Hybrid Data-driven Vigilance Model in Traffic Control Center using Eye-tracking Data and Context Data

Fan Li [1,2*], Lee Ching-Hung [3], Chun-Hsien Chen[1], Li Pheng Khoo[1]

[1]*School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore, 639798*

[2]*Maritime Institute, Nanyang Technological University, Singapore*

[3]*School of Public Policy and Administration, Xi'an Jiaotong University, China*

Abstract:

Vigilance decrement of traffic controllers would greatly threaten public safety. Hence, extensive studies have been conducted to establish the physiological data-based vigilance model for objectively monitoring or detecting vigilance decrement. Nevertheless, most of them using intrusive devices to collect physiological data and failed to consider context information. Consequently, these models can be used in a laboratory environment while cannot adapt to dynamic working conditions of traffic controllers. The goal of this research is to develop an adaptive vigilance model for monitoring vigilance objectively and non-intrusively. In recent years, with advanced information and communication technology, a massive amount of data can be collected from connected daily use items. Hence, we proposed a hybrid data-driven approach based on connected objects for establishing vigilance model in the traffic control center and provide an elaborated case study to illustrate the method. Specifically, eye movements are selected as the primary inputs of the proposed vigilance model; Bagged trees technique is adapted to generate the vigilance model. The results of case study indicated that (1) eye metrics would be correlated with the vigilance performance subjected to the mental fatigue levels, (2) the bagged trees with the fusion features as inputs achieved a relatively stable performance under the condition of data loss, (3) the proposed method could achieve better performance than the other classic machine learning methods.

Keywords: internet of things, traffic control center, vigilance detection, data-driven, eye movements

1. Introduction:

Traffic control centers provide navigation assistance, advice, and information services to traffic operators, viz. pilots, drivers, and masters. One of the main tasks of traffic controllers is monitoring traffic condition and identifying the potential unsafe issues. It has been found that human's cognitive energy depleted fast when they were working on the prolonged, monotonous and boring monitoring tasks [1]. Moreover, early researchers had also pointed out that monitoring misses occurred frequently due to momentary distractions [2], which is common in traffic control operations. For example, air traffic controllers have to answer calls from pilots during monitoring; Vessel traffic controllers have to check vessel information lists periodically. Owing to these factors, traffic controllers always suffer from vigilance decrement [3], which generally refers that operator performance tends to degrade over time [4]. In general, vigilance decrement can be assessed by counting the number of critical signals missed and the time lapsed in reacting to critical signals. It greatly threats public safety. Statistics of Europe showed that more than 10% of traffic accidents were caused by vigilance decrement [5].

To combat this issue, both proactive strategy and reactive strategy had been proposed [6, 7]. The proactive strategies reduce the possibility of errors occurring because of human fatigue and reduce the likelihood of human fatigue itself. These strategies include using a second operator as a second set of eyes, developing clear procedures for monitoring operators, drinking coffee or tea and increasing vigilance with vertical vibration [8]. Although proactive strategies have been proposed and even implemented in some fields, such as air traffic management and rail operations, vigilance decrement continued to occur [8]. Stephen and Jonathan [8] argued that it is impossible to keep vigilance in practice. Hence, besides proactive methods, extensive studies have been conducted to monitor or detect vigilance decrement in real time [5, 9]. These studies have given rise to reactive strategies that alert operators once vigilance decrement occurs. Physiological parameters including brain dynamics, blood flow velocity, heart rate, eye movements were commonly investigated to develop tools for monitoring operator's vigilance [10].

Among these recent physiological studies, eye movements have received a revival interest owing to the following reasons. Firstly, the pupil center corneal reflection technique enabled remote, non-intrusive eye tracking [11]. Thus, eye movements can be monitored and measured continuously during extended cognitive tasks. Furthermore, decades of scientific evidence indicated that there could be a link between sustained attention and eye movements [7]. For instance, some studies have shown that blink frequency and blink duration increased with vigilance decrement [3]. Finally, the human eyes possess an essential role as they are one of the primary and first input media during the monitoring tasks. By tracking the eye movements, cognitive states, visual patterns, and information processing can be monitored and analyzed [11].

A body of previous research studies investigated the correlations between eye movements and vigilance decrement as well as indicated the possibility to develop an eye movement-based detector. Nevertheless, limited studies leverage these correlations to detect vigilance decrement automatically. Moreover, they focused on collecting human physiological data and paid limited attention to the dynamic working conditions of traffic controllers. Nowadays, with the development of advanced information and communication technology (ICT), the internet of things (IoT), and artificial intelligence (AI), data about human states and working condition can be easily collected and analyzed to proactively supporting human decision-making, learning, and action. For example, Akhavian and Behzadan [12] have successfully used mobile sensors and machine learning classifiers to recognize construction equipment activities in a working environment. Considering the above research gaps and knowledge, the objective of this research is developing a hybrid data-driven vigilance model to monitor vigilance in real time using eye movements and context data. With this novel vigilance model, the objectives of (1) automatically and effectively detecting vigilance decrement, (2) being adapted in an empirical working place and (3) proactively alarming operators, could be achieved.

There are several challenges in achieving this objective. Firstly, the relationship between eye movements and vigilance decrement is not clear. There are some contradictions in the existing pieces of literature. Saito [13] did not find any significant quantitative changes in saccadic eye movement in five hours of eye-tracking tasks. Nevertheless, Bodala [14] found that the saccade amplitude and saccade velocity greatly decreased with vigilance decrement. The contradictions in the existing literature hinder the development of eye movement-based vigilance model. Secondly, sensors like eye trackers always lose track, resulting in data loss. How to maintain the performance of vigilance model under the condition of data loss requires more studies. Thirdly, the experiment settings and subjects movements always induce noises into eye movements [15]. As a result, the performance of vigilance model could be affected by the noisy eye movements.

Aiming at achieving the objective of establishing a data-driven vigilance model and dealing with the above challenges, this research proposed decentralized feature extraction, graph-based feature selection, and two-track bagged trees. The rest of the paper is organized as follows: Section 2 gives an introduction of the related works. Section 3 introduces the framework of generating the data-driven vigilance model and the key parts to support the processes of monitoring vigilance. A practical case study was illustrated in Section 4. The authors conclude with a discussion of the potential industrial applications, contributions, limitations and future works in Section 5.

2. Theoretical Background

In this section, works related to eye movement-based vigilance detection, collecting data from connected physical objects and data-driven model for human cognitive states are summarized.

2.1 Eye-tracking data and vigilance decrement

There are several studies conducted on observing human's eye movement on fatigue-inducing tasks or vigilance test to prove the potential of eye movements as an indicator of vigilance.

