Author(s): Zuo, Xin; Zhang, Chi; Cong, Fengyu; Zhao, Jian; Hämäläinen, Timo

Title: Mobile Phone Use Driver Distraction Detection Based on MSaE of Multi-Modality Physiological Signals

Year: 2024

Version: Accepted version (Final draft)

Please cite the original version:

Zuo, X., Zhang, C., Cong, F., Zhao, J., & Hämäläinen, T. (2024). Mobile Phone Use Driver Distraction Detection Based on MSaE of Multi-Modality Physiological Signals. IEEE Transactions on Intelligent Transportation Systems, 25(11), 17650-17665. https://doi.org/10.1109/tits.2024.3416382

Mobile Phone Use Driver Distraction Detection Based on MSaE of Multi-modality Physiological Signals

Xin Zuo, Chi Zhang, Fengyu Cong, Jian Zhao, and Timo Hämäläinen

Abstract—**Driver distraction, a major cause of traffic crashes, is reported to reduce driving performance and be detected with vehicle behavioral features. It also induces physiological responses. Time and frequency-domain features of physiological signals have been used to study distraction, but they are susceptible to residual noise and tend to overlook complexity. Moreover, the resampling problem arises while analyzing physiological signals at multiple time scales. This paper proposes a novel framework based on multiscale entropy on absolute time scales (MSaE) and bidirectional long short-term memory (BiLSTM) network to mine the distraction information in multi-modality physiological signals and detect distraction automatically. Firstly, an entropy-based resampling method is adopted to find the suitable downsampling rates of electroencephalography (EEG), electrocardiogram (ECG), and electromyography (EMG). Then, calculating entropy with absolute time scales instead of relative time scales in a sliding window is utilized to explore the fluctuations of each signal while distraction. Afterward, ReliefF is selected from conventional feature selectors to identify the optimal feature set for each signal. Finally, BiLSTM with time dependency is designed to detect driver distraction with the selected feature set. The results illustrate significant distinctions in the MSaE of multiple physiological signals between normal and distracted driving. Additionally, MSaE, superior to traditional features, is selected as the most discriminative feature for each signal in distraction mining. Furthermore, the accuracy is further improved by about 8%, incorporating multi-modality features rather than vehicle behavioral features. This study indicates the potential of employing various signals to understand and detect driver distraction effectively.**

Index Terms—**Driver distraction, EEG, ECG, EMG, MSaE, multi-modality analysis**

Manuscript received 26 October 2023; revised 1 May 2024; accepted 15 June 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 61703069 and Grant 62001312, in part by the Fundamental Research Funds for the Central Universities under Grant DUT21GF301 and in part by the Science and Technology Planning Project of Liaoning Province under Grant 2021JH1/10400049. The Associate Editor for this article was S. Mumtaz. *(Xin Zuo and Chi Zhang contributed equally to this work.)(Corresponding author: Chi Zhang; Jian Zhao)*.

This work involved human subjects or animals in its research. Approval of all the ethical and experimental procedures and protocols was granted by the Ethics Committee, Liaoning Normal University.

Xin Zuo and Fengyu Cong are with the School of Biomedical Engineering, Faculty of Medicine, Dalian University of Technology, Linggong Road #2, Dalian 116024, China, and also with the Faculty of Information Technology, University of Jyväskylä, Mattilanniemi 2, Jyväskylä FIN-40014, Finland (e-mail: zuoxin93@foxmail.com; cong@dlut.edu.cn). Fengyu Cong is also with the Key Laboratory of Social Computing and Cognitive Intelligence (Dalian University of Technology), Ministry of Education, China.

Chi Zhang is with the School of Biomedical Engineering, Faculty of Medicine, Dalian University of Technology, Linggong Road #2, Dalian 116024, China (e-mail: chizhang@dlut.edu.cn).

Jian Zhao is with the School of Automotive Engineering, Faculty of Vehicle Engineering and Mechanics, Dalian University of Technology, Linggong Road #2, Dalian 116024, China (e-mail: jzhao@dlut.edu.cn).

Timo Hämäläinen is with the Faculty of Information Technology, University of Jyväskylä, Mattilanniemi 2, Jyväskylä FIN-40014, Finland (e-mail: timo.t.hamalainen@jyu.fi).

Digital Object Identifier 10.1109/TITS.2024.3416382

I. INTRODUCTION

TRAFFIC accidents are happening worldwide with increasing frequency in recent years. It will lead not only to severe economic losses but also to injuries and deaths. According to the report of the National Highway Traffic Safety Administration (NHTSA), there are 3,308 lives taken away in vehicle crashes in 2022 because of distracted drivers in the United States, and using mobile phones while driving can greatly increase the potential for traffic accidents [1]. It is also reported in a survey of the World Health Organization (WHO) that using mobile phones would significantly increase drivers' exposure to the risk of car crashes by about four times [2]. Drivers' workload can be increased, and their attention can be diverted from the task of driving to unrelated activities while talking or texting with mobile phones in the driving process. In this condition, it will slow down the reaction time of drivers and make it difficult to keep the right forward direction, then leading to accidents. Therefore, it is necessary to detect mobile phone use induced driver distraction in time and alert drivers to observe the surroundings and focus on driving to reduce such traffic accidents.

A variety of signals have been used to evaluate driver state and detect driver distraction. The most commonly used signals in the existing research are obtained by non-intrusive sensors, including Global Position System (GPS), external and internal cameras [3], [4]. These sensors can easily collect the vehicle behavioral signals (like steering wheel angle, acceleration, speed, etc.) and driver visual images (like head movement, eye fixation, driving posture, etc.). Driver distraction and situation awareness were evaluated in [5] by analyzing the lane keeping and vehicle velocity data, and the performance of different age, gender, and driving experience groups were compared. Eye movement data was used to study the relationship between response performance and cognitive workload while drivers were talking with passengers in [6]. Although the vehicle behavioral signals are easy to collect and analyze, they are subject-dependent and sensitive to traffic conditions and weather [7]. What's more, driver visual images can be influenced by facial occlusion, illumination, and personal driving habits easily, even if they can reflect drivers' attention shift intuitively [8].

There is also research studying driver distraction utilizing intrusive sensors to acquire physiological signals like electroencephalography (EEG), electromyography (EMG), and electrocardiogram (ECG). EEG is the record of the real internal electrical potentials originated by the central nervous system (CNS). It can reflect the brain activities in real-time with high temporal resolution and is the most widely used signal among all physiological signals to measure mental status [9]. For instance, a deep learning approach was proposed in [10] to detect driver distraction on the basis of the temporal and spatial information of EEG signals and examine the contributions of temporal and spatial information. ECG is one kind of record of autonomic nervous system (ANS) activity and is usually used to measure the heart rate. Research has shown that heart rate changes along with mental state and can be used to analyze driver state [11]. EMG, induced by the alteration of muscle tension, is also used in the literature to detect driver states. Since the equipment in cars, like the pedals and steering wheel, needs to be controlled by the human body while driving, the tension of muscles can be used as an indicator to predict the driving state [12]. A muscle computer interface is designed to reduce distraction appearance by sensing and decoding EMG signals associated with different finger postures while distraction in [13]. In sum, physiological signals record the real electrical activities of the human internal state timely, which provides a more accurate and faster way to explore the internal state alterations.

Traditional time-domain and frequency-domain features of physiological signals are widely used to study driver distraction in the literature. Although the temporal and spectral dynamic changes of physiological responses can be reflected in the fluctuations of these features, the rhythmicity and stability of different physiological signals alter along with driver status while driving as well [14]. In this context, the complexity of physiological signals is usually overlooked. Entropy features like Shannon entropy, fuzzy entropy (FE) and sample entropy (SE) have been used to study the complexity information. Nonetheless, physiological signals are usually contaminated by breath, blinks, body movements, etc., especially in driving experiments. Residual noise is retained after preprocessing and has an effect on results since noise contamination cannot be absolutely eliminated by preprocessing [15]. Thus, multiscale entropy (MSE) is proposed to manifest the complexity of signals by calculating entropy in several scales, which can reduce the residual noise effects and improve the robustness of the results [16]. However, how their complexity varies between distraction and normal state needs to be further studied, and the resampling problem appears if analyzing signals at multiple time scales. What's more, it is necessary to analyze signals in high temporal resolution so as to detect and study the subtle changes in driver state in time. But higher resolution will result in larger data size and lower computational efficiency. So, it is vital to mine for the valuable information about driver status using complexity features of physiological signals with optimal sampling frequency.

Recent studies have demonstrated that none of the signals can provide sufficient information to estimate driver mental state and that incorporate signals can compensate for each other [17]–[19]. Additionally, along with the development of sensor technology and less intrusive wearable equipment, multi-modality signals are more and more widely used to study driver distraction. Multi-modality features from speech, facial expression, and car signals were extracted to detect driver distraction in [20], and the results show that the predictive accuracy increases by adding more modalities features. A lightweight framework is proposed and validated in [21] to explore the comprehensive detection of driver distraction by fusing multi-sensor data from drivers, vehicles, and GPS positions. A multi-modal fusion network for detecting driver distraction with acoustic signal, visual signal, EMG, and car motion signals was introduced in [22]. The results show that multi-modality signals can greatly enhance the performance of the proposed network.

Moreover, machine learning (ML) and deep learning (DL) algorithms have been utilized to detect driver distraction. Most of these algorithms make decision on the current time step based on the current input information that is not in accordance with the real driving activity. In fact, driving is a long-lasting and interactive activity among the driver, the vehicle, and the environment. Drivers make decisions on the basis of both the current received information and the previously received information and the upcoming events [23]. Hence, it is necessary to detect driver distraction considering the contextual information. In recent years, it has been proved that long short-term memory (LSTM) network can memorize the long and short-term information in sequential data and can achieve a better performance than conventional classifiers in diagnosing depression and anxiety [24]. Furthermore, bidirectional LSTM (BiLSTM) has also been successfully applied in many fields like sleep staging and text classification with the advantages of learning the two directional (i.e., forward and backward) contextual information in signals [25], [26].

This paper proposes a driver distraction detection framework based on multiscale entropy on absolute time scales (MSaE) of multi-modality physiological signals (i.e., EEG, ECG, and EMG signals) considering the contextual dependency in signals. The main contributions of this study are as follows:

1) We propose to use MSaE for exploring the complexity of physiological signals to reduce the residual noise effect and to eliminate the resampling problem when analyzing physiological signals at multiple time scales.
2) We propose to analyze driver distraction with multi-modality signals and find the most discriminative multimodal feature set by feature selection.

3) We validate the effectiveness of the selected multimodal feature set in driver distraction detection utilizing BiLSTM that can learn the time dependency in the data.

The remainder of the paper is organized as follows. Section II lists relevant research on the estimation of driver distraction. Section III describes the experimental details and the gathered multi-modality signals. In Section IV, the adopted methodologies are introduced. The findings of our study are presented and discussed in Section V. Finally, Section VI summarizes the paper.

