



Research article

Driver fatigue detection using PPG signal, facial features, head postures with an LSTM model

Lu Yu ^{a,*}, Xinyi Yang ^a, Hengjian Wei ^a, Jianguo Liu ^b, Bo Li ^a^a School of Traffic and Transportation Engineering, Dalian Jiaotong University, Liaoning, Dalian, 116028, China^b School of Civil and Resource Engineering from University of Science and Technology Beijing, Beijing, 100083, China

ARTICLE INFO

Keywords:

Driver fatigue
PPG signals
LSTM model
Facial feature
Head posture

ABSTRACT

Background and objective: Human fatigue is a major cause of road traffic accidents. Currently widely used fatigue driving detection methods are based on eyelid closure, vehicle information or physiological parameter detection. However, the detection of each single feature has certain limitations. Which in turn affects the accuracy of detection and the possibility and efficiency of prediction.

Methods: This paper introduces a novel driver fatigue detection framework that leverages facial features, head pose, and PPG signals to establish a fatigue detection model. To validate this approach, a real-road driving experiment was conducted, resulting in the acquisition of multi-source feature signal data from 30 drivers. Utilizing a method for locating 68 facial landmarks, we extracted 2D facial and 3D head feature parameters. Additionally, five-dimensional heart rate variability (HRV) features were extracted from PPG signals. These ten-dimensional features were fused to construct a fatigue driving dataset. Subsequently, a Long Short-Term Memory (LSTM) network model for fatigue detection was established and optimized using four optimization algorithms: Momentum, Rmsprop, Adam, and SGD. For comparison, Decision Tree (DT), Random Forest (RF), and Bidirectional LSTM (BiLSTM) models were also evaluated. Within the dataset, 2880 samples were designated as the training set, while 720 samples served as the test set.

Results: Adam's optimized LSTM fatigue detection model is the most effective, with a model accuracy of 97.36 %, precision of 97.4 %, recall of 97.4 %, and F1 of 0.97. It shows that the model can provide a more timely and accurate prediction and warning for drivers who are already fatigued.

1. Introduction

According to the National Bureau of Statistics of China (NBS), the total number of traffic accidents in China in 2021 was 273,000, resulting in 62,000 deaths [1]. These accidents occur due to a variety of factors such as drivers, vehicles, roads, traffic conditions, etc. The causes and severity of these accidents depend on a variety of factors. Analysis of accident causation reports shows that driver fatigue is the most important factor among multiple factors, which is responsible for more than 90 % of personal safety accidents [2]. NHTSA (National Highway Safety Administration) announced that driver fatigue caused about 1.9 % of driving fatalities (697) to occur in 2019. In another statistical report by NHTSA, it was announced that in 2017, 91,000 traffic police officers reported that crashes

* Corresponding author.

E-mail address: yulubaobeihao@163.com (L. Yu).

related to driver fatigue resulted in about 50,000 injuries and 800 deaths. Research data from the Chinese Automobile Federation Association (CAFA) proves that drivers can become fatigued after driving for more than 2.4 h that fatigued driving occurs 3–5 times per month, and that drivers with more than 3 years of driving experience have been involved in at least 1–2 collisions caused by fatigued driving. Fatigued driving has become one of the most important causes of traffic accidents, which poses a great threat to road safety. Definition of fatigue driving is a physiological and psychological stress effect produced by the human body when at least one of the three conditions occurs, namely, drivers driving continuously for more than a threshold time [3], driving under sleep deprivation [4], and driving under excessive physical exertion [5], and the behavioral characteristics of this stress effect are manifested as reduced attention, slow reaction, and errors in judgment (see Table 1).

Current fatigue driving detection methods include subjective detection, unimodal feature detection, and multimodal feature detection. Subjective detection is the assessment of fatigue using various questionnaires and subjective fatigue rating scales. The disadvantage of this method is that individual differences are very large. And it takes a period to complete the measurement. This length of time misses the possibility of preventing accidents. Usually, three types of data are used in the literature to design unimodal or multimodal fusion fatigue driving detection systems: (1) Vehicle information [6] (2) Vision-based information such as eyelid opening and closing, mouth aspect ratio [7] (3) Physiological data such as EEG [7], ECG [8]. The pattern of generation and development of human fatigue can be expressed in behavior. The early stage is manifested by a significant increase in the frequency of yawning and nodding, middle and late stages will show significant lateral vehicle displacement, drowsiness, and other physiological symptoms. Therefore, it is inaccurate to use vehicle data information to calculate and measure driving fatigue before warning the driver. The advantage of information from the visual system is that it is easy to access. The disadvantage is that data stability is affected by lighting conditions. When the driver wears glasses or sunglasses, the accuracy of the detection decreases dramatically. In summary, combining physiological parameters with facial feature information has been shown to yield highly reliable driver fatigue detection methods. Furthermore, existing research has demonstrated the potential of multi-modal feature fusion in fatigue driving detection [9–13].

PPG is a relatively recent approach for measuring human physiological information. Its principle involves utilizing photoelectric signal conversion to trace variations in blood flow within blood vessels, thereby enabling the estimation of physiological parameters such as pulse, blood oxygen saturation, and heart rate. This serves as the theoretical foundation for PPG technology in assessing fatigue. Medical professionals employ this technology to predict the onset of cardiovascular and cerebrovascular diseases. The merits of PPG include its affordability, portability, high efficiency, and non-invasive nature. Recent studies have shown innovative applications of PPG. For instance, Lee et al. [14] integrated PPG and electrocardiogram (ECG) signals through a fuzzy kernel clustering algorithm, extracting feature vectors via wavelet transform. Erick Javier Argüello-Prada et al. [15] proposed a method for estimating respiratory rate using PPG, demonstrating its significant contribution to this measurement. S. M. Taslim Uddin Raju et al. [16] introduced a convenient blood pressure measurement technique leveraging deep learning for feature extraction, revealing that PPG alone can facilitate non-invasive blood pressure estimation. Yiming Zhang et al. [17] further advanced this field by estimating continuous blood pressure trends across diverse populations using single-channel PPG, enhanced by a BiGRU attention network that captures the temporal dependencies within PPG waveforms.

This article proposes a method for identifying driving fatigue. Its technical core lies in the integration of three types of data information: facial features, head posture, and heart rate characteristics. PPG sensor technology is utilized for the collection of heart rate characteristic data, while video images are employed for the acquisition of facial and head feature data. There are three main technical challenges in this method.

1. Division and integration of data units. Facial and head posture data originate from videos, which require processing into image information. PPG data, on the other hand, is time-series data. It is difficult to integrate facial, head, and PPG data within the same dimension and cycle.
2. Difficulty in measuring experimental data. Most existing publicly available datasets are driving fatigue data measured through simulation driving devices. There is a scarcity of real-world driving datasets. This article conducted real-world driving experiments to obtain a real-world driving dataset.
3. Selection of a fatigue detection model. Given the large amount of feature data collected in this article, choosing a machine learning model with high efficiency, superior performance, and accuracy becomes a challenge.

This discussion is applicable to traffic safety scenarios and is translated into professional English with formal and specialized terminology.

The research presented in this paper was conducted in real-world driving scenarios, leveraging sensors worn by drivers to collect human physiological signals. We constructed a database of driver fatigue states and analyzed the changing patterns within the database under varying degrees of driver fatigue. The aim is to identify characteristic signs of human fatigue and provide timely

Table 1
Background questionnaire for experimental personnel.

Personnel characteristics	Average value	Standard deviation	range	proportion
Age	32	1.15	23–47	–
Driving experience	8	0.99	5–12	–
Annual mileage traveled	90000	110.5	80000–120000	–
Sex ratio	–	–	–	5/1

warnings. Our study provides both theoretical and practical foundations for the future establishment of a predictive and warning system for drivers' hazard perception.

2. Experimental setup

2.1. Date collection and study procedure

2.1.1. Location

The experimental site was a section of a real-world highway located in Dalian, Liaoning Province, China, specifically the highway connecting Dalian to Shenyang. This section of the road features long straight stretches and two lanes in each direction. It is equipped with standard traffic facilities and signaling devices. The starting point of the experimental section is the G15 Shenyang-Haikou Expressway entrance (in the direction of Shenyang), and the endpoint is the G1501 Shenyang Ring Expressway toll station, covering a total distance of 310 km, with a testing driving time of approximately 3 h (see Fig. 1).

2.1.2. Device

The experimental equipment consists of a PPG signal sensor (MAX30101) for collecting human physiological signals and an infrared camera, as shown in Fig. 2. The MAX30101 sensor is a highly integrated physiological parameter monitoring sensor that combines infrared and red LED light sources with an optical detector [18]. The MAX30101 sensor can monitor heart rate and blood oxygen saturation in real time. Its advantages include dual-wavelength measurement, low-power design, strong integration, high-precision measurement capabilities, and digital interfaces. These merits make it an ideal choice for health monitoring devices. In this experiment, the MAX30101 sensor is used to acquire crucial physiological data on heart rate variability. The facial data for this experiment is captured by an infrared camera, which has the advantage of emitting infrared light to obtain infrared facial images, even in dark environments.

2.1.3. Participants

According to the survey report on Chinese truck drivers, the current age of truck drivers is concentrated between 36 and 45 years old, accounting for 48.7 % of the surveyed truck drivers [19]. Among the surveyed drivers, males account for a high proportion of 98.2 %, indicating that truck driving is a predominantly male occupation. The experimental participants were recruited from various sectors of society, mostly male truck drivers, with extensive experience in long-distance driving. The following are the eligibility requirements for participants:

- (1) They must possess a valid Chinese driver's license.
- (2) They should not have any physical conditions that may cause them to become ill due to driving tests, such as motion sickness, dizziness, epilepsy, or other adverse reactions, and vice versa. A total of 30 participants (25 males and 5 females) were recruited for the driving test. The participants ranged in age from 23 to 47 years old, with an average age of 32 years and a standard deviation of 1.15 years.



(a) Experimental route satellite map (b) Highway traffic conditions map (c) View from the passenger seat perspective

Fig. 1. Real way map. (a) Experimental route satellite map (b) Highway traffic conditions map (c) View from the passenger seat perspective.