The percentage of eye closure (PERCLOS) is one of the most widely accepted vigilance indexes in the literature [5]. Besides PERCLOS, blinks received a lot of research attention, too [3, 5]. In Bergasa and Sotelo [5], they used PERCLOS and blink frequency to characterize a driver's level of vigilance. Bodala [14] found that the saccade amplitude and saccade velocity decreased with vigilance decrement. Pupil diameter, pupil eccentricity, and pupil velocity were found to be correlated with the vigilance performance in the study of McIntire et al. [16]. Among these eye movements, fixations received the least attention. In 2017, Bodala et al.[7] proposed a new parameter called 'fixation score,' which was dependent on the distance of fixation position to the target area of interest. They found that the 'fixation score' decrease with the vigilance decrement and indicated the possibility of using fixations to detect vigilance decrement. To enhance the understanding of eye metrics and vigilance, in this study, we investigated the possibility of establishing an eye movement-based vigilance model.

2.2 Internet of things

The term *Internet of things* (IoT) came to public attention in September 2003 [17]. It refers to connecting physical objects with sensing, computing and communication capabilities to realize machine-to-machine learning and communication [17]. The concept of IoT has been widely accepted and fostered a great number of applications in several fields, such as smart cities, and manufacturing schedule [18-20]. Nevertheless, Guo et al. [21] suggested that the traditional view of IoT focused on a thing-oriented perspective and missed the "harmonious" interaction between human and IoT. Moreover, Zhong et al. [22] indicated that the next-generation internet should promote the interaction between human, society and smart objects. To improve the human-IoT interaction, Guo et al.[21] proposed opportunistic IoT which senses human behaviors and is affected by human behaviors.

With the advanced application of IoT and development of human-IoT interaction, a considerable amount of product-generated data, as well as user-generated data which refers to those acquired from the online survey, comments, and questionnaires could be collected.

2.3 Data-driven model of human cognitive states

With the fast development of IoT and AI techniques as well as machine learning, big data analysis can be efficiently and effectively conducted. Hence, data-driven techniques have received increasing interests from various fields [23]. Nevertheless, to the best of authors' knowledge, limited research studies related to data-driven vigilance model could be found. Hence, we reviewed works related to the data-driven model of human cognitive states, including fatigue, mental workload, cognitive distraction, and vigilance. In general, data-driven models can be classified into four types based on their data source:

Behavioral data: Operators changed their behavioral patterns with their cognitive states. For example, Liang and Lee[24] found that drivers made steering less smooth under the condition of cognitive distraction. Fairclough and Graham [25] found that tired drivers made larger steering wheel movement and fewer steering wheel reversals. These changes allow researchers to measure human cognitive states. Pimenta et al. [26] employed a neural network to classify fatigue based on behaviors. They extracted performance indicators of fatigue from the human-computer interactions. Attributes such as key-down time, the time between keys, mouse velocity, and mouse acceleration were appropriate indicators for the continuous classification of mental fatigue.

Physiological data: Many physiological variables were studied to measure human cognitive states. In general, electroencephalographic (EEG) activities, heart activities, and eye movements are the most commonly used physiological indicators of cognitive states. Lin et al. [10] proposed a generalized EEG-based self-organizing neural fuzzy system to predict the vigilance level of drivers. The results of their experimental work showed that it was feasible to estimate the subject’s reaction times based on 1-s EEG power spectra. Based on the effects of fatigue on HRV, Patel et al. [27] presented an artificial intelligence system to detect early onset of fatigue in drivers.

Context data: Context data such as environmental data and working condition data were usually used to develop biomathematics models in the past. As early as 1982, Borbély [28] proposed a two-process model of human fatigue. Processes S and C are used in this model. Process S presents the effects of sleeping and wake-up time on the level of human fatigue. Process C considers the interactions between the circadian rhythm and sleep-wake independent. Over time, researchers found other data such as sleep inertia and light exposure should be included in the fatigue model [29]. However, the performance of these models heavily depends on domain knowledge.

Hybrid data: Some researchers combined data from different sensors to improve the performance of the data-driven model. Zhang et al.[30] proposed a driver workload estimation model using both vehicle data and eye-movement data. Sarter et al. [31] assessed pilots' monitoring strategies and performance by combining behavioral and eye-tracking data. Ji et al. [32] incorporated the visual parameters and the complex contextual information to predict human fatigue.

Unlike the traditional biomathematical model of human performance, the data-driven model would not heavily rely on domain knowledge and could efficiently deal with the increasing amount of sensing data. Despite the above data-driven models, few works discuss a systematic approach to establish vigilance model based on data gleaned from the advanced connected products, especially in the traffic management area.

2.4 AI techniques used in vigilance modeling

AI approaches to vigilance modeling are mostly based on linear regression, decision tree and support vector machine [33, 34]. Among them, the support vector machine is the most widely used method. Nevertheless, it cannot thoroughly address the problem of great diversity in eye-tracking data [35]. To overcome this challenge, in this study, it is proposed to use bagged trees method.

The bagged predictor was first proposed in 1996 [36]. It is a method for manipulating training data to generate multiple predictors and aggregating results of multiple predictors. It attempts to assemble multiple trees so as to improve model performance. A learning system can be used to construct a predictor by taking a bootstrap sample from the training data. Multiple predictors can be constructed by generating repeatedly and training the bootstrap sample from the original training data. All predictors constructed would indicate the class of the instance and determine the final class.

Comparing with other AI techniques, the bagged trees approach can perform much better in analyzing data with substantial classification noise. Breiman [36] indicated that '*If perturbing the learning set can cause significant changes in the predictor constructed, then bagging can improve accuracy*.' This advantage makes it very suitable for eye-movement-based vigilance detection, as the main challenge of establishing eye-movement data-driven vigilance model is noises. Hence, the concept of bagged trees is adopted in this study.

3. Eye-movement Data-driven Vigilance Model

To establish the eye-movement data-driven vigilance model, three phases were conducted. In the first phase, the platform for collecting and analyzing data was built. In the second phase, the detail procedures for selecting features from collected data were conducted. The vigilance levels were predicted based on the selected features in the third phase. In the following section, the three phases were described.