II. RELATED WORK

For detecting driver distraction, the popular way is to first extract features from signals and then input the features into classification models. There are various features used to excavate distraction information existing in different signals in the literature. Time-domain features like mean value and standard deviation of vehicle behavioral signals are usually calculated to evaluate changes in drivers' performance due to distraction. In [27], the mean and standard deviation of speed, acceleration, steering wheel, and lane deviation data were selected as indicators to identify truck driver distraction. In [28], the mean speed, lateral position standard deviation, and the mean angular acceleration of the steering wheel were calculated to analyze the differences between normal and distracted driving. They found that steering wheel indicators are adequate to distinguish distraction and non-distraction. Many studies adopt features based on eye movement (e.g., fixation duration, saccade frequency, and gaze distribution) of driver visual images to explore the useful information while distraction. In [23], not only were the vehicle behavioral features extracted, but also the fixation duration and blink frequency of the eye data were calculated and fed into the Bayesian Networks (BNs) model to detect driver distraction. [29] analyzed the glance behavior of drivers and demonstrated that it serves as a good measure of drivers' attention shift, which makes it possible to detect distraction. As for physiological signals, most existing research utilizes time-domain and frequency-domain features to explore the internal distraction information. [30] extracted the power, standard deviation, and mean absolute values of four EEG rhythms to compare the performance of various classification methods detecting driver distraction. [31] analyzed all possible sub-bands of the ECG spectrum and obtained their mean, standard deviation, and power features to identify driver distraction. A driver inattention detection system using the statistical and spectral features of EMG and ECG was presented in [32], and results showed that EMG and ECG signals can be adopted to detect inattention. These features can indicate the dynamic changes of the physiological signals in the time domain and frequency domain while drivers are distracted, but the complexity information is neglected to some extent.

Physiological signals are recordings of the electrical activities derived from the human body [33]. Extensive amounts of complexity information about human states hide within the fluctuations of signals [34], especially EEG. Since EEG generally records the brain activities of the whole scalp with multi-electrodes rather than in one region. Plenty of valuable information might be overlooked if we do not consider the complexity features. Entropy is a measure of the disorder or randomness of a system and can be selected to manifest the complexity of physiological signals [35]. In this case, entropy features have drawn researchers' attention to use them in estimating human states recently. FE and SE of EEG and electrooculogram (EOG) signals were implemented in [14] to analyze and classify different sleep stages. The results are slightly better than those of the existing methods. An epileptic seizure detection model was trained in [36] by employing differential entropy (DE) and peak-magnitude of root mean square ratio (PRMS) of EEG. The results show that the extracted features could distinguish normal and abnormal seizure EEG with superior prediction accuracy. A machine learning (ML) model was built in [37] to estimate the visual interest level with MSE of EEG. The results indicate that MSE can quantify EEG complexity to reflect human visual interest. Apart from the advantages, there are still challenges in exploring the complexity of physiological signals applying entropy features. Artifacts cannot be completely eliminated whatever preprocessing approaches are employed [15]. Hence, the calculated entropy features (like FE, SE, and DE) involve the residual noise in the complexity of signals, which results in the poor robustness of these features. Besides, researchers have proved that the value of entropy could be affected by the sampling rate, e.g., the value of SE will obviously increase while signals are downsampled [38], [39]. Compared to the above single scale entropy features, MSE is better in robustness as it is calculated at multiple time scales to decrease the effect of residual white noise [16]. Nevertheless, the resampling problem arises when calculating entropy in different time scales, and the compressing or stretching effect occurs if the sampling rate of signals is changed [40]. Moreover, to detect the subtle changes in the driver state in time, signals in high temporal resolution are needed. But a rise then appears in data size and computational cost. To reduce the extensive, time-consuming calculation, most research applies the downsampling approach prior to analysis. Under the circumstances, how to find the appropriate downsampling frequency and manifest the complexity of physiological signals can be challenging.

As for distraction detection, a variety of classification algorithms have been used in the literature. Deep learning (DL), a branch of ML, has a better understanding of large data with deep neural networks than traditional ML [41]. Recent research has pointed out that DL outperforms many conventional ML approaches for time series classification and appears to be the most promising algorithm for classifying temporal signals [42]. A multimodal emotion recognition method adopting deep belief network (DBN) was presented and validated that better recognition performance could be obtained in subtle emotion expressions with DBN in [43]. Discrete dynamic Bayesian (DDB) network was utilized to predict the severity of driver distraction in continuous videos, and a fuzzy system was also implemented to detect multi-class distractions into three severity levels in [44]. Driving is a continuous work with the characteristics of long duration and context dependent. Decisions on each action are made upon the current state as well as many previous states, and it has been reported that context dependent is vital for detecting human mental states [23]. Therefore, adding memory to a neural network can learn more information about driver distraction from time series and improve classification performance [45]. Recurrent neural network (RNN) can memorize the previous states of sequential data and decide the output according to the history for a short duration time series [46]. LSTM, an extension of RNN, achieves to keep both long-term and short-term dependency by adding four gates in each cell [47]. It overcomes the vanishing gradient problem of RNN and has been utilized to detect mental status. A DL model dealing with LSTM for predicting the development of Alzheimer's Disease (AD) is introduced in [48]. It is demonstrated that the performance of LSTM is superior to existing algorithms for the future state prediction of AD. Driver features were fed into a LSTM model to derive the corresponding driver state identifier in [49]. Then, the identifier and contextual information were utilized to develop a multi-class driver distraction risk assessment model that considers the driver, vehicle as well as environment data. Recently, BiLSTM has been validated that it can improve the performance of one directional LSTM by adding a second layer to extend the traditional LSTM networks [25]. As not only the conventional forward but also the backward time dependency in signals can be learned with it, the process of making decisions is closer to the actual driving activity. For example, a vehicle behavior recognition model was developed in [50] with features of vehicle trajectories based on BiLSTM and achieved to classify different behaviors.

In this paper, we propose a framework for driver distraction detection based on MSaE of multi-modality physiological signals captured in simulated driving environment. Our method is to determine the appropriate downsampling rates of multi-modality physiological signals with an entropy-based method, to exploit the complexity variations of each signal while distraction with the MSaE feature, and to learn the long-term contextual information in multimodal features utilizing BiLSTM. A simulated driving experiment is designed to induce driver distraction and then executed to collect multi-modality signals. After that, MSaE is extracted to investigate drivers' mental states, and a variety of features (including entropy features, time-domain, and frequency-domain features) are calculated for comparison. In addition, different feature selectors are compared to provide a reference for feature reduction and fusion. Finally, the selected optimal features are fed into a BiLSTM model to explore the time dependent relationship in them and detect driver distraction. Thereafter, a comparative analysis of the proposed method and four close-match studies is performed to validate its feasibility and efficacy.

III. EXPERIMENT

The experiment is designed to arouse driver distraction by mobile phone use task in simulated driving environment and collect multi-modality signals to be used for driver distraction detection.

A. Subjects

This study was reviewed and approved by the Ethics Committee, Liaoning Normal University. A total of sixty physically and mentally healthy subjects participated in the experiment. All of them are right-handed and have normal or corrected to normal vision and normal hearing. Driving license, driving experience as well as using mobile phone experience are also required. Besides, they are instructed not to smoke and consume any tea, alcohol, coffee, or medicine a day before the experiment. The experiment content was introduced to all participants, and written informed consent was signed previous to the experiment.

B. Experiment Design

The experiment was conducted in the laboratory at Liaoning Normal University. The Xuan Love QJ-3A1 car driving simulator was used to simulate and display a low-traffic rural road with three lanes and gather vehicle behavioral signals. It is equipped with an interactive visual system, simulated cockpit, electronic control system, customized software, and accessory equipment, as shown in Fig. 1(a).

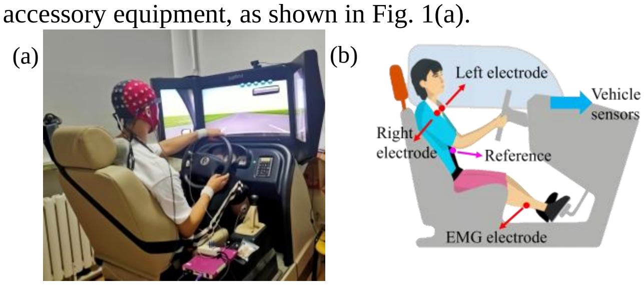

Fig. 1. Experiment apparatus. (a) The driving simulator and electrode cap. (b) The diagram of ECG, EMG, and vehicle sensors.

TABLE I
DETAILED INFORMATION OF THE EXPERIMENT SETUP

Equipment	Purposes
Xuan Love QJ-3A1 driving simulator	Simulate driving environment, collect and record vehicle data
ANT Neuro amplifier	Record physiological signals
64-channel electrode cap	Collect EEG signal
Chest lead III electrodes	Collect ECG signal
Electrode in soleus muscle	Collect EMG signal
Mobile phone	Induce distraction
DELL desktop	Save data
EEGO software	Record physiological signals
MATLAB R2023b	Analyze data

As physiological signals can reflect the internal changes of driver states more accurately and timely, a 64-channel electrode cap following the international 10-20 system was utilized to acquire EEG signals. Besides, we also gathered ECG and EMG signals simultaneously during the experiment. ECG signal was captured with chest lead III as shown in Fig. 1(b). The EMG electrode was placed along the fiber of the soleus muscle in the right lower limb. The sampling frequency of physiological signals was set as 2000 Hz. ANT Neuro amplifier and EEGO software were used to record the EEG, ECG, and EMG signals. To evaluate how distraction affects driving performance, the vehicle behavioral signals were collected in the experiment as well, and the sampling rate was kept at 1 Hz.

In this experiment, driver distraction was designed to be induced by a mobile phone use task in which drivers are forced to keep talking on the phone during the distraction process. After the experiment, the dataset was saved in a DELL desktop and analyzed with MATLAB R2023b. The details of the experimental set are concluded in Table I.

C. Procedure

There is a practice phase prior to the formal experiment to familiarize subjects with simulated driving until they understand the experiment thoroughly. The formal experiment is divided into six blocks with a duration of about one hour. Each block lasts for about 10 minutes, consisting of approximately 7 minutes of normal driving at the beginning and then around 3 minutes of distracted driving. Subjects are told to pay full attention to the road in the normal driving process while the experimenter calls and keeps talking with the subjects in distracted driving. Detailed protocol of the experiment is shown in Fig. 2. Vehicle behavioral signals and physiological signals are recorded during the experiment.

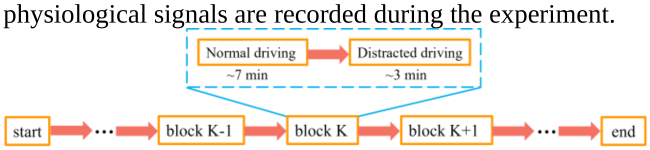

Fig. 2. Protocol of the experiment.

IV. METHODOLOGY

After the experiment, multi-modality signals were gathered and used to detect driver distraction, which is mainly consisted of three parts: feature extraction, feature selection, and distraction detection as shown in Fig. 3.

For EEG, ECG, and EMG, the collected signals were first preprocessed to remove artifacts. Then, entropy features, time-domain features as well as frequency-domain features were extracted from each modal signal. The mean value and standard deviation of the vehicle behavioral signals were calculated to study how driving performance changes with driver state. Each feature was normalized to [0,1] to avoid the influence of numerical value. To avert the high computational cost and overfitting problem induced by the obtained high dimensional feature set, five commonly used feature selectors were employed to find the most discriminative features for each signal. Finally, the optimized feature set was fed into BiLSTM to detect driver distraction, and four other classifiers were also applied for comparison.

A. Preprocessing and Feature Extraction

In this subsection, we introduce how the gathered multi-modality signals were preprocessed and what kinds of features were extracted in this study.

4) **EEG Analysis**

We first extracted the alpha rhythm of EEG employing wavelet decomposition with 4 levels, as it has been validated that alpha rhythm is highly correlated with distraction [51]. Wavelet transform (WT) has been widely used in analyzing non-stationary time series with the advantages of differentiating time series consisting of different frequencies and retaining both time and frequency information of the signals [52]. It decomposes signals with shifted and scaled versions of a mother wavelet $\psi(t)$ and a scaling function $\phi(t)$. The discrete mother wavelet can be mathematically expressed as

$$\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^{-j}t-k) \qquad k,j \in Z \qquad (1)$$

where k is used to determine the position of $\psi(t)$. j is used to determine its width and height. t and Z stand for time and integer, respectively. The original signal $S(t)$ can be defined as

$$S(t) = \sum_k s_j(k)\phi_{j,k}(t) + \sum_k d_j(k)\,\psi_{j,k}(t) \qquad (2)$$

where $s_j(k)$ and $d_j(k)$ represent the approximate and detailed coefficients at the j level. In our study, we selected the commonly used db6 wavelet as the mother wavelet.

