159 (3). The flying cadets were organized to wear the BIOPAC MP150 ECG sensor. The sampling frequency was set to 2048 HZ. Afterward, they performed an airfield traffic pattern mission in accordance with the flight instructions. The airfield traffic pattern mainly includes departure leg, crosswind leg, downwind leg, base leg, and final leg (as shown in Figure 5). Participants were required to complete the five-stage mission at one time during the experiment.

(2) *Experimental design*. This experiment is roughly divided into 5 steps, and the specific steps are as follows:

(1) Some participants were introduced about the experimental procedure and the equipment. Afterwards, the equipment was checked and adjusted before the experiment. Each of them could perform 2-3 flight training before the formal experiment to adapt to the model display interface and cockpit operation. They also need to carefully read the six-dimensional description of NASA-TLX and the instructions for filling out the form.

(2) ECG sensor is applied in the experimental participant. After that, they perform the airport traffic mode according to instructions of the simulated platform vision system and conduct psychological tests.

(3) When flight participants complete each side of the airfield traffic pattern, the NASA-TLX scale-level pop-up will automatically appear on the simulated visual displayer. After completing the questionnaire according to the contents of the scale, they continued the next flight. The mission requires 15–20 minutes for a complete experiment.

(4) After each participant completes the airfield traffic pattern and fills the gauge, replace the object. The next one wears ECG sensor according to the same position and then repeats steps (3)-(4).

(5) After all the participants performed the experiment, the local NASA-TLX scale data and ECG data were saved in the computer for subsequent analysis.

FIGURE 4: Schematic of the psychological demand dimension of the NASA-TLX scale.
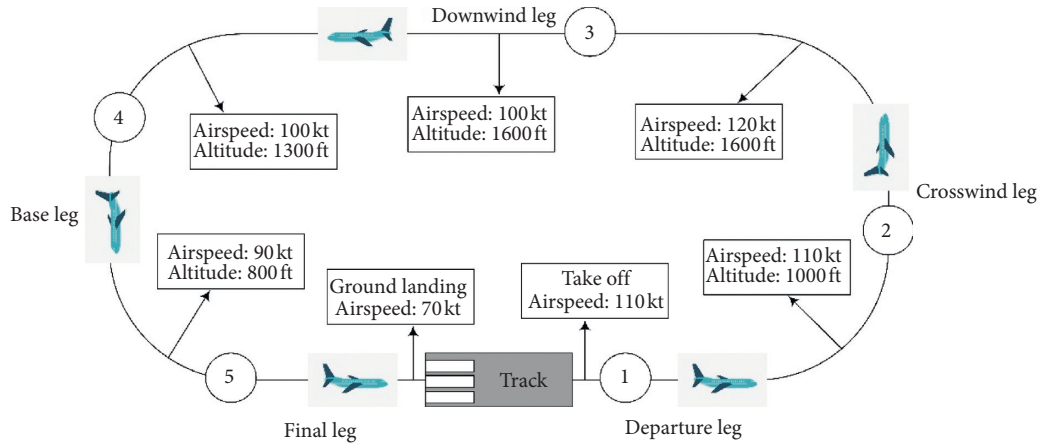

FIGURE 5: Cessna 172 flight schematic of the airfield traffic pattern.

Due to the long time consumption of this experiment, in order to avoid the effect of the experimental time difference, it was divided in two periods with similar physiological cycle of human body, namely, 9 : 00–12 : 00 and 14 : 00–17 : 00, the whole process remained quiet. The specific process is indicated in Figure 6.

3. Results

3.1. Data Analysis

3.1.1. Psychological Data Analysis.
The psychological data analysis flow is revealed in Figure 7.

The NASA-TLX total score (recorded as Total) counting method is roughly divided into two steps. Firstly, the score for each dimension is calculated based on the object's score and the weight value for each dimension of the NASA-TLX scale (the product of the weight value and the score). Secondly, scores for each dimension are summed, and they are divided by 15 at the end. According to the graphical calculation method, total scores of NASA-TLX during the airfield traffic pattern of 100 participants were obtained. Due to space limitations, the NASA-TLX total score calculation process for a flying cadet is shown as an example in Figure 8. NASA-TLX score calculation process is similar to that of other participants during the flight and will not be described here.

According to Figure 8, it can be got:
Total = $(15 \times 4 + 5 \times 1 + 12 \times 2 + 18 \times 5 + 10 \times 2 + 8 \times 1)/15 = 13.8$.

In order to verify whether NASA-TLX scale can distinguish different levels of cognitive load, the validity of cognitive load measurements on each side of the airfield traffic pattern is analyzed in this paper. In other words, it is validated whether there is a significant difference between the measured values on each side of the airfield traffic pattern based on the F test. By analyzing the difference between the measured values of departure leg and the other sides, it is known that there are significant differences between the measured values of the off-field edge and the other four edges, which are ($F_{(1,38)} = 228.85$, $p < 0.05$), ($F_{(1,38)} = 209.17$, $p < 0.05$), ($F_{(1,38)} = 322.14$, $p < 0.05$), and ($F_{(1,38)} = 166.54$, $p < 0.05$), respectively. The other four sides' analyzing processes are similar to the departure leg and will not be described here. From all the analysis results, the measured values on each side are valid. Hence, NASA-TLX subjective measurement has high validity and can be used to measure cognitive load size.

Statistical analysis was performed on NASA-TLX scale data on each side of the airfield traffic pattern. The analysis results are displayed in Table 1 and Figure 9.

In this paper, the radar chart is used to represent the cognitive load of each flight phase in six dimensions (as displayed in Figure 10). The higher brain power requirements are base leg, final leg, and departure leg. The higher physical strength is the downwind leg and the departure leg. The higher time requirements are the final leg, the base leg, and the departure leg. The poor performance levels are the final leg and the base leg. When the participants are on the departure leg and the base leg, they need the more effort. When they are on the final leg and the base leg, they perform the higher more frustration.

The analysis of the psychological index data shows that in the case of good weather, the cognitive load values of each flight stage from the high to the low are the final leg, the
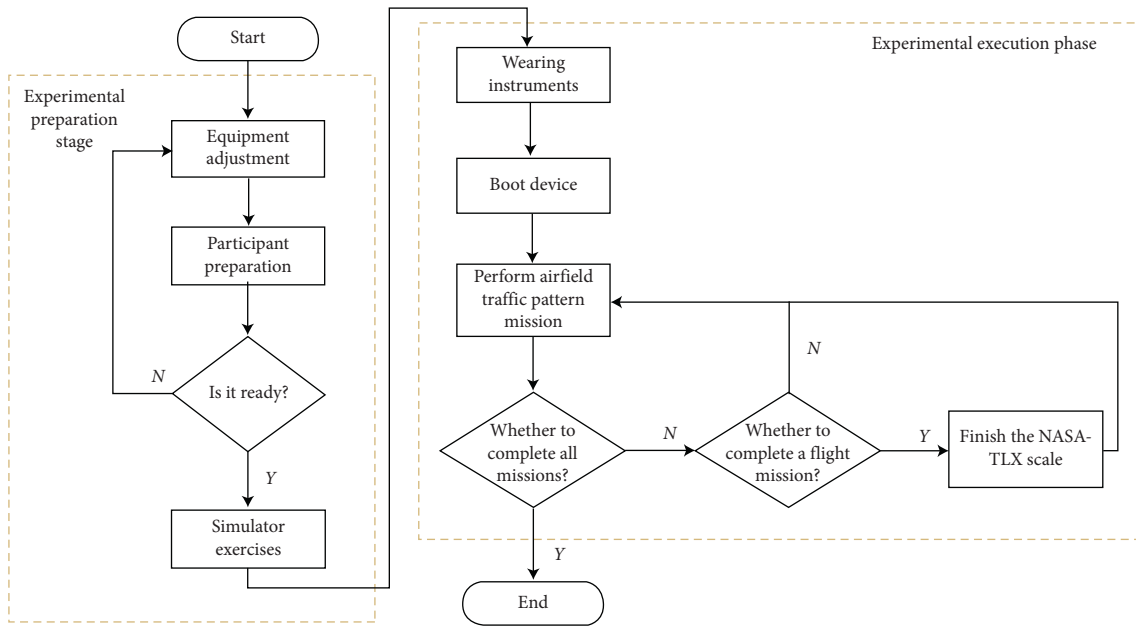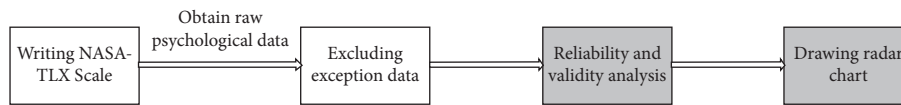
Figure 6: Experimental flow diagram.