3.1 Platform development

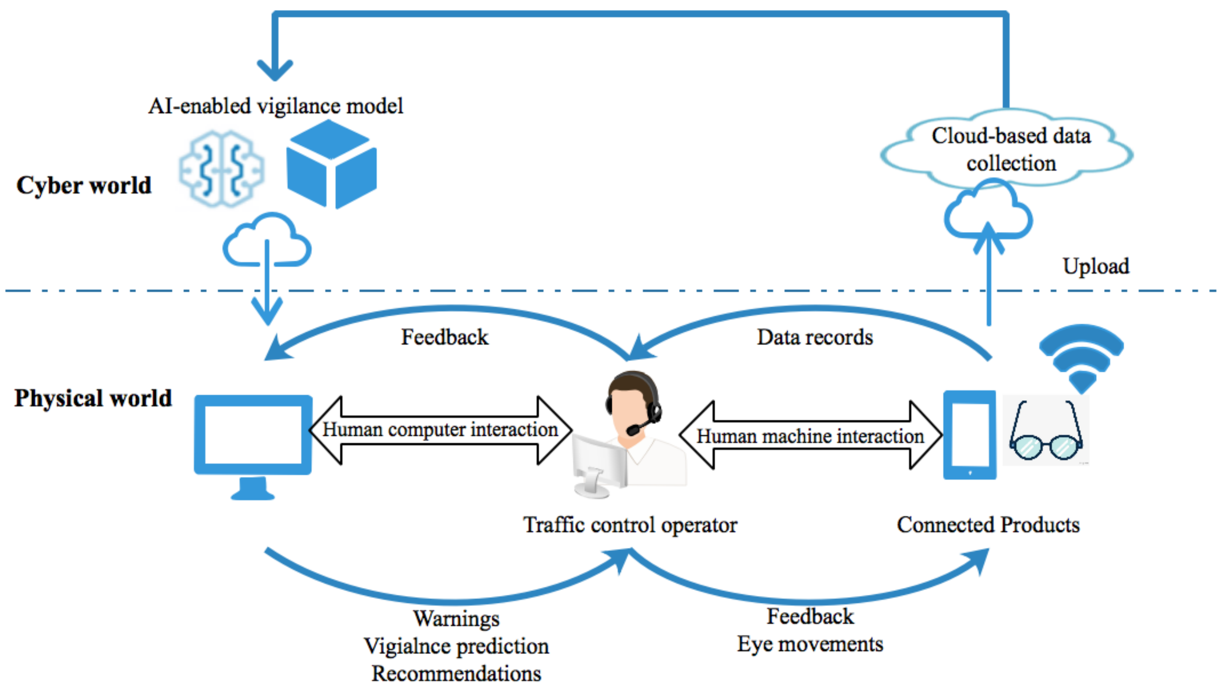

Figure 1: Framework of eye-movement data-driven vigilance model platform

The platform has two main parts: (1) physical sensors for recording data and providing information to users; (2) the cyber part for data processing and modeling vigilance.

As for physical sensors: the data are collected with connected objects including the eye tracker, smartphone and computers. These objects could provide three kinds of information: eye movements, human-computer interaction, and user sensing data. The regular commercial eye trackers provide gaze or eyelid movements with timestamps. The events of fixations, saccades and blinks can be extracted from the raw data of eye trackers using dispersion-based threshold algorithms or velocity threshold algorithms [37, 38]. The computer records data of human-computer interaction, such as the inputs of keyboard and mouse movement. Hence, a set of behavior data could be generated. For example, Pimenta et al. [26] had extracted performance indicators of fatigue from the human-computer interactions. Besides these product-generated data, user sensing data include online feedback, subjective rates, and complaints are collected.

In the cyber world, all the collected data are uploaded to the cloud for data analysis. Three kinds of information, including eye movements, vigilance performance, and context information are collected from several sensors. The enormous amount of data causes several challenges in detecting vigilance levels. First, the connected sensors sometimes are under the risk of losing track so as to result in data loss. Second, it is a challenge to extract efficient features for forming a large and sufficient amount of data. Third, previous studies have shown that eye-tracking data has great noises. Considering these challenges and knowledge mentioned in Section 2, the authors proposed the following two modules, namely cloud-based data collection and bagged tree-based vigilance detection. The cloud-based data collection module pre-process collected data and extracts features based on correlation analysis. The extracted features are used as inputs of the bagged trees-based vigilance detection module, which generate vigilance levels as outputs. These two modules are discussed in the next section.

3.2 Cloud-based data collection

The cloud-based data collection is composed by two parts, i.e. (a) decentralized feature extraction and (b) graph-based feature selection, which is constituted by (b-1) graph establishment, (b-2) correlation analysis and (b-3) feature selection. More detailed explanations are addressed as follows.

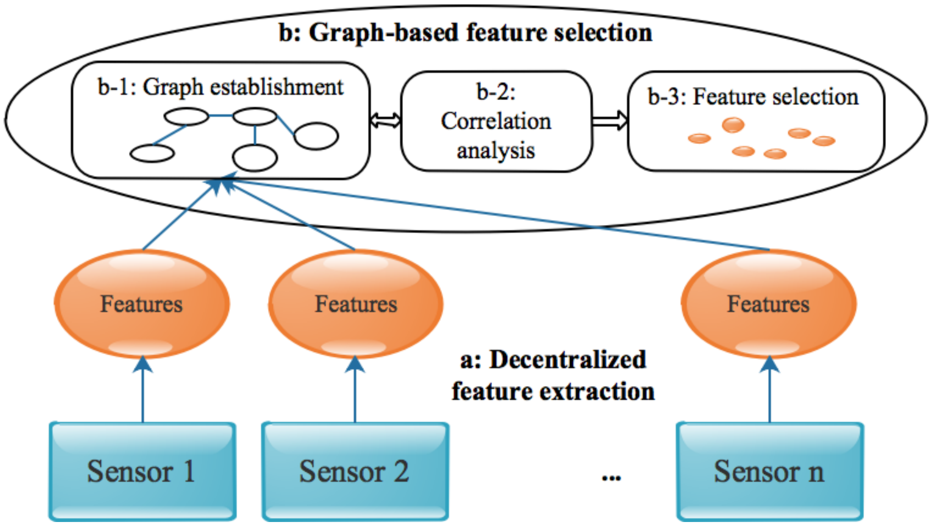

Figure 2: Illustration of cloud-based data collection

(a) Decentralized feature extraction: Sensor-generated data are preprocessed separately and then uploaded to the cloud. In such a way, feature extraction occurs locally on each sensor (Figure 2). An extracted feature is represented as $F_k = \{f, s\}$, where f is the value of the feature, and s is the label of the sensor. The label of the sensor depends on its collected information. This method benefits from local observations and reduces cloud computation time. The critical advantage of decentralized feature extraction is that it is survivable to the loss of sensing nodes. In general, the eye trackers always lost track due to light exposure and subjects' nature shake. If the eye tracker lost track, a part of eye movement data would be lost. Hence, we proposed the decentralized feature extraction to minimize the effects of data loss on extracting other features.