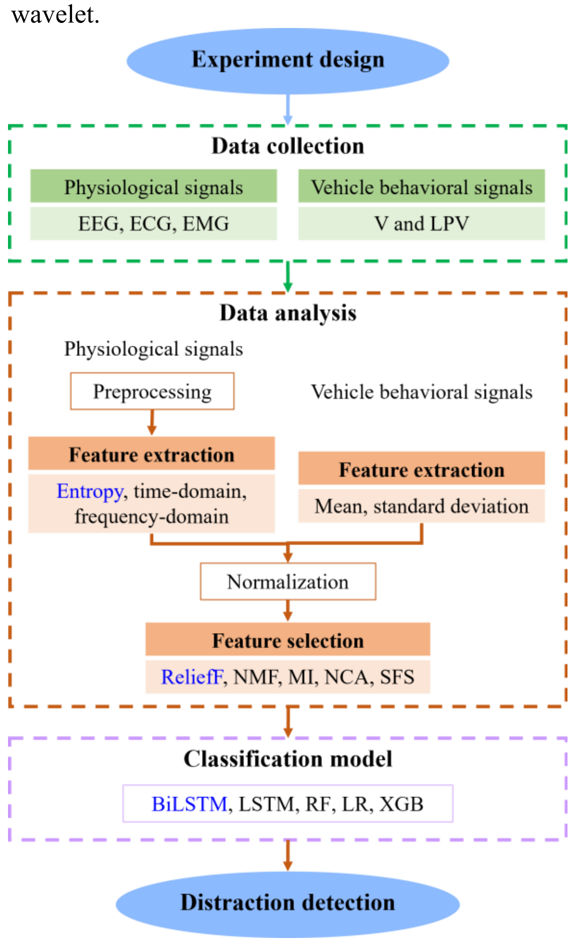

Fig. 3. Flowchart of this study.

Then, to reduce the time-consuming feature calculation of EEG, the signals need to be downsampled. We applied an entropy-based approach to find the appropriate downsampling frequency, which is mainly on the basis of MSE and MSaE.

MSE, developed in [53], can quantify the complexity of signals over various time scales as an extension of SE. MSE algorithm can be briefly summarized as coarse-graining and SE calculation.

a) For the given EEG signal $\{x_1, \ldots, x_i, \ldots, x_N\}$, consecutive coarse-grained time series $\{y^{(\tau)}\}$ at time scales τ should be constructed at first by averaging the τ data points in successive non-overlapping windows (see Fig. 4). The element of $\{y^{(\tau)}\}$ can be calculated by the following equation

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i \qquad 1 \le j \le N/\tau \qquad (3)$$

where N is the length of the given EEG signal. N/τ is the length of each coarse-grained time series. j is the jth value of the coarse-grained signal.

b) SE is then calculated for coarse-grained time series. As a measure of the complexity of signals, it has been adopted in EEG analysis and is detailed introduced in [54].

Previous studies have illustrated that the values of single scale entropy features are related to sampling rate, and stretching or compressing effect appears in MSE with sampling frequency changes [38]–[40]. In the present study, we explored an extension of MSE (i.e., MSaE) that replaces the traditional relative time scale τ with absolute time scale λ with unit second,

$$\lambda = \tau \cdot T_s = \tau / f_s \qquad (4)$$

where T_s and f_s are the sampling period and sampling rate of the original EEG, respectively.

In this way, the coarse-graining process can be described as constructing the coarse-grained time series by averaging the data points every λ seconds in successive non-overlapping windows as shown in Fig. 4. There should be noticed that the time duration t_D of EEG keeps the same before and after downsampling [40], i.e., t_{D_d}=t_D. The rest of the procedure to calculate MSaE is the same as MSE. Hence, MSaE feature can be obtained by calculating SE of the coarsely-grained time series under absolute scale factors.

Now, the entropy-based approach to determine the proper downsampling frequency for the given EEG signal X can be illustrated as:

a) Calculate MSE at different time scale τ and draw the entropy-scale curve to find peak scale τ_{peak}, the scale that MSE reaches its maximum value (see Fig. 5). High correlation between MSE value and time scale τ_{peak} has been demonstrated in [55].

b) Calculate the corresponding peak time λ_{peak} on the basis of (4).

c) Since downsampling has no effect on time duration, peak time after downsampling λ_{peak_d} is the same as before downsampling, i.e., λ_{peak_d}= λ_{peak}.

d) According to (4), the downsampling rate f_{s_d} can be expressed as

$$f_{s\_d} = \frac{\tau_{peak\_d}}{\lambda_{peak\_d}} = \frac{\tau_{peak\_d}}{\lambda_{peak}} \qquad (5)$$

where τ_{peak_d} is the peak scale after downsampling.

e) Finally, we can get f_{s_d} for a given peak scale after downsampling τ_{peak_d}.

In the present study, we calculated the entropy-scale curve of alpha rhythm as shown in Fig. 5. τ_{peak} is 60 and λ_{peak} is 0.03 seconds. Let τ_{peak_d} = 6 referring to the recommended scale range in [56], then the appropriate downsampling rate is calculated as 200 Hz.

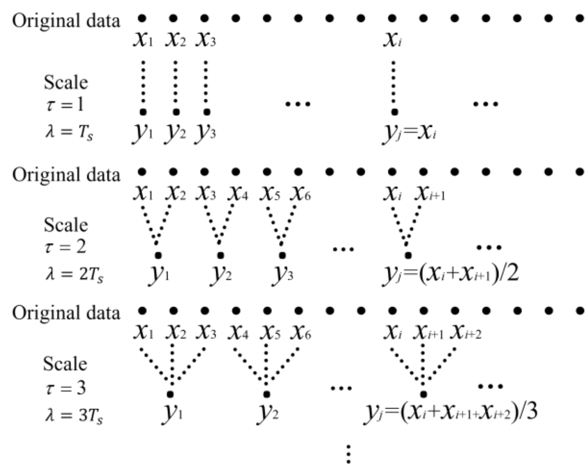

Fig. 4. Schematic illustration of the coarse-graining process.

After the alpha rhythm was downsampled to 200 Hz, we employed a wavelet-based technique for artifacts removal. The approximate and detailed coefficients mentioned above show the correlation between the mother wavelet and the signal, and larger coefficients appear if artifacts exist. These coefficients can be reduced with a threshold and have been used in the analysis of driver fatigue [57]. The threshold is expressed as

$$T_m = mean(C_m) + 2 \times std(C_m) \qquad (6)$$

where C_m is the wavelet coefficient at the mth level. If the coefficient's value is larger than the predefined threshold it is halved. Then the signal without artifacts can be obtained with the new coefficients.

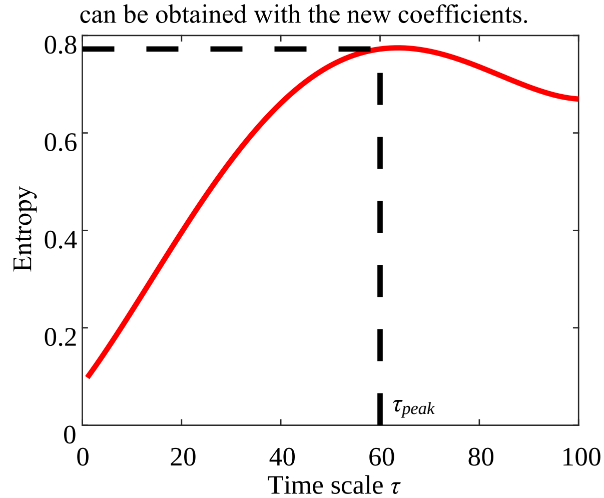

Fig. 5. Entropy-scale curve of the original alpha rhythm with f_s=2000 Hz.

Except for MSaE and MSE, 11 other features of alpha rhythm from the time domain, frequency domain, and complexity were extracted for comparison. Time-domain features include mean value (Mean), standard deviation (STD), skewness, and kurtosis. The amplitude spectrum (AMP) and power spectrum density (PSD) were calculated from the perspective of frequency. As for complexity features, we chose to use DE, FE, approximate entropy (AE), Rényi entropy (RE), and fractal dimension (FD). The detailed algorithms are listed in [35].

2) **ECG Analysis**

ECG signal is the record of the electrical activities of heart, which contains various artifacts such as device, movement, and breath interferences [58]. In order to eliminate these artifacts, it is necessary to preprocess the collected ECG signal. Firstly, the downsampling rate was calculated to decrease the extensive calculation according to the entropy-based resampling approach mentioned above. In this way, the resampling rate of 256 Hz was obtained with a scale of 6. Then, a fourth-order Butterworth band-stop filter of 50 Hz was designed to remove power line noise, and a band-pass filter of 0.7 Hz to 40 Hz was applied to eliminate the baseline drift caused by movements and breath. After preprocessing, the ECG signal can be used to analyze the activities of the heart.

The MSaE and MSE features of ECG in a time window of 5 seconds with 4 seconds overlapping were applied to explore the complexity of ECG. Besides, existing research has pointed out that heart rate variability (HRV) derived from ECG is a reliable indicator for analyzing heart activities [59], [60]. By R-peaks (heart beats) detection, the HRV indicators can be extracted. In this study, we adopted a dynamic threshold approach according to [61] to locate R peaks in each block. Thereafter, 8 HRV features from both the time domain and frequency domain were extracted in the 60-second sliding window with 59 seconds overlapping. The time-domain features are the number of heart beats per minute (BPM), the mean distance (IBI) and standard deviation (SDNN) of R-R intervals, the standard deviation (SDSD) and the root mean square (RMSSD) of successive differences between adjacent R-R intervals. The frequency-domain features involve low-frequency (LF, 0.04-0.15 Hz) spectral component, high-frequency (HF, 0.16-0.5 Hz) component, and the ratio of LF and HF (LF/HF). A detailed description of these measures can be found in [59], [62].

3) **EMG Analysis**

The data was firstly band-pass filtered with a fourth-order Butterworth band-pass filter and the low and high cutoff frequencies were set as 20 Hz and 500 Hz, respectively in accordance with [63]. A band-stop filter was then utilized to remove the interference of the power line, after which a sliding window of 125 milliseconds was employed to average the filter EMG data. Finally, to observe the changes of EMG before and after distraction, we extracted 8 features within a window length of 60 seconds and 59 seconds overlapping. The features are MSaE, MSE, AE, the mean absolute value of amplitude (MAV), the root mean square of EMG amplitude (RMS), zero crossing (ZeroCross), mean power frequency (FMEAN), median frequency (FMED). More details of these measures are described in [64], [65].

4) **Vehicle Behavioral Signals Analysis**

The velocity (V) and lane position variability (LPV) signals were analyzed to validate whether driving performance does alter with distraction. We performed statistical analysis for the signals so as to explore how behavioral alterations happen between normal driving and distracted driving. The mean value and standard deviation of the data were calculated with the same sliding window as ECG and EMG.

B. Feature Selection

After preprocessing and feature extraction of the physiological signals, we get a feature set with 31 kinds of features with a dimension of 811. These indicators are in different numerical ranges, which has an effect on the comparison among them. Thereupon, each feature is normalized to [0,1] applying (7) to avoid the numerical influence.

$$y = \frac{(y_{max}-y_{min})*(x-x_{min})}{x_{max}-x_{min}} + y_{min} \qquad (7)$$

where y is the normalized value, x is the original value, x_{max} and x_{min} are the maximum and minimum values of the original feature, respectively. y_{max} and y_{min} stand for the corresponding normalized maximum and minimum values separately.