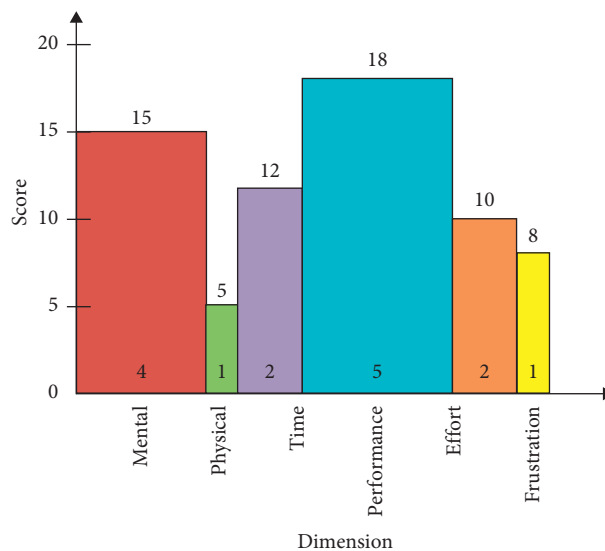Figure 7: Psychological data analysis flow diagram.

Figure 8: NASA-TLX score data chart.

departure leg, the base leg, the crosswind leg, and the downwind leg. The task on the final leg is the most difficult and the task on the downwind leg is the easiest.

3.1.2. ECG Data Analysis. HRV is an indicator of ECG signal. When the sympathetic nerve is excited, HRV increases; when the parasympathetic nerve is done, HRV

decreases [21]. In order to obtain the time-frequency indicators of the ECG signals, it is necessary to perform pretreatment, feature extraction and time-frequency analysis on the ECG signals, as exhibited in Figure 11.

(1) *Pretreatment of ECG signals.* For noise such as baseline drift mixed in the ECG signal, it can be removed using a high-pass filter with a cutoff

TABLE 1: Statistical analysis results of total scores of NASA-TLX scales.

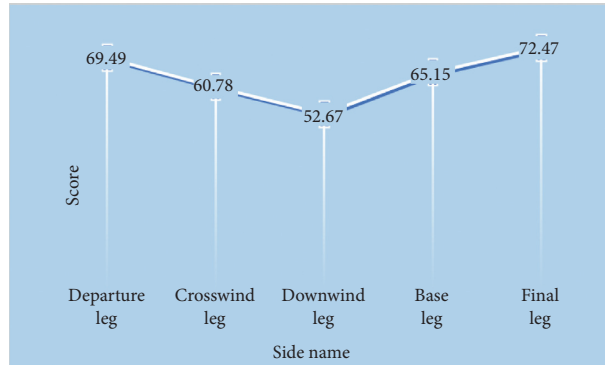| Analysis index airfield traffic pattern	Average value	Standard deviation	Number of samples	95% confidence interval
Departure leg	69.49	5.23	500	(65.39, 73.14)
Crosswind leg	60.78	5.13	500	(57.83, 61.01)
Downwind leg	52.67	4.78	500	(49.64, 56.78)
Base leg	65.15	4.95	500	(64.63, 71.77)
Final leg	72.47	5.41	500	(69.56, 78.64)

FIGURE 9: The scores of NASA-TLX.
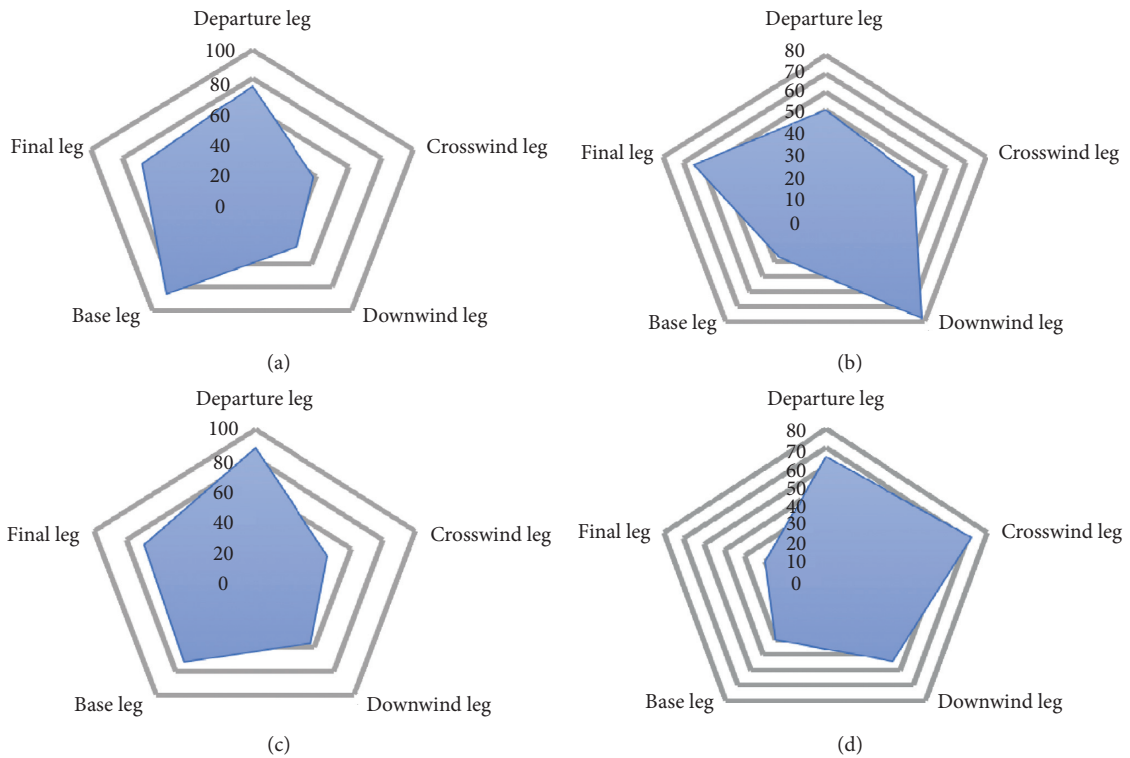

(a)

