Heart Rate Variability-Based Driver Drowsiness Detection and Its Validation With EEG

Koichi Fujiwara , *Member, IEEE*, Erika Abe , *Member, IEEE*, Keisuke Kamata, Chikao Nakayama, Yoko Suzuki, Toshitaka Yamakawa , *Member, IEEE*, Toshihiro Hiraoka , *Member, IEEE*, Manabu Kano , *Member, IEEE*, Yukiyoshi Sumi, Fumi Masuda, Masahiro Matsuo, and Hiroshi Kadotani

Abstract—*Objective:* Driver drowsiness detection is a key technology that can prevent fatal car accidents caused by drowsy driving. The present work proposes a driver drowsiness detection algorithm based on heart rate variability (HRV) analysis and validates the proposed method by comparing with electroencephalography (EEG)-based sleep scoring. *Methods:* Changes in sleep condition affect the autonomic nervous system and then HRV, which is defined as an RR interval (RRI) fluctuation on an electrocardiogram trace. Eight HRV features are monitored for detecting changes in HRV by using multivariate statistical process control, which is a well known anomaly detection method. *Result:* The performance of the proposed algorithm was evaluated through an experiment using a driving simulator. In this experiment, RRI data were measured from 34 participants during driving, and their sleep onsets were determined based on the EEG data by a sleep specialist. The validation result of the experimental data with the EEG data showed that drowsiness was detected in 12 out of 13 pre-N1 episodes prior to the sleep onsets, and the false positive rate was 1.7 times per hour. *Conclusion:* The present work also demonstrates the usefulness of the framework of HRV-based anomaly detection that was originally proposed for epileptic seizure prediction. Significance: The proposed method can contribute to preventing accidents caused by drowsy driving.

Index Terms—Drowsy driving detection, heart rate variability analysis, electroencephalography, anomaly detection, multivariate statistical process control.

I. INTRODUCTION

THE risk of traffic accidents in drowsy drivers is estimated to be four to six times higher than in awake drivers [1].

According to a study by Gottlieb *et al.*, the risk of traffic accident occurrence increases regardless of the drivers' subjective sleepiness when they have sleep apnea or their sleep duration is insufficient [2]. In order to prevent accidents caused by drowsy driving, a driver-assistance system that detects drowsy driving and provides a warning would be effective.

In sleep medicine, electroencephalography (EEG) recording is necessary for sleep scoring because sleep onsets and sleep stages are defined based on EEG [3]. Although EEG-based drowsiness detection methods have been developed [4]–[7], it is difficult to record EEG accurately during driving since EEG recording is intolerant to motion artifacts and puts significant restrictions on the body. Thus, various types of driver drowsiness detection systems that do not use EEG have been developed [8].

Driver face image analysis and vehicle travel data analysis are used for detecting driver drowsiness [9]–[14]; however, these methods require installing special devices in a vehicle, such as a camera for face image acquisition or a data logging device for accessing vehicle travel data.

Instead of installing devices in a vehicle, physiological information other than EEG can be used for drowsiness detection if drivers agree to wear a sensor that measures their physiological signals. Changes in sleep condition affect the autonomic nervous system (ANS) as well as cardiac activities [15], and cardiac signals can be used for drowsiness detection. Chui *et al.* proposed a drowsiness detection method based on an electrocardiogram (ECG) taken from drivers [16]. In addition, some researchers have analyzed photoplethysmography (PPG) signals for detecting drowsy driving [17]. Although they reported that their proposed methods were able to achieve good performance, it is difficult to obtain good ECG or PPG signals stably due to motion artifacts. Besides, these methods would require a heavy computational load because the sampling rate of ECG is usually more than several hundred Hz.

Heart rate variability (HRV), which is the RR interval (RRI) fluctuation in an ECG, is a well-known physiological phenomenon which reflects activities of ANS [18]. Fujiwara *et al.* proposed an epileptic seizure prediction algorithm utilizing HRV analysis [19]. HRV changes before an epileptic seizure because changes in cardiovascular regulation begin ten minutes to several seconds before seizure onsets [20], [21]. In addition, several studies have reported changes in HRV associated with sleep stage transitions [22]–[24].

In the present work, a new HRV-based driver drowsiness detection algorithm is proposed by utilizing the framework of HRV-based epileptic seizure prediction [19]. Abnormalities in the HRV data of drivers are monitored by multivariate statistical process control (MSPC), which is a well-known anomaly detection algorithm used in manufacturing industries [25]–[27]. The proposed algorithm is simpler than previous HRV-based methods [28]–[30] because the number of HRV features it uses is just eight and MSPC is a linear method. Driving simulator experiments were performed to validate the proposed method, in which EEG-based sleep scoring by a sleep specialist was used as a reference.

II. RELATED WORKS

HRV-based drowsiness detection methods have been proposed. Vicente *et al.* proposed a drowsiness detection method that uses HRV analysis and linear discriminant analysis (LDA) [28]; however, the method uses ECG-derived respiratory information in addition to HRV, and ECG signal analysis is still needed. Although Li *et al.* proposed a drowsiness detection system that utilizes the support vector machine (SVM), their system utilizes driver face images as well as HRV [29]. A neural network (NN)-based drowsiness detection model was developed by Patel *et al.* which uses the power spectral density (PSD) of RRI fluctuation as input variables of the NN model [30]. Their method would require a large amount of computational resources because the NN model is complicated and its number of input features is 900. A simple methodology for detecting drowsy driving should be developed for realizing a wearable drowsy driving detection system.

III. METHOD

Although EEG measurement is necessary for detecting sleep onsets [3] in sleep medicine, it is difficult to measure EEG during driving. The proposed algorithm adopts HRV instead of EEG, and EEG-based sleep scoring is used as the reference for the proposed algorithm. This section explains EEG-based sleep scoring and HRV briefly and proposes an HRV-based drowsiness detection algorithm.

A. EEG-Based Sleep Scoring

Sleep consists of REM (rapid eye movement sleep) and NREM (non-REM sleep), which is categorized into three levels: N1, N2, and N3 [3]. N1 is also called transitional sleep or light sleep. According to the sleep scoring manual [31], sleep stages are discriminated based on the 30-second epoch-based EEG scoring method. The N1 onset (sleep onset) is defined by the epoch in which α wave (8–13 Hz) activity is attenuated and replaced by low-amplitude, mixed-frequency activities that occupy more than 50% of the epoch.

Drivers may feel drowsiness shortly before N1, which causes mild cognitive dysfunction, and some researchers have attempted drowsy EEG identification [32]. On the other hand, falling asleep directly contributes to traffic accidents. N1 usually occurs between wakefulness and deeper sleep stages.

During N1, the muscles are still active, the eyes open and close moderately, and persons can be easily awakened by a sensory stimulus. Thus, driver drowsiness should be detected prior to the N1 onset (sleep onset), when a driver can be easily wakened by a stimulus.

It is noteworthy that we cannot define a sleep onset with the accuracy of less than 30 seconds because sleep scoring is based on the 30-second EEG epoch-based method.

B. Heart Rate Variability Analysis

The R wave is the highest peak on an ECG, and the RR interval (RRI) [ms] is defined as the interval between an R wave and the next R wave. HRV is the fluctuation of RRI, which is a physiological phenomenon reflecting ANS activities. Thus, HRV analysis has been used for monitoring stress, and cardiovascular disease [33], [34].

