A Real-Time Fatigue Driving Recognition Method Incorporating Contextual Features and Two Fusion Levels

Wei Sun, Xiaorui Zhang, Srinivas Peeta, Xiaozheng He, and Yongfu Li

Abstract— **Though experimental results have shown a strong correlation between contextual features and the driver's fatigue state, contextual features have been applied only offline to evaluate a driver's fatigue state. This paper identifies three of the most effective contextual features, i.e., continuous driving time, sleep duration time, and current time, to facilitate the real-time (online) recognition of fatigue state. By applying gray relational analysis, the three contextual features, together with the most effective facial and vehicle behavior features, are introduced in a two-level fusion structure to improve fatigue driving recognition. In the first level of fusion, labeled the feature-level fusion, three separate multiclass support vector machine (MCSVM) classifiers are used for the three feature sources, i.e., contextual features, driver's facial features, and vehicle behavior features, to fuse information. These three MCSVM classifiers output probabilities as inputs for the three real-time dynamic basic probability assignments (BPAs) at the second level of fusion, labeled decision-level fusion. These BPAs, and the fusion result of the previous time step, are fused in the decision-level fusion based on the Dempster–Shafer evidence theory. This includes modifying the BPAs to accommodate the decision conflict among the different feature sources. Field experiments show that the proposed recognition method can outperform the single-fatigue-feature method and the single-source fusion-based method.**

Manuscript received July 9, 2016; revised December 22, 2016; accepted March 24, 2017. Date of publication April 24, 2017; date of current version December 7, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61304205, Grant 61502240, Grant 61304197, and Grant 41301037; in part by the Natural Science Foundation of Jiangsu Province under Grant BK20141002; in part by the College Students Practice Innovation Training Program of Nuist under Grant 201610300231 and Grant 201610300048; and in part by the U.S. Department of Transportation through the NEXTRANS Center, the USDOT Region 5 University Transportation Center. The Associate Editor for this paper was Q. Ji. *(Corresponding author: Wei Sun.)*

W. Sun is with the School of Information and Control, Nanjing University of Information Science and Technology, Nanjing 210044, China, and also with the Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: sunw0125@163.com).

X. Zhang is with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China, and also with the Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044, China.

S. Peeta is with the School of Civil Engineering, Purdue University, West Lafayette, IN 47906 USA, and also with the NEXTRANS Center, Purdue University, West Lafayette, IN 47906 USA.

X. He is with the Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA.

Y. Li is with the College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2017.2690914

Index Terms— **Fatigue driving, contextual features, multi-class support vector machine classifier, Dempster–Shafer evidence theory.**

I. Introduction

FATIGUE driving is a primary reason for many traffic accidents, casualties, and property losses. Hence, there is the need to effectively recognize driver's fatigue driving state to improve travel safety by leveraging technological advances. In the literature, many factors have been used to recognize fatigue driving, including drivers' physiological state, facial expression, and vehicle operation/running conditions [1].

Existing fatigue driving recognition methods can be divided into two categories: single-source information and multi-source information. The recognition methods based on single-source information have inevitable limitations in acquiring reliable and robust data. For example, physiology based methods [2]–[4] need to place multiple electrodes on a driver's skin for physiological signals. The direct contact can cause an uncomfortable and annoying feeling for drivers. Facial-features based methods [5]–[10] recognize fatigue driving by analyzing facial expression changes, such as eye closure duration, blinking, yawning, or eyelid/gaze movement, through a real-time monitoring system. However, external interferences, such as lighting change, sudden head movement, and darkness at night, can reduce the recognition accuracy. Vehicle behavior based methods recognize fatigue driving by detecting vehicle operational conditions, such as the lane departure degree [11], [12], or the variation of steering wheel angle [13], [14]. However, a travel environment with blurred lane markings may impede the applicability of these methods.

To overcome the limitations of single-source information methods, fatigue driving recognition methods have been proposed based on information from multiple sources. These methods contain two main steps: selection of fatigue features, and construction of recognition model, which play important roles in improving fatigue driving recognition.

In general, incorporating as many fatigue features as possible from multiple information sources can improve fatigue driving recognition [15]. However, too many features for a recognition model can cause heavy computational burden, which impedes the timeliness of fatigue driving recognition [15]. Therefore, only the most effective features should be selected to enable a practical fatigue recognition method.

1524-9050 © 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

Laboratory experiments suggest that driver's contextual information, such as sleep time, sleep quality, continuous driving time, working time, and temperature in cab, can be used to reflect driver's fatigue states [16], [17]. However, these contextual features have been used previously only to explain fatigue cause. They could serve as candidates in the selection of effective fatigue features to develop fatigue driving recognition methods.

Existing fatigue recognition models are mainly based on artificial neural network [18], support vector machine [19] and dynamic Bayesian network [20]. These models adopt a single-level structure, wherein the fatigue features are fused simultaneously. Therefore, they may not work properly when a certain fatigue feature measurement fails due to unpredictable disturbances, such as sudden lighting changes, missing sensor signals, etc., in a real-world travel environment.

Some fatigue recognition models [21]–[23] adopt a two-level fusion structure based on Dempster-Shafer evidence theory (D-SET). However, they cannot resolve the evidence conflict among multiple pieces of evidence caused by unpredictable disturbances, due to the following inherent methodological limitations. First, the deterministic basic probability assignment (BPA) used in the two-level fusion models is not accurate enough for each feature source. While BPA can be determined dynamically based on Takagi-Sugeno fuzzy neural network (T-SFNN) model [23], training the T-SFNN model requires a large number of samples to cover various fatigue states. A small number of samples could reduce the accuracy of fatigue driving recognition. As obtaining a large-size training sample is difficult, these models are not suitable for online fatigue driving recognition over a long period. Second, with more evidence sources, evidence conflict is inevitable during evidence fusion when external disturbances exist. Without a correction technique to deal with the evidence conflict, existing models cannot effectively produce accurate fatigue recognition results that can accommodate real-world travel environment changes.

In summary, though some methods apply multi-source information to improve the efficiency of fatigue driving recognition, they have the following limitations. First, though contextual features have a strong correlation with driver's fatigue state, they have not been considered for real-time fatigue driving recognition. Second, real-time fatigue driving recognition cannot be improved by simply increasing the number of incorporated fatigue features. Too many fatigue features entail a heavy computational burden and data redundancy. Third, fatigue recognition models lack modeling flexibility to deal with the evidence conflict among features caused by disturbances from travel environment changes.

To address the aforementioned limitations, we propose a fatigue driving recognition method based on multi-feature identification and two-level fusion of multi-source information. Three fatigue feature sources, driver's contextual features, facial features and vehicle behavior features, are considered. The most effective fatigue features are selected through a feature identification process based on grey relational analysis (GRA) that reduces computational burden and precludes potential data redundancy. The potential evidence conflict is eliminated through a modified BPA that is based on the multi-class support vector machine (MCSVM) classifier with a correction technique that considers conflicting evidence. The method proposed in this paper can outperform the methods based on percentage of eye closure (PERCLOS) [24]–[26] in terms of reliability. In particular, when external interferences occur, such as lighting change and sudden head movement, PERCLOS- based methods may not work as they rely totally on a single fatigue feature.

The contributions of this study are threefold. First, contextual features related to fatigue driving, i.e., continuous driving time, sleep duration time, and current time, are selected for the proposed fatigue driving recognition model, leading to improved fatigue driving recognition. Second, a GRA-based method is proposed to select the most effective fatigue features, which can eliminate potential data redundancy. Third, a MCSVM classifier is used to fuse multiple fatigue features from different sources. The MCSVM classifier relies on only a small set of training samples, which not only enhances the feature-level fusion but also enables a dynamic BPA of each feature source in the decision-level fusion. An assignment correction technique for the dynamic BPA is proposed to resolve the evidence conflict caused by disturbances from travel environment changes.

The remainder of this paper is organized as follows. Section II describes the fatigue feature measurements that will be used to recognize driver's fatigue state. Section III presents the fatigue feature identification method, in which grey relational analysis is introduced to investigate the correlation between fatigue driving and the measured fatigue features. Section IV proposes a two-level recognition model, which consists of a feature-level fusion based on the MCSVM classifier and a decision-level fusion based on the improved D-SET. Section V presents a case study using field data collected on the Nanjing-Zhenjiang expressway in China, to illustrate the effectiveness of the proposed method. The final section provides some concluding comments.

II. Preliminaries

A previous study [23] used facial features and vehicle behavior features to characterize the fatigue state of drivers. Facial features considered in [23] include blinking frequency (BF), eye-closed duration (ECD), mean of eye-opened level (MEOL), and yawning frequency (YF). Vehicle behavior features considered include the percentage of non-steering (PNS), standard deviation of steering-angle (SDSA), frequency of abnormal lane deviation (FALD), and standard deviation of vehicle speed (SDVS). This paper includes these features as candidate fatigue features. Please refer to [23] for the real-time measurement of these candidate fatigue features.

Some studies in physiology have shown that factors such as sleep condition, driving time, and environmental temperature in cab have a significant correlation with driver's fatigue state [27]. Therefore, fatigue features derived from driver's contextual information can be used to enhance the fatigue driving recognition. Four contextual features have been shown to have a strong correlation with driver's fatigue

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

state [16], [17], [27]: continuous driving time (CDT), air temperature in cab (ATC), sleep duration time (SDT), and current time (CT). As these features are easy to measure in real time, this study selects these four features as candidate contextual features. They are defined as follows:

(i) Continuous driving time (CDT): Let t_{CDT} denote the CDT, which can be measured by a vehicle-mounted stopwatch.