(b)
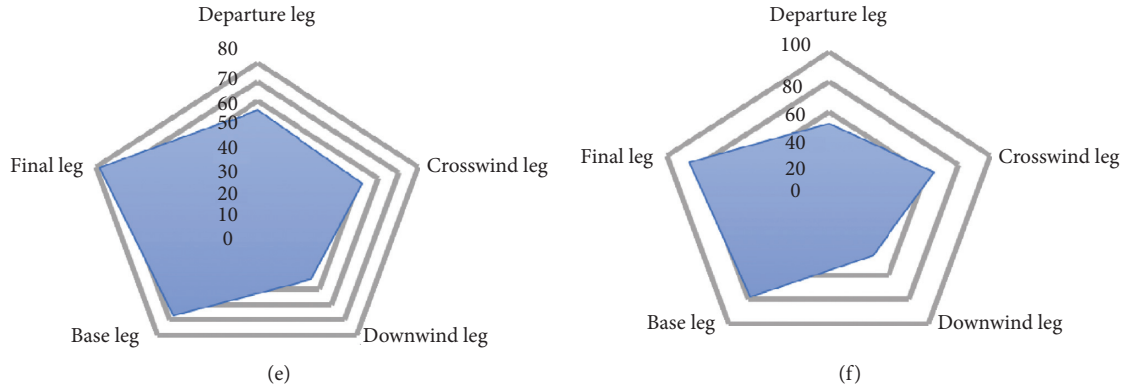

(c)

(d)

FIGURE 10: Continued.

FIGURE 10: Radar chart of cognitive load: (a) psychological need radar chart; (b) physical need radar chart; (c) time required radar chart; (d) performance level radar chart; (e) effort level radar chart; (f) frustration level radar chart.
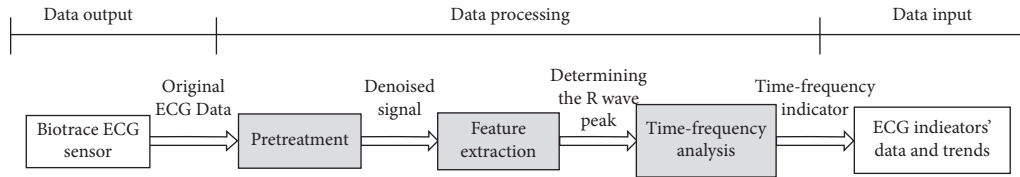

FIGURE 11: Theanalysis flow diagram of ECG data.

frequency of 0.5 Hz. The power frequency interference is filtered by a 50 Hz notch filter [22]. Random noise caused by respiratory wave noise, human motion, and electrode polarization is not considered here because of its low content.

On the basis of filtering out the noise such as distortion drift and power frequency interference, this paper uses the wavelet transform principle to eliminate the myoelectric interference mixed in the ECG signal. Here, the db5 wavelet function is applied to perform 8-scale decomposition on the original ECG signal. Then, based on the decomposition coefficient of each layer, unbiased likelihood estimation is adopted to select the threshold of each layer. Finally, ECG signal can be reconstructed to effectively remove the noise based on the wavelet reconstruction principle. In this paper, SNR and MSE are utilized to judge the denoising effect:

$$\text{SNR} = 101g\left(P_s - P_n\right), \quad (6)$$

$$\text{MSE} = \sum_{i=1}^{N} \frac{\left(Y_t - Y_t'\right)^2}{N}, \quad (7)$$

where P_s is the effective power of the signal; P_n is the effective power of the noise; Y_t is the actual value at time t; and Y_t' is the output value at time t.

According to formulas (6) and (7), SNR = 118.7143 and MSE = 6.1788e–04. The calculated results show that the denoising effect is good. ECG signal preprocessing is displayed in Figure 12.

(2) *R wave feature extraction based on the threshold method*. A complete ECG cycle is a sequence of two adjacent P waves. However, since the amplitude of the R wave is large and has the largest energy, the R-R interval replacement is commonly applied in the practical applications, namely, the distance between adjacent R waves.

In this paper, based on the signal noise reduction, a threshold method is used to extract the R wave. The average value of the high-frequency portion is set to be a threshold value 1 and one half of the threshold value 1 is assigned to a threshold value 2. When the signal value is greater than threshold 1, the signal position is recorded as L_1. When the signal value is less than threshold 2, the signal position is recorded as L_2. It can be seen that there must exist an R wave peak between L_1 and L_2.

(3) *Time-frequency analysis of HRV*. Time-frequency analysis of HRV is divided into two parts which are time-domain and frequency-domain analysis. Time-domain analysis is based on the calculation results of the R-R interval sequence and other statistical indicators for quantitative analysis. The description and calculation formulas of each time-domain indicator can be found in Table 2. The quantitative
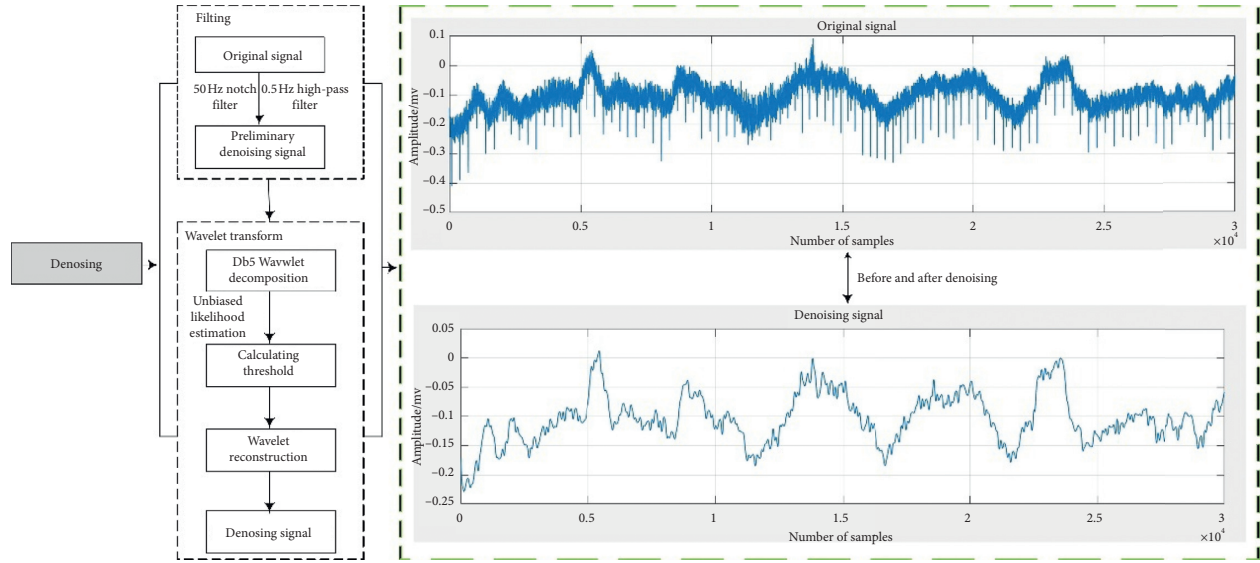
FIGURE 12: ECG signal preprocessing.

TABLE 2: HRV time-domain indicators.

Name	Unit	Description	Calculation formula
MEAN	ms	RR interval mean	$\text{MEAN} = \sum_{i=1}^{N} (\text{RR}_i/N)$
SDNN	ms	Standard deviation of all RR intervals	$\text{SDNN} = \sqrt{\sum_{i=1}^{N} ((\text{RR}_i - \overline{\text{RR}})^2/N)}$
RMSSD	ms	Interneighbor RR interval rms	$\text{RMSSD} = \sqrt{\sum_{i=1}^{N} ((\text{RR}_{i+1} - \text{RR}_i)^2/N)}$
$p\text{NN}_{50}$	%	The ratio of the RR interval is greater than 50 ms	$p\text{NN}_{50} = (\text{NN}_{50}/\text{NN}) \times 100\%$

analysis of HRV time-domain indicators on each side of the airfield traffic pattern are described in Table 3.