(b) Graph-based feature selection: Feature graph is established based on all collected features and their inter-relations. Following the label s, all these features belonging to vigilance performance are selected and represented as V. The inter-relations between any F_k and V is defined as a_k. We selected the features which have close relations with V first. And then select other features which have close relations with the selected features. The inter-relations between any F_k

and F_l are defined as $r_{k,l}$. r is the correlation coefficient. Features are selected using the Algorithm 1.

$$r_{k,l} = corr(F_k, F_l) \tag{1}$$

$$a_k = max\ \{corr(F_k, V)\} \tag{2}$$

Algorithm 1: Graph-based feature selection

Input	All features

	for *k =1:N.* % *N* is the amount of features
	    If $a_k > threshold\ 1$
	    $F_k \rightarrow SF$
	end
	for $F_k \in SF$
	    for *l*=1:N
	    if $R_{k,l} > \ threshold\ 2$
	    $F_l \rightarrow SF$
	end

Output	*SF*: Selected features

3.3 Bagged trees-based vigilance detection

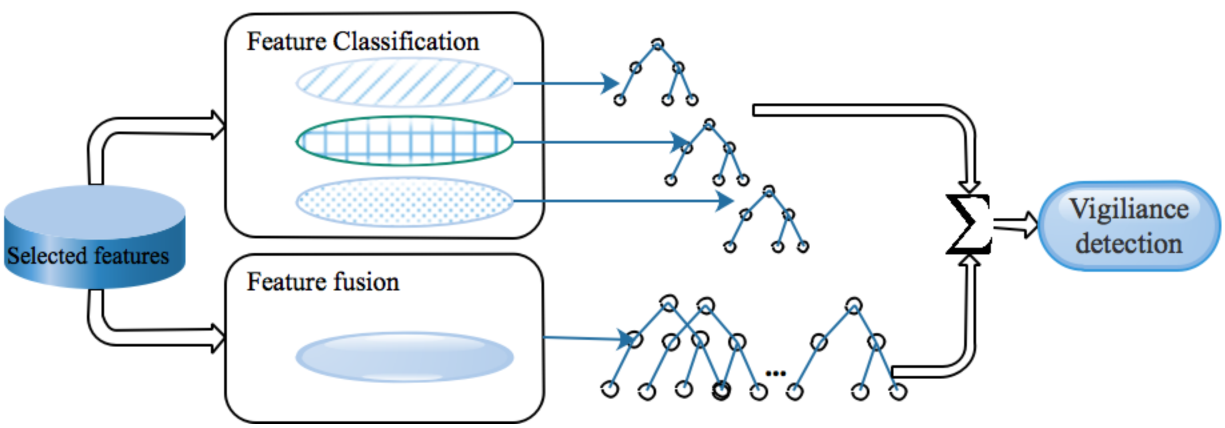

Figure 3: A proposed approach for establishment of vigilance model

The eye movement data quality is usually affected by missing data and noisy data. Hence, the author proposed a two-track bagged trees approach to establish the vigilance model. Figure 3 illustrated the proposed approach. The selected features are duplicated and then analyzed separately. One part of copied features is classified into several types and generate several decision trees independently. The other duplicated features are fused and generate several decision trees using the algorithm of bagged trees. All these decision trees determined the vigilance state. In general, under the condition of data loss, one or two types of features are lost. Hence, the vigilance model generated by fusion data could not provide good performance. The novel method generates vigilance model in two tracks: fusion data and classified data. Thus, the vigilance model is expected to maintain a relatively stable performance under the condition of data loss.

Algorithm 2: Two-track bagged trees

Input	SF
Training	Duplicate $SF \rightarrow \{SF, SF^c\}$
	$SF \rightarrow \{SF^1, SF^2, \dots, SF^M\}$ % classify features based on s
	for i=1:M
	$SF^i \rightarrow Tr^i$
	end
	$SF^c \rightarrow Bagged\ Tr$
Test	
Input	F
	$F \rightarrow \{F^1, F^2, \dots, F^B\}$ % classification
	If B<M
	$V = \sum Tr(F^i)^i$
	else
	$V = Bagged\ Tr\ (F)$
	end
Output	V: Vigilance performance

In order to build A trees for bagging, we generated A copies of data set $T = \{T_1, T_2, \dots T_A\}$ from the training data set SF using Sampling with replacement. Based on each of the data set BT_a, a decision tree Tr^{*a} could be built using Classification and Regression Tree (CART) algorithm.

$$Bagged\ Tr\ (F) = \sum Tr^{*a}\ (F)/A \tag{3}$$

4. Case study

We experimented to collect data and illustrated the proposed approach. Twenty students of Nanyang Technological University (12 males and 8 females) aged 20-27 years old were recruited to participate in this case study. Participants received $20 per hour of compensation for their time and efforts. They have normal or corrected to normal vision. People using both prescription glasses and contact lens were permitted to participate in this experiment. None of them suffer from insomnia. The experiment was approved by the Institutional Review Board of Nanyang Technological University.

The flow of the experiment started with the subjects received a verbal briefing and signed the consent form in the brief session. Then, the participant was instructed to rate their fatigue score with the Samn-Perelli Mental Fatigue Scale [39] and do the vigilance test for 5 minutes. Moreover, the subjects have to monitor vessel traffic conditions to identify potential conflicts for 45 minutes. Next, the vigilance test and traffic monitoring task would be done again before the last mental fatigue assessment and the last vigilance test. In total, each participant did three vigilance tests and two fatigue induce-tasks. This experiment lasted around 120 minutes.

4.1 Platform development

The sensors used in this experiment were smartphones, a laptop, and a Tobii eye tracker. Participants rated their fatigue levels using Google form with their smartphone. The Samn-Perelli Mental Fatigue Scale was adopted in this study. Moreover, participants did choice reaction time tasks with a 13-inch laptop as shown in Figure 4. Participants monitored the display and pressed target key with a visual cue from the screen. When a visual cue present, the participants should press a target key as soon as possible. The visual cues include vowel letters (A, I, U, E, O) and consonant letters (K, L, T, V, X). One of the letters was presented randomly on display. The possible locations of visual cue are presented in Figure 5.

The Tobii Pro X3-120 (as shown in Figure 4) that captures eye movement at 120 Hz was attached on the laptop. To ensure precision and traceability of the Tobii Pro X3-120, subjects were required to sit 50 cm away from the laptop and the eye tracker. The eye tracker analysis software, Tobii studio came along with the eye tracker, enabling users to easily extract data and various standard parameters such as fixations, saccades, blinks and pupil diameters with the I-VT methods. The eye tracker is portable and light. One downside of the equipment/accompanying software is that it cannot differentiate between pure blinks and loss of data (caused by sudden head movement or light reflection). Therefore, parameters related to blinks were not considered in this study.

Figure 4: The Tobii X3-120

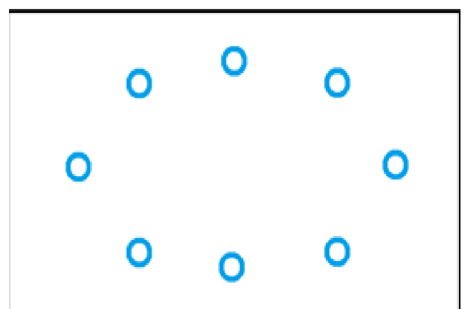

Figure 5: Possible locations of visual cues

4.2 Cloud-based data collection

(a) Decentralized feature extraction: Features were collected following the concept of decentralized feature extraction. Smartphones recorded the time, the scores of the Samn-Perelli Mental Fatigue Scale. They were divided into three classes: Class 1 (first and second scale; no to little fatigue), Class 2 (third and fourth scale; moderate fatigue), and Class 3 (fifth and sixth scale; high fatigue), and then uploaded to cloud database. The laptop uploaded time on task, time on test, and reaction time to cloud database after participants' response. The Tobii X3-120 recorded the

gaze points, and send to Tobii studio. Then the Tobii studio uploaded eye-tracking features and recording time to cloud database. Based on the recording time of the three sensors, all the features were integrated.