Although there are two types of HRV features–linear features and nonlinear features–this work uses the former, simply because the extraction of nonlinear features requires a long-term RRI measurement for stable calculation [35], which is not appropriate for real-time applications like drowsy driving detection. The linear HRV features are classified into time domain features and frequency domain features [18].

The following time domain features can be calculated from the original RRI data [18].
- **MeanNN**: Mean of RRI.
- **SDNN**: Standard deviation of RRI.
- **RMSSD**: Root means square of the difference of adjacent RRI.
- **Total Power (TP)**: Variance of RRI.
- **NN50**: The number of pairs of adjacent RRI whose difference is more than 50 ms within a given length of measurement time.

Frequency domain features cannot be extracted since the raw RRI data are not sampled at equal intervals. Thus, the raw RRI data are interpolated by using spline and resampled at equal intervals. The following frequency domain features can be obtained from the power spectrum density (PSD) of the resampled RRI data, and the PSD can be calculated by using Fourier analysis or an autoregressive (AR) model [18].
- **LF**: Power of the low-frequency band (0.04 Hz–0.15 Hz) in a PSD. LF reflects both the sympathetic and parasympathetic nervous system activities.
- **HF**: Power of the high-frequency band (0.15 Hz–0.4 Hz) in a PSD. HF reflects the parasympathetic nervous system activity.
- **LF/HF**: Ratio of LF to HF. LF/HF expresses the balance between the sympathetic nervous system activity with the parasympathetic nervous system activity.

The guideline recommends that RRI is measured for two to five minutes for HRV analysis, and the sampling rate of ECG should be more than 200 Hz for precise R wave detection [18].

A precise RRI sensor is needed in order to realize an HRV-based drowsy driving detection system. Although the Holter monitor is generally used for measuring ECG outside hospitals, its use in daily life is difficult since the Holter monitor

requires operation skills. Many types of wearable devices such as smartwatches have PPG sensors that can also be used for pulse detection; however, it is notably difficult for PPG to derive RRI precisely enough to carry out HRV analysis [36]. A wearable RRI sensor developed by Yamakawa *et al.* easily measures accurate RRI based on ECG. If an HRV-based drowsy driving detection algorithm can be implemented in such a device, a wearable drowsy driving detection system can be realized.

Changes in HRV associated with sleep stage transitions have been reported [22]–[24]. Bonnet and Arand reported that heart rates vary depending on sleep latency [38]. Chua *et al.* showed that there is a correlation between changes in subjective sleepiness and HRV through sleep deprivation experiments [39]. Since HRV alteration begins prior to a sleep onset [23], drowsy driving may be detected by monitoring HRV.

C. Drowsy Driving Detection

In drowsy driving detection, the awake data and the drowsy data are regarded as normal data and anomalous data, respectively. To build an accurate discriminant model by using both the awake data and the drowsy data, a sufficient amount of drowsy data needs to be collected from drivers. However, in practice, collecting such drowsy data is more difficult than the awake data. Thus, drowsy driving detection is formulated as an anomaly detection problem, in which a model is developed from the awake data only.

Fujiwara *et al.* developed an epileptic seizure prediction algorithm based on multivariate statistical process control (MSPC) to detect abnormalities in HRV [19]. MSPC detects a sample that does not follow the major trend in the modeling data as an anomaly based on principal component analysis (PCA), which has been widely used as fault detection and identification technique in multivariate processes [25]–[27]. Since drowsy driving detection is a similar problem to epileptic seizure prediction, we use MSPC.

The proposed algorithm discriminates between driver statuses of 'awake' and 'drowsy,' where 'drowsy' means that the driver is close to or in N1. In the proposed method, eight HRV features described in Section III-B are used and their abnormalities by sleepiness are monitored using MSPC. The detail of MSPC is explained in the Appendix.

For HRV feature extraction, a rectangular moving window whose window size is three minutes is used. Li *et al.* compared the one-minute and the three-minutes windows in HRV extraction, and they reported that the latter was better for drowsiness detection [29]. Time domain features are extracted from the raw RRI data. For frequency domain feature extraction, the raw RRI data need to be arranged at equal intervals. The raw RRI data are interpolated by using the third-order spline, and the interpolated RRI data are resampled at 4 Hz. An AR model of order 40 is used to calculate the PSD [19].

The proposed drowsy driving detection algorithm is described in Algorithm 1, in which $\boldsymbol{y}^{\{i\}}$ is the awake RRI data recorded from the ith driver and I is the number of drivers. First, awake HRV features are extracted from $\boldsymbol{y}^{\{i\}}$ in Steps 1–3. In the

---

Algorithm 1: Drowsy Detection Preparation.

1: **for all** i such that $1 \leq i \leq I$ **do**
2:     Extract the ith driver awake HRV feature $\tilde{\boldsymbol{X}}^{\{i\}}$ from the ith driver awake RRI data $\boldsymbol{y}^{\{i\}}$.
3: **end for**
4: Merge matrixes $\tilde{\boldsymbol{X}}^{\{1\}}, \ldots, \tilde{\boldsymbol{X}}^{\{I\}}$ into one matrix $\tilde{\boldsymbol{X}}$.
5: Preprocess $\tilde{\boldsymbol{X}}$, which is referred to as $\boldsymbol{X}$.
6: Derive $\boldsymbol{\Sigma}_R$ and $\boldsymbol{V}_R$ from $\boldsymbol{X}$ as Eq. (1)
7: **for all** i such that $1 \leq i \leq I$ **do**
8:     Define the control limits of the T^2 and Q statistics for the ith driver, $\bar{T}^{2\{i\}}$ and $\bar{Q}^{\{i\}}$.
9: **end for**

---

proposed method, eight HRV features are adopted as input variables. The extracted HRV features are merged into one matrix in Step 4. Then, in Step 5, the merged matrix $\tilde{\boldsymbol{X}}$ is preprocessed for model construction. There are various preprocessing methods, and thus an appropriate method should be chosen by taking account of the characteristics of the problem and the data. In this work, each column of $\tilde{\boldsymbol{X}}$ is standardized so that each HRV feature has zero mean and unit variance. In Step 6, the singular value matrix $\boldsymbol{\Sigma}_R$ and the loading matrix $\boldsymbol{V}_R$ are derived from the preprocessed HRV feature matrix $\tilde{\boldsymbol{X}}$. In other words, PCA is applied to $\tilde{\boldsymbol{X}}$, and the correlations among eight HRV features are modeled. In this step, the number of principal components R has to be selected appropriately to realize precise drowsiness detection. The next step is to calculate the T^2 and Q statistics and to define their control limits.

There is considerable individual variability in HRV. Changes in HRV are different for every person, which changes with age, and the variation of the T^2 and Q statistics is also different for every person. Hence, the control limits have to be determined for each driver in Steps 7–9. The control limits can be determined as the $\alpha\%$ confidence of each driver.