Frequency-domain analysis of HRV is based on the frequency characteristics of ECG signal, which decomposes the ECG signal into different frequency bands and calculates the energy it has in a unit of time. Thereby, quantitative analysis of the power of ECG signal in each frequency band is implemented. Depending on the frequency, HRV power spectrum can be mainly divided into three types of spectral components (as shown in Table 4). And, variations in each component can reflect changes in the activity of the autonomic nervous system.

In this paper, based on fast-Fourier transform principle, time-series of ECG signals are converted to the frequency domain and power spectrum of each side of the airfield traffic pattern is obtained. ECG power spectrum curve of the departure leg is displayed in Figure 13. The area enclosed by the power spectrum and the coordinates in each frequency band is numerical power of the signal in the frequency band. Therefore, energy characteristics of each frequency band are extracted according to the power spectrum to quantitatively analyze the frequency-domain characteristics of HRV. According to formula (8), the obtained low-frequency power and high-frequency power are standardized, which is convenient for obtaining their ratios:

$$x_i = \frac{(t_i - \mu_i)}{\Upsilon_i}, \quad i = 1, 2, 3, 4, \tag{8}$$

where t_i is a nonnormalized raw data value and μ_i and Υ_i are the mean and standard deviation of the original data, separately.

Since the frequency-domain analysis process of the flying cadets' ECG signal during the completion of other four-sided missions is similar to that of the departure leg, it is not described here. Results of the HRV frequency-domain indicators on each phase of the airfield traffic pattern obtained by frequency-domain analysis are depicted in Table 5. The indicators include normalized low-frequency power LF_{norm}, normalized high-frequency power HF_{norm}, and the ratio of low-frequency to high-frequency power (LF/HF).

According to the data in Table 3 and Table 5, the variation curves of the indicators in the time-domain and frequency-domain of the ECG signal are plotted on the airfield traffic pattern, as exhibited in Figure 14.

According to the statistical analysis of ECG data, indicators including MEAN, SDNN, LF, and (LF/HF) have the maximal value at the final leg and the minimal value at the downwind side. RMSSD and HF indicators have a maximum at the downwind side and a minimum at the final leg.

Results of the comprehensive psychological indicators and ECG indicators show that when the cognitive load

TABLE 3: HRV time-domain indicator on each side of the air field traffic pattern.

Time-domain indicator airfield traffic pattern	MEAN	SDNN	RMSSD	pNN$_{50}$
Departure leg	810.53	171.84	33.26	19.56
Crosswind leg	720.92	150.93	47.45	9.23
Downwind leg	783.67	162.34	38.69	13.41
Base leg	697.74	126.54	49.42	8.16
Final leg	856.49	180.26	29.73	24.74

TABLE 4: HRV spectral components.

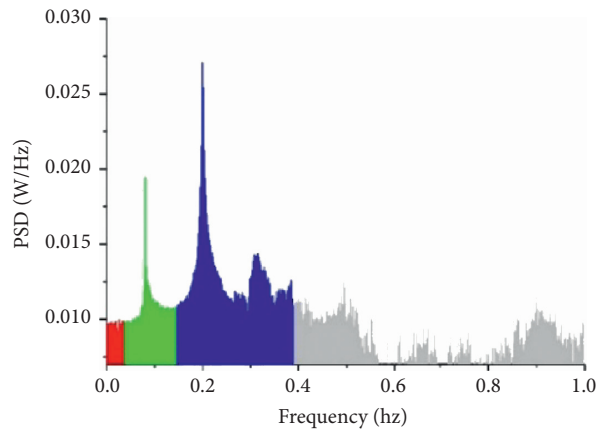| Name	Abbreviation	Frequency range
Very low frequency	VLF	$<0.04\,\mathrm{H_Z}$
Low frequency	LF	$0.04–0.15\,\mathrm{H_Z}$
High frequency	HF	$0.15–0.4\,\mathrm{H_Z}$

FIGURE 13: Departure leg ECG power spectrum.

TABLE 5: Results of HRV frequency-domain indicators on each phase of the airfield traffic pattern.

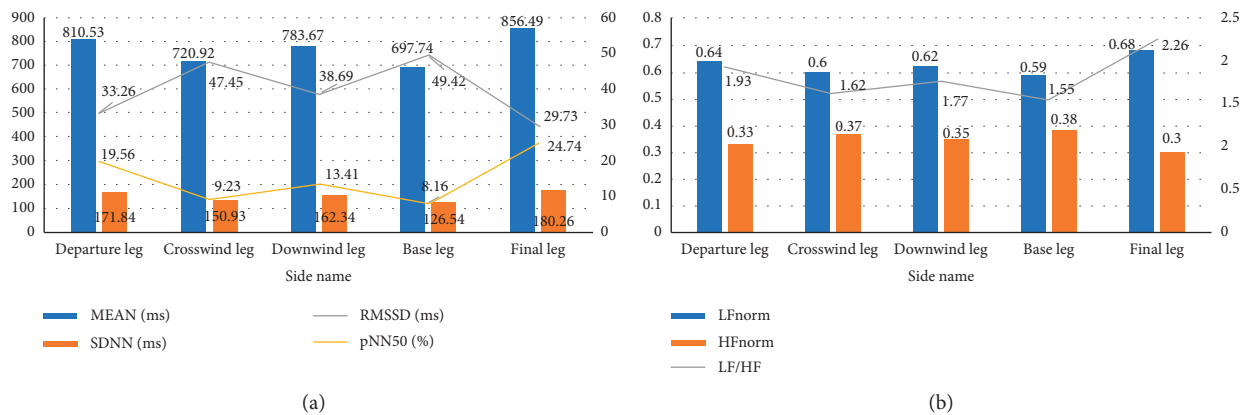| Frequency-domain indicator airfield traffic pattern	LF$_{norm}$	HF$_{norm}$	(LF/HF)
Departure leg	0.64	0.33	1.93
Crosswind leg	0.60	0.37	1.62
Downwind leg	0.62	0.35	1.77
Base leg	0.59	0.38	1.55
Final leg	0.68	0.30	2.26

FIGURE 14: ECG indicators change chart. (a) Time-domain indicator change chart; (b) frequency-domain indicator change chart.

increases, MEAN, SDNN, LF, and (LF/HF) will do accordingly. Meanwhile, RMSSD and HF indicators will have a tendency to decrease.

3.2. Correlation Analysis on Psycho-Physiological Factors.

Since it is difficult to rank cognitive load by a fixed threshold, it is done according to the total score of the NASA-TLX subjective scale. Therefore, the cognitive load is divided into three levels: high (score higher than 70), medium (score between 60 and 70), and low (score below 60). It can be observed from Figure 9 that total scores on each side of the airfield traffic pattern are 69.49, 60.78, 52.67, 65.15, and 72.47. Thence, the cognitive load levels on each side of the airfield traffic pattern are divided into the cognitive load: the final leg is high, the crosswind leg, the departure leg, and base leg are medium grade and the downwind leg is low grade.