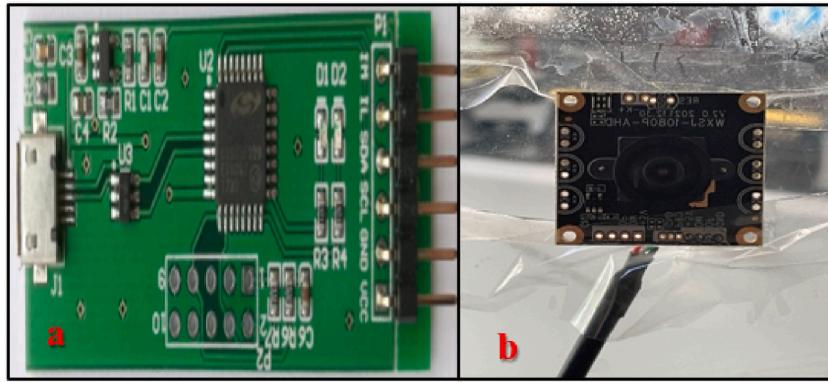


Fig. 2. MAX30101 sensor (a) and infrared camera (b).

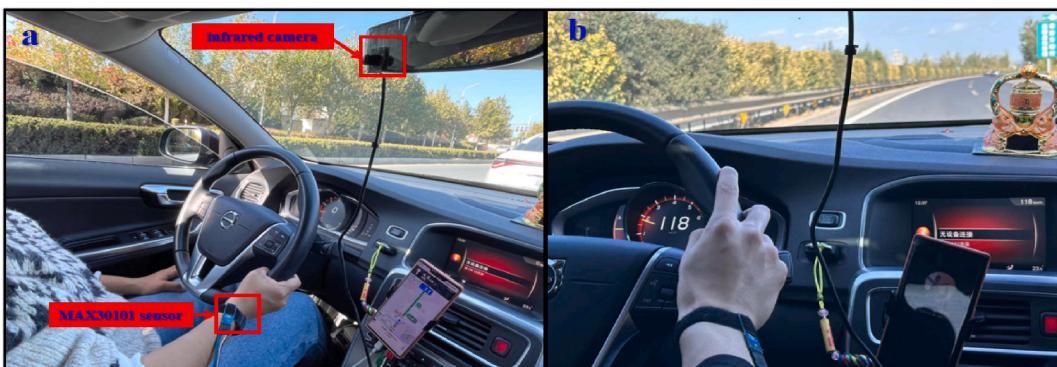
2.1.4. Experimental procedure

- (1) Experimental Preparation. To ensure the authenticity of the experiment, the drivers' natural circadian rhythms were maintained unaltered. On the day of the experiment, drivers were prohibited from consuming alcohol, tea, and coffee in the morning. A highly experienced professional was seated in the passenger seat, acting as an observer who constantly monitored the driver's state to promptly take control of the vehicle in case of emergencies, thereby preventing traffic accidents and safeguarding the driver's personal safety. Zhendong Lan et al. have found that drivers are more prone to fatigue between 14:00pm and 16:00pm compared to 10:00am to 12:00am [20]. Consequently, the driving experiments were conducted during the most fatigue-prone period for drivers, from 14:00pm to 17:00pm, lasting for 3 consecutive hours without interruption.
- (2) Commencement of the experiment. The experimenter wore the MAX30101 sensor on their wrist, which continuously recorded the participant's PPG signals. Simultaneously, a camera was activated to record the driving video from the cab throughout the experiment. The participant drove continuously for 3 h, adhering to normal traffic rules. Every 10 min, an evaluator in the passenger seat filled out a fatigue assessment scale based on verbal inquiries and recorded the responses.
- (3) End of the Experiment. The fatigue evaluation scales filled out by the experimental participants, as well as the PPG signals and facial video signals recorded over 3 h, were collected and organized. A fatigue-driving dataset was established. The actual situation in the cab during the experiment is illustrated in Fig. 3.

2.2. Determination of fatigue levels

In this study, a comprehensive assessment of the driver's fatigue status was conducted using a subjective evaluation method that combines self-evaluation and third-party evaluation (see Fig. 4). This study refers to the Karolinska Sleepiness Scale to construct the "Driver Fatigue Status Scale" (see Table 2 for details). This scale divides the driver's fatigue status into three levels: Level 1, Level 2, and Level 3. Each level is accompanied by corresponding descriptions of physical and mental fatigue manifestations, providing important references for self-evaluation and third-party evaluation.

The determination of fatigue levels consists of two parts: the driver's subjective self-evaluation L_1 and the evaluation by others L_2 . L_1 is obtained through verbal inquiries about the driver's driving condition during the experiment, while L_2 is determined by the



(a) Equipment Installation

(b) Driving Process

Fig. 3. Driving scenario. (a) Equipment installation (b) Driving process.

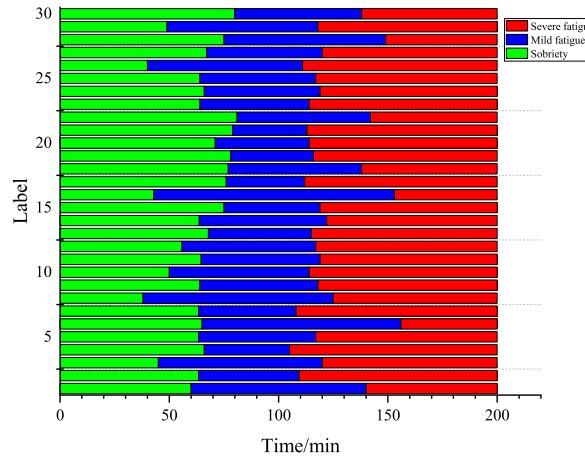


Fig. 4. Distribution of fatigue state of experimental personnel.

observer after the experiment through analyzing the experimental driver's driving condition in video recordings and referring to the fatigue level classification table. When the driver is in a state of deep fatigue, their consciousness may not be sufficiently clear to evaluate their own fatigue status, thus the fatigue level L_1 obtained through the experimental driver's subjective self-evaluation cannot be used as the final evaluation criterion. Therefore, to obtain the most accurate fatigue state of the driver, this study sets the weight of the experimental driver's subjective self-evaluated fatigue level L_1 to 0.3 and the weight of the fatigue level L_2 obtained by the observer through observing the driver's driving condition to 0.7 when determining the final fatigue level. This means that when there is a discrepancy between the two evaluations, the observer's evaluation is more likely to be adopted. The calculation formula for the final evaluated fatigue level L is as follows:

$$L = \text{int}(0.3 \times L_1 + 0.7 \times L_2)$$

L — Final evaluated fatigue level value

L_1 — Self – evaluated fatigue level value

L_2 — Fatigue level value evaluated by others

int — Integer rounding function.

The final fatigue level distribution of 30 drivers, determined through the calculation formula, is illustrated in the figure. As can be seen from the figure, six drivers already exhibited mild fatigue within the first hour of the experiment, while another six drivers developed mild fatigue after 60 min. By the 150th minute of the experiment, only three drivers had not reached a severe fatigue state, while the remaining 27 drivers had all entered a severe fatigue state.

3. Methodology

3.1. PPG signal preprocessing

The actual PPG signals measured are often characterized by relatively amplitudes small, necessitating amplification through an amplifier. Furthermore, during the measurement process, these signals are susceptible to interference from various sources of motion, including hand movements, respiratory rates, mechanical vibrations, ambient light, and electromyographic signals, leading to the presence of noise in the signals [21]. To extract data that can represent human cardiovascular information and heart rate variability components, it is imperative to preprocess the PPG signals.

The frequency of human pulse signals lies within the 0.1–20 Hz band, with the majority of energy concentrated near the fundamental frequency of the pulse wave (approximately 1 Hz), accounting for approximately 95 % of the total energy. Frequency components within this range primarily reflect the characteristics of the cardiac pressure wave [22]. To filter the acquired pulse signals, we employed a Butterworth digital filter implemented in MATLAB, whose transfer function is detailed in Equations (3-1).

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_N z^{-N}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_N z^{-N}} (3-1)$$

In the equation, b_0, b_1, \dots, b_N and a_0, a_1, \dots, a_N represent the coefficients of the filter. The filtering frequency range is set between 0.1 Hz and 20 Hz, with a sampling frequency of 100 Hz. A high-pass digital filter with a cutoff frequency of 0.1 Hz is applied to eliminate baseline drift caused by respiration and finger pressure variations. The changes during data preprocessing are illustrated within the red