Before driver drowsiness monitoring starts, the initial RRI data of a driver have to be stored for more than the window size W to calculate HRV features. After the initial RRI data collection, driver drowsiness can be monitored by following Algorithm 2. $y[t] \in \Re$ denotes the tth RRI and t is the number of sampling from the monitoring start. τ is a time counter variable, and C denotes the binary driver status $C = \{\mathcal{A}, \mathcal{D}\}$ where $\mathcal{A}$ and $\mathcal{D}$ are 'awake' and 'drowsy,' respectively. That is, $\neg \mathcal{A} = \mathcal{D}$ and vice versa. In Step 5, the extracted HRV feature $\tilde{x}$ is preprocessed in the same manner in Algorithm 1.

To realize accurate drowsy driving detection, it is crucial to decrease false positives. In drowsy driving detection, false positives are mainly caused by ECG artifacts, which significantly affect the T^2 and Q statistics. Hence, the driver status is determined as 'drowsy' only when either the T^2 or Q statistic continuously exceeds its control limit for more than the predefined period $\bar{\tau}$. Conversely, to change the status from 'drowsy' to 'awake,' both statistics have to continuously stay below their control limits for more than $\bar{\tau}$. In Steps 7–14, the driver status is discriminated. A warning is given to the driver when the algorithm detects drowsiness, that is, $C = \mathcal{D}$.

Algorithm 2: Drowsy Driving Detection.

1: set $\tau[0] \longleftarrow 0$, $C[0] \longleftarrow \mathcal{A}$.
2: **while do**
3:    Collect the newly measured tth RRI $y[t]$.
4:    Extract HRV feature $\tilde{\boldsymbol{x}}[t]$.
5:    Preprocess $\tilde{\boldsymbol{x}}[t]$, which is denoted as $\boldsymbol{x}[t]$.
6:    Calculate the tth T^2 and Q statistics, $T^2[t]$ and $Q[t]$
       from $\boldsymbol{x}[t]$ by using (4) and (3).
7:    **if** $((T^2[t] > \bar{T}^2 \vee Q[t] > \bar{Q}) \wedge (C[t-1] = \mathcal{A}))$
          $\vee((T^2[t] \leq \bar{T}^2 \wedge Q[t] \leq \bar{Q}) \wedge (C[t-1] = \mathcal{D}))$
          then
8:       $\tau[t] = \tau[t-1] + y[t]$.
9:    **else**
10:      $\tau[t] = 0$.
11:   **end if**
12:   **if** $\tau[t] \geq \bar{\tau}$ **then**
13:      $C[t] = \neg\, C[t-1]$ and $\tau[t] = 0$.
14:   **end if**
15:   Wait until the next RRI data $y[t+1]$ is measured.
16: **end while**

D. Participants

The inclusion criteria for the participants were non-professional drivers, with a valid driving license. The exclusion criterion was having a chronic illness that may affect HRV such as cardiovascular disease, arrhythmia, epilepsy, or sleep disorders. These experiments and analyses were approved by the Research Ethics Committee of the Graduate School of Science and Technology, Kumamoto University. Written informed consent was obtained from each participant prior to the experiments.

IV. RESULT

This section evaluates the performance of the proposed drowsy driving detection algorithm through an application of real RRI data obtained from an experiment using a driving simulator.

A. Data Collection

The RRI data and the EEG data were collected from experiment participants (drivers) while they drove a virtual vehicle on a simulator. Participants A, ..., Z, α, ..., θ, 25 males and nine females aged 18–36 years (mean 22.7 years) participated in this experiment. In order to avoid effects on HRV, participants were instructed not to take alcohol, caffeine or smoke for one night before the experiments. Before the experiment, the participants' driving careers, health statuses, mediations, and sleep habits were checked by means of a questionnaire. The questionnaire collects age, sex, weight, height, occupation, medical history (hypertension, diabetes, cardiovascular disease, epilepsy, etc.), meditation, and habits about breakfast, exercise, sleep, caffeine, and smoke. In addition, we asked the sleep time of the previous night. As a result, all participants were healthy and took enough sleep (>6 hours), and there were no rejected participants.
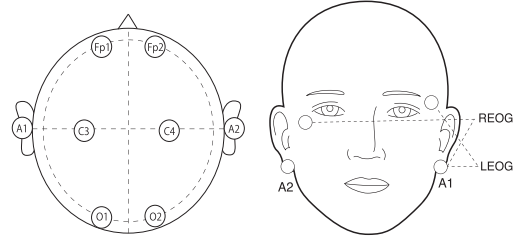

Fig. 1. Electrode allocations for sleep scoring: EEG (left) and EOG (left).

In this work, participants drove a driving simulator constructed based on a commercial racing simulator (GRAN TURISMO 5, Sony Interactive Entertainment Inc.), which was used in some researches of driving physiology [40]–[42]. The simulator equips with an LCD display, a steering, an accelerator, and a brake pedal. Thus, it can simulate driving operation. Before experiments, we confirmed that EEG during driving can be measured precisely enough for sleep scoring in this simulator.

The participants drove twice on a course that simulated a nighttime, monotonous highway loop line for 1.5 hours in a dark room, resting and taking lunch for an hour between the two trials. There were no other vehicles and it took about ten minutes to cycle the loop line at 80 kilometers per hour. The first and second trials started from around 11 am and after lunch, respectively. This setting was determined in consideration of avoiding participants' excessive fatigue due to extended experiments as well as with the expectation that the participants may become drowsy.

The EEG data during driving were recorded for sleep scoring using a digital EEG recording system (Grapevine, Ripple), whose sampling frequency was 1,000 Hz. Although the International 10–20 system is a standard scalp electrode allocation of EEG recording, the number of electrodes was reduced in this work with reference to polysomnography (PSG) tests performed in sleep laboratories. Fig. 1 (left) shows the adopted electrode allocation, in which Fp1, Fp2, C3, C4, O1, O2 were EEG electrodes and earlobes A1 and A2 were for reference. This electrode allocation is enough for sleep scoring. Electrooculogram (EOG) was also recorded during driving for making sleep scoring easy. The EOG electrode allocation is shown in Fig. 1 (right) and a left earlobe A1 was for reference. The RRI data were obtained for HRV analysis by using a wearable RRI sensor [37]. In addition, participant video during driving was recorded for confirming participant behaviors after the experiments.

Since artifacts were generated when participants moved during driving, data in which either the RRI data or the EEG data were contaminated with strong artifacts were eliminated before analysis. A sleep specialist certified by the Japanese Society of Sleep Research determined sleep onsets of participants by visual check of the EEG data based on the 30-second epoch-based scoring method recommended in the sleep scoring manual [31].