Although both ECG and psychological indicators can be used to identify cognitive load, it is difficult to completely distinguish the difference in cognitive load with a single indicator. And, because the degree of consistency and degree of correlation of the indicators are different, it is not possible to directly identify the cognitive load levels of different flight periods with the selected indicators. Therefore, according to Pearson correlation analysis method, the correlation coefficient between each index (including 7 ECG features and 1 psychological characteristic index) is calculated. The Pearson correlation matrix is shown in Table 6.

It can be seen from Table 6 that the correlation between psychological index Total and ECG indicator is weak. The degree of correlation amongLF, HF, (LF/HF), and other indicators is not significant, and the characteristics of the changes are inconsistent. Therefore, Total, LF, HF, and (LF/HF)indicators are selected as the characteristic indicators for evaluating the cognitive load of flying cadets.

3.3. Cognitive Load Identification Model.

In this paper, a cognitive load identification model is established based on LSTM-RNN. The model structure is shown in Figure 15.Total, LF, HF, and (LF/HF) are input variables in the model, and each variable needs to be standardized according to formula (8). The pilot's cognitive load level is a model output variable.

3.3.1. Parameter Determination.

The total number of experimental participants is 100 in this paper. ECG and psychological data of each flying cadet is obtained on each side of the airfield traffic pattern. Thus, total of sample data is 500. Among them, 80% is randomly selected as a training sample and the remaining 20% is used as a test sample. Thence, 400 sample data are trained and 100 sample data are tested. MAE and RMSE were selected as indicators for test error analysis. These two indicators are calculated according to the following formulas:

$$\text{MAE} = \sum_{t=1}^{N} \frac{|Y_t - Y_t'|}{N}, \tag{9}$$

$$\text{RMSE} = \sqrt{\sum_{t=1}^{N} \frac{(Y_t - Y_t')^2}{N}}, \tag{10}$$

where N is the number of samples, Y_t is the actual value at time t, and Y_t' is the model output value at time t.

A three-layer LSTM-RNN neural network is established in the paper, including input layer, hidden layer, and output layer. Since the input variable includes 4 feature indicators, the number of input layer neurons is 4. Only one of the output variables is the cognitive load level, so the number of neurons in the output layer is determined to be 1.

The method for determining the number of hidden layer neurons is as follows: firstly, a single hidden layer of LSTM-RNN is built and the fixed training parameters are set (learn rate set to be 0.001, batch size set to be 10, time step set to be 20, and number of iterations set to be 50). The training of the same number of iterations is performed under different neuron numbers, and the number of neurons with the smallest error is selected as the number of neurons in the first hidden layer. Secondly, the number of hidden layers increases gradually, and the above process is repeated until the degree of error reduction does not change. Namely, the number of hidden layers and neurons is determined. The error calculation method is shown in formulas (9) and (10). The test errors for the single hidden layer, the second hidden layer, and the third hidden layer are exhibited in Figure 16.

As can be clearly seen from the Figure 16(a), it is the most desirable when the number of neurons in the first hidden layer is selected to be 6, with the minimal error. Similarly, it is appropriate to set the number of the second and third hidden layer neurons to be 6 and 5, respectively, from the Figures 16(b) and 16(c). In summary, when the model has only one hidden layer, the optimal neural network structure is 4-6-6-1. When the model has two hidden layers, the optimal structure is 4-6-6-5-1. As depicted in Figure 17, test errors of the two structures are almost the same. Since the training time of 4-6-6-1 is smaller, the network structure of this model is 4-6-6-1.

The method of determining the parameters such as the learning rate, the time step, the batch size, and the number of iterations is the same as one of determining the number of hidden layer nodes. Based on the determined network structure, test errors are compared by changing the size of the parameters. The test error is calculated according to the formulas (9) and (10), and the corresponding parameter size is the best when the selection error is minimal. The test error varies with the value of each parameter as displayed in Figure 18.

The number of sample data used in the model for each training process is the batch size. Under normal

TABLE 6: Pearson correlation matrix.

	Indicator category	MEAN	SDNN	RMSSD	$p\text{NN}_{50}$	LF	HF	(LF/HF)	Total
MEAN	Pearson correlation	1	0.951*	−0.993**	0.975**	0.972**	−0.985**	0.984**	0.432
	Significant (bilateral)	—	0.013	0.001	0.005	0.006	0.002	0.002	0.467
SDNN	Pearson correlation	0.951*	1	−0.940*	0.899*	0.899*	−0.920*	0.979**	0.311
	Significant (bilateral)	0.013	—	0.017	0.038	0.038	0.027	0.004	0.611
RMSSD	Pearson correlation	−0.993**	−0.940*	1	−0.974**	−0.953*	0.972**	−0.986**	−0.456
	Significant (bilateral)	0.001	0.017	—	0.005	0.012	0.006	0.002	0.440
$p\text{NN}_{50}$	Pearson correlation	0.975**	0.899*	−0.974**	1	0.988**	−0.994**	0.938*	0.622
	Significant (bilateral)	0.005	0.038	0.005	—	0.002	0.001	0.018	0.262
LF	Pearson correlation	0.972**	0.899*	−0.953*	0.988**	1	−0.997**	0.923*	0.579
	Significant (bilateral)	0.006	0.038	0.012	0.002	—	0.000	0.025	0.306
HF	Pearson correlation	−0.972**	0.889*	0.972*	−0.994**	−0.997**	1	−0.947*	−0.562
	Significant (bilateral)	0.002	0.027	0.006	0.001	0.000	—	0.015	0.324
(LF/HF)	Pearson correlation	0.984**	0.979**	−0.986**	0.938*	0.923*	−0.947**	1	0.346
	Significant (bilateral)	0.002	0.004	0.002	0.018	0.025	0.015	—	0.568
Total	Pearson correlation	0.432	0.311	−0.456	0.622	0.579	−0.562	0.346	1
	Significant (bilateral)	0.467	0.611	0.440	0.262	0.306	0.324	0.568	—

*Significantly correlated at the 0.05 level (both sides); **significantly correlated at the 0.01 level (both sides).
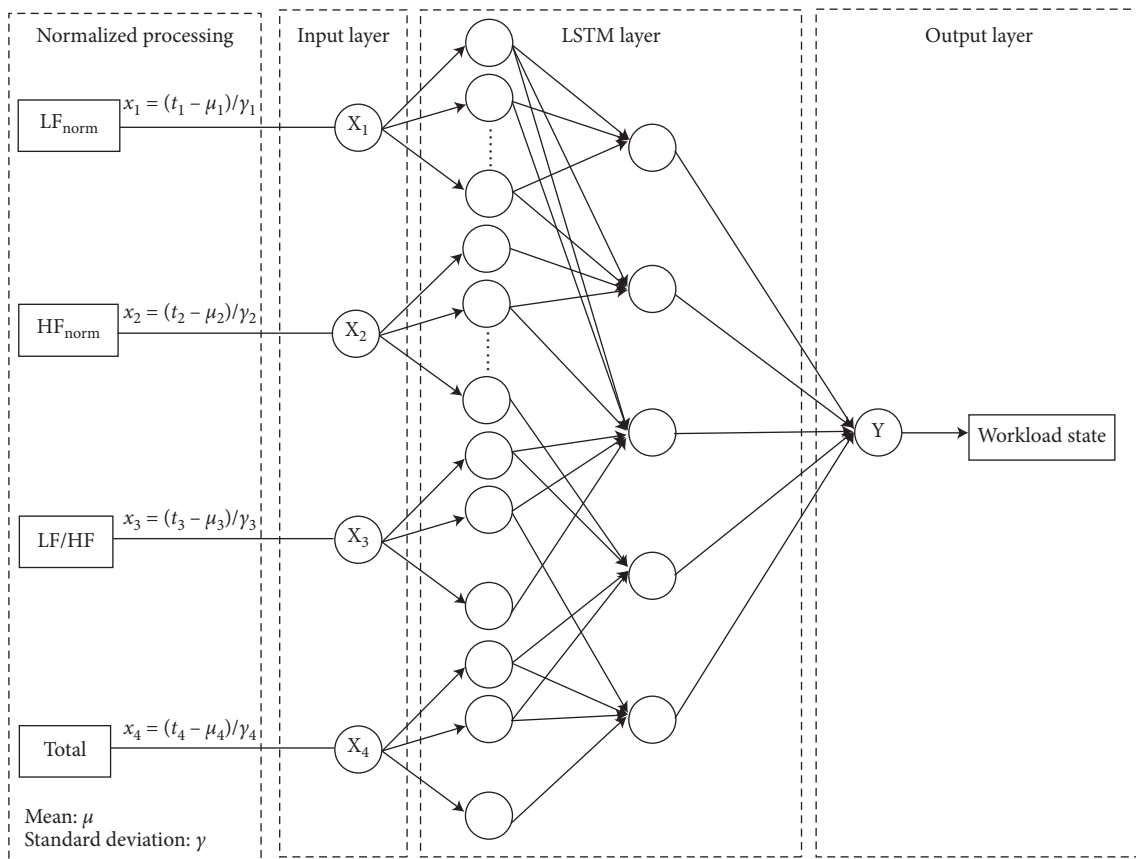