(ii) Air temperature in cab (ATC): Let t_{ATC} denote the ATC, which can be detected in real time by a vehicle-mounted thermometer.

(iii) Sleep duration time (SDT): Let t_{SDT} denote the SDT of a driver, which reflects the sleep situation of a driver before starting to drive. $t_{\mathrm{SDT}} = T_{\mathrm{SS}} - T_{\mathrm{A}}$, where T_{SS} is the time when a driver starts to go to sleep, and T_{A} is the time when the driver wakes up after having a rest. T_{SS} and T_{A} can be measured by a sleep tracker worn by the driver.

(iv) Current time (CT): Variable t_{CT} denotes the CT of the day, which can be measured by a vehicle-mounted clock.

III. Fatigue Feature Selection

Efficiency and reliability are two contradictory factors for a real-time fatigue driving recognition model. From an efficiency perspective, if a recognition model incorporates fewer fatigue features, it can compute faster. By contrast, from a reliability perspective, if a recognition model incorporates a larger number of fatigue features, it can be more reliable. This is because if a sensor fails in a multiple-source based recognition system due to unpredictable disturbances, a reliable recognition result can be obtained through other sensors based on data fusion algorithms [15], [20], [21]. This illustrates a tradeoff related to the number of fatigue features selected. It motivates the need to evaluate the effectiveness of the fatigue features presented in Section II, and to select the most effective ones as inputs to the fatigue recognition model.

A. Fatigue State Assessment

Previous studies focus on subjective assessments of fatigue driving by observing phenomena such as a driver's facial expression and operational behavior, and by his/her self-evaluation of fatigue. As fatigue state assessment based on EEG [28] is objective and has higher reliability, we introduce a comprehensive method for assessing fatigue state by combining subjective and objective assessments, to evaluate and determine driver's fatigue states and to provide ground truth data for model calibration and verification.

The comprehensive assessment method is composed of four steps: (i) observer assessment, (ii) objective assessment, (iii) self-assessment, and (iv) comprehensive assessment. We adopt the 9-point Karolinska sleepiness scale (KSS) table [29], [30], for self-assessment, instead of the 7-point Stanford Sleepiness Scale table used in [23], to enable a more reliable assessment of fatigue state.

During the self-assessment, every participant evaluates his/her fatigue state according to the current physical, physiological and psychological situations. Note that the observer and objective assessment processes based on EEG detection must be implemented simultaneously to cross-check the fatigue state. The self-assessment process based on the KSS table should be carried out within one minute after the other two assessments are accomplished.

B. Correlation Analysis Based on GRA

To avoid the heavy computational burden caused by including many features, only features that are highly correlated with fatigue driving are adopted in the proposed model. The GRA method is effective in determining the critical elements that significantly influence certain defined objectives [31]. Compared with other methods, such as Pearson test [23], it has advantages such as small sample size, no restriction of functional form, no requirement for independence or normal distribution, and fewer computations. Hence, it has been extensively used for relevance analysis in various disciplines [31]. Here, the GRA method is used to perform correlation analysis between fatigue features and fatigue state, and to select the most effective fatigue features. The detailed steps are as follows:

(i) Based on the fatigue feature measurement and fatigue state assessment, the collected samples are viewed as a group of discrete data sequences, denoted by $X_0(j)$ and $X_i(j)$, $i = 1, 2, \cdots, M$, $j = 1, 2, \cdots, N$, where $X_0(j)$ represents a parent data sequence and $X_i(j)$ represents a child data sequence. We further define fatigue level of the sample data according to their fatigue state evaluated in Section III.A. Here, $M = 12$ represents the twelve candidate fatigue features presented in Section II. $X_0(j)$ represents the fatigue level value of the jth sample data and $X_i(j)$ denotes the ith fatigue feature value of the jth sample data, where $X_1(j), X_2(j), \cdots, X_{11}(j)$ and $X_{12}(j)$ denote the measurement values of BF, ECD, MEOL, YF, PNS, SDSA, FALD, SDVS, CDT, ATC, SDT, and CT of the jth sample data, respectively.

(ii) Sample data are normalized as:

$$\bar{X}_0(j) = \frac{X_0(j)}{\sum_{j=1}^{N} X_0(j)/N}, \quad \bar{X}_i(j) = \frac{X_i(j)}{\sum_{j=1}^{N} X_i(j)/N} \tag{1}$$

where $i = 1, 2, \cdots, M$ and $j = 1, 2, \cdots, N$.

(iii) The grey relational coefficient between $\bar{X}_0(j)$ and $\bar{X}_i(j)$ is calculated as:

$$\psi\ (\bar{X}_0(j), \bar{X}_i(j)) = \frac{\min\limits_i \min\limits_j \left|\bar{X}_0(j) - \bar{X}_i(j)\right| + \rho \max\limits_i \max\limits_j \left|\bar{X}_0(j) - \bar{X}_i(j)\right|}{\left|\bar{X}_0(j) - \bar{X}_i(j)\right| + \rho \max\limits_i \max\limits_j \left|\bar{X}_0(j) - \bar{X}_i(j)\right|} \tag{2}$$

where ρ is a coefficient in [0, 1], which can be determined through experiments and is used to better distinguish between the parent and child data sequences.

(iv) The grey relational coefficient is transformed into a grey relational density $\gamma_j(0, i)$:

$$\gamma_j(0, i) = \frac{\psi(\bar{X}_0(j), \bar{X}_i(j))}{\sum_{j=1}^{N} \psi(\bar{X}_0(j), \bar{X}_i(j))}, \quad i = 1, \cdots, M, \quad j = 1, \cdots, N. \tag{3}$$

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

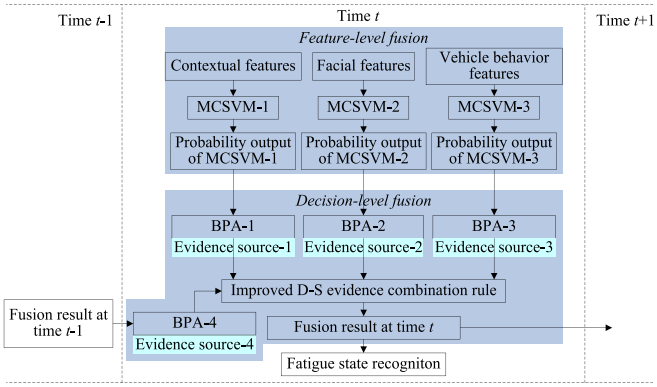


Fig. 1. Flowchart of fatigue driving recognition.

(v) After the grey relational density $\gamma_j(0,i)$ is obtained, the entropy $E(i)$ of the grey relational coefficient of each child sequence is obtained:

$$E(i) = \frac{-\sum_{j=1}^{N} \gamma_j(0,i)\,\ln\gamma_j(0,i)}{\ln(N)}, \quad i = 1, 2, \cdots, M \tag{4}$$

where the term of $-\sum_{j=1}^{N} \gamma_j(0,i)\ln\gamma_j(0,i)$ expresses the grey entropy between the parent sequence $\bar{X}_0(j)$ and the child sequences $\bar{X}_i(j)$, and $\ln(N)$ is the maximum grey entropy, with $0 \le E(i) \le 1$.

(vi) By multiplying the entropy $E(i)$ and the average grey relational coefficient of the child sequence $\bar{X}_i(j)$, we obtain the grey relational degree (GRD):

$$D(i) = \frac{E(i)}{N} \sum_{j=1}^{N} \psi(\bar{X}_0(j), \bar{X}_i(j)) \tag{5}$$

where $D(i) \in [0, 1]$ is the GRD between the child and parent sequences. A larger value of $D(i)$ indicates a higher degree of correlation between the child and parent sequences. Therefore, the GRD can be used to identify the effectiveness of fatigue features.

IV. Two-Level Fusion Model

We propose a two-level fusion model based on multi-source information for fatigue driving recognition, which consists of the feature-level fusion based on the MCSVM and the decision-level fusion based on D-SET. The proposed model structure is illustrated in Fig. 1. In the feature-level fusion, the contextual, facial, and vehicle behavior features are the inputs of three MCSVM classifiers, MCSVM-1, MCSVM-2, and MCSVM-3, respectively, to produce dynamic BPAs in real time for the decision-level fusion. The outputs of the three MCSVM classifiers are inputs for the decision-level fusion based on D-SET.

In the decision-level fusion, the outputs of MCSVM-1, MCSVM-2 and MCSVM-3 are viewed as three pieces of evidence. The fusion result from the previous time step is the fourth evidence, and these four pieces of evidence are fused using an improved evidence combination rule. In the improved evidence combination, a correction technique is proposed for the BPAs to resolve evidence conflicts.

A. Feature-Level Fusion

Support vector machine (SVM) is based on statistical learning theory for classification, and has the advantage of requiring only a small sample set compared to artificial neural network and Bayesian based classification methods. The basic SVM classifier can enable binary classification, but cannot achieve multi-class classification [32]. To classify fatigue feature data into multiple class labels, we propose a feature-level fusion method based on MCSVM classifier. A sigmoid function is introduced to carry out the probability classification of the data with two class labels based on the basic SVM classifier. Then, the "one-against-one" strategy [33] is introduced to compute multi-class probability classification for the data with multiple class labels.