As a result of sleep scoring, 12 participants were scored in N1 during driving. The data 15 minutes before and 5 minutes

TABLE I
PARTICIPANT DEMOGRAPHICS AND COLLECTED EPISODES

| Participant	Sex	Age	Awake episodes	Pre-N1 episodes	Participant	Sex	Age	Awake episodes	Pre-N1 episodes
A	male	22	A1, A2	Ad1	R	male	24	R1, R2	
B	male	36	B1 – B3	Bd1	S	male	18	S1, S2	
C	male	23	C1, C2	Cd1	T	male	19	T1 – T3	Td1
D	female	34	D1 – D4		U	male	23	U1, U2	Ud1
E	male	22	E1 – E3		V	male	21	V1, V2	
F	male	19	F1 – F3		W	male	21	W1 – W3	
G	male	22	G1, G2		X	female	36	X1 – X6	
H	female	22	H1 – H3	Hd1	Y	female	20	Y1 – Y3	
I	female	21	I1, I2		Z	male	23	Z1, Z2	Zd1
J	female	22	J1 – J3		α	female	22	$\alpha1 - \alpha3$	
K	male	21	K1, K2	Kd1	β	female	21	$\beta1, \beta2$	βd1
L	male	21	L1, L2	Ld1	γ	male	23	$\gamma1 - \gamma3$	
M	male	22	M1, M2		δ	male	24	$\delta1 - \delta5$	
N	male	21	N1, N2	Nd1, Nd2	ϵ	male	19	$\epsilon1, \epsilon2$	
O	male	20	O1 – O4		ζ	female	22	$\zeta1 - \zeta3$	
P	male	22	P1, P2	Pd1	η	male	23	$\eta1, \eta2$	
Q	male	21	Q1 – Q4		θ	male	21	$\theta1, \theta2$	
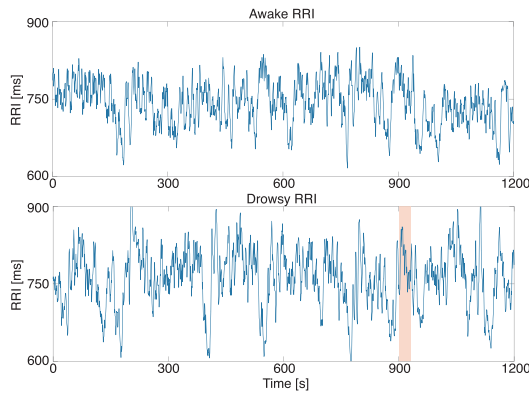
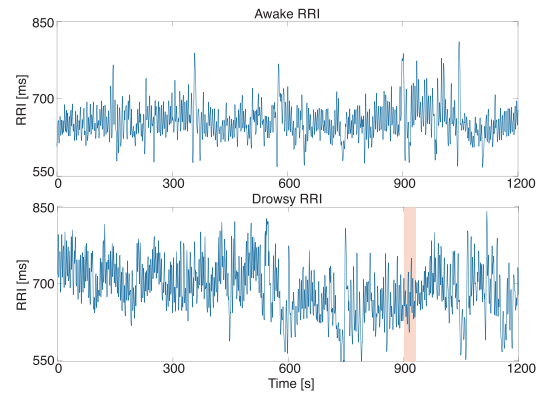Fig. 2. RRI data of L2 (top) and Ld1 (bottom).

Fig. 3. RRI data of N2 (top) and Nd2 (bottom).

after sleep onset were stored as pre-N1 episodes, following the report by Jurysta *et al.* on sleep stage transition that cardiac activities precede 9–20 minutes (mean 12 minutes) before EEG changes [23]. In addition, the data that were not scored as sleep were clipped as awake episodes.

Consequently, we collected 13 pre-N1 episodes named Ad1, Bd1, …, Zd1 and β d1, and 91 awake episodes named A1, A2, …, θ 1, and θ 2, which are shown in Table I. Their total lengths of awake and pre-N1 episodes were 66.8 and 4.3 hours, respectively.

B. RRI Data and HRV Features

The raw RRI data of participants L and M in awake and drowsy periods are shown in Figs. 2 and 3. In these figures, an orange colored band denotes the N1 epoch. Eight HRV features described in Section III-B were extracted. Figs. 4–7 are the HRV features extracted from the RRI data shown in Figs. 2 and 3. Although frequency domain features seemed to change before sleep onset in Figs. 5 and 7, a similar fluctuation is also observed in the awake HRV features in Figs. 4 and 7.

These results show that it is difficult to detect drowsiness by monitoring changes in respective HRV features, and suggest that multiple HRV features should be monitored simultaneously.
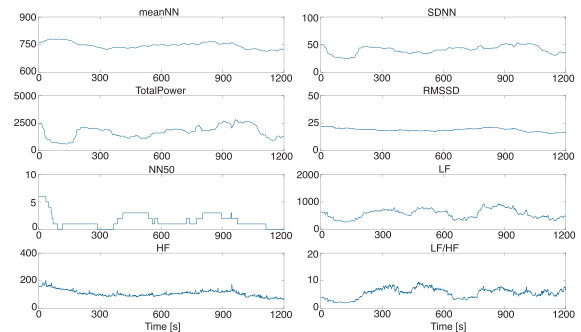

Fig. 4. HRV features derived from L2.

C. Drowsy Detection Preparation

Driver drowsiness detection was prepared according to Algorithm 1. Table II shows 34 awake episodes used for modeling. The total recorded length of the analyzed episodes was 26.7 hours.

All HRV features calculated in Section IV-B were used as inputs. In MSPC, the number of retained principal components R was determined so that the cumulative proportion reached more than 90%, and $R = 3$. The control limits of the T^2 and Q statistics were defined for each participant so that they represent
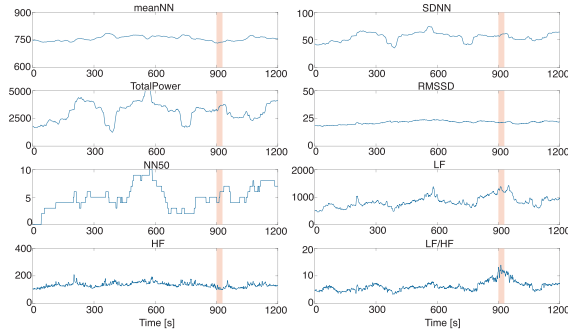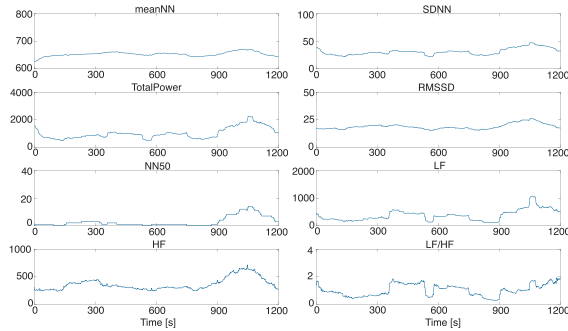
Fig. 5. HRV features derived from Ld1.
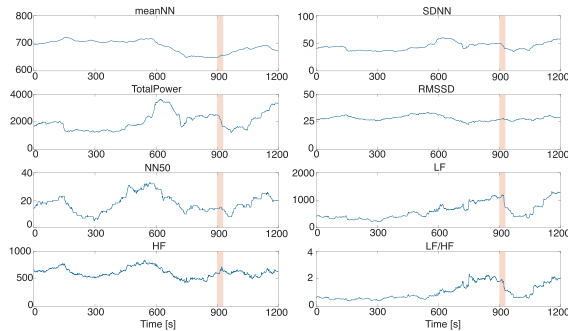

Fig. 6. HRV features derived from N2.