FIGURE 15: Cognitive load identification model based on LSTM-RNN.

circumstances, the test error will decrease with the increase of batch size and the downward trend will be slower. Therefore, the batch size is determined to be 10 according to the actual training environment. It can be seen from Figures 18(a) and 18(b) that the test error is the smallest when the learning rate is 0.001 and the time step is 40. As can be seen from Figure 18(c), the training error becomes smaller as the number of iterations increases. When the number of iterations is less than 30, the training error decreases rapidly. As the number of iterations increases, the training error decreases gradually. When the number of iterations reaches 80 or so, the training error tends to
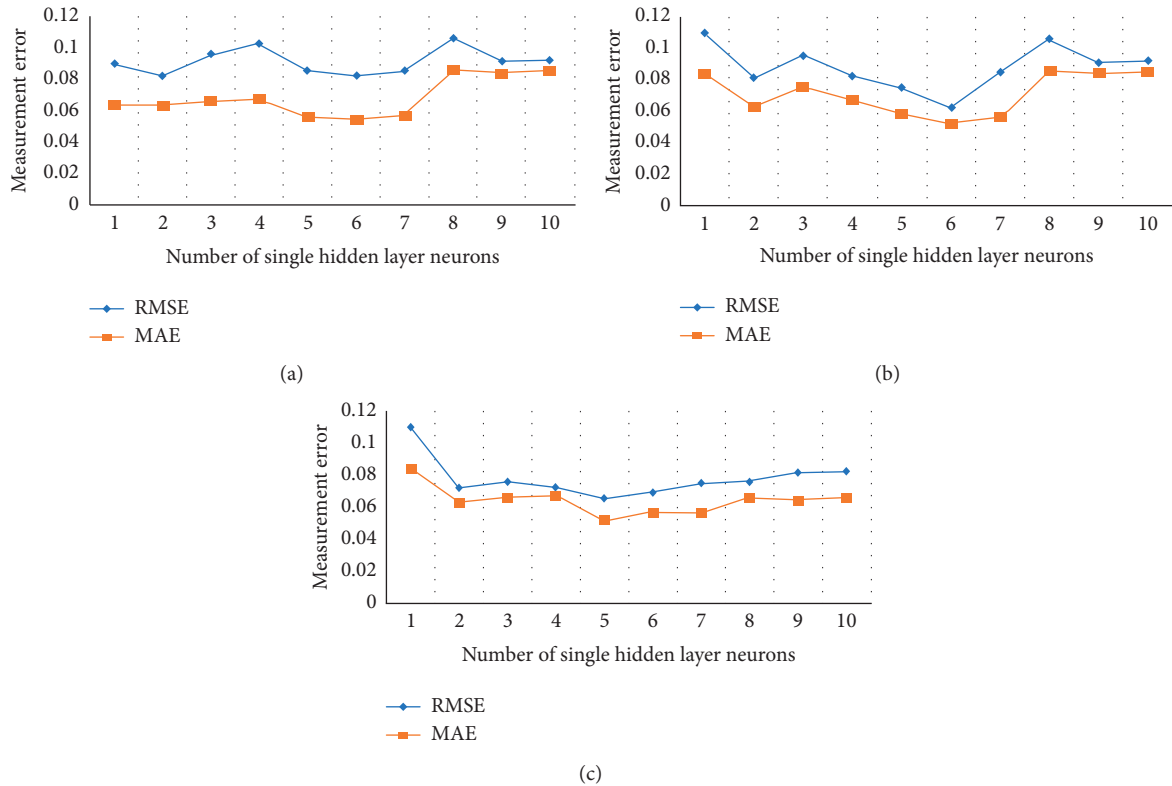
(a)

(b)

(c)

FIGURE 16: Training error with the number of neurons. (a) Single hidden layer neuron change chart. (b) Second hidden layer neuron change chart. (c) Third hidden layer neuron change chart.
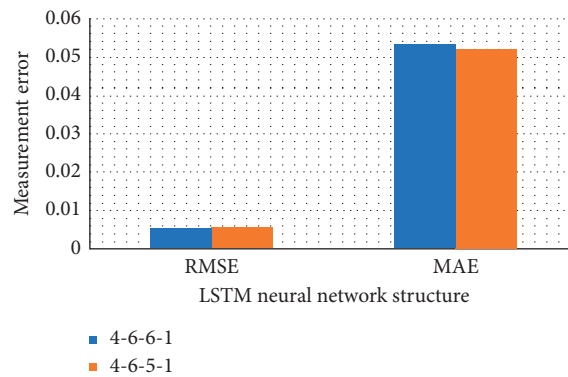

FIGURE 17: Comparison of two network structures.

converge. Therefore, the number of iterations selected in this experiment is 100. The remaining parameters include learning rate which is set to be 0.001 and time step with 40.

3.3.2. LSTM-RNN Performance. After the model training is completed, the test set is substituted into the model to obtain the test result. At the same time, this paper uses a 5-fold cross-validation method for model evaluation. The confusion matrix of the training set and the cognitive load level identification result on the test set of the model is shown in Figure 19. The accuracy of the training set and the test set was 79.9% and 77.6%, respectively. There was also no

systematic bias in the sensitivity (true positive rate) and specificity (true recessive rate) of the training set and test set.