Since the dimension of the extracted feature set for distraction detection is high, it is time-consuming and may induce overfitting problem if all of them are put into the classifier. On this account, it is necessary to fulfill the feature selection procedure to reduce features' redundancy and improve efficiency. Five commonly used feature selectors were recruited in this study to find out the most discriminative features [66]-[70]. They are briefly illustrated below.

1) **ReliefF Algorithm**

ReliefF algorithm is an efficient feature weighting algorithm that assigns the features with different weights in agreement with the correlation between features and categories [66]. The greater the weight of the feature, the better it is for classification. It explores the correlation by searching for the nearest neighbor of each feature from all categories utilizing Manhattan Distance. As the feature weight vector is stable and ReliefF can do with noise and incomplete problems, it is first selected for feature selection in our study.

2) Non-negative Matrix Factorization (NMF)

NMF is an unsupervised algorithm processing high dimensional data with matrix factorization. It searches for non-negative basis and coefficient matrix to parts-basely represent a non-negative data matrix, which retain the discriminative information of the data. Then, reduce the dimension of the data matrix. Extensive research has proved that NMF is effective in dealing with high-dimensional data for feature selection in the fields of physiology and neuropsychology [67].

3) Mutual Information (MI)

MI measures the amount of information that one variable has about the other and provides a way to estimate the dependency between features and categories. The feature with a larger MI value is more discerning to the corresponding category. With the advantages of detecting nonlinear relationships between variables and analyzing multidimensional variables, MI has been a popular approach for feature selection [68].

4) Neighborhood Component Analysis (NCA)

NCA is a non-parametric method for selecting features with the goal of maximizing the prediction accuracy of classification. It optimizes the leave-one-out classification by means of the first-order nearest neighbor and then obtains the weight of each feature [69]. By analyzing the weight vector with a threshold, the most relevant feature subset can be determined. Besides, the result is barely affected by the increase in the number of irrelevant features.

5) Sequential Forward Selection (SFS)

SFS is a bottom-up search process that begins with an empty feature set and subsequently adds features on the basis of minimizing the mean square error [70]. The process continues until the performance achieves the highest, and the corresponding feature subset is regarded as the selected features. It is widely used for dimension reduction due to its simplicity and saving computing time.

D. Classification

In order to find out the most discriminative features for distraction detection, we recruited random forest (RF) to compare the performance of different feature selectors. After that, the new set of features coming from the best feature selector will be input into a BiLSTM classifier to detect driver distraction.

1) Random Forest

RF is a kind of bagging model for classification consisting of many randomized decision trees in accordance with ensemble learning technique. Each tree has its local decision, and the final decision is predicted by the majority votes of all trees. The tree is trained with a bootstrap aggregating technique that randomly selects the subsets with replacements from the original training data [71]. Therefore, the stability and robustness of the classifier is increased. In addition, it can also estimate the importance of the inputs. Considering the two merits, it is adopted to evaluate the most efficient feature selector among ReliefF, NMF, MI, NCA, and SFS in the present study.

2) BiLSTM

LSTM, a variant of RNN, addresses the vanishing gradient problem and can keep both long and short-term memory of a longer time series. A memory cell configuring with four neural layers is used to learn long-term dependency and store valuable information in the context (see Fig. 6). The LSTM cell utilizes three so-called gates (i.e., forget gate, input gate, and output gate) to control the information propagation in the network instead of directly overwriting the cell information in RNN [47].

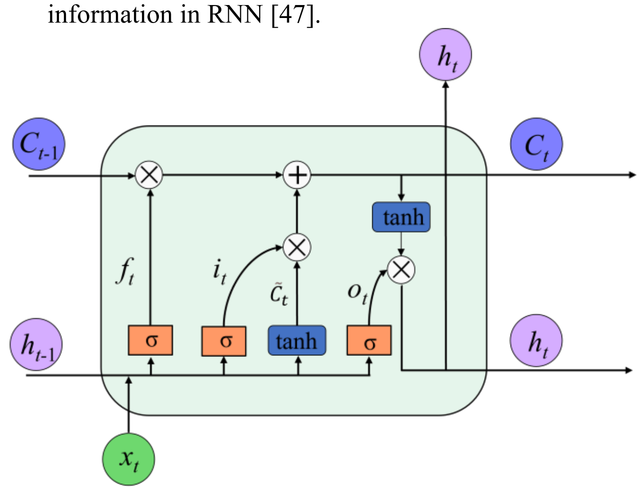

Fig. 6. The basic architecture of LSTM memory cell.

The input data is transferred to every gate by a sigmoid activation function. The forget gate is first used to determine whether there is information that needs to be discarded from the inner cell state (8). Then, the input gate controls the amount of new information that can be stored in the cell following three steps. Step 1 involves a sigmoid layer exploring any updated information using (9). After that, a tanh layer creates $\widetilde{C}_t$, which is the new candidate cell according to (10). The new cell state C_t substitutes the old cell state C_{t-1} with (11) in the last step. Finally, the output h_t is determined by the output gate by (12) and (13).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{8}$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{9}$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{10}$$
$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{11}$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{12}$$
$$h_t = o_t * \tanh(C_t) \tag{13}$$

where x_t is the feature of time t. σ and tanh are the sigmoid and tanh activation functions in the cell, respectively. W, b, and h separately indicate the weight, bias, and hidden state of each gate.

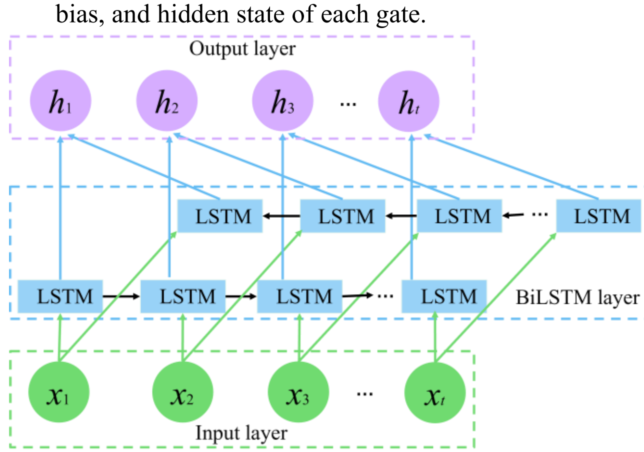

Fig. 7. The basic topological structure of BiLSTM.

For driving activity, it is beneficial to classify the current state by accessing both the past and future information. BiLSTM, introducing a second LSTM layer, fulfills to memorize the forward as well as the reverse contextual information in sequence data. Thus, it is more like the process of the driver making decisions. The basic structure of a BiLSTM network is shown in Fig. 7.

V. RESULTS AND DISCUSSION

Driver distraction has an effect on drivers from multiple aspects and threatens driving safety. It not only causes physiological changes in drivers but also affects driving performance, which are manifested in the dynamic changes in physiological and vehicle behavioral signals [72]. The dynamic variations in the physiological signals and the distinctions in driving performance between normal and distraction are here analyzed and discussed. A driver distraction detection model with the ability of memorizing long-term context information is proposed based on the extracted features. The results of single-modality features are compared with those of multi-modality features to validate whether taking advantage of multi-modality signals is conducive to exploring more characteristics of driver distraction. We also compared our approach with traditional single-directional LSTM and four close-match methods in previous studies to prove its feasibility and effectiveness.

A. Analysis of EEG Signal

A total of 13 features were calculated from the alpha frequency band in the present work, and we mainly focus on the MSaE feature to evaluate the complexity of EEG because of its advantages mentioned above. After determining the appropriate downsampling rate and scale parameter, the MSaE feature was extracted in a sliding time window of 5 seconds with 4 seconds overlapping.

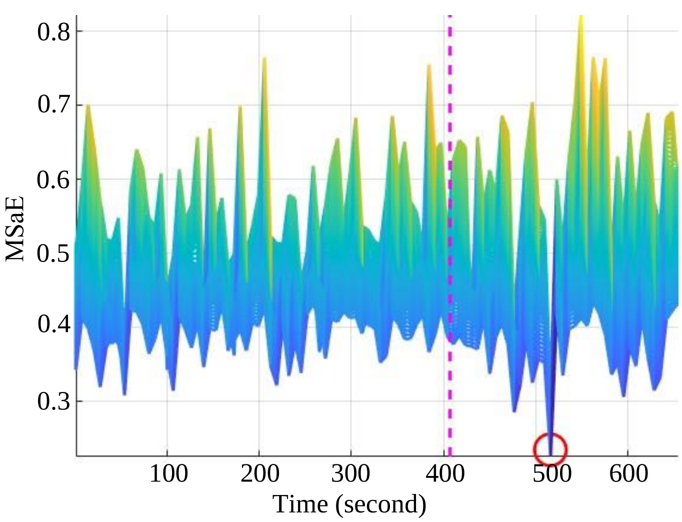

Fig. 8. The MSaE result of alpha rhythm. The magenta dash line shows the onset of using mobile phone, and the red circle is the minimum MSaE value.

As shown in Fig. 8, the waveform of MSaE fluctuates gently in the normal driving process, while it is more obvious in the distracted driving process. This phenomenon indicates that the complexity of the alpha frequency band is less stable while distraction. Besides, a minimum value appears in a while after the onset of the distraction task, and the trough of the waveform is apparently lower than the average MSaE value. It represents that the complexity of alpha rhythm arrives at its minimum at this point. Drivers need to allocate most of the brain resources to pay full attention to the surroundings and keep driving safely in normal driving [73]. In this case, a driver's attention is focused, and brain activity is vibrant with a much stable activation degree. The complexity of EEG then changes slightly, which embodies the relatively stable MSaE waveform. Different from the normal driving process, drivers' alertness reduces because of the distraction task, they also have to handle multiple tasks (i.e., driving and the distraction task) in the distraction process, the activation degree of brain activity is thereby not as stable as normal driving. As a result, the fluctuation of MSaE is evident. What's more, it normally takes a while for drivers to shift their attention to distraction task and be distracted to the maximum [72]. The brain activity decreases to the lowest under this condition. Therefore, a trough appears in MSaE sooner after the onset of the task.

B. Analysis of ECG Signal

Two entropy features, five time-domain features, and three frequency-domain features of ECG were extracted to find out the dynamic fluctuation of heart rate during distraction. Fig. 9 shows the obtained MSaE feature.

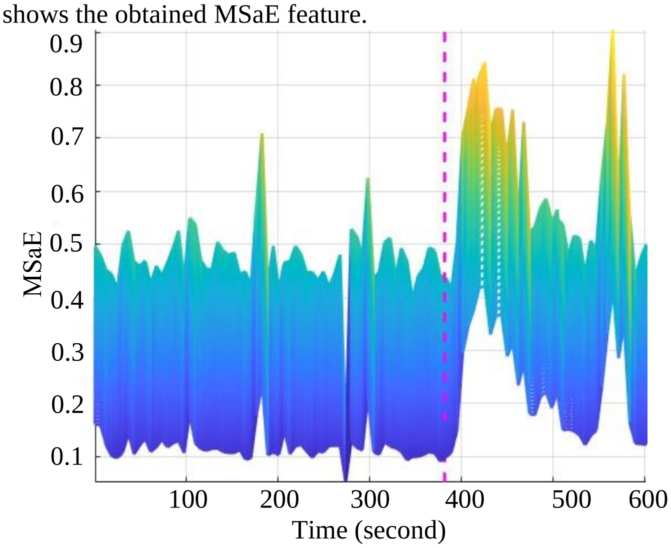

Fig. 9. The MSaE result of ECG. The magenta dash line shows the onset of the distraction task.