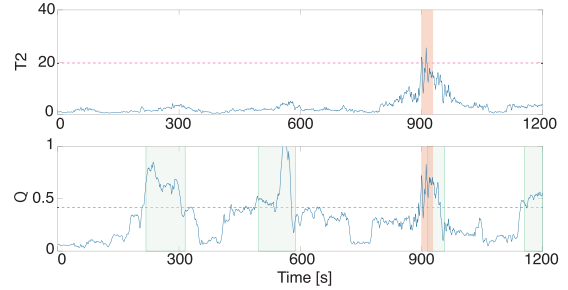

Fig. 7. HRV features derived from Nd2.

Fig. 8. Detection result of pre-N1 episode Ld1 (orange band: N1 epoch, green band: detected drowsy period).
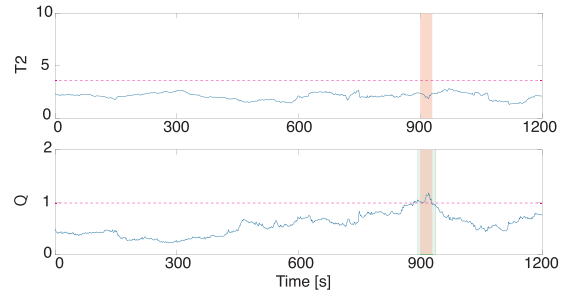
Fig. 9. Detection result of pre-N1 episode Nd2.

TABLE II
AWAKE EPISODES USED FOR MODELING

Participant	Episode	Length [h]	Participant	Episode	Length [h]
A	A1	0.92	R	R1	1.50
B	B1	0.87	S	S1	1.50
C	C1	0.30	T	T1	0.53
D	D1	0.75	U	U1	0.73
E	E1	1.20	V	V1	0.35
F	F1	0.47	W	W1	0.94
G	G1	0.85	X	X1	0.12
H	H1	0.25	Y	Y1	0.55
I	I1	0.32	Z	Z1	1.45
J	J1	0.85	α	$\alpha 1$	0.12
K	K1	0.68	β	$\beta 1$	0.48
L	L1	1.48	γ	$\gamma 1$	0.39
M	M1	1.50	δ	$\delta 1$	0.28
N	N1	0.95	ϵ	$\epsilon 1$	0.26
O	O1	1.03	ζ	$\zeta 1$	0.98
P	P1	0.57	η	$\eta 1$	1.46
Q	Q1	0.63	θ	$\theta 1$	1.47
			Total		26.7

the 90% confidence limits. Although the 99% or the 95% confidence limits are usually adopted in MSPC for suppressing false positives, this research used the 90% confidence limits because it is important to prevent erroneous drowsiness detection from the viewpoint of safety. The parameter $\bar{\tau}$ was determined as ten seconds according to [19].

D. Drowsy Driving Detection Results

All of the drowsy and awake episodes that were not used for modeling were monitored by Algorithm 2. The numbers of validated awake and pre-N1 episodes were 57 and 13, respectively. The total length of the validated awake episodes was 40.1 hours. Here, drowsiness detection success means that drowsiness is detected from 15 minutes before to just before a sleep onset.

From the pre-N1 episode results, the Q statistic detected 12 out of 13 pre-N1 episodes excluding episode Bd1. On the other hand, the T^2 statistic detected 8 out of 13 pre-N1 episodes excluding episodes Ad1, Ld1, Nd2, Pd1, and Zd1. As a result, the sensitivity of the T^2 and Q statistics are 62% and 92%, and the mean and the standard deviation of the first drowsiness detection time by the T^2 and Q statistics were 484 ± 383 and 642 ± 401 seconds before sleep onsets, respectively.

Detection results of pre-N1 episodes Ld1 and Nd2 are shown in Figs. 8 and 9, in which horizontal dashed lines express the control limits of the T^2 and Q statistics. Orange and green colored bands denote the N1 epochs scored by the sleep specialist and drowsy periods detected by the proposed method, respectively. According to Algorithm 2, driver drowsiness is detected only when either T^2 or Q statistic exceeds its control limit continuously for more than $\bar{\tau} = 10$ seconds. Although the T^2 statistic
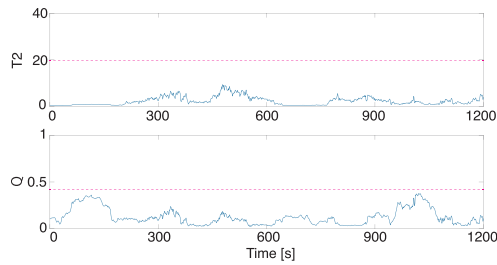
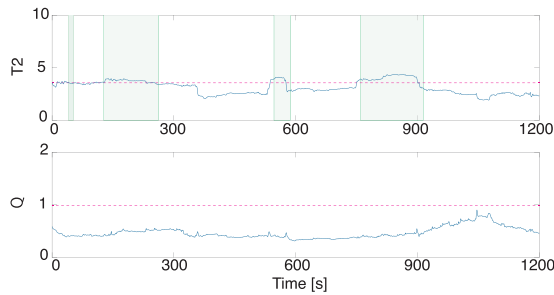Fig. 10. Detection result of awake episode L2.

Fig. 11. Detection result of awake episode N2.

around 900 seconds in Fig. 8 exceeded its control limit; it was not detected as drowsiness because it did not exceed its control limit for more than ten seconds continuously. Figs. 10 and 11 show detection results of awake episodes L2 and N2. There were no false positives in episode L2 while a false positive occurred according to the T^2 statistic in episode N2.

Table III summarizes the number of false positives (#FP) and false positive (FP) rates, which are defined as #FP per hour. The #FP in all awake periods (total 40.1 hours) by the T^2 and Q statistics were 105 and 70, and the FP rates were 2.6 and 1.7 times per hour, respectively.

V. DISCUSSION

In the experiment, we collected HRV data from a total of 34 participants, and drowsiness detection was based on the T^2 and Q statistics. In order to confirm the validity of these statistics, sample powers were calculated. The sample powers of the T^2 and Q statistics were 0.09 and 0.71, respectively. Thus, the sample size was not insufficient for the T^2 statistic; however, the sample size was adequate when the Q statistic was used for drowsiness detection. In fact, the drowsiness detection of the Q statistics was higher than the T^2 statistic in the experiment.

The collected data in the experiment consist of a total of 66.8 hours of awake episodes and a total of 4.3 hours of pre-N1 episodes. This unbalanced ratio of awake to pre-N1 episodes justifies the adoption of the anomaly detection framework of MSPC, in which a drowsiness detection model is developed from the awake episodes only.