In this study, we merged the data using a 5-second time window. All the features belonging to the same time window were merged as one data piece. In total, we collected 2400 pieces of 5-second data for this study. Table 1 shows the 16 features extracted from the sensors. In this study, the local gaze features which depend on the visual cues were not studied because a general vigilance decrement model is expected to be established.

Table 1: Features collected by decentralized feature extraction

Feature Type	Parameters	Descriptions
Eye metrics	1. Saccade peak velocity	The maximum velocity of the gaze points belonging to the saccade.
	2. Saccade mean velocity	The mean velocity of the gaze points belonging to the saccade.
	3. Fixation duration	Duration in milliseconds of a fixation
	4. Saccade duration	Duration in milliseconds between two subsequent fixations
	5. Saccade amplitude	Distance in pixels between two subsequent fixations.
	6. Standard pupil left	The standardized diameter of the left pupil.
	7. Standard pupil right	The standardized diameter of the right pupil.
	8. Saccade number	Counts of saccades in a period.
	9. Fixation dispersion	Root mean square of the distances from each fixation to the average fixation position
	10. Fixation number	Counts of fixations in a period.
	11. Fixation quality	The standard deviation of positions in pixels of gaze points belonging to the fixation.
	12. Fixation velocity standard deviation	The standard velocity deviation of gaze points belonging to the fixation.
	13. Fixation Mean velocity	The mean velocity of gaze points belonging to the fixation.
Context information	14. Mental fatigue	Rated by the Samn-Perelli Mental Fatigue Scale.
	15. Time on test	1 min, 2 mins, 3 mins, 4 mins
	16. Time on task	0 hour, 1 hour, 2 hours

(b) Graph-based feature selection: To select features from all these original features, graph-based feature selection was conducted.

(b-1) Graph establishment: Figure 6 was depicted based on the previous study [40]; it presents the relationship between all the features and vigilance performance. As mentioned in Section 4.1, the features that have close correlations with vigilance performance are selected first and other features were continued to be selected. A great number of studies have confirmed the high correlations between time, mental fatigue and vigilance performance. In contrast, the correlations between eye metric and vigilance performance were usually questioned. Hence, the time on test, time on task, and mental fatigue were selected first. Algorithm 1 was adopted to select other features from eye metrics.

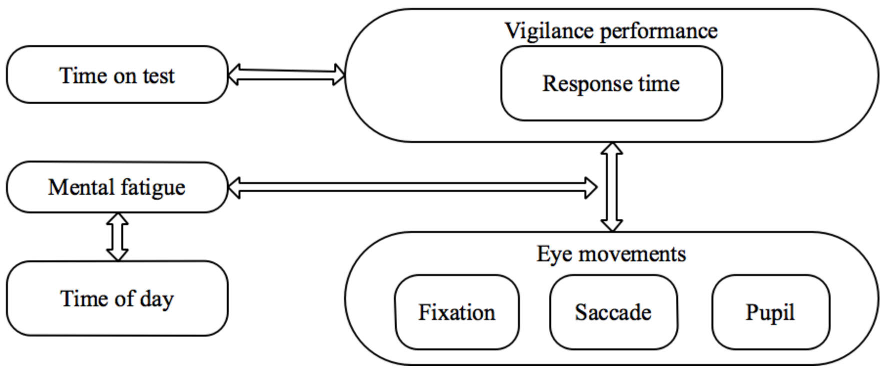

Figure 6: The established graph for correlation analysis

(b-2) Correlation analysis: According to the previous study [40], we found that the correlations between saccade peak velocity and vigilance performance were affected by mental fatigue. Hence, we hypothesized that eye metrics would be correlated with the vigilance performance subjected to the mental fatigue level as H1.

The collected eye-tracking data pieces were categorized into three types, namely low mental fatigue, moderate mental fatigue, and high mental fatigue depending on the score of the Samn-Perelli Mental Fatigue Scale. Pearson correlations were performed (separately for low, moderate, and high mental fatigue levels) to relate response time and eye metrics. T-test was conducted to test the null hypotheses that there are no correlations between eye metrics and response time. The Pearson Correlation and significance are shown in Table 2. The results indicate that in low mental fatigue, saccade mean velocity, saccade amplitude, fixation number, and fixation velocity are correlated with vigilance performance. In Moderate mental Fatigue, saccade mean velocity, fixation duration, and saccade number, fixation mean velocity are correlated with vigilance performance. In high mental fatigue, only fixation duration shows no correlation with the vigilance performance. The correlations between eye metrics were tested using Pearson correlations, too.

Table 2: Correlation analysis

Variables correlated with	**Reaction time**					
	Low Mental Fatigue		Moderate Mental Fatigue		High Mental Fatigue	
	r	*p*	*r*	*p*	*r*	*p*
1. Saccade peak velocity	0.057	0.211	-0.053	0.0245	0.128**	0.002
2. Saccade mean velocity	0.153**	0.001	-0.118**	0.010	0.163**	0.000
3. Fixation duration	0.03	0.514	-0.164**	0.000	0.071	0.083
5. Saccade amplitude	0.137**	0.003	-0.089	0.051	0.143**	0.000
6. Standard pupil left	-0.016	0.727	0.060	0.188	0.092*	0.025
7. Standard pupil right	0.068	0.141	0.074	0.105	0.088*	0.032
8. Saccade number	-0.027	0.553	0.211**	0.000	-0.151**	0.000
10. Fixation number	-0.14**	0.002	-0.023	0.619	0.108**	0.008
11. Fixation quality	0.036	0.437	0.064	0.160	0.133**	0.001
12. Fixation velocity standard deviation	0.154**	0.001	0.016	0.730	0.105*	0.011
13. Fixation mean velocity	0.077	0.092	0.184**	0.000	-0.088*	0.033

**. Correlation is significant at the 0.01 level (2 tailed)

*. Correlation is significant at the 0.05 level (2 tailed)

(b-3) Feature selection: To use the Algorithm 1 to select features, the threshold of correlation coefficient should be selected. In this case study, it was set as 0.1. Hence, the features that have *r* larger than 0.1 were selected. The results of the data analysis suggested that pupil parameters were not significantly correlated with response time. However, we found that they were closely correlated with fixation and saccade parameters. The detail results of correlations among the eye movement parameter are discussed in the authors' other works. According to Algorithm 1, the

pupil diameters were selected. Moreover, it is found that the correlations between eye movements and vigilance performance were subjected to the level of mental fatigue. Finally, 14 features were selected. The features '4. Saccade duration' and '9. Fixation dispersion' were removed due to low correlations.