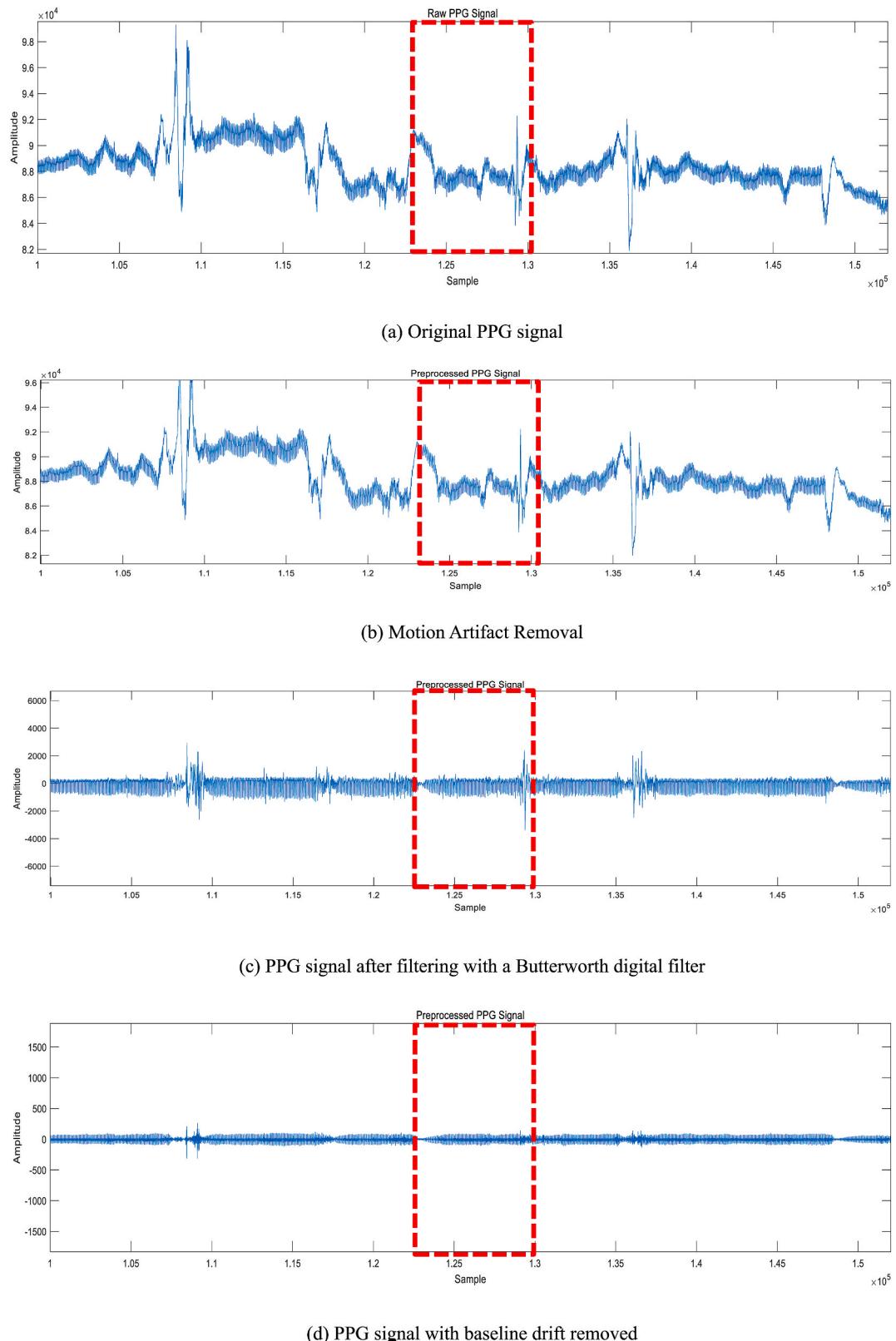


Fig. 5. PPG signal processing.

box in Fig. 5.

3.1.1. HRV feature extraction

Heart rate variability (HRV) is regarded as an objective reflection of deeper-level information regarding an individual's physiological state and fatigue level [23]. HRV serves as a physiological indicator reflecting the activity of the cardiac autonomic nervous system [24]. It quantifies the variability in time intervals between consecutive heartbeats [25]. During the heartbeat process, the interplay between the sympathetic and parasympathetic nervous systems leads to variations in the heartbeat intervals, which HRV reflects [26]. Consequently, HRV can be utilized as a metric to assess autonomic nervous system activity. PPG signals are biological signals acquired through photoplethysmography [27], which records blood flow beneath the skin. Each heartbeat triggers a change in the volume of blood flowing beneath the skin, and PPG signals capture these variations. Therefore, PPG signals are employed to calculate heart rate and heart rate intervals. By conducting HRV analysis on PPG signals, we gain insights into the influence of the autonomic nervous system, enabling the assessment of participants' physiological states.

3.1.2. Time domain analysis

Time-domain analysis of HRV involves employing a series of statistical methods to describe and quantify the variability characteristics of heartbeat intervals (R-R intervals). Commonly used indices include standard deviation of normal-to-normal intervals (SDNN), HRV triangular index, mean of the standard deviations of all RR intervals for all 5-min segments of a 24-h recording (SDANN), and root mean square of the successive differences (RMSSD) between adjacent RR intervals [28]. Among these, the sinusoidal R-R interval, also known as the R-R interval, represents the time interval between two adjacent R waves during sinus rhythm, reflecting the regularity of the cardiac sinus rhythm, as illustrated in Fig. 6 below. In this study, SDNN and RMSSD are extracted as salient features for significant fatigue detection.

SDNN (Standard Deviation of NN intervals) is the standard deviation of all normal-to-normal (NN) intervals throughout the entire period. Its formula is (3-2):

$$SDNN = \sqrt{\frac{1}{N} \sum_{i=1}^N (RR_i - \bar{R})^2} (3-2)$$

In equations (3-2), N represents the total number of normal heartbeats for the driver, and RR_i represents the i th successive sinus R-R interval, which is the average value of the R-R interval in the heart rate feature.

RMSSD (Root Mean Square of Successive Differences) is the square root of the average of the squared differences between adjacent normal R-R intervals. The formula is (3-3):

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_{i+1} - RR_i)^2} (3-3)$$

3.1.3. Frequency domain

HRV frequency-domain analysis, also known as spectral analysis, involves transforming the temporal fluctuations of RR intervals into a superposition of different frequency components in the frequency domain to quantitatively assess the modulatory effect of the sympathetic nervous system on heart rate, thereby precisely evaluating the dynamic balance between sympathetic and vagal nerve activities within the cardiac autonomic nervous system [27] (see Table 2). In HRV frequency-domain analysis, HRV signals are

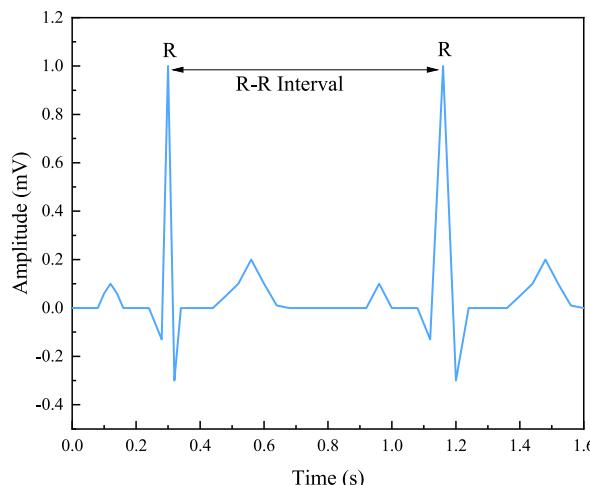


Fig. 6. Illustration of R-R intervals.

typically categorized into the following primary frequency bands, as outlined in [Table 3](#).

Among them, due to the extremely low frequency of the ultra-low frequency (ULF) band and the lack of a fully understood physiological significance for the very low frequency (VLF) band, both require significantly long recording periods for accurate analysis. Consequently, the low frequency (LF) and high frequency (HF) bands are adopted as the extracted features. An increase in LF power reflects enhanced sympathetic activity or decreased parasympathetic activity, whereas an increase in HF power typically indicates enhanced parasympathetic activity, indirectly reflecting the modulatory effect of respiration on heart rate.

For the sequence of N R-R intervals, denoted as $t(n)$, its frequency spectrum is represented by $F(w)$, and the power spectral density is denoted as $P(W)$, as defined in Equations (3-4):

$$P_{(w)} = |F_{(w)}|^2 (3 - 4)$$

LFP (Low Frequency Power) is the integral of the power spectrum obtained by performing frequency transformation on the time series of the N R-R intervals in the heart rate feature, specifically in the low frequency range (0.04Hz–0.15 Hz).

$$LFP = \sum P_{(w)} \quad 0.04 < w < 0.15 (3 - 5)$$

LFP/HFP is the ratio of LFP to HFP.

$$\frac{LFP}{HFP} = \frac{LFP}{HFP} (3 - 6)$$

We randomly selected heart rate variability data from five subjects for analysis: Subject 3, Subject 8, Subject 15, Subject 22, and Subject 28. Heart rate variability refers to the variability of the interval between heartbeats, which is quantified by calculating the standard deviation of the heartbeat intervals. The calculated values of SDNN, RMSSD, and LFP/HFP are shown in [Fig. 7](#). As driving duration increases, drivers' fatigue levels gradually intensify, resulting in a downward trend in SDNN and RMSSD, while LFP/HFP exhibits an opposite trend, increasing in value compared to the awake state under fatigue conditions. This is because fatigue can lead to disturbances in the autonomic nervous system, which primarily controls heart rate variability. An increase in sympathetic nerve activity and a decrease in parasympathetic nerve activity lead to an increase in heart rate and a decrease in heart rate variability.

3.2. Facial feature extraction

3.2.1. Facial feature point localization based on Dlib library

The Dlib library is an open-source data toolkit developed in C++ that encompasses numerous machine learning algorithms and models. When compared to TensorFlow and PyTorch, Dlib exhibits strong versatility and superiority in image processing, facial feature extraction, classification, and comparison [29–31]. In this study's experiments, the Python interface of Dlib was utilized to analyze a 3-h full-time domain recorded video using the shape_predictor_68_face_landmarks.dat model library. Through the Dlib-based facial landmark detector, keypoints were acquired, initially constructed from the original image, and then iteratively adjusted using a gradient boosting strategy until they aligned with the actual facial contours. At each stage, the least squares method was employed to calculate regression coefficients. Finally, the locations of these keypoints were plotted onto the image, marking 68 facial landmarks. These landmarks were arranged in a fixed order, yielding critical information such as facial landmark coordinates, face bounding boxes, and face angles, as illustrated in [Fig. 8](#). Concurrently, all 68 keypoints were detected. Utilizing this information, we can discern facial activity states and further derive essential metrics like the eye aspect ratio and mouth aspect ratio (see [Fig. 9](#)).

3.2.2. Eye feature extraction

The Eye Aspect Ratio (EAR) algorithm, introduced by Paulo Augusto de Lima Medeiros et al., in 2022, serves as a tool for detecting blink frequency [32]. By computing the EAR value, an assessment of ocular status can be derived. The specific methodology involves firstly utilizing a 68-point facial landmark algorithm to locate key facial feature points. Subsequently, the vertical-to-horizontal ratio (EAR) of the eyes is extracted based on the variations in distances between these feature points. The combined outcomes of these two steps provide insights into the current pupil condition, enabling the determination of a driver's fatigue level through EAR value fluctuations. A threshold of 0.2 for EAR has been established as a criterion for fatigue detection [33], which is adopted in this study as the standard for assessing fatigue status (see [Fig. 8\(b\)](#)). Specifically, the left eye's facial landmark indices range from 36 to 41, while those of the right eye span from 42 to 47. These left and right eye coordinates are correspondingly mapped to P1 through P6, with their detailed calculation formula presented in Equations (3-7):

$$EAR = \frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2 \|P_1 - P_4\|} (3 - 7)$$

Table 2
Fatigue rating classification scale.