One driving trial was limited to 1.5 hours, and there was a one-hour rest including lunch between two trials, in consideration of the burden of the participants. As a result, 12 out of 34 participants were scored by N1 during driving. Ten out of thirteen pre-N1 episodes were observed in the second trial,

TABLE III
FALSE POSITIVES

Participant	Length [h]	T^2		Q	
		#FP	FP rate	#FP	FP rate
A	0.37	0	0	3	8.1
B	1.00	6	6.0	3	3.0
C	0.17	0	0	0	0
D	1.41	8	5.6	2	1.4
E	0.86	0	0	0	0
F	2.28	8	3.5	6	2.6
G	1.46	8	5.5	0	0
H	1.06	1	0.9	0	0
I	1.42	2	1.4	0	0
J	2.07	0	0	1	0.5
K	0.74	5	6.7	5	6.7
L	0.53	1	1.8	1	1.8
M	1.43	6	4.2	12	8.4
N	0.44	5	11.4	0	0
O	1.83	0	0	0	0
P	0.61	3	4.9	9	14.7
Q	2.18	2	1.0	5	2.3
R	1.11	0	0	1	0.9
S	1.51	4	2.6	1	0.6
T	0.60	3	5.0	0	0
U	0.81	6	7.4	2	2.4
V	0.96	10	10.4	0	0
W	1.91	2	1.0	1	0.5
X	2.02	2	1.0	2	1.0
Y	2.26	4	1.8	0	0
Z	0.28	0	0	0	0
α	0.45	0	0	1	2.2
β	0.46	0	0	1	2.1
γ	0.58	0	0	0	0
δ	2.03	7	3.4	1	0.5
ϵ	0.37	1	2.7	1	2.7
ζ	1.93	2	1	1	0.5
η	1.44	6	4.2	4	2.8
θ	1.47	6	4.1	7	4.8
Total	40.1	105	2.6	70	1.7

which was performed after lunch. Although all participants in the experiment were healthy, sleep loss and daytime sleepiness in the Japanese adult population are common [43]. Thus, 1.5 hours of driving after lunch may have been sufficient to induce drowsiness for some participants.

α waves, which indicate sleep-related brain activities, were observed in correspondence with the increase of the Q statistic. Fig. 12 shows α waves recorded during the awake period, which corresponded to a false positive by the Q statistic in episode X6. According to EOG of this period, the participant repeatedly blinked. Such awake α waves associated with the increase of the Q statistic were observed in episodes J2, M2, P2, $\delta4$ as well as X6. These correspondences between changes in the Q statistic and α wave appearance support the validity of the proposed method.

Microsleep, which is sudden short sleep lasting for a fraction of a second or up to 30 seconds, is a well-known phenomenon in sleep science [44]. Persons with microsleep often remain unaware of it. Its causes are, for example, sleep apnea syndrome, narcolepsy, mental fatigue, and insufficient sleep. Many microsleep identification methods have been proposed, and there is little agreement on which is the best for use at this time [44]. It is possible that the awake α wave appearance associated with an increase of the Q statistic was microsleep, although this is difficult to confirm.
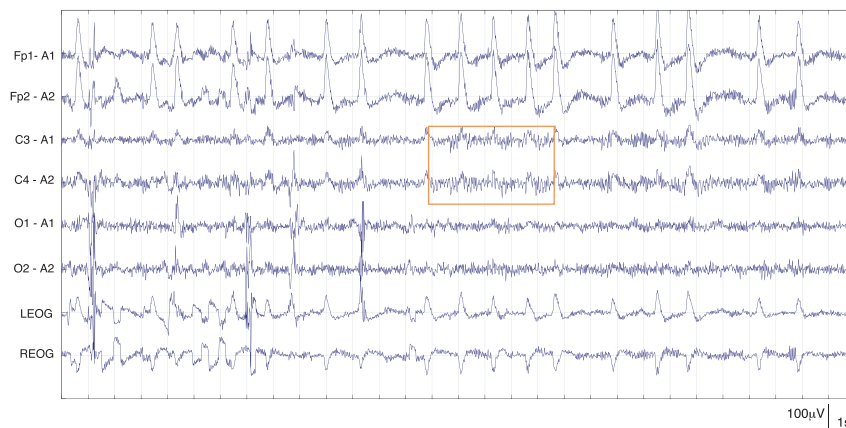
Fig. 12. The awake α waves associated with a false positive by the Q statistic in episode X6.

The occurrence of microsleep may be dangerous particularly in situations that demand constant alertness, such as driving or working with heavy machinery. Microsleep decreases the capability for task execution, which is equivalent to a mild cognitive dysfunction; however, falling asleep directly contributes to the occurrence of traffic accidents, since N1 sleep is equivalent to complete cognitive dysfunction. In fact, the multiple sleep latency test (MSLT), which measures the elapsed time from wakefulness to N1 onset, is correlated to the risk of traffic accidents [45], [46]. Thus, the target of the proposed algorithm is the N1 onset instead of microsleep. Future studies are needed to evaluate whether the proposed methodology for detecting the N1 onset is also applicable to microsleep detection.

According to an evaluation of the EEG data and videos, most false positives occurred in correspondence with participant's motion. The drowsy driving detection model was constructed by using HRV data collected during driving in which participants sat in a seat and rarely moved. Because few HRV data with body motion were contained in the modeling data, fluctuations of HRV caused by body motion in the validation data were detected as false positives. For example, there were eight false positives of the T^2 statistic in participant F, seven of which occurred when his EEG was contaminated with electromyogram (EMG) artifacts caused by body motion.

Xaio *et al.* developed an HRV-based sleep stage scoring method using random forest (RF), which is an ensemble learning technique using multiple decision trees [47]. Their method classifies sleep condition into awake, REM, and NREM and the accuracy of their method was 72%–88%, even though it uses a total of 41 HRV features consisting of nonlinear features, time domain features, and frequency domain features as input variables [48]. It is difficult to compare their sleep stage scoring method with the proposed drowsiness detection algorithm from the viewpoint of performance, because the purpose of their method is to discriminate sleep stages, while that of the proposed method is to detect drowsiness prior to the N1 onset; however, the proposed drowsiness detection method is much simpler than their method since the proposed algorithm uses only eight HRV features and a linear model. This indicates that the framework of HRV-based anomaly detection by MSPC is

useful for drowsiness detection, although it cannot be used for sleep stage scoring.

The sensitivity of the proposed method achieved almost the same level as conventional camera-based drowsiness detection methods [8]; however, most of the latter use facial expression evaluation by referees or subjective evaluation by questionnaires instead of EEG-based sleep scoring for driver sleepiness evaluation. Because it is impossible to detect precise N1 onsets, microsleep, and awake α waves by facial expression evaluation and questionnaires, there is a possibility that previous researchers overlooked such sleep-related phenomena.

It is concluded that the proposed HRV-based drowsy driving detection algorithm is more promising than other conventional methods with respect to accuracy as well as practical use.

VI. Conclusion and Future Work

A driver drowsiness detection method was proposed utilizing the framework of epileptic seizure prediction, by which multiple HRV features are extracted from the RRI data and MSPC monitors abnormalities in the extracted HRV features. The experimental result showed that 12 out of 13 pre-N1 episodes were detected prior to sleep onsets, and the false positive rate was about 1.7 times per hour. The experimental result was discussed from the viewpoint of sleep science. This work demonstrated the usefulness of the framework of HRV-based anomaly detection because it can be applied to driver drowsiness detection as well as epileptic seizure prediction.

Limitations of the study include the properties of the collected experimental data, such as a highly controlled laboratory environment, the limited number of participants, and the fact that all participants were young Japanese persons. Accordingly, more studies are required to confirm our results by using well-matched groups of participants in a real driving environment.