4.3 Bagged trees-based vigilance detection

Following the steps mentioned in Section 3.3, all the selected features were classified into three types, as listed in Table 3. The eye metrics were divided into fixations related parameters, saccades related parameters, and pupil related parameters. The context features involve time on task, time on test and mental fatigue. Feature fusion refers to integrate all selected features as inputs of the bagged trees.

Table 3: Classification of the selected features

Feature types		**Parameters**
Eye metrics	Fixations	Fixation number Fixation duration Fixation quality Fixation mean velocity Fixation velocity standard deviation
	Saccades	Saccade number Saccade amplitude Saccade peak velocity Saccade mean velocity
	Pupil	Standard left pupil diameters Standard right pupil diameters
Context features		Mental fatigue Time on tasks Time on test
Fusion features		Eye metrics Context features

The efficiency of Algorithm 2, namely two-track bagged trees, in addressing data loss was tested in this section. Six conditions of data loss were stimulated. Condition 1, all data were collected, the fusion features were used as inputs. Condition 2 refers to the loss of eye-tracking data. Only context features were used as inputs. In contrary, all context data were lost in Condition 3, so eye metrics were used as inputs. Conditions 4-6 refer to the three different levels of data loss in eye-tracking data. Eye metrics have several parameters; some kinds of parameters would be lost due to hardware limitations. Hence, we simulated the three conditions and named them with their inputs, such as “fixations”, “saccades”, and “pupils”.

The performance of the proposed bagged trees-based vigilance detection model was tested across the six conditions. It was hypothesized as H2 that there is no difference in the performance of vigilance detection model across the six conditions. MATLAB 2017a was used to generate decision trees and bagged trees for vigilance detection. In this study, the Bayesian optimization was used to preliminary select optimal hyperparameters for the bagged trees model. After a preliminary study, the number of trees was set as 1500 and the minimum leaf size was set as 1. Vigilance performance includes response time. Hence, we generated the regression models. 10-fold cross-validations were conducted to verify the models. The root-mean-square error (RMSE) was utilized to assess the performance of the regression models. One-way ANOVA was performed, and then Tukey HSD Test was utilized to do a post hoc test.

The result of one-way ANOVA test showed the one or more conditions achieved significantly high performance ($F_{(5, 54)} = 392.59$, $p < 0.01$). The post hoc test showed that there are significant differences between fusion features, context, eye metrics, and fixations ($p < 0.05$). Nevertheless, the differences between saccades and pupils were not significant.

The results showed that using the context as inputs yielded better performance than using the eye metrics as inputs. Specifically, without the inputs of the context features, the RMSE of the regression models greatly increased (Figure 7). Hence, the authors concluded that the context features contribute more to the vigilance decrement detection than eye metrics. The results could be explained from the above correlation analysis. The context information includes mental fatigue levels which could significantly affect the correlations between eye movements and response time (Table 2). Hence, without the context information, the performance of eye movement-based vigilance model would be impaired. Also, though the correlations between pupil and response time are not significant, the pupil parameters could contribute to the vigilance model. Hence, the proposed Algorithm 1 for feature selection is feasible. In other words, the proposed method could achieve a relatively stable performance under the condition of data loss.

The bagged trees with the fusion features as inputs yielded the best performance with the lowest RMSE (0.119) and the smallest standard deviation (0.0148) for the response time.

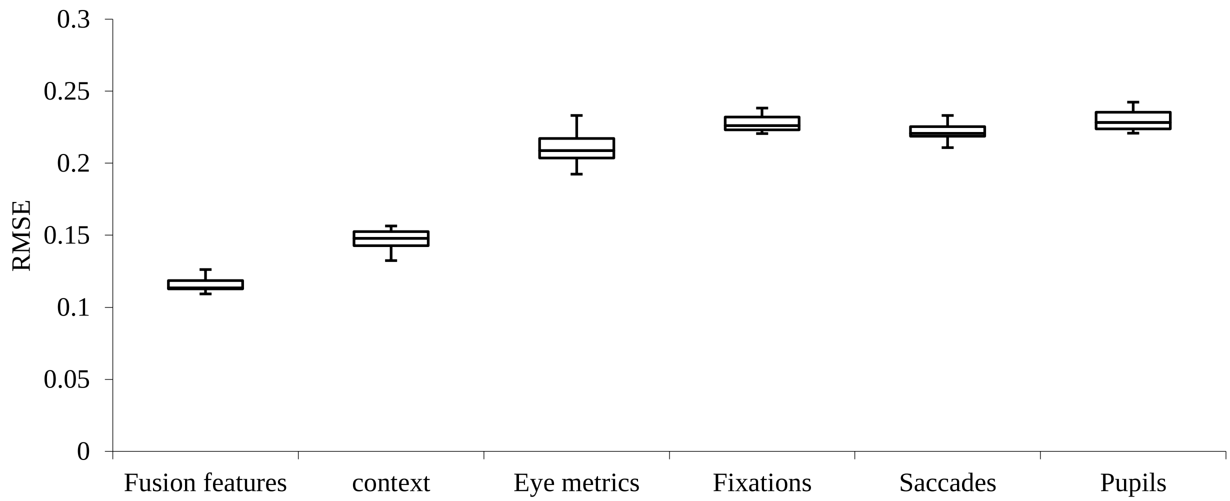

Figure 7: Model performance across different inputs

Table 4: Tukey HSD P-value

	Fusion	Context	Eye metrics	Fixations	Saccades
Fusion					
Context	0.001**				
Eye metrics	0.001**	0.001**			
Fixations	0.001**	0.001**	0.001**		
Saccades	0.001**	0.001**	0.023*	0.516	
Pupils	0.001**	0.001**	0.001**	0.899	0.198

**. Difference is significant at the 0.01 level
*. Difference is significant at the 0.05 level

4.4 Comparison with other methods

To test H3 that the performance of the proposed bagged trees-based vigilance detection model is better than the other three models, a comparison test was conducted. We compared the proposed method with support vector machine, decision trees, and linear regression, as they were the widely used methods for cognitive models. Four models were generated based on the fusion features collected in the case study. Their performance was assessed with the RMSE. 10-fold cross-validations were conducted. One-way ANOVA was used to do the statistical analysis, and Tukey HSD Test was utilized to do post hoc test. The null hypothesis: there is no difference between the performance of the proposed method and other methods.

The F ratio ($F_{(3, 36)} = 834.77$, $p < 0.01$) shown in one-way ANOVA indicated the significance of performance difference. The post hoc test also indicated the significance of the differences between pairs. Considering the results of F-test, Figure 8 presented that our method performed better than other methods significantly. Overall, the performance evaluation result was very positive.