From an overall perspective, the feature of ECG shows an increasing tendency after the onset of the using mobile phone and rises to its maximum value later. The average MSaE value also augments after the task. In addition, the fluctuation of the waveform is more obvious after distraction than before distraction. In normal driving, drivers only focus on driving safely, and their heart rates are relatively regular and stable. On this occasion, the complexity of ECG slightly changes with driving, thus a stable waveform appears with lower values. While in distracted driving, drivers have to perform multiple tasks to satisfy the needs of answering phone and keeping safe at the same time, which leads to their nerves and decreases the regularity and stability of ECG. As a result, an augmentation occurs in the value of MSaE after distraction, and the fluctuation of the waveform gets more obvious. The dynamic alteration of MSaE aligns with the previous study that analyzed ECG-based measures to monitor the physiological changes while distraction [74]. Its results pointed out that entropy features of ECG project higher values while distraction and that the complexity of ECG generates a sign of increasing during distraction. The reason is that multitasking (i.e., driving and secondary task) while distraction increases driver's workload and changes the functional state of the heart [75]. Hence, the complexity of ECG signal augments. The results can also be explained from the biological aspect as reported in [76]. The variation in heart rate is brought about by the blood regulation executed by the heart. During distracted scenarios, more oxygenated blood needs to be constantly supplied to the brain in order to satisfy the demands of multitasking. The increment of oxygen demand then increases blood flow, prompting the heart to contract and relax more frequently with lower rhythmicity and ultimately resulting in the changes in ECG feature.

C. Analysis of EMG Signal

We currently calculated three entropy, three time-domain, and two frequency-domain features for the EMG signal so as to study the differences in muscle state between normal and distracted drivers. The MSaE feature is shown in Fig. 10.

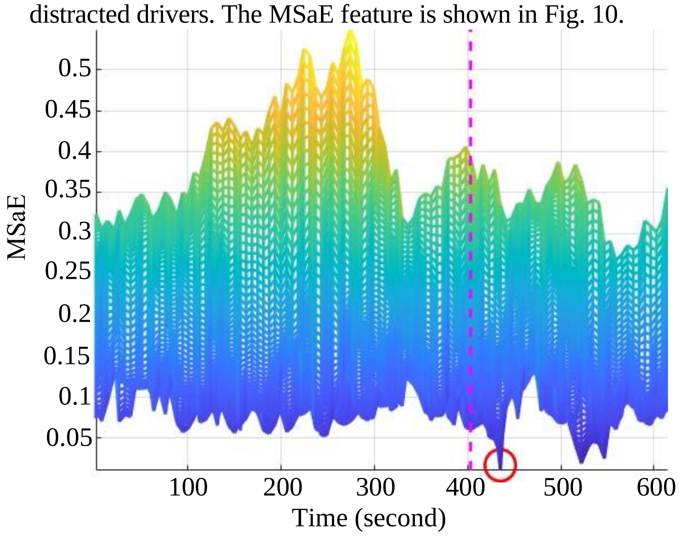

Fig. 10. The MSaE result of EMG. The magenta dash line shows the onset of the distraction task. The red circle is the minimum value.

The values of MSaE show a slightly decreasing trend after drivers start to do the task until the minimum value is reached, which is evidently smaller than the mean value of MSaE in Fig. 10. The reason why it declines and the minimum comes up may be that the muscle tension gets lower at this stage. EMG signal is argued to be connected with human mental states in [77]. High EMG tonus happens if a person is awake, and lower EMG tonus arises while drowsiness. In our study, drivers remain vigilant to ensure driving safety with highly focused attention in normal driving. Consequently, they exhibit strong ability of controlling limbs with their muscles in an active and highly tense state, and then increasing the complexity of the ECG signal. However, drivers' attention is distracted to answering phones after the onset of the distraction task. The limbs control then recedes, giving rise to the lower muscle vitality and tension. In this way, a decline in the complexity of EMG happens. In addition, drivers normally spend a while to be distracted to the maximum [72], thereby, the muscle tension decreases to the lowest under the circumstances resulting in the minimum of MSaE. Moreover, the phenomenon can be further verified by the changes in vehicle behavioral data in Fig. 11.

D. Analysis of Vehicle Behavioral Signals

Statistical analysis was performed for the acquired vehicle behavioral signals to investigate how the driving performance changes after distraction. The mean value and standard deviation were calculated for V and LPV signals and shown in Fig. 11.

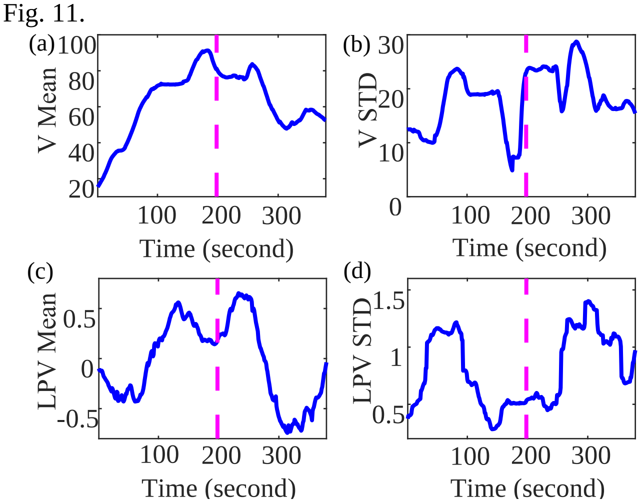

Fig. 11. The features of vehicle behavioral signals. (a) V Mean value. (b) V standard deviation. (c) LPV Mean value. (d) LPV Standard deviation. The magenta dash line shows the onset of the distraction task.

The mean value of velocity shows a downward trend after using mobile phone in Fig. 11(a), which illustrates that drivers prefer to decrease the velocity so as to keep safe after the beginning of the task. The result is in accordance with the existing research that driving speed displays a significant decrease trend when drivers execute secondary tasks [78]. In Fig. 11(c), the amplitude of LPV mean value rises due to the distraction task. It reveals that the lateral distance of the vehicle from the road center line increases after distraction. What’s more, the standard deviation of V and LPV also augments after distraction in Fig. 11(b) and Fig. 11(d). These findings are in line with current driving performance analysis in the literature. The effects of phone use on driving performance were studied in [79], and it is demonstrated that drivers have lower ability of longitudinal control and show higher variations in lane positioning when attention is shifted from driving to task-related work. In our experiment, drivers consume more attention resources on the distraction task, the limbs control as well as the muscle tension are thereby weakened (see Fig. 10). Sequentially, diminished ability in lane and velocity keeping occurs. Therefore, the velocity declines, and the amplitude of LPV increases after distraction. Moreover, greater standard deviations can be observed in V and LPV.

E. Performance Evaluation of Different Feature Selectors

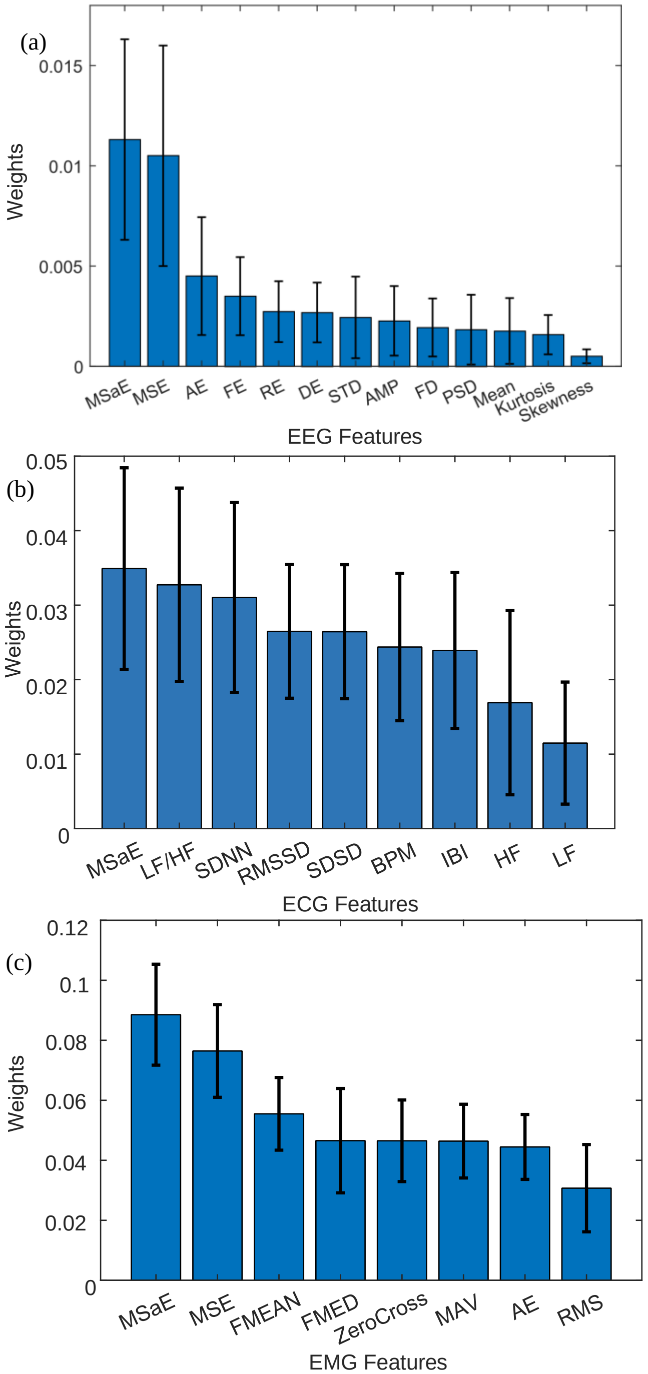

Fig. 12. The mean weights of different features obtained with ReliefF. (a) EEG features. (b) ECG features. (c) EMG features. The error bar shows the standard deviation.

As described in Section IV, there were 31 types of features extracted from EEG, ECG and EMG signals. The dimension of the obtained multi-modality physiological feature set is 811, which may result in the overfitting problem. Hence, these features of each modality signal were fed into five feature selectors after normalization to explore the most discriminative features.

Two steps were adopted in the procedure of EEG feature selection. Firstly, the three most discriminative electrodes were searched among all the 61 electrodes for each EEG feature. Then, the best feature for each electrode could be identified by analyzing the weights of the thirteen features (see Fig. 12(a)). Different from EEG, the electrode selection step was skipped for ECG and EMG features, and one best feature was chosen for each modal signal. Fig. 12 shows the feature weights for each signal. It is clear that the weight of MSaE ranks the first among all features followed by MSE for each modal physiological signal utilizing ReliefF. Since MSaE and MSE are calculated in multiple time scales, they can explore more valuable information about distraction than traditional single scale features, and the residual noise effect can also be reduced by multiple scale calculation [53]. Additionally, MSaE can remain unchanged after resampling by using absolute time scales in coarse-graining instead of relative time scales used in MSE [40]. Hence, MSaE outperforms MSE and other single scale features. The best feature selected by the feature selectors was then input into a RF classifier to estimate the most effective selector. The corresponding results are shown in Fig. 13. It is clear that the optimal feature selector is ReliefF for the feature set obtained by it displays higher accuracies than those of the other four selectors. On this occasion, the feature set selected by ReliefF was finally utilized to detect whether a driver is distracted or not in the following study.