1) Probability Classification Based on Binary-Class SVM: Instead of predicting the binary label, we apply a posterior class probability for the classification of fatigue state. To solve the problem, we introduce a sigmoid function to calculate the posterior probability $P(y = 1\,|\,f)$. [34]:

$$P(y = 1\,|\,f) \approx P_{U,V}(f) = \frac{1}{1 + \exp(U \cdot f + V)} \tag{6}$$

where $f = f(x) = \sum_{i=1}^{l} y_i \alpha_i^* k(x_i, x) + b^*$, which is the decision function of binary classification. Discriminant function $sgn(f(x))$ is used to predict the binary label of test sample x. The radical basis function (RBF) is selected as kernel function of the SVM. For the parameter calculation related to the binary classification based on basic SVM, please refer to [35]. For the optimal parameter $z^* = (U^*, V^*)$ in (6), please refer to [36].

2) Probability Classification Based on Multi-Class SVM: The method in the previous section is suitable only for binary classification. For the multi-class classification problem, "one-against-all" [37] and "one-against-one" [33] strategies are two effective methods. However, compared to the "one-against-all" strategy, the "one-against-one" strategy performs better in terms of training time in practice [33]. Therefore, we propose a new probability classification method based on the basic SVM classifier and "one-against-one" strategy. The detailed procedure is as follows:

(i) Given the data set $\{(x_i, y_i)\}_{i=1}^{l}$, $x_i \in R^n$, which is a fatigue feature vector with n elements, and $y_i \in \{1, 2, \cdots, k_F\}$ is the class label of x_i. Here, k_F is the number of evaluated fatigue states.

(ii) Construct N_M binary-class SVM classifiers using the method described in [35], where $N_M = k_F(k_F - 1)/2$.

(iii) In the N_M SVM classifiers, for the binary-class classifier constructed for the sample data of the ith and jth classes, the posterior probability of x belonging to the ith class can be calculated as:

$$p_{i,j}(x) = \frac{1}{1 + \exp(U^* \cdot f_{i,j} + V^*)} \tag{7}$$

where $f_{i,j}$ is the estimate of the decision function $f(x)$ determined by the ith class and jth class sample data.

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

(iv) For the classification of three fatigue states, the final posterior probability of x belonging to the ith fatigue state can be calculated as:

$$\bar{p}_i(x) = \frac{\sum_{j=1, j\neq i}^{k_F} p_{i,j}(x)}{\sum_{r=1}^{k_F}\sum_{j=1, j\neq r}^{k_F} p_{r,j}(x)}, \quad i = 1, 2, \cdots, k_F. \tag{8}$$

B. Decision-Level Fusion

1) Dynamic Basic Probability Assignment: Generally, after the BPA is subjectively assigned by experts according to their experience, it is unable to dynamically change the assignment according to external travel environment, which will degrade fatigue driving recognition.

To overcome this limitation, a dynamic BPA assignment method is proposed, where the BPAs of three fatigue feature sources are assigned dynamically by three MCSVM classifiers, according to real-time fatigue feature measurements. In particular, we let $m(A_i) = \bar{p}_i(x)$, where $m(A_i)$ is the BPA of the ith state in the D-SET, $i = 1, 2, \cdots, k_F$.

2) Improved Evidence Combination: Compared with other statistical inference methods, D-SET is closer to the human perception and reasoning process, and can fuse multi-source information to infer the results with some degree of certainty [38]. However, evidence conflict is inevitable during evidence fusion. To resolve this problem, this section proposes an improved evidence fusion method by introducing a conflict factor to judge conflicting evidence and a conflict scale factor to eliminate the conflict.

Denote discernment frame as $\Theta = \{A_1, A_2, \cdots, A_{N_H}\}$, evidence set as $e = \{e_1, e_2, \cdots, e_{n_E}\}$, $m_1, m_2, \cdots$, and m_{n_E} as mass functions on Θ, $m_1(\cdot)$, $m_2(\cdot)$, $\cdots$, and $m_{n_E}(\cdot)$ as the BPAs on Θ, N_H as the number of hypotheses, and n_E as the number pieces of evidence. In this paper, N_H also equals the number of fatigue states. The Dempster combination rule for every evidence is as follows [39]:

$$m(A_k) = \begin{cases} \dfrac{1}{1-K}\displaystyle\sum_{\cap A_j = A_k}\prod_{1\leq i\leq n_E} m_i(A_j), & \forall A_k \in \Theta, A_k \neq \emptyset \\ 0, & A_k = \emptyset \end{cases} \tag{9}$$

where $K = \sum_{\cap A_j = \emptyset}\prod_{1\leq i\leq n_E} m_i(A_j)$ denotes the conflict degree among all pieces of evidence.

In this paper, four pieces of evidence from different information source are combined at each time step, where three pieces are from three fatigue feature sources and the fourth is from the fusion result based on D-SET in the previous time step.

In (9), if $K = 1$, then the n_E pieces of evidence completely conflict to each other. In the case, the rule of (9) fails and cannot be used to obtain a correct result. When $K \to 1$, the n pieces of evidence have a higher conflict, and an illogical result may be produced by this rule. Therefore, we need to identify whether conflict will happen before conducting evidence combination. If the conflict happens, we need to make a correction for the conflicting evidence.

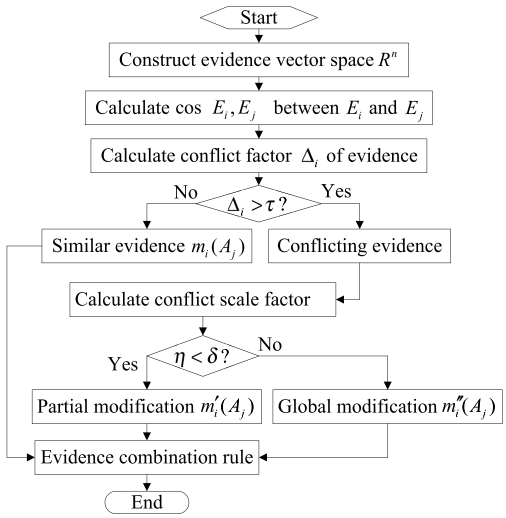


Fig. 2. Flowchart of the improved evidence combination.

The conflicting evidence can also contain some valuable information for evidence fusion. If the conflicting evidence piece is removed randomly, it will result in a loss of information [40]. Therefore, we propose an improved evidence combination method that combines both the similar evidence piece and conflicting evidence piece to obtain a more reasonable inference. The evidence combination method is illustrated in Fig. 2.

Here, we define a conflict factor based on the average similarity between a certain evidence piece and other evidence pieces to identify the conflicting evidence piece. The detailed procedure is as follows:

(i) Construct a vector space R^{N_H} with N_H dimensions according to the number of fatigue states. Let $E_k = \left[m_k(A_1), \cdots, m_k(A_{N_H})\right]^{\mathrm{T}}$ be an evidence vector in the space R^{N_H} for all $1 \leq k \leq n_E$. The similarity between evidence vectors of E_i and E_j is defined as:

$$S_{i,j} = \frac{E_i^{\mathrm{T}} E_j}{\left[\left(E_i^{\mathrm{T}} E_i\right)\left(E_j^{\mathrm{T}} E_j\right)\right]^{\frac{1}{2}}} \tag{10}$$

where $E_i^{\mathrm{T}} E_j = \sum_{l=1}^{N_H} m_i(l)\, m_j(l)$, $i, j = 1, 2, \cdots, n_E$, $l = 1, 2, \cdots n_H$.

(ii) The average similarity $\bar{S}_i$ between evidence piece E_i and other evidence pieces is measured by: $\bar{S}_i = \sum_{j=1, i\neq j}^{n_E} S_{i,j}/(n_E - 1)$.

(iii) Conflict factor Δ_i of evidence piece i is defined as: $\Delta_i = (\beta - \bar{S}_i)/(\beta - \alpha)$, where $\alpha = \min_{1\leq i\leq n_E} \bar{S}_i$, and $\beta = \max_{1\leq i\leq n_E} \bar{S}_i$. When $\Delta_i = 0$, it indicates that evidence piece E_i is identical with other evidence pieces and has no conflict with them. When $\Delta_i = 1$, it indicates that E_i is entirely different from other evidence pieces, and has the largest conflict with them. Therefore, based on the conflict factor, the evidence pieces can be divided into two categories: similar evidence pieces and conflicting evidence pieces. For evidence m_i, if the conflict factor $\Delta_i \leq \tau$, then evidence m_i is regarded as a similar evidence piece; if conflict factor $\Delta_i > \tau$, then evidence m_i is regarded as a conflicting evidence piece, where

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

$0 \leq \tau \leq 1$ is a threshold set, which is determined through a sequence of numerical experiments based on field data.

To overcome evidence conflict, we modify the original BPAs of the similar evidence pieces and conflicting evidence pieces by defining a conflict scale factor. The conflict scale factor η indicates the proportion of conflicting evidence pieces to all evidence pieces in the discernment frame Θ, which can be expressed as: $\eta = n_c/n_E$, where n_c is the number of conflicting evidence pieces. The BPA modification procedure is as follows:

(i) When $\eta < \delta$, a partial modification is made for only conflicting evidence pieces: $m_i'(A_j) = S_i^{(A)} \cdot \hat{m}_i(A_j)$, where $\hat{m}_i(A_j)$ represents the BPA of the $i$thconflicting evidence piece and $S_i^{(A)}$ is the absolute similarity of $i$thconflicting evidence piece formulated as: $S_i^{(A)} = \bar{S}_i/\beta$.