The results can be further explained in two aspects. On the one hand, the correlations between eye-tracking data and vigilance performance were relatively low although they are significant. Consequently, the traditional linear regression cannot achieve good performance in detecting vigilance impairment using eye-tracking data. On the other hand, the high classification noises of eye-tracking data induce great challenges in applying support vector machine. As a result, both support vector machine and linear regression cannot provide a good result in this case study. In contrast, decision tree provided relatively better performance than those two methods. By bagging hundreds and thousands of decision trees, the proposed two-track bagged trees achieved the best performance.

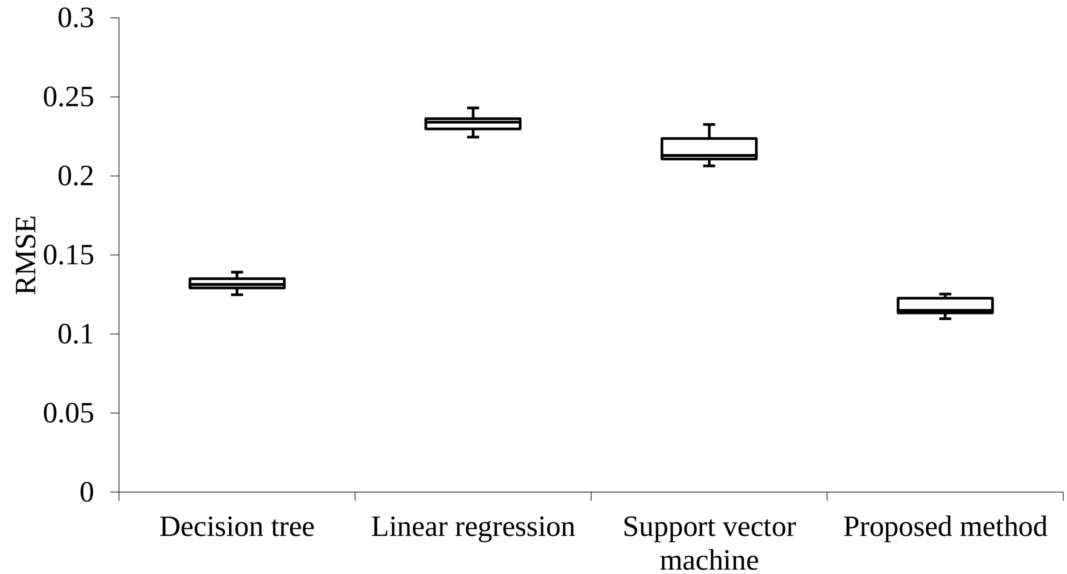

Figure 8: RMSE of vigilance models

4.5 Discussion

The case study illustrated the process of establishing data-driven eye movement-based vigilance model, the procedures of selecting eye metrics, and demonstrated the efficiency of the proposed two-track bagged trees in dealing with eye-tracking data loss. In Section 4.2, the H1 was tested. The results indicated that the correlations between eye metrics and vigilance performance subjected to fatigue levels. Specifically, more eye parameters have correlations to vigilance performance with the increase of fatigue level. Nevertheless, the correlations are low although they are significant.

The low correlations may be caused by the following factors: first, great variances of eye metrics. As mentioned in Section 2.1, eye-tracking data has long been found as useful indicators of vigilance performance, while the correlations between them were questioned in some studies. Based on the authors' previous study and this study, we found that the eye-tracking data at the risk of great variances due to noises, track loss, and long recording-time effects. Second, a large sample size used in this study. We collected 2400 pieces of data for data analysis. The correlations are normally low when the sample size is large due to the introduction of noises. Third, we made scatter plots after correlation analysis and found non-linear correlations. In conclusion, the significant results indicated the existing of correlations; the low index indicated that the correlation might be non-linear. In this study, we focused on realizing the detection of vigilance performance even the data quality is low. Hence, we proposed to use bagged trees instead of linear regression method to model vigilance performance under the condition of data loss.

In Section 4.3, the H2 was tested to investigate the effects of data loss. Under six conditions of data loss, the proposed method showed significantly different performance. Although the proposed method cannot maintain the performance under the condition of context data loss, it can maintain relatively stable performance under the condition of partial eye-tracking data loss. The H3 was tested in Section 4.4. Comparing with three classical methods, linear regression, decision tree, and support vector machine, the proposed method can achieve significant lower RMSE. Considering the results obtained, we can deduce that the proposed method can well conquer the problem of eye-tracking data loss, classification noises, and the non-linear correlations between eye-tracking data and vigilance performance.

5. Conclusion

In the traffic control center, vigilance decrement is a normal phenomenon that has a great negative impact on human performance. It could occur at any time. This study intends to propose a data-driven method that proactively detects vigilance decrement using connected physical objects and artificial intelligence techniques. The main contributions are summarized as below:

First, a novel framework of developing an eye movement data-driven vigilance model was proposed. Graph-based feature selection and distributed model generation are proposed to improve computational efficiency and maintain model performance under the condition of data loss. Two-track bagged trees are proposed to generate vigilance models. As such, the research gaps about eye movement-based detection of vigilance decrement and alarming proactive mechanism can be bridged by this new model.

Second, a clarification of the correlations was conducted between eye movements and vigilance decrement. An observed trend is the increasing regression coefficient of the eye movements against the response time in different mental fatigue classes. The mentioned relationship is a new finding, as the authors had only come across literature that discussed the existence of the relationship between eye movements and response time (i.e., statistically significant) alone, without considering its influence under different states of mental fatigue.

Third, a practical vigilance model was achieved based on eye movements and context information. It enables objectively and unobtrusively detection of vigilance decrement. Moreover, with the consideration of context information, the vigilance model could be adapted to the dynamic conditions of the traffic controller and could achieve better performance than eye movement-based model.

Fourth, a pioneer work of applying IoT techniques in human factors study was demonstrated. We enriched the literature of IoT-application with eye movement data collection in the traffic control field. We made a moving-forward step of using data generated from IoT and human-IoT interaction to establish a vigilance model. The human-IoT scenario design could be highly adapted in an empirical working place and supporting dynamic working conditions of traffic controllers.

The proposed framework adopts screen-based eye trackers to collect eye movements, so the vigilance performance of operators can be monitored objectively and non-intrusively by the proposed approach. The limitations to our study stemmed from the unclear data classification and limited data collected in the laboratory condition. In the field of transportation, many data could be collected from embedded automated identification systems, such as vehicle location and vehicle speed. Hence, in future studies, the authors would like to propose a classification scheme for human-centric data in the vision of IoT. Moreover, more data would be collected for the practical case study.

Reference

[1] M. Masoudian and H. Razavi, "An investigation of the required vigilance for different occupations," *Safety Science*, 2018.