4. Discussion

Cognitive load is an important factor affecting the pilot's behavior. It is not only a significant prerequisite to identify the pilot's cognitive load under the complex environment for realizing active vehicle safety controlling but also an important content for preventing flight accident. Through a large number of simulated flight experiments, it can be clearly obtained for the physiological and psychological data
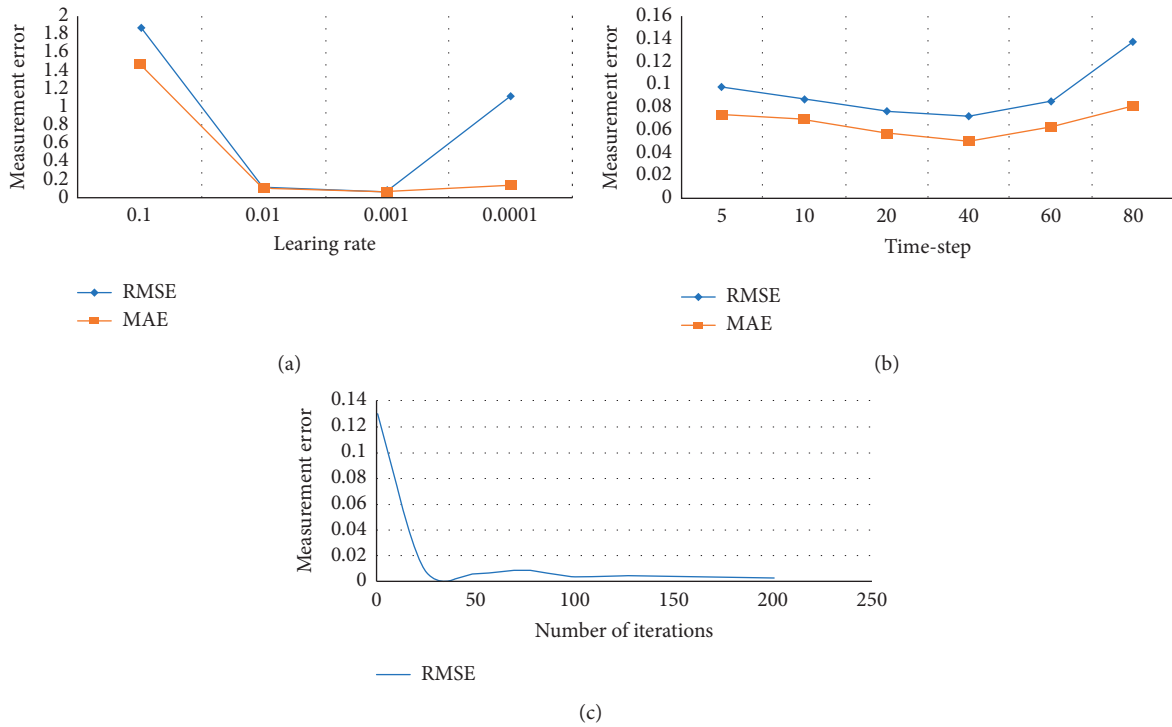
Figure 18: Test error variation chart of each parameter. Test error with (a) learning rate change chart, (b) time-step change chart, and (c) iteration number change chart.
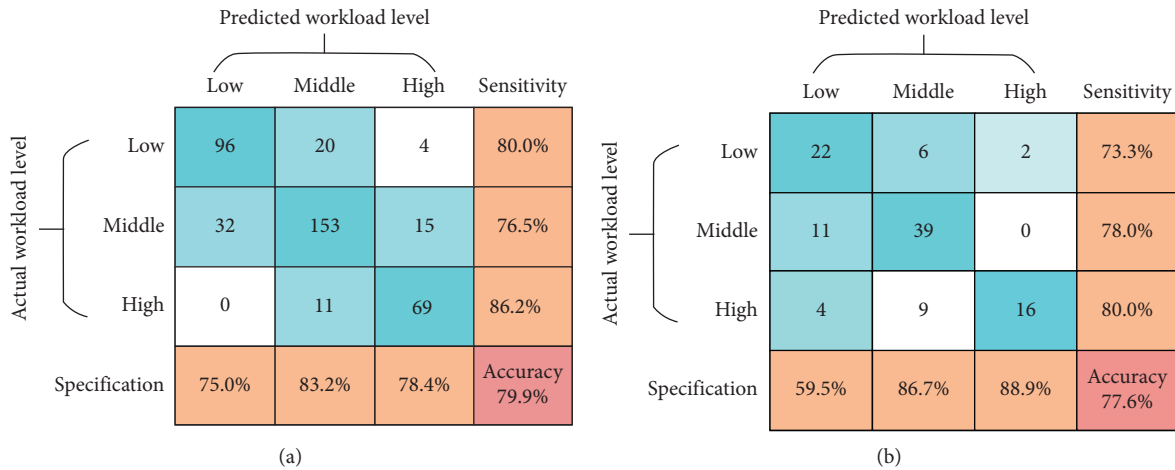

Figure 19: The confusion matrix of model classification result. (a) Training set identification accuracy (b) Test set identification accuracy.

of the pilot during the airfield traffic pattern. In this paper, the correlation analysis is performed on the basis of the preprocessing of each characteristic data, and the cognitive load identification model is established using LSTM-RNN neural network structure.

Different from the traditional RNN network structure, LSTM-RNN neural network presented in this paper has great advantages. As main reflection in the ability to process longer time-series data, the degree of fitting to the data is

strong. As exhibited in Figure 20, compared the traditional RNN neural network, support vector machine and other models or only using physiological indicators to identify the cognitive load, and the model adopted in this paper is more accurate.

Due to the strong correlation between the model input variables, it is easy to cause unnecessary errors in the identification results. In order to avoid this situation, physiological and psychological indicators have been
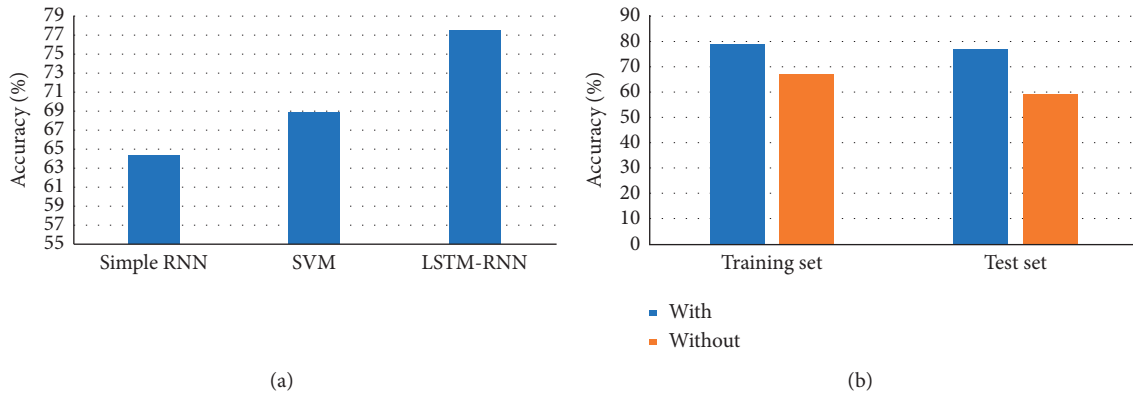
Figure 20: Model accuracy comparison. (a) LSTM-RNN unit performance (b) Impact of the multisource data.

screened through Pearson correlation analysis. Finally, four independent indicators were selected as input variables of the model.

The output of the cognitive load identification model is the cognitive load level of flying cadets. Because of the difficulty in analyzing the cognitive load, this paper combines the actual scores according to the total score of NASA-TLX scale. Results of model identification and NASA-TLX scale test were compared, proving that the model is accurate.