(ii) When $\eta \geq \delta$, a global modification is made for all evidence pieces: $m_i''(A_j) = S_i^{(R)} \cdot m_i(A_j)$, where $m_i(A_j)$ represents the BPA of the ith evidence piece and $S_i^{(R)}$ is the relative similarity of the ith evidence piece formulated as: $S_i^{(R)} = \bar{S}_i/\sum_{j=1}^{N_H} \bar{S}_j$. According to the analysis above, the original BPA is updated as $\tilde{m}_i(A_j)$:

$$\tilde{m}_i(A_j) = \begin{cases} m_i(A_j) & \Delta_i \leq \tau \\ m_i'(A_j) & \Delta_i > \tau, \eta < \delta \\ m_i''(A_j) & \Delta_i > \tau, \eta \geq \delta \end{cases} \tag{11}$$

Based on the updated BPAs, the improved Dempster evidence combination rule is given by:

$$m(A_k) = \begin{cases} \dfrac{1}{1-\tilde{K}} \displaystyle\sum_{\cap A_j = A_k} \prod_{1 \leq i \leq n_E} \tilde{m}_i(A_j), & \forall A_k \in \Theta, \ A_k \neq \emptyset, \\ 0, & A_k = \emptyset, \end{cases} \tag{12}$$

where $\tilde{K} = \sum_{\cap A_j = \emptyset} \prod_{1 \leq i \leq n_E} \tilde{m}_i(A_j)$.

3) Decision Making: After the evidence fusion, we formulate a decision rule to draw the final conclusion on the fatigue state. Let $m(A_F) = \max\{\tilde{m}(A_k), A_k \in \Theta\}$, $m(A_S) = \max\{\tilde{m}(A_k) | A_k \in \Theta \backslash A_F\}$ represent the largest and the second largest probability values, respectively. The proposed decision is determined as follows:

$$\begin{cases} m(A_F) - m(A_S) > \varepsilon_{T1} \\ m(A_F) > \varepsilon_{T2} \end{cases} \tag{13}$$

where ε_{T1} and ε_{T2} are pre-specified thresholds [39]. If the fusion result satisfies the decision rule (13), then A_F is adopted as the final fatigue driving state. If the fusion result at current time step does not satisfy decision rule (13), the driver's fatigue state at the previous time step is adopted to ensure the program reliability considering that driver's fatigue state would not change over a relatively short time.

V. Field Experiments

Field experiments were conducted on the Nanjing-Zhenjiang expressway in China to investigate the performance of the proposed fatigue driving recognition model.

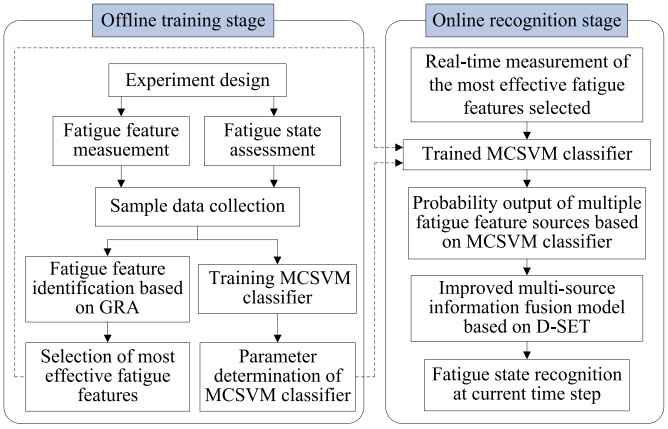


Fig. 3. Experimental stages of fatigue driving recognition.

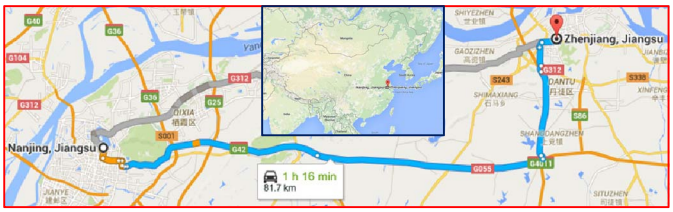


Fig. 4. Experiment route shown on Google map.

A. Two Experiment Stages

These experiments are divided into two stages: the offline training stage and the online recognition stage, as shown in Fig. 3. The offline training stage first constructs an experiment platform and designs a detailed experimental procedure. Second, it conducts fatigue feature measurements and classifies the fatigue state of each fatigue feature. Third, based on the collected sample data, including fatigue features and corresponding fatigue states, the offline training stage identifies the effectiveness of these fatigue features using the GRA method and selects the most effective fatigue features. Fourth, based on the collected sample data, the offline training stage trains the MCSVM classifier and obtains the optimal parameter values.

The online recognition stage first performs real-time measurement of the most effective fatigue features selected in the offline training stage. Second, as input parameters, these most effective fatigue features are fused by the trained MCSVM classifiers to generate real-time probability outputs for further decision-level fusion. Third, the online recognition model fuses the four pieces of evidence (see Fig. 1) using the improved D-SET fusion method. Fourth, the driver's fatigue state is determined using the set decision rule based on the final fusion result, and the effectiveness of the proposed method is validated.

B. Experiment Design

Fig. 4 highlights the experiment location, the Nanjing-Zhenjiang expressway, which is 81.7 km long. Six men and four women participated in the experiments. Their ages ranged from 24 to 34 years, and they had more than three years of driving experience each. Alcohol, tea, coffee, drugs, or any

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

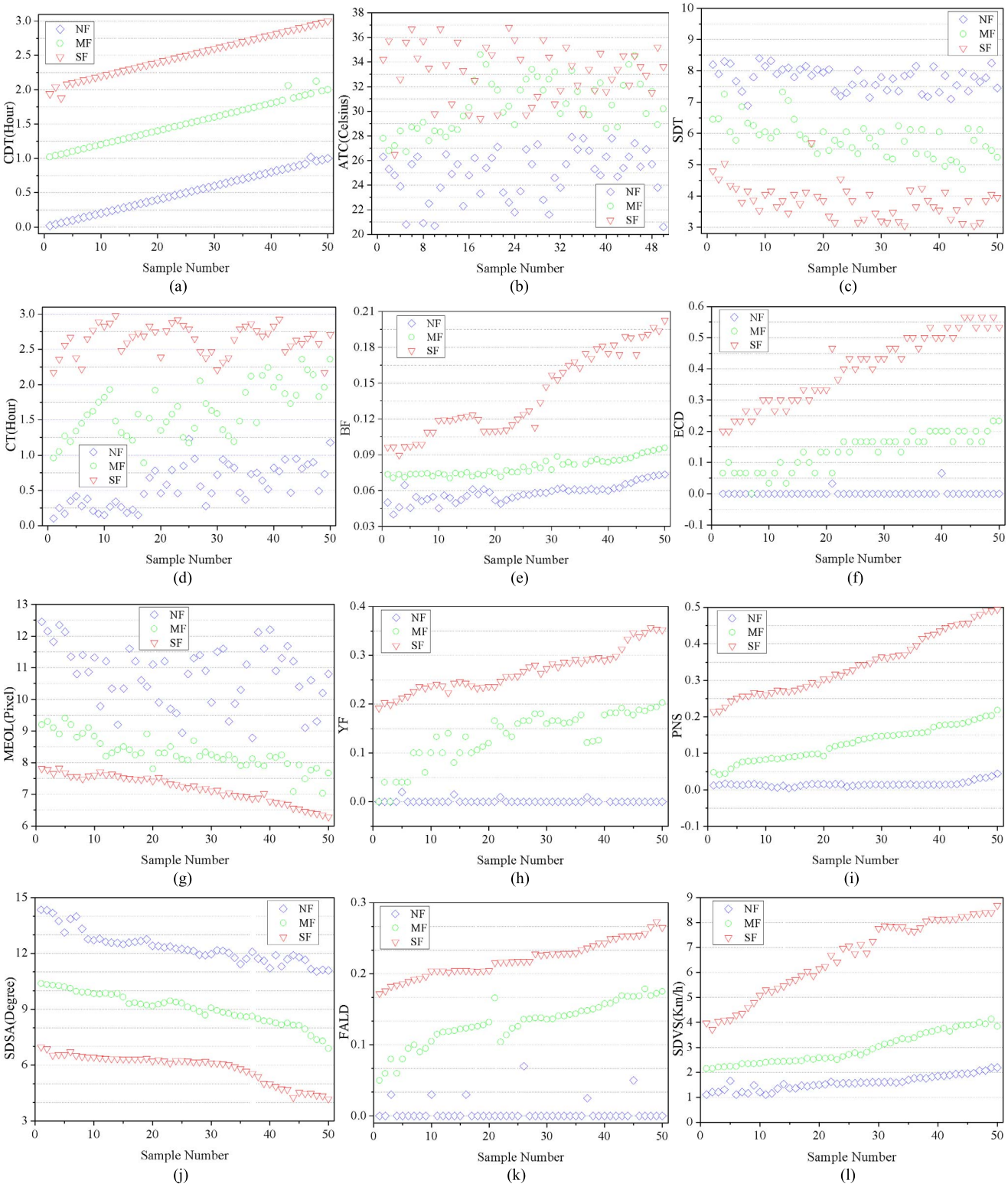


Fig. 5. Fatigue feature measurements: (a) CDT. (b) ATC. (c) SDT. (d) CT. (e) BF. (f) ECD. (g) MEOL. (h) YF. (i) PNS. (j) SDSA. (k) FALD. (l) SDVS.

drinks that could cause excitement to the nervous system were prohibited for 24 hours before the experiments. All experiments were performed after informed consent of the procedures was received from all participants.