Label	Fatigue state	Criteria for classification
1	Sobriety	Lively and energetic
2	Mild fatigue	Occasional yawning, blinking more frequently
3	Severe fatigue	Eyelids take longer to close and begin to doze off

Table 3

Signal frequency domain division.

Frequency Band	Range	
Ultra-Low Frequency (ULF)	0–0.0033 Hz	Variations in this frequency band correlate with prolonged physiological regulatory mechanisms, notably thermoregulation and hormonal fluctuations.
Very Low Frequency (VLF)	0.0033–0.04 Hz	Additionally, these variations are intimately linked to blood pressure regulation and the mechanical activities of the heart.
Low Frequency (LF)	0.04–0.15 Hz	The underlying mechanism of these variations is commonly attributed to the concerted action of the sympathetic and parasympathetic nervous systems.
High Frequency (HF)	0.15–0.4 Hz	Furthermore, this frequency band serves as an indicator of the modulatory influence of respiration on heart rate dynamics.

3.2.3. Mouth feature extraction

In this study, the Mouth Aspect Ratio (MAR) algorithm is employed to detect the frequency of yawning in participants [34]. The key facial landmarks around the mouth are captured using Dlib and OpenCV. The degree of mouth openness, represented as the aspect ratio of the mouth (MAR), is calculated based on the horizontal and vertical coordinates illustrated in Fig. 8. Specifically, the vertical coordinates 51, 53, 57, and 59 correspond to M2, M3, M5, and M2 (repeated as M5 was likely a typo for M6 or another distinct point, assuming consistency with facial landmark notation), while the horizontal coordinates 49 and 55 are associated with M1 and M4, respectively. Fig. 8(a, c) visually depicts these relationships. The detailed calculation formula is presented in Equations (3-8) as follows:

$$MAR = \frac{\|M_2 - M_6\| + \|M_3 - M_5\|}{2 \|M_1 - M_4\|} (3 - 8)$$

The data analysis of five randomly selected participants is presented in the figure. As shown, with the prolongation of driving time, the driver's fatigue level gradually intensifies, resulting in a downward trend in both EAR and MAR values. Among them, Subject 22's EAR value reached 0.2 at 68 min, which is considered the critical threshold for EAR-based fatigue detection. When evaluating fatigue using MAR values, there is no fixed threshold in different experimental scenarios. The physical meaning of the MAR value is the ratio of the vertical length to the horizontal length of the mouth. Therefore, frequent changes in MAR values represent frequent mouth opening and closing, which in this experiment correspond to the driver's yawning behavior, indicating the onset of driver fatigue. As depicted in Fig. 8(b), the MAR values of the drivers exhibited significant fluctuations within the first hour of the experiment, stabilized after 80 min, and then entered a second phase of frequent changes after 160 min. This suggests that the experimental drivers frequently yawned and entered a state of mild fatigue at the beginning of the experiment, followed by a decrease in yawning frequency, and then again started yawning frequently after 2 h. Clearly, the facial features of driving fatigue are diverse and cannot be solely determined by the number of eyelid closures and yawns. Rather, a comprehensive evaluation of multiple fatigue indicators should be employed.

The analysis of the entire driving video recorded by the camera was conducted in a data analysis environment of Python 3.6, OpenCV, and Dlib 19.7. The Python runtime environment utilized Anaconda. High-definition cameras were employed to record the driver's driving status. The driver's facial movement video stream was extracted, and the video stream was iteratively processed for frame reading. The debugging results are shown in Fig. 10, demonstrating smooth and clear video footage. The use of 68 landmarks accurately described the facial key point information, verifying the effectiveness of the camera module in capturing the dynamic facial data of the driver.

3.3. Head pose-based feature extraction

When a driver exhibits fatigue while driving, they may display frequent nodding, lowering of the head, shaking of the head, and other abnormal driving postures. Therefore, abnormal head angles can be used to determine whether a driver is in a state of fatigue driving [35]. Head pose refers to the direction of the head relative to the global coordinate system. The direction of head movement of ordinary people is generally limited to three degrees of freedom. These three angles are Pitch, Roll, and Yaw. A schematic diagram of the motion directions of these three angles is shown in Fig. 11. According to publicly reported data from a research paper [36], the maximum range of rotation angles for the human head around the x-axis is $[-60.4^\circ, 69.6^\circ]$, around the y-axis is $[-90^\circ, 75^\circ]$, and around the z-axis is $[-40.9^\circ, 36.3^\circ]$.

To obtain head pose characteristics, the conversion between the world coordinate system, camera coordinate system, and image plane coordinate system is completed based on the mapping relationship [37]. This paper employs the Perspective-n-Points (PnP) method, which utilizes the correspondence between 2D image feature points and their corresponding 3D model coordinates. The projection relationship is calculated to obtain the head Euler angles. The diagram illustrating the effect of head pose estimation is shown as follows. The 12 red edges are redrawn by reprojecting the 8 vertices of the cube in the world coordinate system using the head Euler angle rotation matrix and translation vector. It can display the Euler angles of the head pose in real time. Analysis is conducted on the three angle data: Pitch, Roll, and Yaw, as shown in Fig. 11. A spatial coordinate system is established for these three angle data, with the origin of the spatial coordinate system marked as a reference. A fatigue level chart is set up, where the color of the fatigue level chart gradually changes from purple to dark red, indicating that the degree of head shaking deviation gradually increases from slight to severe. When the fatigue level is purple, the head deviation amplitude is small. Over time, as the deviation increases, the color

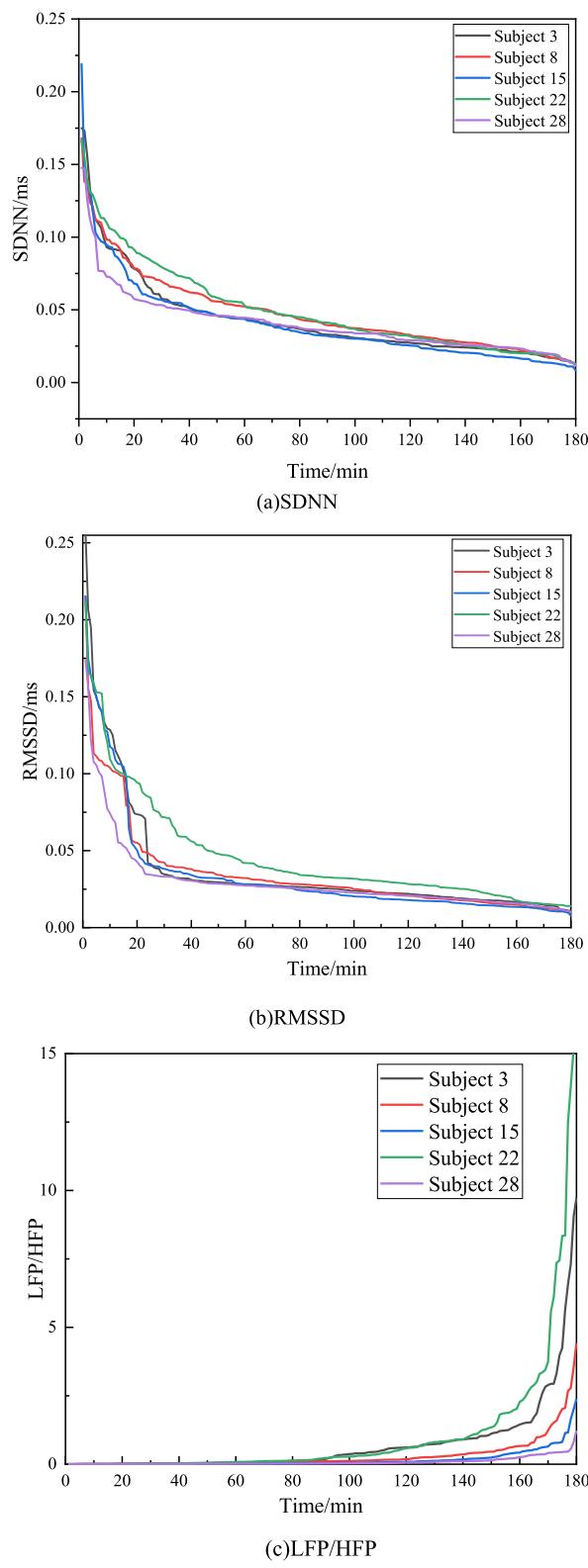


Fig. 7. Variation diagram of characteristic parameters of HRV.