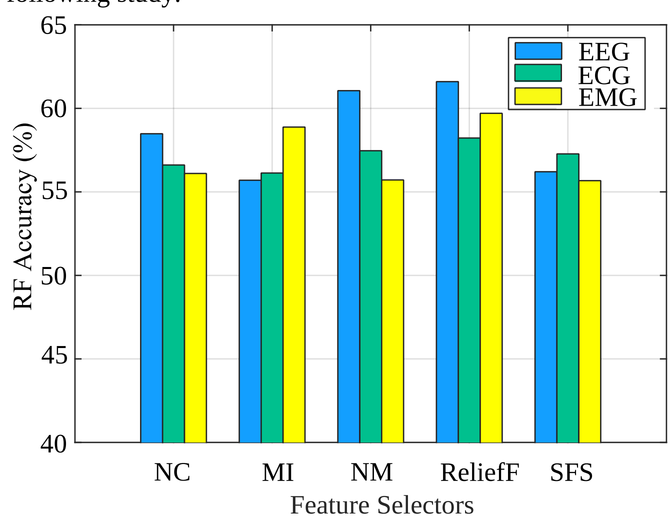

Fig. 13. The Performance of different feature selectors.

F. The Classification Results

After acquiring the optimal feature selector and the most discriminative feature for each kind of physiological signal, the new feature set was fed into a BiLSTM model to detect distraction. There are two categories in our experiment i.e., normal and distraction. The input feature matrix was divided into training set and testing with a ratio of 8:2. In the present work, we conducted a comparative analysis of physiological signals in the context of distraction detection. Specifically, their performance was evaluated not only when employing the most effective feature to detect distraction, but also when using all extracted features as inputs for classification. Besides, we further investigated whether different signals can mutually provide complementary information by adopting the multi-modality feature set for distraction detection. Moreover, the performance of our proposed method was compared with that of LSTM and four closely matched research studies.

Table II presents the results of the BiLSTM model with physiological features, and Table III lists the results of vehicle features. In Table II, "MSaE" is the selected best feature for each kind of physiological signal, and "MSE" is for comparison with MSaE. "All features" stands for all the extracted features of ECG and EMG signals, while it means all the features extracted from the selected electrodes of the EEG signal. "ALL" is the feature set consisting of EEG, ECG, and EMG features. "Multi-modality" in Table III represents the multi-modality feature set comprising not only physiological features but also vehicle features.

TABLE II
THE MEAN ACCURACIES OF DIFFERENT PHYSIOLOGICAL FEATURES USING BILSTM (%)

	EEG	ECG	EMG	ALL
MSaE	76.12	68.27	65.49	79.89
MSE	72.96	68.33	63.87	76.45
All features	62.75	66.5	65.28	62.31

In Table II, MSaE, with accuracies of 76.12%, 68.27%, and 65.49% for EEG, ECG, and EMG separately, performs better than MSE in detecting distraction for each physiological signal. Both of them contribute more than the other features in detecting distraction, as can be inferred from Fig. 12. It is due to the mitigated impact of residual noise on exploring the complexity of physiological signals by multiple scale calculation and because MSaE remains unaffected by resampling unlike MSE [16], [40]. Besides, MSaE of EEG contributes a higher accuracy of 76.12% by contrast with that of ECG and EMG when only one feature was used for classification. A similar conclusion can be drawn when inputting MSE to detect distraction. Since EEG is motivated by CNS, it can respond more promptly with higher resolution to changes in mental state than other signals generated by ANS [80]. In addition, the EMG feature exhibits the lowest accuracies in the single feature conditions, with an accuracy of 65.49% for MSaE and 63.87% for MSE. The performance of the ECG feature lies between EEG and EMG features while employing one feature to detect distraction. Moreover, detecting distraction with the selected MSaE feature outperforms that with multiple features for each signal, which demonstrates the necessity of feature selection. The results are consistent with previous studies [81], [82]. Their research has clarified that high-dimensional feature set not always brings with high accuracy in classification tasks, and that it is necessary to filter out redundant features and select the proper relevant features for saving time and avoiding over-fitting. When utilizing MSaE of the three modalities of physiological signals as inputs simultaneously, the model performance was further enhanced, reaching the peak accuracy of 79.89%. The accuracy of BiLSTM with multi-modality feature set also increased under the condition of selecting the MSE feature of each signal. Furthermore, when adding the multi-modality physiological features to vehicle features, the detection accuracy is promoted again with an augment of about 8% compared to vehicle features in Table III. Coinciding with previous studies, different signals manifest human mental states from a variety of aspects, they can compensate for each other and provide more valuable and sufficient information of distraction, thus promoting the prediction accuracy [20], [22]. Additionally, BiLSTM is slightly better than the traditional LSTM when detecting driver distraction as shown in Table III. BiLSTM has the ability of memorizing the forward and backward long-term dependency among features, so more sufficient information about driver distraction can be learned than traditional LSTM contributing to a higher accuracy in distraction detection.

TABLE III
THE MEAN ACCURACIES OF VEHICLE AND MULTI-MODALITY FEATURES (%)

	Vehicle	Multi-modality
LSTM	72.61	81.27
BiLSTM	74.72	82.86

Table IV shows the comparison results of the proposed method (i.e., multimodal MSaE+BiLSTM) and previous studies in detecting driver distraction utilizing our dataset. It is obvious that the proposed method with an accuracy of 82.86% in our study outperforms the other four methods in driver distraction detection. In [83], the mean value and absolute gradient (Abs) of LPV, ECG, and respiration rate were extracted to detect whether a driver uses a cellphone or not based on logistic regression (LR). For comparison, we calculated the two kinds of features from the gathered signals and an accuracy of 63.75% was achieved with LR. As for the approach in [84], six features of EEG and contextual data were employed to detect driver distraction with a RF classifier in simulated environment. But this approach did not perform so well while recruiting our data. There are two reasons for explanation of the results. Firstly, MSaE can explore the complexity of signals from multiple time scales and more valuable information about distraction can be represented than traditional single scale features [53]. This can also be validated in the feature selection procedure (see Fig. 12). Secondly, with the advantages of memorizing two directional information flow, BiLSTM is more similar to the decision-making process of drivers and is beneficial for reducing the train and test error of the model [85]. We also calculated various complexity features, time-domain features and frequency-domain features used in [86] to develop an extreme gradient boosting (XGB) classifier for distraction classification and got an accuracy of 68.23%. As we know, there may be feature redundancy for a high dimensional feature set, which may induce the overfitting problem and affect the classification performance [66]. Besides, XGB is an ensemble algorithm based on gradient boosted decision trees developed and cannot learn the contextual dependency among features. On the above two bases, our method is superior to the algorithm employed in [86]. When compared to our previous research in [87], the accuracy of the current approach is about 4% higher recruiting MSaE rather than MSE, which can also be illustrated in Table II. This is because MSaE mitigates the residual noise effect on feature calculation by multiple scale calculation, and its value can remain unaffected by resampling [16].

TABLE IV
COMPARISON WITH CLOSE MATCH STUDIES USING OUR DATASET (%)

Reference	Features	Classifiers	Accuracy
[83]	Mean, Abs	LR	63.75
[84]	PSD, SE, ZeroCross, RMS, Mean, Skewness.	RF	64.09
[86]	Complexity features, time-domain and frequency-domain features.	XGB	68.23
[87]	MSE	BiLSTM	78.41
Our work	MSaE	BiLSTM	82.86

In general, the highest accuracy is obtained by the MSaE feature of EEG when employing one feature of single modality physiological signal. In addition, MSaE is superior to the other single or multiple features of physiological signals adopted in our present study. Furthermore, the performance of BiLSTM can be enhanced if features of multiple modalities signals are simultaneously recruited.

The results in our present work indicate that MSaE can potentially exploit the discriminating distraction information of physiological signals and that multiple modalities features can enhance the performance of driver distraction detection. Despite the reliability and superiority of the proposed framework, limitations still exist in the current study. We explored the features of EEG with few electrodes, which to some degree ignores the spatial information of EEG and the interaction between electrodes from various perspectives. Besides, we used the features of each signal to detect driver distraction regarding each model signal as independent variable. But whether there is an interrelationship among different modalities of signal still needs to be further studied. What's more, our current work only studied driver distraction induced by using a mobile phone. But there are also many factors that can lead to driver distraction such as eating and talking with passengers. So, it is valuable to study how different kinds of driver distraction affect drivers and then to develop a comprehensive model for various driver distraction detection.

VI. CONCLUSION

This paper proposed a driver distraction detection method based on MSaE of multi-modality physiological signals. It has validated that the MSaE of multiple physiological signals in distracted driving is apparently different from that in normal driving. Besides, MSaE is better for analyzing the complexity of physiological signals than other conventional features employed in previous research. What's more, the detection accuracy is further improved by adopting features of multi-modality signals compared to single modality feature. The results confirm that incorporated signals supply complementary useful information of driver distraction. When a driver is distracted, physiological and behavioral changes come into being. The MSaE value of the EEG signal descends to a trough soon after the task onset. The heart rate tends to be less regular and stable as the distraction task goes on. Moreover, the ability of muscle and vehicle control also recedes, which can be manifested by the lower complexity of EMG, decreased velocity, and rising variability of V and LPV.

In future work, we will mine the distraction information of EEG on the basis of the brain functional network to improve the ability of exploring the spatial dependencies between electrodes. In addition, we will explore whether interrelationships exist in different signals and how to use their dependency in distraction detection. Furthermore, we will include more distracting tasks in the experiment design and study how to detect different types of distraction accurately.

ACKNOWLEDGMENT

We gratefully acknowledge the financial support from the National Natural Science Foundation of China (grant number: 61703069 and 62001312), the Fundamental Research Funds for the Central Universities (grant number: DUT21GF301) and the Science and Technology Planning Project of Liaoning Province (grant number: 2021JH1/10400049).

REFERENCES

[1] "Distracted driving," National Highway Traffic Safety Administration, Washington, USA. 2022. [online]. Available: https://www.nhtsa.gov/risky-driving/distracted-driving
[2] "Road traffic injuries," World Health Organization, Geneva, Switzerland, Dec. 2023. [online]. Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries
[3] E. Ohn-Bar, S. Martin, A. Tawari, and M. M. Trivedi, "Head, eye, and hand patterns for driver activity recognition," in *2014 22nd Int. Conf. Pattern Recognit., Stockholm, Sweden,* Aug. 2014, pp. 660–665.
[4] F. Yan, M. Liu, C. Ding, Y. Wang, and L. Yan, "Driving style recognition based on electroencephalography data from a simulated driving experiment," *Front. Psychol.*, vol. 10, 2019, [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01254
[5] A. Aksjonov, P. Nedoma, V. Vodovozov, E. Petlenkov, and M. Herrmann, "A method of driver distraction evaluation using fuzzy logic: Phone usage as a driver's secondary activity: Case study," in *2017 XXVI Int. Conf. Inf., Commun. Autom. Technol. (ICAT)*, Sarajevo, Bosnia and Herzegovina, Oct. 2017, pp. 1–6. doi: 10.1109/ICAT.2017.8171599.
[6] G. Tillman, D. Strayer, A. Eidels, and A. Heathcote, "Modeling cognitive load effects of conversation between a passenger and driver," *Atten. Percept. Psychophys.*, vol. 79, no. 6, pp. 1795–1803, 2017, doi: 10.3758/s13414-017-1337-2.
[7] J. Healey and R. Picard, "SmartCar: detecting driver stress," in *Proc. 15th Int. Conf.Pattern Recognit., ICPR-2000*, Barcelona, Spain, Sep. 2000, pp. 218–221. vol.4. doi: 10.1109/ICPR.2000.902898.
[8] B. Zhang, W. Chang, and X. Li, "Fatigue driving detection based on spatial-temporal electroencephalogram features and parallel neural networks," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 23, no. 2, pp. 315, 2023.

[9] H. Zeng *et al.*, "EEG classification of driver mental states by deep learning," *Cogn. Neurodyn.* Vol. 12, pp. 597–606, 2018, https://doi.org/10.1007/s11571-018-9496-y

[10] G. Li, W. Yan, S. Li, X. Qu, W. Chu, and D. Cao, "A temporal–spatial deep learning approach for driver distraction detection based on EEG signals," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 4, pp. 2665–2677, 2022, doi: 10.1109/TASE.2021.3088897.