Pilot's physiological and psychological characteristics in complex environments are explored in this paper, and pilots' cognitive load based on these characteristics is studied. It makes up for the lack of pilots' microbehavior study in the field of flight traffic safety and provides a basis for further research on pilots' cognitive load. Nevertheless, there are still some deficiencies in this study: (1) there is a general link between physiological signals and cognitive load. This article only deals with and screens various indicators of ECG signals. Physiological indicators such as EEG signals, skin electrical signals, and vision are also sensitive to changes in cognitive load. Because of the energy limitation, this article will not be considered for the time being. (2) This paper measures the psychological characteristics of flying cadets based on the NASA-TLX scale. The flying cadets fill in according to their true feelings and thus have certain subjectivity. (3) Due to differences in individual characteristics, the workload perception level of the same mission may result in different results depending on the individual between the flight participants. (4) Some external factors will bring a certain amount of error for the experiment results. For example, there are differences in the technical level between the simulated flight and the real flight in the flight simulation environment, which have an impact on the experimental results. In order to achieve higher identification accuracy, the analysis of input variables affecting the pilot's cognitive load identification model from multiple perspectives is the focus of follow-up research. More detailed pilot's differences need to be further explored. The influence of subjective and external factors on the experiment results should be studied in subsequent studies.

5. Conclusion

This paper explores the cognitive load of pilot based on their physiological and psychological indicators. A cognitive load identification model for pilot based on LSTM-RNN was established. By considering the correlation and saliency between the indicators, the four indicators are finally determined as the input of the identification model. These include three characteristics of the ECG signal (low-frequency power LF, high-frequency power HF, and their ratio (LF/HF)) and one psychological characteristic indicator (Total). It was verified by low-level simulator experiments, and the model achieved 79.9% training set accuracy and 77.6% test set accuracy.

In this paper, various indicators of ECG signals are significantly correlated with cognitive load. In addition, the cognitive load level of pilots on each side of the airfield traffic pattern is also different. Among them, the load on the final leg is the largest and on the downwind leg it is the smallest. Indicators such as RR interval MEAN, RR interval SDNN, LF, and (LF/HF) have maximum values at the final leg and minimum values at the downwind leg; RMSSD and HF indicators have a maximum at the downwind leg and a minimum at the final leg. In summary, as the cognitive load increases, MEAN, SDNN, LF, and (LF/HF) will increase; in the same time, RMSSD and HF indicators will decrease.

In this paper, the cognitive load of pilot was accurately identified. It also explores their cognitive load changes during each phase of the airfield traffic pattern mission. Through the analysis of the dynamic characteristics of pilots under different cognitive load in complex flight environment, there was a comprehensive understanding of the psychological and physiological characteristics of the flying cadet. This study provides a good foundation for further research on the identification of the pilot's cognitive load and provides a basis for pilots' microbehavior analysis. The present study helps the pilots' ability to cope with missions that overcome load and ensure safe flight. It is of great practical significance to research the future active safety warnings for flight, pilot's microbehavior assessment, and flight safety.

Data Availability

The part of data used to support the findings of this study is included within the article. The data contain personal privacy information of the participants in the experiment. In order to protect the privacy of the participants, data are not disclosed without permission.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Job, *Air Disaster*, Aerospace Publication, Fyshwick, Australia, 1994.

[2] S. Shappell, D. A. Wiegmann, and A. Reason, "The human factors analysis and classification system," *Human Factors and Aerospace Safety*, vol. 1, pp. 59–86, 2001.

[3] R. Carretta, "USAF Pilot selection and classification systems," *Aviation Space & Environmental Medicine*, vol. 60, no. l, pp. 46–49, 1989.

[4] L. Ben and O. H. David, "The shape of aviation psychology: a review of 4 articles published in the first 5 years of the international journal of aviation psychology," *The International Journal of Aviation Psychology*, vol. 10, no. 1, pp. 1–11, 2000.

[5] J. Gregory, "Feist. The past and future of psychology of science," *Review of General Psychology*, vol. 10, no. 2, pp. 92–97, 2006.

[6] J. B. Noel, K. W. Bauer, and J. W. Lanning, "Improving pilot mental workload classification through feature exploitation and combination: a feasibility study," *Computers & Operations Research*, vol. 32, no. 10, pp. 2713–2730, 2005.

[7] R. Gray, J. Gaska, and M. Winterbottom, "Relationship between sustained, orientated, divided, and selective attention and simulated aviation performance: training & pressure effects," *Journal of Applied Research in Memory and Cognition*, vol. 5, no. 1, pp. 34–42, 2016.

[8] H. Mansikka, K. Virtanen, D. Harris, and P. Simola, "Fighter pilots' heart rate, heart rate variation and performance during an instrument flight rules proficiency test," *Applied Ergonomics*, vol. 56, pp. 213–219, 2016.

[9] T. Andrew and H. G. Robert, "Level of operator control and changes in heart rate variability during simulated flight maintenance," *The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 4, pp. 682–698, 1995.

[10] A. Jonathan and G. Rob, "Flying under pressure: effects of anxiety on attention and gaze behavior in aviation J," *Journal of Applied Research in Memory and Cognition*, vol. 3, no. 2, pp. 63–71, 2014.

[11] A. Tjolleng, K. Jung, W. Hong et al., "Classification of a driver's cognitive workload levels using artificial neural network on ECG signals," *Applied Ergonomics*, vol. 59, pp. 326–332, 2017.

[12] G. J. Wang, H. L. Wu, Y. C. Sun et al., "Optimization design of human-machine display interface based on cognitive characteristics," *Aircraft Design*, vol. 37, no. 3, pp. 55–67, 2017.

[13] B. Gianluca, A. Laura, V. Giovanni et al., "Measuring neurophysiological signals in aircraft pilots and car drivers for theassessment of mental workload, fatigue and drowsiness," *Neuroscience and Biobehavioral Reviews*, vol. 44, pp. 58–75, 2014.

[14] S. Scannella, V. Peysakhovich, F. Ehrig, E. Lepron, and F. Dehais, "Assessment of ocular and physiological metrics to discriminate flight phases in real light aircraft," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 60, no. 7, pp. 922–935, 2018.

[15] H. M. Antonio, M. Damien, M. Nadine et al., "Cardiovascular correlates of emotional state, cognitive workload and time-on-task effect during a realistic flight simulation," *International Journal of Psychophysiology*, vol. 128, pp. 62–69, 2018.

[16] X. Wanyan, D. Zhuang, Y. Lin, X. Xiao, and J.-W. Song, "Influence of mental workload on detecting information varieties revealed by mismatch negativity during flight simulation," *International Journal of Industrial Ergonomics*, vol. 64, pp. 1–7, 2018.

[17] U. Manawadu, T. Kawano, S. Murata, M. Kamezaki et al., "Multiclass classification of driver perceived workload using Long short-term memory based recurrent neural network," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, 2018.

[18] H. Kim, M. Park, C. W. Kim, and D. Shin, "Source localization for hazardous material release in an outdoor chemical plant via a combination of LSTM-RNN and CFD simulation," *Computers & Chemical Engineering*, vol. 125, no. 9, pp. 476–489, 2019.

[19] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): results of empirical and theoretical research," *Advances in Psychology*, vol. 52, pp. 139–183, 1988.

[20] Civil Aviation Administration of China, *Rules for the Certification of Pilots and Ground Instructors of Civil Aircraft*, Civil Aviation Administration of China, Beijing, China, 2014.

[21] E. B. Lin, M. Haske, M. Smith et al., *Wavelet Analysis of ECG Signals, Multiscale Signal Analysis And Modeling*, pp. 105–110, Springer, New York, NY, USA, 2013.

[22] L. Gustavo, P. Nicolas, O. Tobias et al., "P wave detection and delineation in the ECG based on the phase free stationary wavelet transform and using intracardiac atrial electrograms as reference," *Biomed. Technik/Biomedical Engineering*, vol. 61, no. 1, pp. 7–56, 2016.