The experiments were conducted from 12:00 PM to 3:00 PM on March 15th, 2015 for all participants. Previous studies suggest that drivers are more easily fatigued during this time period of a day [3]. To avoid traffic accidents, the experiments were conducted on road sections with few vehicles. In addition, to ensure safety, an experienced driver assistant was asked to sit in the front passenger seat to warn the participant or execute the brake in an emergency.

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

C. Fatigue State Classification

Based on the comprehensive assessment method of fatigue state in Section III.A, we propose an accurate classification of three driver fatigue states: Non-fatigue (NF), Moderate fatigue (MF) and Severe fatigue (SF), respectively. For the detailed steps, please refer to our previous study [23].

Note that in the self-assessment step, fatigue is rated into three states, NF (1-4 points), MF (5-7 points), and SF (8-9 points), based on the scores from the KSS table.

D. Field Data Collection

Based on the measurements described in Section II, twelve fatigue features were measured. 1,000 data points were collected, each of which included the twelve fatigue features. They were separated into two sets. The training set included 600 data points for model calibration, and the testing set included 400 data points for model verification. Further, based on the proposed comprehensive assessment method, the measurements of every fatigue feature were divided into three groups according to their corresponding fatigue states (NF, MF, and SF). We select 150 data from the 1,000 data illustrated in Fig. 5. For each fatigue feature, 50 data belong to the NF group, 50 data belong to the MF group, and the rest belong to the SF group.

As shown by Fig. 5, for some fatigue features, it is difficult to classify driver's fatigue state according to the distribution of the measured value. For example, the data for the NF state should have larger values than the data for the MF state, and data for the MF state should be higher than those for the SF state for ATC, MEOL and PNS, as shown in Figs. 5(b), (g), and (i), respectively. However, some data for the NF state have smaller values than those for the MF state. Therefore, the effectiveness of these measured fatigue features need to be verified for a reliable fatigue driving recognition.

E. Fatigue Feature Selection

This step identifies the fatigue features, and selects the most effective fatigue features. First, the GRA is used to calculate the GRD between every fatigue feature and fatigue driving state. Then, according to the GRD rank, the fatigue features with a larger GRD value are selected as the most effective fatigue features. For determining the GRD, we identify the fatigue level for the collected sample data as follows: (i) if the fatigue state of the sample is NF, the fatigue level is quantified as 1; (ii) if the fatigue state of the sample is MF, the fatigue level is quantified as 2; (iii) if the fatigue state of the sample is SF, the fatigue level is quantified as 3. We randomly select 800 sample data from the 1,000 collected data points to perform the correlation analysis. The GRD values of the twelve fatigue features are shown in Fig. 6.

As shown in Fig. 6, nine fatigue features (SDT, BF, CDT, ECD, SDSA, YF, FALD, SDVS, and CT) have GRD values over 0.80, and the other three (PNS, MEOL, and ATC) have values less than 0.80. This indicates that these nine fatigue features have more significant correlation with fatigue driving, and are selected as the most effective fatigue features. The other three are excluded. In future research, we can further analyze whether inter-correlation exists among the most effective fatigue features selected in the paper to identify more effective fatigue features.

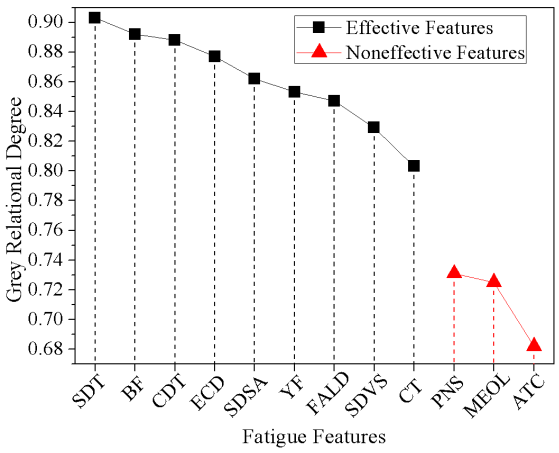


Fig. 6. GRD between fatigue features and fatigue driving.

F. Model Verification

1) Feature-Level Fusion Result: The feature-level fusion fuses the most effective fatigue features and generates dynamic BPAs for the decision-level fusion. From the fatigue feature identification results, three contextual features (SDT, CDT and CT), three facial features (BF, ECD and YF) and three vehicle behavior features (SDSA, FALD and SDVS) are fused as input parameters for MCSVM-1, MCSVM-2 and MCSVM-3 to provide BPAs for the decision-level fusion, respectively. We select 400 sample data from the training set to implement the feature-level fusion, and use six of the fusion results as examples in Fig. 7. In the feature-level fusion, the fusion result of every MCSVM classifier is a vector with three elements. Let $\{\bar{p}_i^t(A_1), \bar{p}_i^t(A_2), \bar{p}_i^t(A_3)\}$ $i = 1, 2, 3$ denote the fusion result of the ith MCSVM classifier at time t. The three elements, $\bar{p}_i^t(A_1)$, $\bar{p}_i^t(A_2)$, and $\bar{p}_i^t(A_3)$, denote the probabilities of the sample belonging to the three fatigue states, NF, MF and SF, respectively. In Fig. 7, they are denoted by three vertical bars.

2) Decision-Level Fusion Result: Based on the improved evidence fusion method, the four pieces of evidence, including the three fatigue features and the fusion result for the previous time step, are fused at the decision-level fusion. The fusion result for the previous time step is shown in Fig. 8(a), and the fusion result for the current time step is shown in Fig. 8(b).

The decision-level fusion result is also a vector with three elements. Let $\{m^t(A_1), m^t(A_2), m^t(A_3)\}$ denote the fusion result at time t. The three elements, $m^t(A_1)$, $m^t(A_2)$, and $m^t(A_3)$, denote the probabilities of the sample belonging to the three fatigue states, NF, MF and SF, respectively.

Fig. 9 illustrates the difference between the maximum and the second maximum of the three vector elements in the feature-level fusion and the decision-level fusion. The difference values are much smaller in the feature-level fusion than in the decision-level fusion. For example, for the 5$^{\text{th}}$ sample, in the MCSVM-1 classifier, $\bar{p}_1^t(A_1) = 0.483$, $\bar{p}_1^t(A_2) = 0.315$, $\bar{p}_1^t(A_3) = 0.202$; the difference between the maximum (0.483)

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

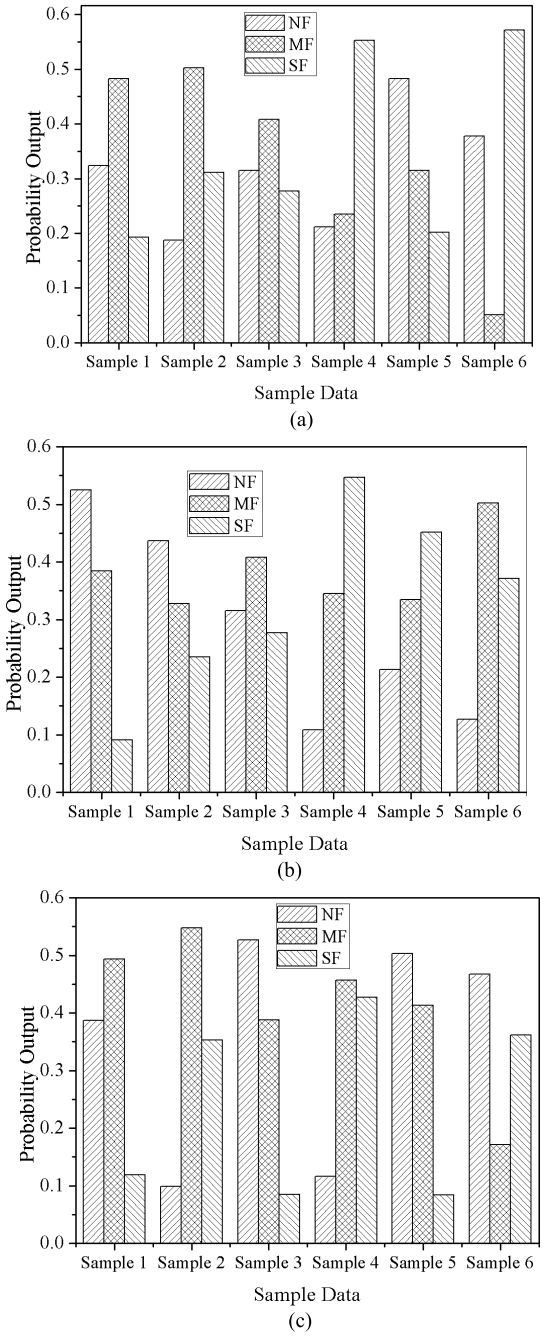


Fig. 7. Feature-level fusion results for: (a) Contextual features, (b) Facial features, and (c) Vehicle behavior features.

and the second maximum (0.315) values is 0.168. Similarly, for the MCSVM-2 classifier, the difference between the maximum 0.452 ($\bar{p}_2^t(A_3)$) and the second maximum 0.335 ($\bar{p}_2^t(A_2)$) is 0.117, and for the MCSVM-3 classifier, the difference between the maximum 0.503 ($\bar{p}_3^t(A_1)$) and the second maximum 0.413 ($\bar{p}_3^t(A_2)$) is 0.090. However, after the decision-level fusion, the difference value between the maximum 0.603 ($m^t(A_1)$) and the second maximum 0.294 ($m^t(A_2)$) is 0.309, which is much larger than those based on the feature-level fusion. Hence, this enhancement based on the two-level fusion process makes the recognition of fatigue state easier and more credible.