[11] S. Ahn, T. Nguyen, H. Jang, J. G. Kim, and S. C. Jun, "Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and FNIRS data," *Front. Hum. Neurosci*. Vol. 10, no. 219, 2016, https://doi.org/10.3389/fnhum.2016.00219

[12] S. Haufe, M. Treder, M. Gugler, M. Sagebaum, G. Curio, and B. Blankertz, "EEG potentials predict upcoming emergency brakings during simulated driving," *J. Neural Eng.*, vol. 8, no. 5, pp. 056001, 2011, doi: 10.1088/1741-2560/8/5/056001

[13] R. N. Khushaba, S. Kodagoda, D. Liu, and G. Dissanayake, "Muscle computer interfaces for driver distraction reduction," *Comput. Meth. Programs Biomed.*, vol. 110, no. 2, pp. 137–149, 2013, doi: https://doi.org/10.1016/j.cmpb.2012.11.002.

[14] Z. Zhang *et al*., "Efficient sleep classification based on entropy features and support vector machine classifier," *Physiol. Meas.*, vol. 39, Oct. 2018, doi: 10.1088/1361-6579/aae943.

[15] J. Urigüen and B. Zapirain, "EEG artifact removal – State-of-the-art and guidelines," *J. Neural Eng.*, vol. 12, p. 031001, Apr. 2015, doi: 10.1088/1741-2560/12/3/031001.

[16] J. Gao, J. Hu, F. Liu, and Y. Cao, "Multiscale entropy analysis of biological signals: a fundamental bi-scaling law," *Front. Comput. Neurosci.*, vol. 9, 2015, [Online]. Available: https://www.frontiersin.org/articles/10.3389/fncom.2015.00064

[17] K. Ryu and R. Myung, "Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic," *Int. J. Ind. Ergon.*, vol. 35, no. 11, pp. 991–1009, 2005, doi: https://doi.org/10.1016/j.ergon.2005.04.005.

[18] B. Mehler, B. Reimer, J. F. Coughlin, and J. A. Dusek, "Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers," *Transp. Res. Rec.*, vol. 2138, no. 1, pp. 6–12, 2009.

[19] D. He, B. Donmez, C. C. Liu, and K. N. Plataniotis, "High cognitive load assessment in drivers through wireless electroencephalography and the validation of a modified N-back task," *IEEE Trans. Hum. Mach. Syst.*, vol. 49, no. 4, pp. 362–371, 2019, doi: 10.1109/THMS.2019.2917194.

[20] Y. Du, A. Black, L.-P. Morency, and M. Eskenazi, "*Multimodal polynomial fusion for detecting driver distraction*," in Interspeech 2018, Hyderabad,India, Sep. 2018, pp. 611-615. doi: 10.21437/Interspeech.2018-2011.

[21] G. Lechner *et al*., "A lightweight framework for multi-device integration and multi-sensor fusion to explore driver distraction," in *Adv. Inf. Syst. Eng.*, 2019, pp. 80–95, Springer International Publishing.

[22] Y. Zhang, Y. Chen, and C. Gao, "Deep unsupervised multi-modal fusion network for detecting driver distraction," *Neurocomputing*, vol. 421, pp. 26–38, 2021, doi: https://doi.org/10.1016/j.neucom.2020.09.023.

[23] Y. Liang, J. D. Lee, and M. L. Reyes, "Nonintrusive detection of driver cognitive distraction in real time using bayesian networks," *Transp. Res. Rec.*, vol. 2018, no. 1, pp. 1–8, Jan. 2007, doi: 10.3141/2018-01.

[24] W. Xie *et al*., "Multimodal fusion diagnosis of depression and anxiety based on CNN-LSTM model," *Comput. Med. Imaging Graph*, vol. 102, no. 102128, 2022, doi: https://doi.org/10.1016/j.compmedimag.2022.102128

[25] Y. Zhang *et al*., "Sleep stage classification using bidirectional LSTM in wearable multi-sensor systems," IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Paris, France, 2019, pp. 443-448, doi: 10.1109/INFCOMW.2019.8845115

[26] G. Liu, and J. Guo, Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, vol. 337, pp. 325–338, 2019, https://doi.org/10.1016/j.neucom.2019.01.078.

[27] W. Zhang and H. Zhang, "Research on distracted driving identification of truck drivers based on simulated driving experiment," *IOP Conf. Ser. Earth Environ. Sci.*, Feb. 2021, pp. 012039. doi: 10.1088/1755-1315/638/1/012039.

[28] G. Kountouriotis, P. Spyridakos, O. Carsten, and N. Merat, "Identifying cognitive distraction using steering wheel reversal rates," *Accid. Anal. Prev.*, vol. 96, Jul. 2016, doi: 10.1016/j.aap.2016.07.032.

[29] K. Kircher and C. Ahlström, "Evaluation of methods for the assessment of minimum required attention," *Accid. Anal. Prev.*, vol. 114, pp. 40-47, May 2018, doi: 10.1016/j.aap.2017.03.013.

[30] V. Alizadeh and O. Dehzangi, "The impact of secondary tasks on drivers during naturalistic driving: Analysis of EEG dynamics," in *2016 IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Rio de Janeiro, Brazil, Dec. 2016, pp. 2493–2499. doi: 10.1109/ITSC.2016.7795957.

[31] S. V Deshmukh and O. Dehzangi, "ECG-based driver distraction identification using wavelet packet transform and discriminative kernel-based features," in *2017 IEEE Int. Conf.Smart Comput. (SMARTCOMP)*, Hong Kong, China, May 2017, pp. 1–7. doi: 10.1109/SMARTCOMP.2017.7947003.

[32] A. Sahayadhas, K. Sundaraj, M. Murugappan, and R. Palaniappan, "A physiological measures-based method for detecting inattention in drivers using machine learning approach," *Biocybern. Biomed. Eng.*, vol. 35, no. 3, pp. 198–205, 2015, doi: https://doi.org/10.1016/j.bbe.2014.12.002.

[33] E. A. Mattar, H. J. Al-Junaid, and K. N. Al-Mutib, "Electroencephalography features extraction and deep patterns analysis for robotics learning and control through brain-computer interface," in *2019 Int. Conf.Innov. Intell. Inform., Comput., and Technol. (3ICT)*, Sakhier, Bahrain, Nov. 2019, pp. 1–6. doi: 10.1109/3ICT.2019.8910277.

[34] Z. Gao *et al.*, "An adaptive optimal-Kernel time-frequency representation-based complex network method for characterizing fatigued behavior using the SSVEP-based BCI system," *Knowledge-Based Syst.*, vol. 152, pp. 163–171, 2018, doi: https://doi.org/10.1016/j.knosys.2018.04.013.

[35] W. Lv *et al*., "Applying Entropy to Understand Drivers' Uncertainty during Car-following.," in *Proc. the joint meeting 12th Int. Conf. Meas. Behav.6th Semin. Behav. Meth.*, Krakow, Poland, Oct. 2021.

[36] M. Ravi Kumar and Y. Srinivasa Rao, "Epileptic seizures classification in EEG signal based on semantic features and variational mode decomposition," *Cluster Comput.*, vol. 22, no. 6, pp. 13521–13531, 2019, doi: 10.1007/s10586-018-1995-4.

[37] M. Fraiwan, M. Alafeef, and F. Almomani, "Gauging human visual interest using multiscale entropy analysis of EEG signals," *J. Ambient. Intell. Humaniz. Comput.*, vol. 12, no. 2, pp. 2435–2447, 2021, doi: 10.1007/s12652-020-02381-5.

[38] P. Raffalt, J. Mccamley, W. Denton, and J. Yentes, "Sampling frequency influences sample entropy of kinematics during walking," *Med. Biol. Eng. Comput.*, vol. 57, Nov. 2018, doi: 10.1007/s11517-018-1920-2.

[39] F. Fallahtafti, S. R. Wurdeman, and J. M. Yentes, "Sampling rate influences the regularity analysis of temporal domain measures of walking more than spatial domain measures," *Gait Posture*, vol. 88, pp. 216–220, 2021, doi: https://doi.org/10.1016/j.gaitpost.2021.05.031.

[40] J. Zheng *et al*., "Effects of sampling rate on multiscale entropy of electroencephalogram time series," *Biocybern. Biomed. Eng.*, vol. 43, no. 1, pp. 233–245, 2023, doi: https://doi.org/10.1016/j.bbe.2022.12.007.

[41] K. M. Hossain, Md. A. Islam, S. Hossain, A. Nijholt, and M. A. R. Ahad, "Status of deep learning for EEG-based brain–computer interface applications," *Front. Comput. Neurosci.*, vol. 16, 2023, [Online]. Available: https://www.frontiersin.org/articles/10.3389/fncom.2022.1006763

[42] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: a review," *Data Min. Knowl. Discov.*, vol. 33, no. 4, pp. 917–963, 2019, doi: 10.1007/s10618-019-00619-1.

[43] H. Ranganathan, S. Chakraborty, and S. Panchanathan, "Multimodal emotion recognition using deep learning architectures," in *2016 IEEE Winter Conf. Appl. Comput. Vis.(WACV)*, Lake Placid, NY, USA, Mar. 2016, pp. 1–9. doi: 10.1109/WACV.2016.7477679.

[44] A. Fasanmade *et al*., "A fuzzy-logic approach to dynamic bayesian severity level classification of driver distraction using image recognition," *IEEE Access*, vol. 8, pp. 95197-95207, 2020, doi: 10.1109/ACCESS.2020.2994811.

[45] S. M. Kouchak and A. Gaffar, "Detecting driver behavior using stacked long short term memory network with attention layer," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3420–3429, 2021, doi: 10.1109/TITS.2020.2986697.

[46] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv preprint arXiv:1412.3555, Dec. 2014.

[47] N. Phutela, D. Relan, G. Gabrani, P. Kumaraguru, and M. Samuel, "Stress classification using brain signals based on LSTM network," *Comput.*