The proposed method requires drivers to put some electrodes on the skin before driving, because precise RRI measurement based on ECG is needed for HRV analysis. Since it is burdensome for drivers to attach electrodes before driving, a new type of electrode that is easy to use should be developed. Tsukada *et al.* developed a new wearable textile electrode using a
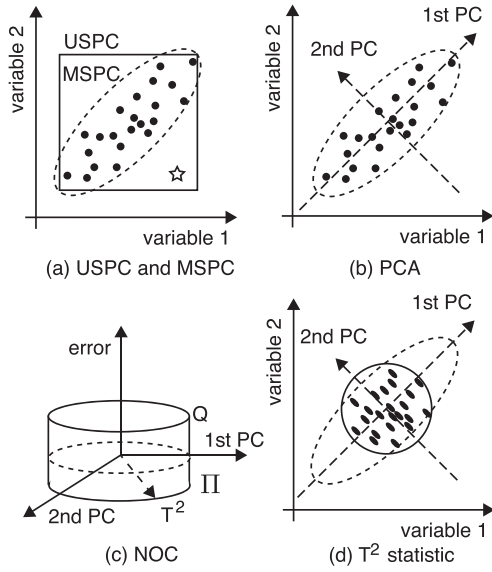
Fig. 13. Schematic diagram of MSPC.

conductive fiber [49], and a smart shirt woven with textile electrodes has been developed for ECG measurement. Therefore, it will be easy for drivers to use the proposed HRV-based drowsy driving detection method when the smart shirt becomes available. In addition, the proposed algorithm can be easily implemented into mobile computers, such as a smartphone, since the computational load of the proposed method is much lighter than the camera-based methods that use real-time video analysis.

In future works, additional experimental data must be collected to improve the drowsiness detection performance, and the system under development will be tested in a real driving environment.

APPENDIX
MULTIVARIATE STATISTICAL PROCESS CONTROL (MSPC)

The appendix explains multivariate statistical process control (MSPC) used in the drowsiness detection algorithm.

The proposed drowsiness detection algorithm described in Section III-C detects driver drowsiness as anomalies in HRV. The simplest way of detecting anomalies is to check whether or not all variables are within their upper and lower bounds. This simple method is called univariate statistical process control (USPC); it is also known as control charts and Shewhart charts. When multiple variables are monitored simultaneously, the nominal region of USPC becomes rectangular as shown in Fig. 13(a). USPC is intuitive and easy-to-use, and therefore has been widely used in various fields including the manufacturing industry. However, it cannot detect an anomaly that does not satisfy normal correlation among variables. In Fig. 13(a), for example, USPC cannot detect the anomaly ☆, which does not follow a positive correlation between variables 1 and 2, because it is located within the normal rectangular area of USPC. If the normal ellipsoid area defined by the dashed line is used instead of the normal rectangular area, the anomaly ☆ can be detected. This example demonstrates that the correlation among variables

should be taken into account for detecting anomalies that do not follow the major trend in the data.

In MSPC, the correlation among variables is modeled by using principal component analysis (PCA) [50], which finds linear combinations of variables that describe major trends in a dataset as shown in Fig. 13(b). Let us assume a normal data matrix $X \in \Re^{N \times M}$ whose nth row is the nth sample $x_n \in \Re^M$, wherein samples are mean-centered and scaled appropriately, and M and N denote the number of variables and samples, respectively. Such a matrix X can be factorized by singular value decomposition (SVD) as follows:

$$X = U \Sigma V^T$$
$$= \begin{bmatrix} U_R & U_0 \end{bmatrix} \begin{bmatrix} \Sigma_R & 0 \\ 0 & \Sigma_0 \end{bmatrix} \begin{bmatrix} V_R & V_0 \end{bmatrix}^T \quad (1)$$

where U, Σ, and V are the left singular matrix, the diagonal matrix whose diagonal elements are singular values, and the right singular matrix, respectively. Using Eq. (1), the matrix factorization of X by PCA is defined as:

$$X = T_R V_R^T + E \quad (2)$$

where $T_R \in \Re^{N \times R}$ is the score matrix and $T_R \equiv U \Sigma$. $E \in \Re^{N \times M}$ is an error matrix. The column of V_R spans the subspace Π which expresses the correlation among variables as shown in Fig. 13(c). $R(\le M)$ is the number of principal components retained in the PCA model.

The T^2 statistic is used for monitoring anomalies in Π, which is defined as follows:

$$T^2 = \sum_{r=1}^{R} \frac{t_r^2}{\sigma_{t_r}^2}$$
$$= x^T V_R \Sigma_R^{-2} V_R^T x \quad (3)$$

where σ_{t_r} denotes the standard deviation of the rth score t_r and x is a newly measured sample. Since the T^2 statistic is the Mahalanobis distance, which is defined as the distance normalized by the standard deviations of the scores, it defines a circular nominal region as shown in Fig. 13(d). Hence, the sample is close to the mean of the modeling data when the T^2 statistic is small.

The Q statistic is defined as the squared distance between the sample and Π as

$$Q = \sum_{m=1}^{M} e_m^2 = \sum_{m=1}^{M} (x_m - \hat{x}_m)^2$$
$$= x^T (I - V_R V_R^T) x. \quad (4)$$

The normal operating condition (NOC) is defined using the T^2 and the Q statistics [51]. Fig. 13(c) shows the image of NOC in MSPC. Since the monitored subspaces by two statistics are orthogonal to each other, NOC can be considered as a cylinder and the control limits of the T^2 and Q statistics correspond to its diameter and height, respectively. MSPC usually detects an anomaly when either the T^2 or Q statistic exceeds the predefined control limit. Thus, the control limits of the T^2 and Q statistics have to be determined carefully, which can be set as

α% confidence limits, and usually, the 99% or 95% confidence limits are adopted.

In order to detect anomalies by MSPC, the number of principal components R also has to be appropriately determined. The number of principal components R can be determined so that the cumulative proportion of principal components reaches a predefined value, such as 80% or 90%.

REFERENCES

[1] S. G. Klauer *et al.*, "The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data," *National Highway Traffic Safety Administration*, 2006.

[2] D. J. Gottlieb *et al.*, "Sleep deficiency and motor vehicle crash risk in the general population: a prospective cohort study," *BMC Med.*, vol. 16, no. 1, 2018, Art. no. 44.

[3] M. H. Kryger *et al.*, *Principles and Practice of Sleep Medicine*. 6th ed., New York, NY, USA: Elsevier, 2017.

[4] H. Su and G. Zheng, "A partial least squares regression-based fusion model for predicting the trend in drowsiness," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 38, no. 5, pp. 1085–1092, 2008.

[5] C.-T. Lin *et al.*, "A real-time wireless brain–computer interface system for drowsiness detection," *IEEE Trans. Biomed. Circuits Syst.*, vol. 4, no. 4, pp. 214–222, Aug. 2010.

[6] F.-C. Lin *et al.*, "Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 9, pp. 2044–2055, Sep. 2012.

[7] A. Hashemi *et al.*, "Time driver's drowsiness detection by processing the eeg signals stimulated with external flickering light," *Basic Clin. Neurosci.*, vol. 5, no. 1, pp. 22–27, 2014.

[8] S. Kaplan, "Driver behavior analysis for safe driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3017–3032, Dec. 2015.

[9] R. O. Mbouna *et al.*, "Visual analysis of eye state and head pose for driver alertness monitoring," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1462–1469, Sep. 2013.

[10] J. Jo *et al.*, "Detecting driver drowsiness using feature-level fusion and user-specific classification," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1139–1152, 2014.