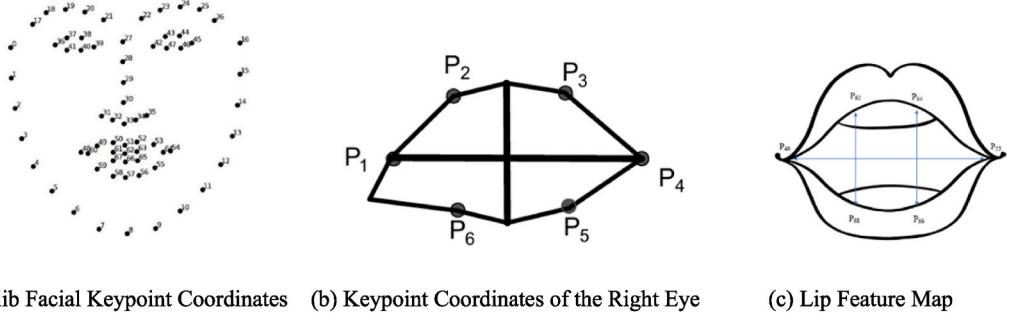
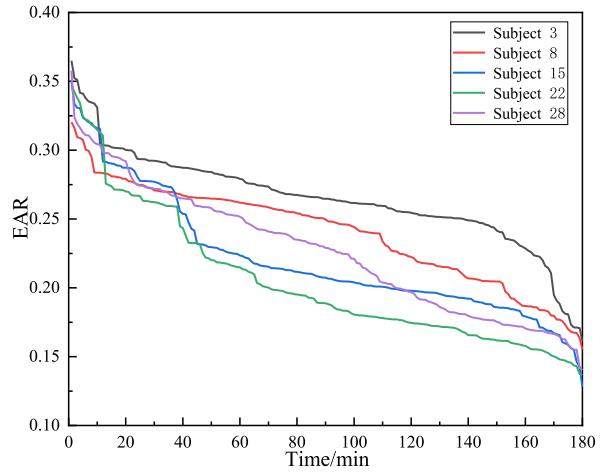
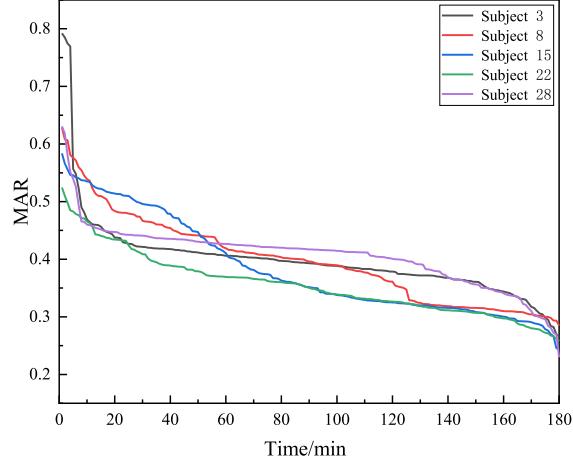


Fig. 8. Facial signal features. (a) Dlib facial keypoint coordinates (b) keypoint coordinates of the right eye (c) lip feature map.



(a) EAR



(b) MAR

Fig. 9. Trend chart.

gradually changes. When the fatigue level changes from green to the red range, the area of the red zone increases, indicating that the majority of participants' head shaking amplitude gradually increases, entering a state of mild fatigue. When the fatigue level is displayed as dark red, the amplitude of the head deviation angle has significantly increased compared to before, indicating that the participant has entered a severe state of fatigue.

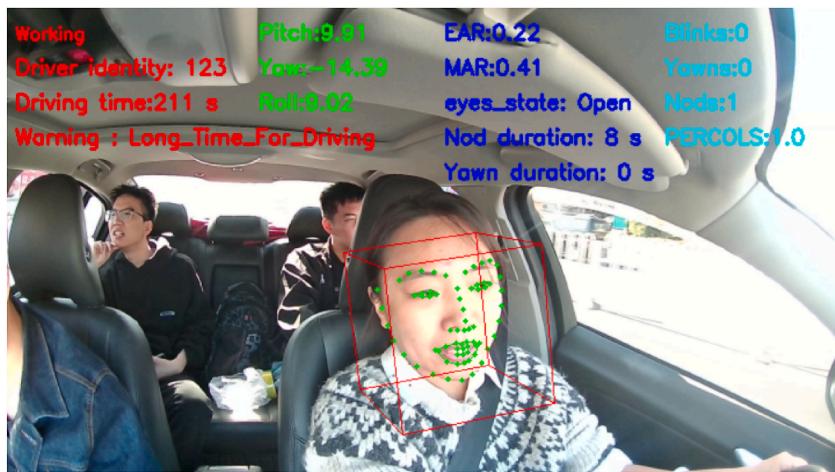


Fig. 10. Facial feature extraction.

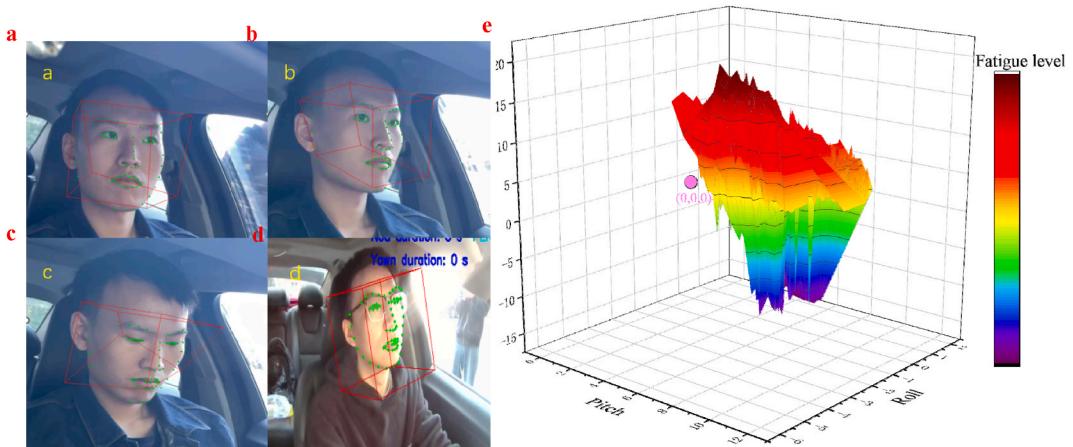


Fig. 11. Head pose variations - left turn (a), straight ahead (b), downward tilt (c), upward tilt (d), and trend chart of fatigue level (e).

Table 4
Characteristics of driver fatigue indicators.

Feature	Category	Interpretation	Frequency	Software	Average	Standard Deviation
SDNN	HRV	R-R Interval Standard Deviation		MATLAB	0.040232	0.027404
RMSSD	HRV	Root Mean Square of Successive Differences in R-R Intervals		MATLAB	0.036144	0.031369
LFP	HRV	Low-Frequency Power	0.04–0.15 Hz	MATLAB	1741.525	21463.33
HFP	HRV	High-Frequency Power	0.15–0.4 Hz	MATLAB	3182983.096	75292498.93
LFP/ HFP	HRV	LF/HF Ratio	–	MATLAB	0.361152	1.343138
EAR	Aspect Ratio of Eye Region	–	–	Anaconda	0.249267	0.042268
MAR	Degree of Mouth Openness	–	–	Anaconda	0.408107	0.069956
Pitch	Head Pose	–	–	Anaconda	7.124329	2.219085
Roll	Head Pose	–	–	Anaconda	-0.49596	1.93471
Yaw	Head Pose	–	–	Anaconda	4.142303	8.438131

3.4. Dataset creation

The facial information and heart rate data of drivers are captured by cameras and PPG signal sensors, respectively. Both serve as quantifiable metrics for assessing fatigue states. However, their inherent differences in data length and structure—with facial information manifested as a series of image frames and heart rate data as temporal sequences—introduce complexities in fusing these two modalities. To integrate multiple sources of information with disparate structures, two primary levels of information fusion are typically utilized: feature-level and decision-level [38–40]. In this paper, we adopt the feature-level information fusion approach, integrating heart rate variability features with driver facial features and head pose. The specific feature indicators are detailed in [Table 4](#), while [Table 5](#) outlines the categorization of fatigue levels. The process of dataset establishment is illustrated in [Fig. 12](#) (see [Table 6](#)).

3.5. Model establishment

3.5.1. LSTM

Hochreiter and Schmidhuber, among others, first introduced the Long Short-Term Memory Network (LSTM) [41], a significant advancement in neural network architectures. While Convolutional Neural Networks (CNNs) have been widely employed in fatigue detection, their fixed-size image input paradigm constrains their flexibility in capturing nuanced features amidst dynamic changes in driving states [42]. The dynamic evolution of driving states often involves intricate interactions across multiple regions, such as the face and eyes, where critical details may not be fully captured by a fixed-sized input window. As an advanced variant of Recurrent Neural Networks (RNNs), LSTM significantly enhances the ability to capture long-term dependencies and retain vital memory information within time-series data through its unique gating mechanism [43]. This paper constructs an LSTM-based fatigue driving detection model that integrates temporal data from heart rate signals, facial features, and head poses as inputs, aiming to substantially elevate the accuracy and real-time responsiveness of fatigue driving detection.

3.5.2. Optimization algorithms

SGD (Stochastic Gradient Descent), which computes the gradient and updates parameters by randomly selecting a single sample, considers only the gradient information from one sample during each parameter update. Its advantages lie in fast parameter updates and ease of implementation. However, since only one sample is used at a time, the gradient estimation may be noisy, leading to significant fluctuations during convergence, vulnerability to outliers, and potentially large oscillations near the optimal solution [44].

Momentum, by introducing the concept of momentum, takes into account not only the current gradient but also previous gradient information when updating parameters. Its benefits include accelerating the training process, particularly in flat or weakly curved regions, and assisting in avoiding local minima [45].

RMSProp, an improvement over the Adagrad optimizer, applies weighted averages to gradients, allowing for smaller updates where gradients vary significantly and larger updates where gradients change slowly. RMSProp can adaptively adjust the learning rate for each parameter, reducing oscillations and instability during training, thereby enhancing training efficiency and stability [46].

Adam (Adaptive Moment Estimation) combines the strengths of RMSProp and Momentum, adjusting the learning rate through computing the exponentially weighted averages of both gradients and squared gradients. It can adaptively tune the learning rate for each parameter based on historical gradient information, enabling rapid convergence with larger learning rates initially and more precise minimization of the loss function with smaller learning rates later. Adam's merits, including adaptive learning rate adjustment, momentum regulation, normalization processing, and prevention of overfitting, make it an outstanding choice for various deep learning tasks [47].

3.5.3. Model construction

The input source for the recognition model comprises the previously extracted fused features. In medical practice, the standard detection duration for heart rate monitoring is typically 15 s, whereas the time step assigned to the fused features is set at 30 s.

If one simply increases the number of LSTM layers without considering the specific situation, it will result in a significant increase in computing power and a noticeable gradient disappearance problem. In most cases, one or two layers of LSTM networks are sufficient for sequential information processing. However, when the number of layers exceeds two, the vanishing gradient phenomenon becomes more pronounced. This leads to a decrease in the update iteration speed and convergence effect of the LSTM layer close to the input layer, resulting in a local optimal dilemma. To address these issues, this article proposes a model consisting of a one-way two-layer stacked LSTM layer with 100 units in the hidden layer. The steps to build a fatigue level classification model based on LSTM network are shown in [Fig. 13](#).