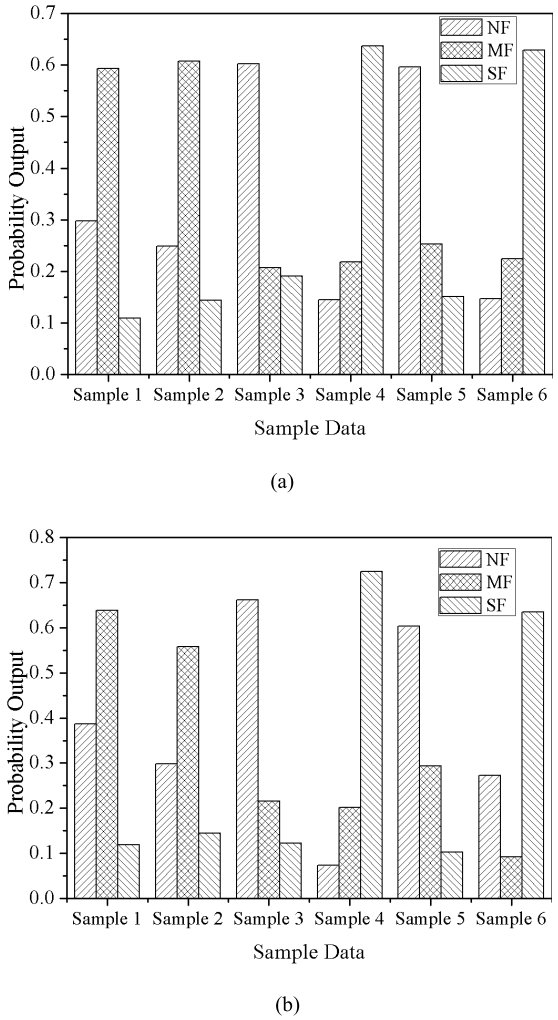


Fig. 8. Decision-level fusion results for: (a) Previous time step, and (b) Current time step.

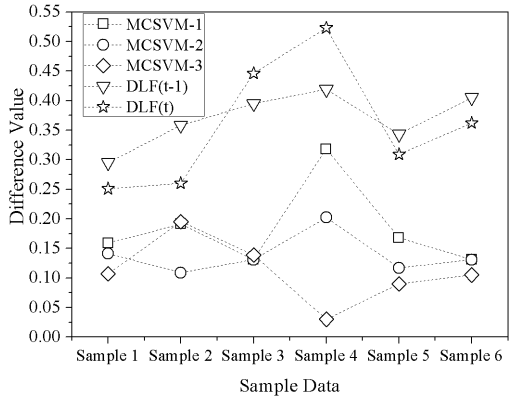


Fig. 9. The difference between the maximum and the second maximum of three vector elements in feature-level fusion and in decision-level fusion.

The recognition results based on the proposed comprehensive assessment method are viewed as ground truth to verify the effectiveness of the proposed method. By analyzing the recognition results based on the training dataset and ground truth data, we specify $\varepsilon_{T1} = 0.25$ and $\varepsilon_{T2} = 0.55$ in (13) to conclude the fatigue state. We find that the proposed

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

recognition method is capable of accommodating external condition changes, especially under the cases when driver nods sharply, lane becomes blurred, lighting becomes dim, or GPS is incapable of receiving location signals. For example, for the 4th sample data, $\bar{p}_1^t(A_3)$ has the maximum value (0.553) in the output of MCSVM-1 classifier, $\bar{p}_2^t(A_3)$ has the maximum value (0.452) for MCSVM-2 classifier, $\bar{p}_3^t(A_2)$ has the maximum value (0.457) for MCSVM-3 classifier, and $m^t(A_3)$ has the maximum value (0.725) in the decision-level fusion at time t. Therefore, from the individual results of the MCSVM-1 and MCSVM-2 classifiers and the decision-level fusion at time t, the driver's fatigue state should be identified as SF. Conversely, based on the fusion result of MCSVM-3, the driver's fatigue state should be identified as MF, which is contradictory. Based on the proposed comprehensive assessment method, we find that the driver's actual fatigue state is evaluated as SF, which is consistent with the recognition result of the decision-level fusion. An analysis of the collected data suggests that the MCSVM-2 classifier was inconsistent because the fatigue features ECD and YF were incorrectly measured due to the driver's suddenly nodding.

Similarly, for the 4th sample data, the recognition results of MCSVM-1, MCSVM-2 and decision-level fusion indicate SF, while that of MCSVM-3 classifier is MF. The video records show that the fatigue feature FALD was incorrectly measured due to the blurred lane and dim lighting, which resulted in the failure of MCSVM-3 classifier. By contrast, the recognition result of fatigue state based on decision-level fusion at current time step is consistent with the result based on the proposed comprehensive assessment. Therefore, although certain fatigue feature measurements are problematic, the proposed fusion model can robustly recognize driver's fatigue state.

For the 6th sample data, the driver's fatigue state is identified as SF, MF, and NF according to the fusion results based on the MCSVM-1, MCSVM-2 and MCSVM-2 classifiers, respectively. However, the fatigue state of the 6th sample data is SF according to the decision-level fusion result shown in Fig. 8(b). The comprehensive assessment verifies that the actual fatigue state is SF, which is consistent with the decision-level fusion result. An analysis of the collected data illustrated that the GPS device did not receive any signal of vehicle position and velocity, and the camera did not capture legible face image of driver because the vehicle had entered a tunnel, which resulted in the recognition failure of the MCSVM-2 and MCSVM-3 classifiers. By contrast, the decision-level fusion was able to obtain correct fatigue recognition by considering the fusion result at the previous time step. This indicates that even when two fatigue feature measurements from different feature sources fail simultaneously due to sudden external travel environment changes, the proposed method may be able to correctly recognize driver's fatigue state by factoring the fusion result from the previous time step.

The performance of the proposed two-level fusion model is investigated based on four sets of fatigue feature measurements from five perspectives: number of incorporated fatigue features (NIFF), accuracy rate (AR), missing rate (MR), false alarm rate (FAR), and average computational time (ACT). Among these, AR, FAR and MR can reflect the reliability of the proposed method and ACT can reflect its efficiency. Here, AR$= (N_{0,0} + N_{1,1} + N_{2,2})/N$, FAR $= (N_{0,1} + N_{0,2})/N$, MR $= (N_{1,0}+N_{2,0})/N$, where N is the number of samples, and $N_{i,j}$ is the number of sample data points recognized as having fatigue state j when the actual fatigue state is i. The four sets of fatigue features are characterized as follows: (i) all 12 fatigue features, (ii) the 9 most effective fatigue features (excluding PNS, MEOL and ATC), (iii) only 8 fatigue features (by excluding contextual features from (ii)), and (iv) the 6 most effective fatigue features (by excluding contextual features, PNS and MEOL from (ii)). The results are summarized in Table I, where CFs indicates contextual features.

TABLE I

PERFORMANCE COMPARISON BASED ON DIFFERENT FATIGUE FEATURE SETS

Fatigue feature selection		NIFF	AR (%)	MR (%)	FAR (%)	ACT (ms)
Including CFs	Without GRA	12	92.8	2.5	2.6	348
	With GRA	9	95.1	2.1	2.2	287
Excluding CFs	Without GRA	8	92.7	2.6	2.5	321
	With GRA	6	94.0	2.2	2.3	259

TABLE II

PERFORMANCE COMPARISON BASED ON DIFFERENT EVIDENCE SOURCES

Evidence source selection	NIFF	NIES	AR (%)	MR (%)	FAR (%)	ACT (ms)
CFs+FFs	6	2	92.7	3.2	3.3	242
CFs+VBFs	6	2	92.6	3.5	3.2	187
FFs+VBFs	6	2	93.1	3.2	3.2	251
CFs+FFs +VBFs	9	3	93.5	3.0	3.1	278

Table I shows that the proposed method performs better not only in terms of reliability but also in terms of efficiency when the contextual features are included and ineffective fatigue features are excluded by performing fatigue feature selection. In particular, while fewer fatigue features are used in the fatigue recognition model, the performance of the proposed method is improved in terms of AR, MR, FAR and ACT. This demonstrates that implementing fatigue feature selection is necessary and effective for real-time fatigue driving recognition. Further, incorporating the contextual features can also enhance the reliability of the recognition method.