Intell. Neurosci., vol. 2022, p. 7607592, 2022, doi: 10.1155/2022/7607592.
[48]X. Hong *et al.*, “Predicting Alzheimer’s disease using LSTM,” *IEEE Access*, vol. 7, pp. 80893–80901, 2019, doi: 10.1109/ACCESS.2019.2919385.
[49] A. Fasanmade *et al.*, "Context-aware quantitative risk assessment machine learning model for drivers distraction." arXiv preprint arXiv:2402.13421 (2024).
[50] Zhu, J.; Sun, K.; Jia, S.; Lin, W.; Hou, X.; Liu, B.; Qiu, G. Bidirectional Long Short-Term Memory Network for Vehicle Behavior Recognition. Remote Sens. 2018, 10, 887. https://doi.org/10.3390/rs10060887
[51]J. Murphy, C. Devue, P. M. Corballis, and G. M. Grimshaw, “Proactive control of emotional distraction: evidence ffrom EEG alpha suppression,” *Front. Hum. Neurosci.*, vol. 14, 2020, [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2020.00318
[52]M. Rhif, A. Ben Abbes, I. R. Farah, B. Martínez, and Y. Sang, “Wavelet transform application for/in non-stationary time-series analysis: A review,” *Applied Sciences*, vol. 9, no. 7, p. 1345, 2019.
[53]M. Costa, C.-K. Peng, A. L. Goldberger, and J. M. Hausdorff, “Multiscale entropy analysis of human gait dynamics,” *Physica A: Stat.Mech. Appl.*, vol. 330, no. 1, pp. 53–60, 2003, doi: https://doi.org/10.1016/j.physa.2003.08.022.
[54]Y.-H. Wang, I.-Y. Chen, H. Chiueh, and S.-F. Liang, “A low-cost implementation of sample entropy in wearable embedded systems: an example of online analysis for sleep EEG,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021, doi: 10.1109/TIM.2020.3047488.
[55]M. Borowiec, A. Rysak, D. N. Betts, C. R. Bowen, H. A. Kim, and G. Litak, “Complex response of a bistable laminated plate: Multiscale entropy analysis,” *Eur. Phys. J. Plus*, vol. 129, no. 10, p. 211, 2014, doi: 10.1140/epjp/i2014-14211-3.
[56]M. Costa, A. L. Goldberger, and C.-K. Peng, “Multiscale entropy analysis of biological signals,” *Phys. Rev. E.*, vol. 71, no. 2, p. 021906, 2005.
[57]S. Kar, M. Bhagat, and A. Routray, “EEG signal analysis for the assessment and quantification of driver’s fatigue,” *Transp. Res. Pt. F-Traffic Psychol. Behav.*, vol. 13, no. 5, pp. 297–306, 2010, doi: https://doi.org/10.1016/j.trf.2010.06.006.
[58]Y. Bai, Y. Guan, and W.-F. Ng, “*Fatigue assessment using ECG and actigraphy sensors*,” in *Proc.2020 ACM Int. Symp. Wearable Comput.*, Virtual Event, Mexico, Sep. 2020, pp. 12-16. doi: https://doi.org/10.1145/3410531.3414308.
[59]S. Huang, J. Li, P. Zhang, and W. Zhang, “Detection of mental fatigue state with wearable ECG devices,” *Int. J. Med. Inform.*, vol. 119, pp. 39–46, 2018, doi: https://doi.org/10.1016/j.ijmedinf.2018.08.010.
[60]C. Zhao, M. Zhao, J. Liu, and C. Zheng, “Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator,” *Accid. Anal. Prev.*, vol. 45, pp. 83–90, 2012, doi: https://doi.org/10.1016/j.aap.2011.11.019.
[61]J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm,” *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, 1985, doi: 10.1109/TBME.1985.325532.
[62]A. Noviyanto, S. M. Isa, I. Wasito, and A. Arymurthy, “Selecting features of single lead ECG signal for automatic sleep stages classification using correlation-based feature subset selection,” *Int. J. Comput.Sci. Iss. (IJCSI)*, vol. 8, Sep. 2011.
[63]C. Sapsanis, G. Georgoulas, and A. Tzes, “EMG based classification of basic hand movements based on time-frequency features,” in *21st Mediterr. Conf. Control Autom.*, Platanias, Greece, June 2013, pp. 716–722. doi: 10.1109/MED.2013.6608802.
[64]S. Pourmohammadi and A. Maleki, “Stress detection using ECG and EMG signals: A comprehensive study,” *Comput. Meth. Programs Biomed.*, vol. 193, p. 105482, 2020, doi: https://doi.org/10.1016/j.cmpb.2020.105482.
[65]K. Mahaphonchaikul, D. Sueaseenak, C. Pintavirooj, M. Sangworasil, and S. Tungjitkusolmun, “EMG signal feature extraction based on wavelet transform,” in *ECTI-CON2010: The 2010 ECTI Int. Conf. Electr. Eng./Electronics, Comput., Telecommun. Inf. Technol.*, Chiang Mai, Thailand, May 2010, pp. 327–331.
[66]Y. Zhang, X. Ren, and J. Zhang, “Intrusion detection method based on information gain and ReliefF feature selection,” in *2019 Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary, July 2019, pp. 1–5. doi: 10.1109/IJCNN.2019.8851756.
[67]C. Peng, Z. Zhang, Z. Kang, C. Chen, and Q. Cheng, “Nonnegative matrix factorization with local similarity learning,” *Inf. Sci. (N Y)*, vol. 562, pp. 325–346, 2021, doi: https://doi.org/10.1016/j.ins.2021.01.087.
[68]G. Doquire and M. Verleysen, “Mutual information-based feature selection for multilabel classification,” *Neurocomputing*, vol. 122, pp. 148–155, 2013, doi: https://doi.org/10.1016/j.neucom.2013.06.035.
[69]N. S. Malan and S. Sharma, “Motor imagery EEG spectral-spatial feature optimization using dual-tree complex wavelet and neighbourhood component analysis,” *IRBM*, vol. 43, no. 3, pp. 198–209, 2022, doi: https://doi.org/10.1016/j.irbm.2021.01.002.
[70]A. Marcano-Cedeño, J. Quintanilla-Domínguez, M. G. Cortina-Januchs, and D. Andina, “Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network,” in *IECON 2010 - 36th Annu. Conf. IEEE Ind. Electron. Soc.*, Glendale, AZ, USA, Nov. 2010, pp. 2845–2850. doi: 10.1109/IECON.2010.5675075.
[71]P. A. Karthick, D. M. Ghosh, and S. Ramakrishnan, “Surface electromyography based muscle fatigue detection using high-resolution time-frequency methods and machine learning algorithms,” *Comput. Meth.s Programs Biomed.*, vol. 154, pp. 45–56, 2018, doi: https://doi.org/10.1016/j.cmpb.2017.10.024.
[72]H. V. Koay, J. H. Chuah, C.-O. Chow, and Y.-L. Chang, “Detecting and recognizing driver distraction through various data modality using machine learning: A review, recent advances, simplified framework and open challenges (2014–2021),” *Eng. Appl. Artif. Intell.*, vol. 115, p. 105309, 2022, doi: https://doi.org/10.1016/j.engappai.2022.105309.
[73]Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, “Driver inattention monitoring system for intelligent vehicles: a review,” *IEEE Trans. Intel. Transp. Syst.*, vol. 12, no. 2, pp. 596–614, 2011, doi: 10.1109/TITS.2010.2092770.
[74]S. V Deshmukh and O. Dehzangi, “Characterization and identification of driver distraction during naturalistic driving: an analysis of ECG dynamics,” in *Adv. Body Area Netw*. 2019, pp. 1–13.
[75]L. Yu, X. Sun, and K. Zhang, “Driving distraction analysis by ECG signals: an entropy analysis,” in *Int. Conf. Inter. Des. Glob. Dev.*, Berlin, Heidelberg, July 2011, pp. 258–264.
[76]S. V. Deshmukh, “Study of online driver distraction analysis using ECG-dynamics,” Ph.D. dissertation, Dept. Comput. Inf. Sci., Univ. Michigan-Dearborn, Dearborn, USA, 2018.
[77]M. B. Kurt, N. Sezgin, M. Akin, G. Kirbas, and M. Bayram, “The ANN-based computing of drowsy level,” *Expert Syst. Appl.*, vol. 36, no. 2, Part 1, pp. 2534–2542, 2009, doi: https://doi.org/10.1016/j.eswa.2008.01.085.
[78]B. Reimer, B. Mehler, Y. Wang, and J. Coughlin, “A field study on the impact of variations in short-term memory demands on drivers’ visual attention and driving performance across three age groups,” *Hum. Factors*, vol. 54, pp. 454–468, May 2012, doi: 10.1177/0018720812437274.
[79]P. Choudhary and N. R. Velaga, “Effects of phone use on driving performance: A comparative analysis of young and professional drivers,” *Saf. Sci.*, vol. 111, pp. 179–187, 2019, doi: https://doi.org/10.1016/j.ssci.2018.07.009.
[80]J. Zhang, Z. Yin, P. Chen, and S. Nichele, “Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review,” *Inf. Fusion*, vol. 59, pp. 103–126, 2020, doi: https://doi.org/10.1016/j.inffus.2020.01.011.
[81]A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang, “Benchmark for filter methods for feature selection in high-dimensional classification data,” *Comput. Stat. Data Anal.*, vol. 143, p. 106839, 2020, doi: https://doi.org/10.1016/j.csda.2019.106839.
[82]M. Ghosh, S. Adhikary, K. K. Ghosh, A. Sardar, S. Begum, and R. Sarkar, “Genetic algorithm based cancerous gene identification from microarray data using ensemble of filter methods,” *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 159–176, 2019, doi: 10.1007/s11517-018-1874-4.
[83]A. MDarzi, S. M. Gaweesh, M. M. Ahmed, and D. Novak, “Identifying the causes of drivers' hazardous states using driver characteristics, vehicle kinematics, and physiological measurements,” *Front. Neurosci.*, vol. 12, no. 568, 2018, https://doi.org/10.3389/fnins.2018.00568.
[84]E.T.M. Beltrán, M.Q. Pérez, S.L. Bernal, G.M. Pérez, and A.H. Celdrán, “SAFECAR: A brain–computer interface and intelligent framework to detect drivers’ distractions,” *Expert Syst. Appl.*, vol. 203, pp. 117402, 2022, https://doi.org/10.1016/j.eswa.2022.117402.
[85]S. M. Kouchak and A. Gaffar, "Using bidirectional long short term memory with attention layer to estimate driver behavior," in 18th IEEE ICMLA, Boca Raton, FL, USA, 2019, pp. 315-320.
[86]P. Angkan et al., "Multimodal brain–computer interface for in-vehicle driver cognitive load measurement: dataset and baselines," *IEEE Trans. Intell. Transp. Syst.*, pp. 1-16, 2024, doi: 10.1109/TITS.2023.3345846.
[87]X. Zuo, C. Zhang, F. Cong, J. Zhao and T. Hämäläinen, "Driver distraction detection using bidirectional long short-term network based on

multiscale entropy of EEG," *IEEE Trans. Intell. Transp. Syst*., vol. 23, no. 10, pp. 19309-19322, Oct. 2022, doi: 10.1109/TITS.2022.3159602.

X**in Zuo** received the B.S. in Naval Architecture and Ocean Engineering from Harbin Engineering University, China, in 2016, and the M.S. in Design and Manufacture of Ship and Ocean structure from Dalian University of Technology, in 2019. She is currently pursuing the Ph.D. degree in Software and Communications Engineering at University of Jyväskylä, Jyväskylä, Finland.

Her research interests include driver distraction, biomedical signal processing.

Chi Zhang received the B.S., M.S., and Ph.D. degree of Science in Engineering from Northeastern University, China, in 2010, 2012 and 2016, respectively.

He is an Associate Professor with the School of Biomedical Engineering, Faculty of Medicine, Dalian University of Technology, Dalian, China. His research interests include biomedical signal processing, brain-computer interface, and artificial intelligence.

Fengyu Cong received the B.S. degree in Power and Thermal Dynamic Engineering and the Ph.D. degree in Mechanical Design and Theory from Shanghai Jiao Tong University, China, in 2002 and 2007, and the Ph.D. degree in Mathematical Information Technology from the University of Jyväskylä, Jyväskylä, Finland, in 2010.

He is a Professor in the School of Biomedical Engineering, Faculty of Medicine, Dalian University of Technology, Dalian, China. His current research interests include brain signal processing, acoustic signal processing, independent component analysis, tensor decomposition, and pattern recognition/machine learning/data mining.

Jian Zhao received the M.S. and Ph.D. degree in Mechanical Engineering from Xidian University, China, in 2006 and 2008, respectively.

He is currently a Professor at the School of Automotive Engineering, Faculty of Vehicle Engineering and Mechanics, Dalian University of Technology, Dalian, China. His current research interests include autonomous driving, MEMS sensors, compliant mechanics, and nonlinear dynamics.

Timo Hämäläinen received the Ph.D. degree in telecommunication from the University of Jyväskylä, Jyväskylä, Finland, in 2002.

In 1997, he joined the University of Jyväskylä, where he is currently a Professor of computer networks. He has more than 25 years research and teaching experience of computer networks. He has led many external funded network management related projects. He has launched and leads master programs with the University of Jyväskylä (SW & Comm. Eng.) and teaches network management related courses. He has more than 200 internationally peer reviewed publications and he has supervised almost 40 Ph.D. dissertation. His research interests include network resource management, IoT, and networking security.