The softmax activation function is employed to convert raw scores into a probability distribution across each class, facilitating

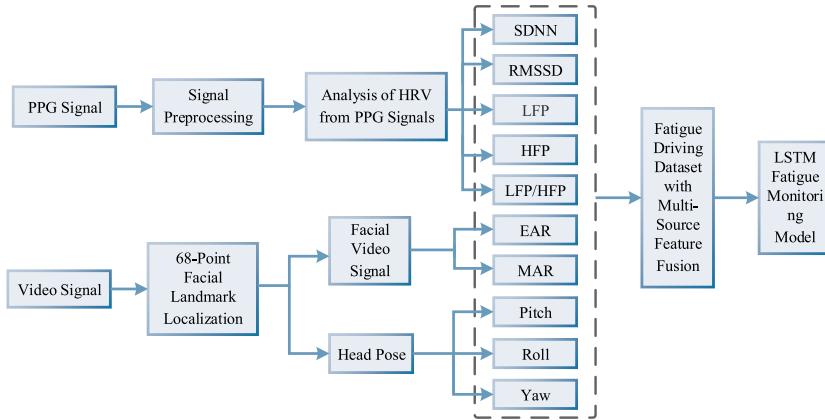
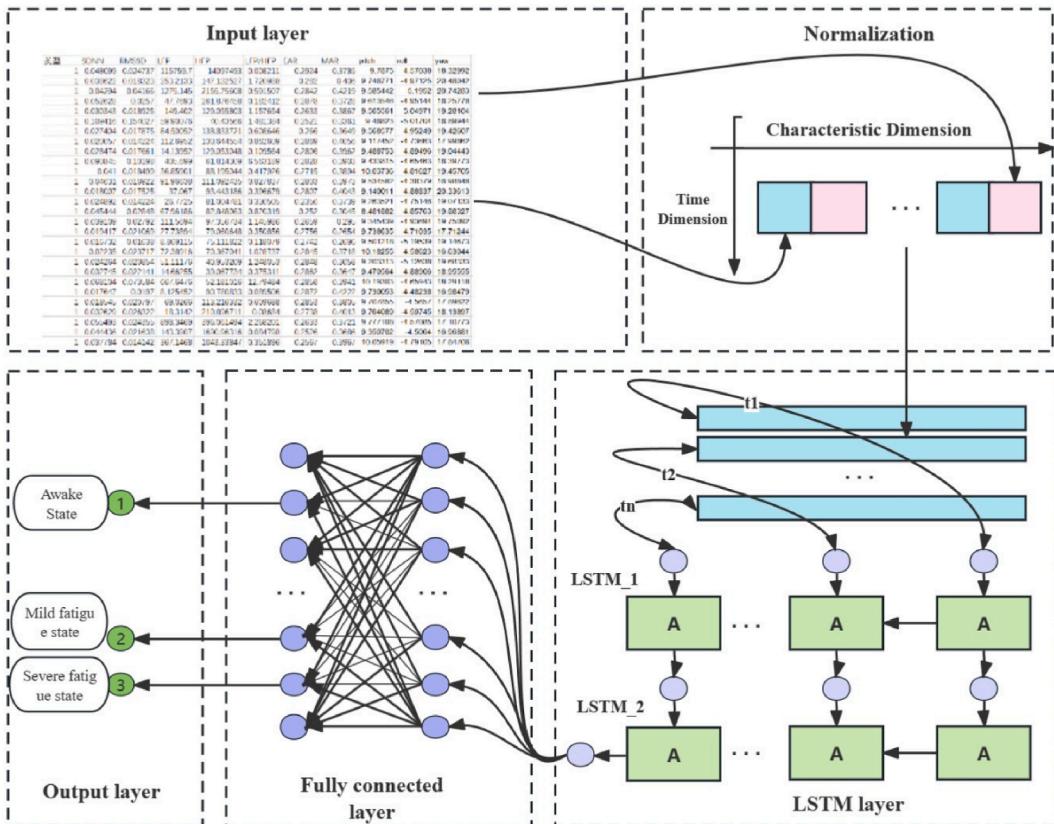
Table 5
Classification of fatigue levels.

Subject	Fatigue State
1	Awake
2	Mild Fatigue
3	Severe Fatigue

Table 6

Performance assessment via confusion matrix.

Optimizer	Tran				Test			
	Accuracy	Precision	Recall	F1 value	Accuracy	Precision	Recall	F1 value
Adam	0.991331	0.991309	0.991358	0.991333	0.977809	0.978141	0.977761	0.977951
Momentum	0.869972	0.869852	0.870381	0.870117	0.833564	0.834857	0.833583	0.83422
Rmsprop	0.753121	0.752643	0.753038	0.75284	0.737864	0.737632	0.737566	0.737599
SGD	0.731623	0.730593	0.732456	0.731523	0.708738	0.707049	0.707974	0.707511

**Fig. 12.** Establishment process of fatigue driving dataset.

classification predictions [48]. Within the entire neural network architecture, the output layer serves as the final stage, responsible for mapping the neural network's outputs onto specific categories or labels for predictive classification tasks. For the purpose of this study, the number of classification categories is set to three, namely: (1) awake state, (2) mild fatigue state, and (3) severe fatigue state.

4. Results and discussion

To implement LSTM network model training and testing on the MATLAB platform, we utilize 10-dimensional fused features as input characteristics, with the driver's fatigue state constituting the target variable. The model's input layer is designed to be 10-dimensional. The model architecture comprises two layers of unidirectional LSTM, each with 100 hidden units. The output layer is configured for classification into three specified categories. In configuring the training options, we designate the solver as "Adam," set the maximum number of iterations to 230,400, specify a learning rate of 0.001, and limit the maximum training epochs to 800. The dataset is partitioned based on an 80:20 ratio, resulting in 80 % for training and 20 % for testing. Data from 30 experimental participants, each contributing 1 h of experimental data segmented every 30 s, yields a total of 3600 data instances. After partitioning, the training set comprises 2880 instances, and the test set includes 720 instances. During the specification of training options, we incorporate four optimization algorithms—Adam, RMSProp, SGD, and Momentum—to optimize and compare their performance, aiming to identify the optimal optimization method.

4.1. Confusion matrix

The confusion matrix serves as a crucial criterion for evaluating the performance of a classifier [49]. As depicted in Figures 14–17, the confusion matrices derived from the training and testing of the LSTM neural network illustrate the classification outcomes. Within the confusion matrix, the FN (False Negative) value signifies instances where samples are incorrectly classified as negative when they are actually positive. Conversely, the FP (False Positive) value represents samples mistakenly identified as positive when they are truly negative. Notably, the confusion matrices of all four algorithms exhibit no signs of overfitting, indicating that the model complexity is appropriately balanced, and the feature selection is effective.

Comparison of the four optimization algorithms yields the adam algorithm to be more advantageous and the model is chosen for specific analysis. The accuracy on the training set is 99.17 %. The accuracy on the test set is 97.36 %. From the confusion matrix diagram, we can calculate the recognition accuracy, recall, and F1 value of the fatigue detection model in three states: awake, mild fatigue, and severe fatigue. As shown in the table. In the model optimized by adam algorithm, the values of FN are 4, 8, 7. 1.7 %, 3.3 %, 2.8 % respectively. The percentage difference is small. Compare the classification accuracy of the training set and test set of the neural network. The difference in accuracy between the two is 7.3 %. The difference is small. It shows that the neural network model has not overfitting phenomenon, has strong generalisation ability, and can be applied to real production life.

4.2. ROC curve

To conduct a comprehensive evaluation of the model's performance, we analyze the ROC (Receiver Operating Characteristic) curve utilizing the data optimized by the Adam algorithm, which exhibited the best testing performance. The ROC curve is illustrated in the accompanying figure. As evident from Fig. 18, the Area Under the Curve (AUC) values, presented in Table 7, consistently exceed 0.9 and approach 1.0. This robust AUC range underscores the model's capability to accurately detect and classify fatigue states, demonstrating its high level of precision and reliability in this task.

4.3. Comparison of models

To demonstrate the superiority of multi-feature fusion, separate training sessions were conducted using single-feature datasets on LSTM networks, alongside Random Forest (RF), Decision Trees (DT), and BiLSTM models. The training outcomes were subsequently compared and analyzed, with the accuracy results presented in Fig. 19. As discernible from the figure, notable discrepancies exist in the accuracy achieved across different feature datasets. Specifically, the training accuracy of single-feature datasets is conspicuously lower than that achieved with the fused feature dataset. This observation conclusively substantiates that integrating multiple features enhances model accuracy, thereby verifying the merits and superiority of multi-feature fusion.

5. Conclusion

One of the fundamental aspects in the development of autonomous driving technology is the monitoring of the driver's status. This is necessary because the driver must be ready to take control of the vehicle in case the automatic system fails. Additionally, monitoring the driver's status can help predict driver fatigue, thereby preventing accidents caused by fatigue. In this study, we have developed a method for fatigue driving detection that involves fusing multiple sources of information. This method also serves as a driver status detection method. Through real driving experiments, we have analyzed the data and drawn the following conclusions:

1. The fusion of multiple sources of information has a significant impact. By combining features related to heart rate variability, facial expressions, and head posture, we are able to comprehensively capture the driver's physiological and behavioral characteristics, leading to improved accuracy in detecting fatigue while driving. The fusion of multiple sources of information has proven to be

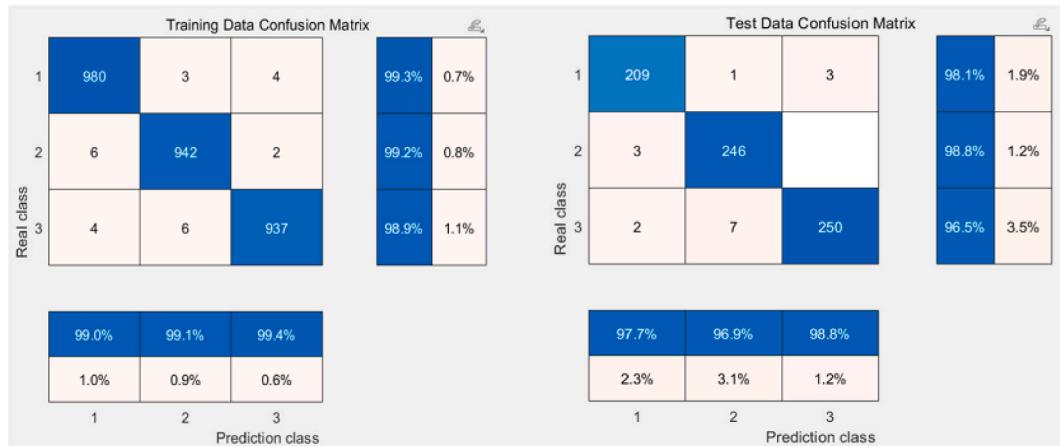


Fig. 14. Adam optimization model training results - training data confusion matrix (a) and test data confusion matrix (b).