The performance of the proposed method based on different evidence sources is summarized in Table II, where NIES denotes the number of incorporated evidence sources, and FFs and VBFs indicate facial features and vehicle behavior features, respectively. Note that, unlike in Table I, the used fatigue features in the feature-level fusion are based only on the most effective fatigue features through feature selection in Table II, and the fatigue state at the previous time step is not considered in the decision-level fusion. By comparing tables I and II, we note that the proposed method obtains better recognition results when the fatigue state at the previous time step is considered in the decision-level fusion.

The performance of a fatigue recognition method applying single-level fusion to different evidence sources is summarized

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

TABLE III

PERFORMANCE COMPARISON OF THE MCSVM CLASSIFIER BASED ON DIFFERENT EVIDENCE SOURCES

Evidence source selection	NIFF	NIES	AR (%)	MR (%)	FAR (%)	ACT (ms)
CFs+FFs	6	2	92.1	3.6	3.4	232
CFs+VBFs	6	2	92.3	3.7	3.5	181
FFs+VBFs	6	2	92.5	3.4	3.4	248
CFs+FFs +VBFs	9	3	91.8	3.2	3.3	276

TABLE IV

PERFORMANCE COMPARISON OF SINGLE FEATURE AND SINGLE-SOURCE FUSION BASED METHODS

Evidence source selection	NIFF	AR (%)	MR (%)	FAR (%)	ACT (ms)
Single feature based (CDT)	1	91.8	3.7	3.8	2
Single feature based (YF)	1	85.9	4.3	4.1	53
Single feature based (SDSA)	1	85.4	4.4	4.0	11
Single-source fusion based (CFs and MCSVM)	3	90.6	3.8	3.7	28
Single-source fusion based (FFs and MCSVM)	3	91.9	3.3	3.6	95
Single-source fusion based (VBFs and MCSVM)	3	91.2	3.4	3.9	44

in Table III. In Table III, all features from different evidence sources are incorporated into the MCSVM classifier, and driver's fatigue state recognition results are outputted through it. A comparison of tables II and III illustrates that the recognition method based on two-level fusion outperforms the one based on single-level fusion in terms of accuracy and reliability.

Table IV compares the proposed method with those based on single feature and single-source fusion in terms of NIFF, AR, MR, FAR and ACT. Based on tables I and IV, the proposed method, which incorporates the most effective fatigue features from four feature sources, outperforms the other methods in terms of AR, MR and FAR. It indicates that the proposed method can provide more reliable and robust recognition results in real-world applications.

In addition, we also investigate the performace of the method that incorporates the contextual features and physiological feature into the MCSVM classifier. Its AR, MR and FAR are 94.0%, 2.9%, and 3.1%, respectively, where the contextual features include CDT, SDT, and CT, and physiological feature is measured by $r_{\alpha,\theta,\beta}$. The α, θ and β represent three wave bands of EEG signal, respectively. For the calculation of the $r_{\alpha,\theta,\beta}$, please see [23].

Finally, we compare the proposed method with the one developed in [23]. From Table I, the AR, MR and FAR of the proposed method are 95.1%, 2.1% and 2.2%, respectively. By contrast, the AR, MR and FAR from the method developed in [23] are 93.8%, 2.3% and 2.8%, respectively. This suggests that the proposed method has better reliability. In addition, from Table IV, the recognition results based on FFs and MCSVM classifier are 91.9%, 3.3% and 3.6% in terms of AR, MR and FAR, respectively, and those based on VBFs and MCSVM classifier are 91.2%, 3.4% and 3.9%, respectively. However, the recognition results based on the FFs and T-SFNN model in [23] are 91.6%, 3.4% and 3.7% in terms of AR, MR and FAR, respectively, and those of the VBFs and T-SFNN model are 90.8%, 3.6% and 4.1%, respectively. These comparisons indicate that the proposed MCSVM classifier outperforms the T-SFNN model in the feature-level fusion applied in [23].

TABLE V

LIST OF ACRONYMS

Acronym	Meaning	Acronym	Meaning
ACT	Average computational time	MEOL	Mean of eye-opened level
AR	Accuracy rate	MF	Moderate fatigue
ATC	Air temperature in cab	MR	Missing rate
BF	Blinking frequency	NF	Non-fatigue
BPA	Basic probability assignment	NIES	Number of incorporated evidence sources
CDT	Continuous driving time	NIFF	Number of incorporated fatigue features
CT	Current time	PERCLOS	Percentage of eye closure
D-SET	Dempster-Shafer evidence theory	PNS	Percentage of non-steering
ECD	Eye-closed duration	RBF	Radical basis function
EEG	Electroencephalography	SDSA	Standard deviation of steering-angle
FAR	False alarm rate	SDT	Sleep duration time
GRA	Grey relational analysis	SF	Severe fatigue
GRD	Grey relational degree	SVM	Support vector machine
KSS	Karolinska sleepiness scale	T-SFNN	Takagi-Sugeno fuzzy neural network
MCSVM	Multi-class support vector machine	YF	Yawning frequency

The analyses heretofore indicate that the proposed fatigue driving recognition method incorporating contextual feature with multi-feature identification and two-level fusion is more reliable, even under ineffective fatigue features or sensor failure.

VI. CONCLUDING COMMENTS

Based on the need to recognize fatigue driving of drivers reliably and robustly, this study proposes a novel fatigue driving recognition method incorporating contextual features with multi-feature identification and two-level fusion. It has the following characteristics. First, contextual features related to fatigue driving are shown to improve fatigue driving recognition. Second, the proposed GRA-based fatigue feature selection method can efficiently identify the most effective fatigue features, which can enhance the efficiency and reliability of the recognition model. Third, a two-level fusion model consisting of feature-level fusion and decision-level fusion is developed. In the feature-level fusion, the most effective fatigue features are fused based on the proposed MCSVM classifier, which enables a dynamic assignment of BPA for each fatigue feature source. In the decision-level fusion based on the D-SET, the evidence conflict among multiple pieces of evidence is resolved during evidence combination and the reliability of the recognition model is enhanced by modifying the BPA and combining the fatigue state identified at the previous time step.

The study offers the possibility of developing more sophisticated fatigue driving recognition methods. First, the proposed method can be extended by identifying and selecting other fatigue features that can reflect the fatigue driving state more

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.

accurately. Second, more evidence sources can be considered and combined through the improved decision-level fusion rule. Finally, more accurate feature-level fusion algorithms can be integrated into the two-level fusion model.

APPENDIX

See Table V.

REFERENCES

[1] S. Kaplan, M. A. Guvensan, A. G. Yavuz, and Y. Karalurt, "Driver behavior analysis for safe driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3017–3032, Dec. 2015.

[2] C. Zhang, H. Wang, and R. Fu, "Automated detection of driver fatigue based on entropy and complexity measures," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 168–177, Feb. 2014.

[3] R. Fu and H. Wang, "Detection of driving fatigue by using noncontact EMG and ECG signals measurement system," *Int. J. Neural. Syst.*, vol. 24, no. 3, p. 1450006, May 2014.

[4] J. Wang, Y. Wu, H. Qu, and G. Xu, "EEG-based fatigue driving detection using correlation dimension," *J. Vibroeng.*, vol. 16, no. 1, pp. 407–413, Feb. 2014.

[5] Q. Ji, Z. Zhu, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1052–1068, Jul. 2004.

[6] T. D'Orazio, M. Leo, C. Guaragnella, and A. Distante, "A visual approach for driver inattention detection," *Pattern Recognit.*, vol. 40, no. 8, pp. 2341–2355, Aug. 2007.

[7] J. Jo, S. J. Lee, K. R. Park, I.-J. Kim, and J. Kim, "Detecting driver drowsiness using feature-level fusion and user-specific classification," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1139–1152, Mar. 2014.

[8] T. Azim, M. A. Jaffar, and A. M. Mirza, "Fully automated real time fatigue detection of drivers through fuzzy expert systems," *Appl. Soft Comput.*, vol. 18, pp. 25–38, May 2014.

[9] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-time system for monitoring driver vigilance," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 63–77, Mar. 2006.

[10] B. Cyganek and S. Gruszczyński, "Hybrid computer vision system for drivers' eye recognition and fatigue monitoring," *Neurocomputing*, vol. 126, pp. 78–94, Feb. 2014.

[11] T. H. Chang, C. S. Hsu, C. Wang, and L. K. Yang, "Onboard measurement and warning module for irregular vehicle behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 3, pp. 501–513, Sep. 2008.

[12] C.-F. Wu, C.-J. Lin, and C.-Y. Lee, "Applying a functional neurofuzzy network to real-time lane detection and front-vehicle distance measurement," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 4, pp. 577–589, Jul. 2012.

[13] A. D. McDonald, J. D. Lee, C. Schwarz, and T. L. Brown, "Steering in a random forest: Ensemble learning for detecting drowsiness-related lane departures," *Human Factors*, vol. 56, no. 5, pp. 986–998, Aug. 2014.

[14] R. Sayed and A. Eskandarian, "Unobtrusive drowsiness detection by neural network learning of driver steering," *Proc. Inst. Mech. Eng. D, J. Automob. Eng*, vol. 215, no. 9, pp. 969–975, Sep. 2001.

[15] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 586–614, Jun. 2011.