[11] B. Cyganek and S. Gruszczynski, "Hybrid computer vision system for drivers' eye recognition and fatigue monitoring," *Neurocomputing*, vol. 126, no. 27, pp. 78–94, 2014.

[12] J. H. Yang *et al.*, "Detection of driver fatigue caused by sleep deprivation," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 4, pp. 694–705, Jul. 2009.

[13] D. Sandberg *et al.*, "Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 1, pp. 97–108, Mar. 2011.

[14] P. M. Forsman *et al.*, "Efficient driver drowsiness detection at moderate levels of drowsiness," *Accident Anal. Prev.*, vol. 50, pp. 341–350, Jan. 2013.

[15] J. Trinder *et al.*, "Autonomic activity during human sleep as a function of time and sleep stage," *J. Sleep Res.*, vol. 10, no. 4, pp. 253–264, 2001.

[16] K. T. Chui *et al.*, "An accurate ECG-based transportation safety drowsiness detection scheme," *IEEE Trans. Ind. Inform.*, vol. 12, no. 4, pp. 1438–1452, Aug. 2016.

[17] B. G. Lee *et al.*, "Real-time physiological and vision monitoring of vehicle driver for non-intrusive drowsiness detection," *IET Commun.*, vol. 5, no. 17, pp. 2461–2469, 2011.

[18] A. J. Camm *et al.*, "Guidelines heart rate variability—standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 115, no. 5, pp. 354–381, 1996.

[19] K. Fujiwara *et al.*, "Epileptic seizure prediction based on multivariate statistical process control of heart rate variability features," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 6, pp. 1321–1332, Jun. 2016.

[20] D. H. Kerem and B. Geva, "Forecasting epilepsy from the heart rate signal," *Med. Biol. Eng. Comput.*, vol. 43, pp. 230–239, 2005.

[21] K. Kato *et al.*, "Earlier tachycardia onset in right than left mesial temporal lobe," *Neurology*, vol. 83, pp. 1332–1336, 2014.

[22] F. Versace *et al.*, "Heart rate variability during sleep as a function of the sleep cycle," *Biol. Psychol.*, vol. 63, no. 2, pp. 146–162, 2003.

[23] F. Jurysta *et al.*, "A study of the dynamic interactions between sleep EEG and heart rate variability in healthy young men," *Clin. Neurophysiol.*, vol. 114, no. 11, pp. 2146–2155, 2003.

[24] E. Tobaldini *et al.*, "Heart rate variability in normal and pathological sleep," *Front. Physiol.*, vol. 4, 2013, Art. no. 294.

[25] P. Nomikos and J. F. MacGregor, "Monitoring batch processes using multiway principal component analysis," *AIChE J.*, vol. 40, no. 8, pp. 1361–1375, 1994.

[26] J. F. MacGregor and T. Kourti, "Statistical process control of multivariate processes," *Control Eng. Pract.*, vol. 3, no. 3, pp. 403–414, 1995.

[27] M. Kano *et al.*, "Comparison of multivariate statistical process monitoring methods with applications to the eastman challenge problem," *Comput. Chem. Eng.*, vol. 26, no. 2, pp. 161–174, 2002.

[28] J. Vicente *et al.*, "Drowsiness detection using heart rate variability," *Med. Biol. Eng. Comput.*, vol. 54, no. 6, pp. 927–937, 2016.

[29] G. Li and W.-Y. Chung, "Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier," *Sensor*, vol. 13, no. 12, pp. 16 494–16 511, 2013.

[30] M. Patel *et al.*, "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert. Syst. Appl.*, vol. 38, no. 6, pp. 7235–7242, 2011.

[31] *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, 2nd ed., Darien, IL, USA: American Academy of Sleep Medicine, 2016.

[32] J. Santamaria and K. H. Chiappa, *The Eeg of Drowsiness*. New York, NY, USA: Demos Medical Pub, 1987.

[33] R. E. Kleiger *et al.*, "Decreased heart rate variability and its association with increased mortality after acute myocardial infarction," *Am. J. Cardiol.*, vol. 59, no. 4, pp. 256–262, 1987.

[34] A. Malliani, "Cardiovascular neural regulation explored in the frequency domain," *Circulation*, vol. 84, no. 2, pp. 482–492, 1991.

[35] J. M. Yentes *et al.*, "The appropriate use of approximate entropy and sample entropy with short data sets," *Ann. Biomed. Eng.*, vol. 41, no. 2, pp. 349–365, 2013.

[36] G. Lu and F. Yang, "Limitations of oximetry to measure heart rate variability measures," *Cardiovasc. Eng.*, vol. 9, no. 3, pp. 119–125, 2009.

[37] T. Yamakawa *et al.*, "Accuracy comparison of two microcontroller-embedded r-wave detection methods for heart-rate variability analysis," in *Proc. APSIPA ASC*, 2015, pp. 1010–1013.

[38] M. H. Bonnet and D. L. Arand, "Activity, arousal, and the MSLT in patients with insomnia," *Sleep*, vol. 23, no. 2, pp. 205–212, 2000.

[39] E. C. Chua *et al.*, "Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation," *Sleep*, vol. 35, no. 3, pp. 325–334, 2012.

[40] M. Ohsuga *et al.*, "Estimation of driver's arousal state using multi-dimensional physiological indices," in *Proc. EPCE*, 2011, pp. 176–185.

[41] S. Bando and A. Nozawa, "Detection of driver inattention from fluctuations in vehicle operating data," *Artif. Life Robot.*, vol. 20, no. 1, pp. 28–33, 2015.

[42] T. Nguyen *et al.*, "Utilization of a combined EEG/NIRS system to predict driver drowsiness," *Sci. Rep.*, vol. 7, 2017, Art. no. 43933.

[43] X. Liu *et al.*, "Sleep loss and daytime sleepiness in the general adult population of Japan," *Psychiatry Res.*, vol. 93, no. 1, pp. 1–11, 2000.

[44] M. J. Sateia, "International classification of sleep disorders-third edition: highlights and modifications," *Chest*, vol. 146, no. 5, pp. 1387–1394, 2014.

[45] F. Pizza *et al.*, "Daytime sleepiness and driving performance in patients with obstructive sleep apnea: Comparison of the MSLT, the MWT, and a simulated driving task," *Sleep*, vol. 32, no. 3, pp. 382–391, 2009.

[46] P. Philip, "Can the MSLT be a useful tool to assess motor vehicle crash risk in sleepy drivers?" *Sleep*, vol. 33, no. 6, pp. 729–730, 2010.

[47] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[48] M. Xiao *et al.*, "Sleep stages classification based on heart rate variability and randomforest," *Biomed. Signal Process Control*, vol. 8, no. 6, pp. 624–633, 2013.

[49] S. Tsukada *et al.*, "Conductive polymer combined silk fiber bundle for bioelectrical signal recording," *PLoS One*, vol. 7, no. 4, 2012, Art. no. e33689.

[50] M. Kano *et al.*, "A new multivariate statistical process monitoring method using principal component analysis," *Comput. Chem. Eng.*, vol. 25, no. 7–8, pp. 1103–1113, 2001.

[51] J. E. Jackson *et al.*, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, 1973.