Fig. 15. RMSProp algorithm optimization model training results - training data confusion matrix (a) and test data confusion matrix (b).



Fig. 16. Momentum optimization model training results - training data confusion matrix (a) and test data confusion matrix (b).

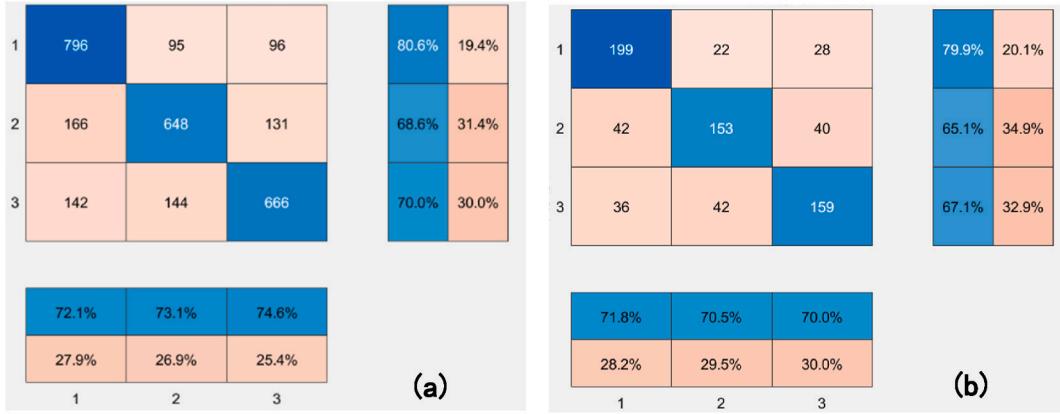


Fig. 17. SGD algorithm optimization model training results - training data confusion matrix (a) and test data confusion matrix (b).

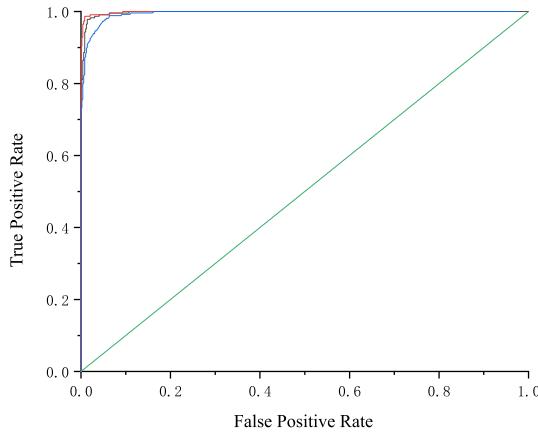


Fig. 18. ROC curve.

Table 7
AUC values.

Driver status	AUC value
awake state	0.9987
Mild fatigue state	0.9962
Severe fatigue state	0.9945

advantageous in identifying fatigue states, providing a more reliable assessment compared to relying on a single source of information.

2. The Effectiveness of LSTM Networks. In response to the limitations of high computational complexity, the relative difficulty in tuning network parameters, and sensitivity to sequence length, we employed a double-layer stacked LSTM network. This network leverages its long-sequence processing capabilities, combined with gating mechanisms and robust generalization ability, to train the model with multi-source information input. Furthermore, we utilized four optimization methods, namely Adam, RMSProp, SGD, and momentum, to refine the model. Comparisons were made with Decision Trees (DT), Random Forests (RF), and Bidirectional LSTM (BiLSTM). The experimental results underscore the robust performance of LSTM networks in modeling time-series data, with Adam optimizer emerging as the optimal choice, effectively capturing driver fatigue patterns. Notably, this network exhibited a high degree of accuracy in classification tasks, further validating its suitability for the intended application.
3. Improved accuracy in identifying fatigue driving conditions: Our fatigue driving detection model demonstrated promising performance in experiments. The accuracy rates for identifying awake state, mild fatigue state, and severe fatigue state were 98.1 %, 98.8 %, and 96.5 % respectively, resulting in an overall recognition accuracy rate of 97.78 %. These findings highlight the potential application of our method in real-world driving situations.

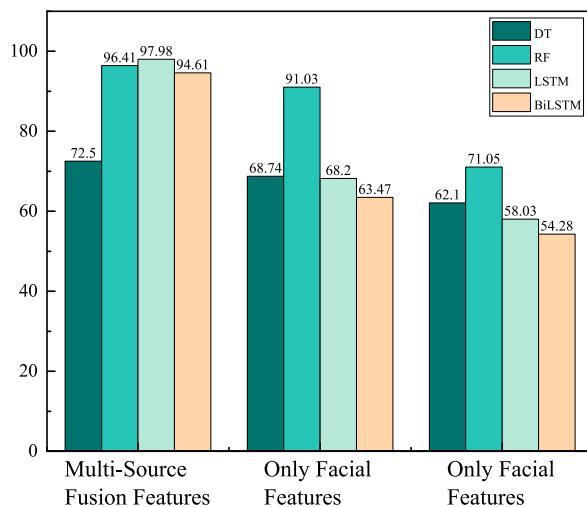


Fig. 19. Comparison of accuracy.

- Creation of experimental data set: Additionally, we have curated a fatigue driving data set that incorporates multi-information fusion of heart rate variability features, facial features, and head posture features. This data set provides valuable resources for future research and applications. Moving forward, our plan is to further enhance the algorithm to accommodate a wider range of actual driving scenarios. By doing so, we aim to promote its implementation in the field of driving safety, thereby reducing the risk of traffic accidents caused by fatigue driving and safeguarding the lives of drivers and other road users.

In the future, we aim to enhance the algorithm to accommodate a wider range of real-life driving scenarios. Furthermore, we intend to apply it to the domain of driving safety, with the goal of mitigating the risk of traffic accidents caused by fatigue driving. This will help safeguard the lives of drivers and other road users.

CRediT authorship contribution statement

Lu Yu: Conceptualization. **Xinyi Yang:** Data curation. **Hengjian Wei:** Software. **Jianguo Liu:** Validation. **Bo Li:** Software.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author, [L.Y.], upon reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

T

Acknowledgments

This work was supported by the Basic scientific research project of Liaoning Provincial Department of Education (Grant No. LJKQZ20222462).

References

- [1] National Bureau of statistics, China Statistical Yearbook, China Statistics Press, Beijing, 2022.
- [2] J.J. Rolison, S. Regev, S. Moutari, et al., What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records, Accid. Anal. Prev. 115 (2018) 11–24, <https://doi.org/10.1016/j.aap.2018.02.025>.
- [3] National Center for Statistics and Analysis, Overview of motor vehicle crashes in 2019: traffic safety facts research note. report no. dot hs 813 060. <https://crashstats.nhtsa.dot.gov/Api/Public/Publication/813060>, 2019.
- [4] National Highway Traffic Safety Administration, Traffic safety facts 2015. <https://crashstats.nhtsa.dot.gov/Api/Public/Publication/812384>, 2015.
- [5] Vanlaar Ward, Herb Simpson, Dan Mayhew, et al., Fatigued and drowsy driving: a survey of attitudes, opinions and behaviors, J. Saf. Res. 39 (2008) 303–309, <https://doi.org/10.1016/j.jsr.2007.12.007>.
- [6] C.X. Zhang, Y.F. Ma, S.Y. Chen, Driving safety performance evaluation method for heavy vehicle drivers based on super-efficiency data envelopment analysis, China J. Highw. Transp. 36 (2023) 326–342, <https://doi.org/10.19721/j.cnki.1001-7372.2023.09.025>.