[16] H. Zhang, X. Yan, C. Wu, and T. Qiu, "Effect of circadian rhythms and driving duration on fatigue level and driving performance of professional drivers," *Transp. Res. Rec., J. Transp. Res. Board*, no. 2402, pp. 19–27, Jul. 2014.

[17] L. Wang and Y. Pei, "The impact of continuous driving time and rest time on commercial drivers' driving performance and recovery," *J. Safety Res.*, vol. 50, pp. 11–15, Sep. 2014.

[18] P. Boyraz, M. Acar, and D. Kerr, "Multi-sensor driver drowsiness monitoring," *Proc. Inst. Mech. Eng. D, J. Automob. Eng.*, vol. 222, no. 11, pp. 2041–2062, Nov. 2008.

[19] S. Li, L. Wang, Z. Yang, B. Ji, and F. Qiao, "A driver fatigue recognition model using multiple physiological features based on support vector machine," *Information*, vol. 15, no. 12, pp. 5321–5328, Dec. 2012.

[20] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model based on information fusion and dynamic Bayesian network," *Inf. Sci.*, vol. 180, no. 10, pp. 1942–1954, May 2010.

[21] B. Cheng, W. Zhang, Y. Lin, R. Feng, and X. Zhang, "Driver drowsiness detection based on multisource information," *Human Factors Ergonom. Manuf. Service Ind.*, vol. 22, no. 5, pp. 450–467, Oct. 2012.

[22] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model using fusion of multiple features," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Waikoloa, HI, USA, Oct. 2005, pp. 1777–1784.

[23] W. Sun, X. Zhang, S. Peeta, X. He, Y. Li, and S. Zhu, "A self-adaptive dynamic recognition model for fatigue driving based on multi-source information and two levels of fusion," *Sensors*, vol. 15, no. 9, pp. 24191–24213, Sep. 2015.

[24] B. Mandal, L. Li, G. S. Wang, and J. Lin, "Towards detection of bus driver fatigue based on robust visual analysis of eye state," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 3, pp. 545–557, Mar. 2017.

[25] T. Pradhan, A. N. Bagaria, and A. Routray, "Measurement of PERCLOS using eigen-eyes," in *Proc. 4th Int. Conf. Intell. Human Comput. Interac. (IHCI)*, Kharagpur, India, Dec. 2012, pp. 1–4.

[26] T. Von Jan, T. Karnahl, K. Seifert, J. Hilgenstock, and R. Zobel, "Don't sleep and drive-VW's fatigue detection technology," in *Proc. 19th Int. Conf. Enhanced Safety Veh.*, Washington, DC, USA, Jun. 2005, pp. 1–12.

[27] A. Williamson and R. Friswell, "Investigating the relative effects of sleep deprivation and time of day on fatigue and performance," *Accident Anal. Prevention*, vol. 43, no. 3, pp. 690–697, May 2011.

[28] S. Kar, M. Bhagat, and A. Routray, "EEG signal analysis for the assessment and quantification of driver's fatigue," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 13, no. 5, pp. 297–306, Sep. 2010.

[29] T. Akerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual," *Int. J. Neurosci*, vol. 52, nos. 1–2, pp. 29–37, Jul. 1990.

[30] J. Krajewski, S. Schnieder, D. Sommer, A. Batliner, and B. Schuller, "Applying multiple classifiers and non-linear dynamics features for detecting sleepiness from speech," *Neurocomputing*, vol. 84, pp. 65–75, May 2012.

[31] W. Sun, X. Zhang, J. Wang, J. He, and S. Peeta, "Blink number forecasting based on improved Bayesian fusion algorithm for fatigue driving detection," *Math. Problems Eng.*, vol. 2015, Jun. 2015, Art. no. 832621.

[32] F. Xie, Y. Peng, H. Yang, and H. Gao, "Data fusion detection model based on SVM and evidence theory," in *Proc. IEEE 14th Int. Conf. Commun. Technol.*, Chengdu, China, Nov. 2012, pp. 814–818.

[33] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Mar. 2002.

[34] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in Large Margin Classifiers*, A. Smola, P. Bartlett, B. Schölkopf, D. Schuurmans eds., Cambridge, MA, USA: MIT Press, Jun. 2000.

[35] L. Li, Z.-P. Gao, and W.-Y. Ding, "Fuzzy multi-class support vector machine based on binary tree in network intrusion detection," in *Proc. Int. Conf. Elect. Control Eng. (ICECE)*, Wuhan, China, Jun. 2010, pp. 1043–1046.

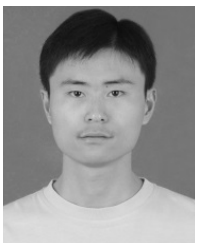
[36] H.-T. Lin, C.-J. Lin, and R. C. Weng, "A note on Platt's probabilistic outputs for support vector machines," *Mach. Learn.*, vol. 68, no. 3, pp. 267–276, 2007.

[37] J.-H. Hong and S.-B. Cho, "A probabilistic multi-class strategy of one-vs.-rest support vector machines for cancer classification," *Neurocomputing*, vol. 71, nos. 16–18, pp. 3275–3281, Oct. 2008.

[38] L. Ai, J. Wang, and X. Wang, "Multi-features fusion diagnosis of tremor based on artificial neural network and D–S evidence theory," *Signal Process.*, vol. 88, no. 12, pp. 2927–2935, Dec. 2008.

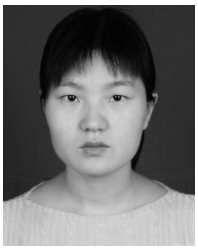
[39] L. Si, Z. Wang, C. Tan, and X. Liu, "A novel approach for coal seam terrain prediction through information fusion of improved D–S evidence theory and neural network," *Measurement*, vol. 54, pp. 140–151, Aug. 2014.

[40] E. Lefevre, O. Colot, and P. Vannoorenberghe, "Belief function combination and conflict management," *Inf. Fusion*, vol. 3, no. 2, pp. 149–162, Jun. 2002.



Wei Sun received the B.S. and M.S. degrees in mechanical manufacture and automation from Henan University of Science and Technology, China, in 2004 and 2006, respectively, and the Ph.D. degree in instrument science and technology from Southeast University, China, in 2010. From 2014 to 2015, he was a Post-Doctoral Researcher with the NEXTRANS Center, Purdue University, USA. He is currently an Associate Professor of Information and Control with Nanjing University of Information Science and Technology. His research interests include driver behavior inference and prediction, computer vision, pattern recognition, and active safety detection of vehicle.

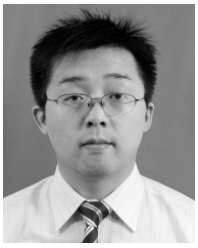
Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.



Xiaorui Zhang received the B.S. and M.S. degrees in mechanical manufacture and automation from Henan University of Science and Technology, China, in 2004 and 2006, respectively, and the Ph.D. degree in instrument science and technology from Southeast University, China, in 2010. From 2013 to 2014, she was a Post-Doc Researcher with the ViDi Center, University of Pennsylvania, Philadelphia, PA, USA. She is currently an Associate Professor of Computer Science and Technology with Nanjing University of Information Science and Technology. Her research interests include virtual reality and human–computer interaction, haptic perception, and pattern recognition.

Srinivas Peeta received the B.S. degree from IIT Madras in 1988; the M.S. degree from California Institute of Technology, Pasadena, CA, USA, in 1989; and the Ph.D. degree in civil engineering from The University of Texas at Austin, Austin, TX, USA, in 1994. He is currently the Jack and Kay Hockema Professor of Civil Engineering and the Director of the NEXTRANS Center with the USDOT Region 5 University Transportation Center, Purdue University, West Lafayette, IN, USA. His research interests broadly span transportation and infrastructure systems. He has served as a Chair of the Transportation Network Modeling Committee of the Transportation Research Board of the National Academies. He serves on the Editorial Advisory Boards of *Transportation Research Part B*, *Intelligent Transportation Systems Journal*, and *Transportmetrica B*. He serves as an Area Editor of *Transport Dynamics for Networks and Spatial Economics*.

Xiaozheng He received the B.S. and M.S. degrees in mathematics from Nanjing University, Nanjing, China, in 2000 and 2003, respectively, and the M.S. and Ph.D. degrees in civil engineering from the University of Minnesota Twin Cities, Minneapolis, MN, USA, in 2007 and 2010, respectively. He is currently an Assistant Professor with the Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. He was a Post-Doctoral Research Associate with the NEXTRANS Center, Purdue University, West Lafayette, IN, USA. His research interests include equilibration modeling, information network analysis, and infrastructure network reliability.



Yongfu Li received the Ph.D. degree in control theory and control engineering from Chongqing University, Chongqing, China, in 2012. Since 2014, he has been a Post-Doctoral Researcher with the NEXTRANS Center, Purdue University, West Lafayette, IN, USA. He is currently an Associate Professor of Control Science and Engineering with the Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include intelligent transportation systems, automotive electronics, and control theory. He has served as reviewer for over 20 international journals or conferences, such as IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEM, *Nonlinear Dynamics*, *Journal of Advanced Transportation*, *International Journal of Dynamics and Control*, the *International Journal of Communication Systems*, IFAC World Congress, and CCC.

Authorized licensed use limited to: DALIAN MARITIME UNIVERSITY. Downloaded on May 29,2021 at 02:35:34 UTC from IEEE Xplore. Restrictions apply.