- [7] B.T. Zhang, W.W. Chang, X.L. Li, Fatigue Driving Detection Based on Spatial-Temporal Electroencephalogram Features and Parallel, vol. 23, 2023, pp. 315–325, <https://doi.org/10.16097/j.cnki.1009-6744.2023.02.033>.
- [8] D.H. Li, Q. Liu, W. Yuan, Relationship between fatigue driving and traffic accident, *J. Traffic Transport. Eng.* 10 (2010) 104–109.
- [9] B. Bakker, B. Zablocki, A. Baker, V. Riethmeister, B. Marx, G. Iyer, C. Ahlström, A multi-stage, multi-feature machine learning approach to detect driver sleepiness in naturalistic road driving conditions, *IEEE Trans. Intell. Transport. Syst.* 23 (5) (2021) 4791–4800, <https://doi.org/10.1109/TITS.2021.3090272>.
- [10] Naurois de, Charlotte Jacobé, et al., Detection and prediction of driver drowsiness using artificial neural network models, *Accid. Anal. Prev.* 126 (2019) 95–104, <https://doi.org/10.1016/j.aap.2017.11.038>.
- [11] Guanglong Du, et al., A multimodal fusion fatigue driving detection method based on heart rate and PERCLOS, *IEEE Trans. Intell. Transport. Syst.* 23 (11) (2022) 21810–21820, <https://doi.org/10.1109/TITS.2022.3176973>.
- [12] Chen He, et al., Fatigue at the wheel: a non-visual approach to truck driver fatigue detection by multi-feature fusion, *Accid. Anal. Prev.* 199 (2024) 107511, <https://doi.org/10.1016/j.aap.2024.107511>.
- [13] Anna Persson, et al., Heart rate variability for classification of alert versus sleep deprived drivers in real road driving conditions, *IEEE Trans. Intell. Transport. Syst.* 22 (6) (2020) 3316–3325, <https://doi.org/10.1109/TITS.2020.2981941>.
- [14] S.X. Cai, C.K. Du, S.Y. Zhou, et al., Fatigue driving state detection based on vehicle running data, *J. Transport. Syst. Eng. Inf. Technol.* 20 (4) (2020) 77–82, <https://doi.org/10.16097/j.cnki.1009-6744.2020.04.012>.
- [15] Argiñelo Prada, Erick Javier, Katherine Daniela Marcillo Ibarra, Kevin Leonardo Díaz Jiménez, The use of successive systolic differences in photoplethysmographic (PPG) signals for respiratory rate estimation, *Heliyon* 10 (2024) e26036, <https://doi.org/10.1016/j.heliyon.2024.e26036>, 4.
- [16] Araf Nishan, et al., A continuous cuffless blood pressure measurement from optimal PPG characteristic features using machine learning algorithms, *Heliyon* 10 (2024) e27779, <https://doi.org/10.1016/j.heliyon.2024.e27779>, 6.
- [17] Ziyi Liu, Yiming Zhang, Congcong Zhou, BiGRU-attention for Continuous blood pressure trends estimation through single channel PPG, *Comput. Biol. Med.* 168 (2024) 107795, <https://doi.org/10.1016/j.combiomed.2023.107795>.
- [18] Jukka-Pekka Sirkkä, Tuukka Panula, Matti Kaisti, Investigating the impact of contact pressure on photoplethysmograms, *Biomed. Eng. Adv.* 7 (2024) 100123, <https://doi.org/10.1016/j.bea.2024.100123>.
- [19] Q. Wu, Caring for the truck driver community to promote the healthy development of the industry: final release of the "China truck driver survey report, China Logist. Purch. (8) (2023) 28–29, <https://doi.org/10.16079/j.cnki.issn1671-6663.2023.08.040>.
- [20] Z. Lan, J. Zhao, P. Liu, C. Zhang, N. Lyu, L. Guo, Driving fatigue detection based on fusion of EEG and vehicle motion information, *Biomed. Signal Process Control* 92 (2024) 106031, <https://doi.org/10.1016/j.bspc.2024.106031>.
- [21] Zihao Hao, et al., PPG heart rate extraction algorithm based on the motion artifact intensity Classification and removal framework, *Biomed. Signal Process Control* 94 (2024) 106287, <https://doi.org/10.1016/j.bspc.2024.106287>.
- [22] Sukesh Rao, K. Adithi, Sanith C. Bangera, An experimental investigation on pulse transit time and pulse arrival time using ecg, pressure and ppg sensors, *Med. Novel Technol. Devices* 17 (2023) 100214, <https://doi.org/10.1016/j.medintd.2023.100214>.
- [23] Tien Anh Hoang, et al., Prognosis value of heart rate variability measured by Camera HRV application in patients after acute myocardial infarction, *Indian Heart J.* (2024), <https://doi.org/10.1016/j.ijhj.2024.07.008>.
- [24] Kulbhushan Chand, Shilpa Chandra, Varun Dutt, A comprehensive evaluation of linear and non-linear HRV parameters between Paced Breathing and Stressful Mental State, *Heliyon* (2024) e32195, <https://doi.org/10.1016/j.heliyon.2024.e32195>.
- [25] Giuseppina Pilloni, et al., Heart Rate Variability (HRV) serves as an objective correlate of distress and symptom burden in multiple sclerosis, *Int. J. Clin. Health Psychol.* 24 (2) (2024) 100454, <https://doi.org/10.1016/j.ijchp.2024.100454>.
- [26] Nancy Gullett, et al., Heart rate variability (HRV) as a way to understand associations between the autonomic nervous system (ANS) and affective states: a critical review of the literature, *Int. J. Psychophysiol.* (2023), <https://doi.org/10.1016/j.ijpsycho.2023.08.001>.
- [27] Pınar Özgen Kavas, et al., Machine learning-based medical decision support system for diagnosing HFpEF and HFrEF using PPG, *Biomed. Signal Process Control* 79 (2023) 104164, <https://doi.org/10.1016/j.bspc.2022.104164>.
- [28] Liang Wu, et al., An optimization study of the ultra-short period for HRV analysis at rest and post-exercise, *J. Electrocardiol.* 63 (2020) 57–63, <https://doi.org/10.1016/j.jelectrocard.2020.10.002>.
- [29] Mohamed Elhoseny, et al., Advanced deep learning for masked individual surveillance, *Int. J. Cognit. Comput. Eng.* (2024), <https://doi.org/10.1016/j.ijcce.2024.07.003>.
- [30] Quqi Chen, Lei Sang, Face-mask recognition for fraud prevention using Gaussian mixture model, *J. Vis. Commun. Image Represent.* 55 (2018) 795–801, <https://doi.org/10.1016/j.jvcir.2018.08.016>.
- [31] J. Robert Theivadas, Suresh Ponnai, VigilEye: machine learning-powered driver fatigue recognition for safer roads, *Measurement: Sensors* 33 (2024) 101186, <https://doi.org/10.1016/j.measen.2024.101186>.
- [32] Paulo Augusto de Lima Medeiros, Gabriel Vinícius Souza da Silva, Felipe Ricardo dos Santos Fernandes, Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam, *Expert Syst. Appl.* 188 (2022) 116073, <https://doi.org/10.1016/j.eswa.2021.116073>.
- [33] Paula Augusto de Lima Medeiros, Gabriel Vinícius Souza da Silva, Felipe Ricardo dos Santos Fernandes, Efficient machine learning approach for volunteer eye-blink detection in real-time using webcam, *Expert Syst. Appl.* 188 (2022) 116073, <https://doi.org/10.1016/j.eswa.2021.116073>.
- [34] Z. Zhang, H. Wang, Q. You, L. Chen, H. Ning, A novel temporal adaptive fuzzy neural network for facial feature based fatigue assessment, *Expert Syst. Appl.* 252 (2024) 124124, <https://doi.org/10.1016/j.eswa.2024.124124>.
- [35] Y. Sun, et al., Design and implementation of a driver fatigue detection system based on information fusion, *Electron. Des. Eng.* 32 (11) (2024) 71–75+81, <https://doi.org/10.14022/j.issn1674-6236.2024.11.015>.
- [36] Y.Y. Jiao, Y.N. Deng, Y. Luo, et al., Driver sleepiness detection from EEG and EOG signals using GAN and LSTM networks, *Neurocomputing* 408 (2020) 100–111, <https://doi.org/10.1016/j.neucom.2019.05.108>.
- [37] Z.X. Zhai, Research on fatigue state recognition of air traffic controllers based on infrared and visible light image fusion, Master's thesis, Civil Aviation Flight University of China (2024), <https://doi.org/10.27722/d.cnki.gzgmh.2024.000155>.
- [38] Arun A. Ross, Govindarajan Rohin, Feature level fusion of hand and face biometrics, *Biometric technology for human identification II* 5779 (2005), <https://doi.org/10.1117/12.606093>.
- [39] Mohammad Haghhighat, Mohamed Abdel-Mottaleb, Wadee Alhalabi, Discriminant correlation analysis: real-time feature level fusion for multimodal biometric recognition, *IEEE Trans. Inf. Forensics Secur.* 11 (9) (2016) 1984–1996, <https://doi.org/10.1109/TIFS.2016.2569061>.
- [40] Firouz Abdullah Al-Wassai, N.V. Kalyankar, Ali A. Al-Zaky, Multisensor images fusion based on feature-level, arxiv preprint arxiv:1108.4098, <https://doi.org/10.48550/arXiv.1108.4098>, 2011.
- [41] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780, <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [42] Sayan Sarkar, Aayushman Ghosh, Schrödinger spectrum and slim CNN architecture-based signal quality estimation for Photoplethysmogram signals, *Biomed. Signal Process Control* 94 (2024) 106240, <https://doi.org/10.1016/j.bspc.2024.106240>.
- [43] Jose Principe, Neil Euliano, Shayan Garani, Principles and networks for self-organization in space–time, *Neural Network.* 15 (8–9) (2002) 1069–1083, [https://doi.org/10.1016/S0893-6080\(02\)00080-1](https://doi.org/10.1016/S0893-6080(02)00080-1).
- [44] Zhuan Zhang, et al., Gradient preconditioned mini-batch SGD for ridge regression, *Neurocomputing* 413 (2020) 284–293, <https://doi.org/10.1016/j.neucom.2020.06.092>.
- [45] X. Li, X. Liu, Accelerated analysis on the triple momentum method for a two-layer ReLU neural network, *J. King Saud Univ.-Comput. Inf. Sci.* 36 (4) (2024) 102016, <https://doi.org/10.1016/j.jksuci.2024.102016>.
- [46] D. Xu, S. Zhang, H. Zhang, D.P. Mandic, Convergence of the RMSProp deep learning method with penalty for nonconvex optimization, *Neural Network.* 139 (2021) 17–23, <https://doi.org/10.1016/j.neunet.2021.02.011>.

- [47] Q. Tong, G. Liang, J. Bi, Calibrating the adaptive learning rate to improve convergence of ADAM, *Neurocomputing* 481 (2022) 333–356, <https://doi.org/10.1016/j.neucom.2022.01.014>.
- [48] X. Han, X. Zhu, W. Pedrycz, A.M. Mostafa, Z. Li, A design of fuzzy rule-based classifier optimized through softmax function and information entropy, *Appl. Soft Comput.* 156 (2024) 111498, <https://doi.org/10.1016/j.asoc.2024.111498>.
- [49] X. Deng, Q. Liu, Y. Deng, S. Mahadevan, An improved method to construct basic probability assignment based on the confusion matrix for classification problem, *Inf. Sci.* 340 (2016) 250–261, <https://doi.org/10.1016/j.ins.2016.01.033>.