[2] W. A. Harrell, "Lifeguards' vigilance: effects of child-adult ratio and lifeguard positioning on scanning by lifeguards," *Psychological reports,* vol. 84, no. 1, pp. 193-197, 1999.

[3] L. K. McIntire, R. A. McKinley, C. Goodyear, and J. P. McIntire, "Detection of vigilance performance using eye blinks," *Applied ergonomics*, vol. 45, no. 2, pp. 354-362, 2014.

[4] L. K. McIntire, R. A. McKinley, C. Goodyear, and J. P. J. A. e. McIntire, "Detection of vigilance performance using eye blinks," vol. 45, no. 2, pp. 354-362, 2014.

[5] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-time system for monitoring driver vigilance," *IEEE Transactions on Intelligent Transportation Systems,* vol. 7, no. 1, pp. 63-77, 2006.

[6] E. C. Poulton, "Increased vigilance with vertical vibration at 5 Hz: an alerting mechanism," *Applied ergonomics*, vol. 9, no. 2, pp. 73-76, 1978.

[7] I. P. Bodala, N. I. Abbasi, Y. Sun, A. Bezerianos, H. Al-Nashash, and N. V. Thakor, "Measuring vigilance decrement using computer vision assisted eye tracking in dynamic naturalistic environments," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 2017, pp. 2478-2481: IEEE.

[8] S. M. Casner and J. W. Schooler, "Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task," *Consciousness and cognition,* vol. 35, pp. 33-41, 2015.

[9] Q. Ji and X. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real-time imaging,* vol. 8, no. 5, pp. 357-377, 2002.

[10] F.-C. Lin, L.-W. Ko, C.-H. Chuang, T.-P. Su, and C.-T. Lin, "Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system," *IEEE*

Transactions on Circuits and Systems I: Regular Papers, vol. 59, no. 9, pp. 2044-2055, 2012.
[11] A. Olsen, "The Tobii I-VT fixation filter," *Tobii Technology,* 2012.
[12] R. Akhavian and A. H. Behzadan, "Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers," *Advanced Engineering Informatics,* vol. 29, no. 4, pp. 867-877, 2015.
[13] S. Saito, "Does fatigue exist in a quantitative measurement of eye movements?," *Ergonomics,* vol. 35, no. 5-6, pp. 607-615, 1992.
[14] I. P. Bodala, J. Li, N. V. Thakor, and H. Al-Nashash, "EEG and eye tracking demonstrate vigilance enhancement with challenge integration," *Frontiers in human neuroscience,* vol. 10, p. 273, 2016.
[15] S. V. Wass, L. Forssman, and J. Leppänen, "Robustness and precision: How data quality may influence key dependent variables in infant eye-tracker analyses," *Infancy,* vol. 19, no. 5, pp. 427-460, 2014.
[16] L. K. McIntire, J. P. McIntire, R. A. McKinley, and C. Goodyear, "Detection of vigilance performance with pupillometry," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, 2014, pp. 167-174: ACM.
[17] "EPC Symposium (2003) Inaugural EPC Executive Symposium. http://xml.coverpages.org/EPCSymposium200309.html. Accessed 5 July 2010," 2003.
[18] M. Wang, R. Y. Zhong, Q. Dai, and G. Q. Huang, "A MPN-based scheduling model for IoT-enabled hybrid flow shop manufacturing," *Advanced Engineering Informatics*, vol. 30, no. 4, pp. 728-736, 2016.
[19] C.-H. Lee, Y.-H. Wang, and A. J. Trappey, "Service design for intelligent parking based on theory of inventive problem solving and service blueprint," *Advanced Engineering Informatics,* vol. 29, no. 3, pp. 295-306, 2015.
[20] Y.-H. Wang, C.-H. Lee, and A. J. Trappey, "Service design blueprint approach incorporating TRIZ and service QFD for a meal ordering system: A case study," *Computers & Industrial Engineering,* vol. 107, pp. 388-400, 2017.
[21] B. Guo, D. Zhang, Z. Wang, Z. Yu, and X. Zhou, "Opportunistic IoT: Exploring the harmonious interaction between human and the internet of things," *Journal of Network Computer Applications,* vol. 36, no. 6, pp. 1531-1539, 2013.
[22] N. Zhong *et al.*, "Research challenges and perspectives on Wisdom Web of Things (W2T)," *The Journal of Supercomputing,* vol. 64, no. 3, pp. 862-882, 2013.
[23] Z.-j. Zhang, L. Gong, Y. Jin, J. Xie, and J. Hao, "A quantitative approach to design alternative evaluation based on data-driven performance prediction," *Advanced Engineering Informatics,* vol. 32, pp. 52-65, 2017.
[24] Y. Liang and J. D. Lee, "Combining cognitive and visual distraction: Less than the sum of its parts," *Accident Analysis Prevention,* vol. 42, no. 3, pp. 881-890, 2010.
[25] S. H. Fairclough and R. Graham, "Impairment of driving performance caused by sleep deprivation or alcohol: a comparative study," *Human factors,* vol. 41, no. 1, pp. 118-128, 1999.
[26] A. Pimenta, D. Carneiro, J. Neves, and P. Novais, "A neural network to classify fatigue from human–computer interaction," *Neurocomputing,* vol. 172, pp. 413-426, 2016.

[27] M. Patel, S. Lal, D. Kavanagh, and P. Rossiter, "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert systems with Applications,* vol. 38, no. 6, pp. 7235-7242, 2011.
[28] A. A. Borbély, "A two process model of sleep regulation," *Human neurobiology*, 1982.
[29] M. E. Jewett and R. E. Kronauer, "Interactive mathematical models of subjective alertness and cognitive throughput in humans," *Journal of biological rhythms,* vol. 14, no. 6, pp. 588-597, 1999.
[30] Y. Zhang, Y. Owechko, and J. Zhang, "Driver cognitive workload estimation: A data-driven perspective," in *Intelligent Transportation Systems*, 2004, pp. 642-647: Citeseer.
[31] N. B. Sarter, R. J. Mumaw, and C. D. Wickens, "Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data," *Human factors,* vol. 49, no. 3, pp. 347-357, 2007.
[32] Q. Ji, P. Lan, and C. Looney, "A probabilistic framework for modeling and real-time monitoring human fatigue," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and humans*, vol. 36, no. 5, pp. 862-875, 2006.
[33] F. Tango and M. Botta, "Real-time detection system of driver distraction using machine learning," *IEEE Transactions on Intelligent Transportation Systems,* vol. 14, no. 2, pp. 894-905, 2013.
[34] K.-Q. Shen, X.-P. Li, C.-J. Ong, S.-Y. Shao, and E. P. Wilder-Smith, "EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate," *Clinical Neurophysiology,* vol. 119, no. 7, pp. 1524-1533, 2008.
[35] R. Kamalian, E. Yeh, Y. Zhang, A. M. Agogino, and H. Takagi, "Reducing human fatigue in interactive evolutionary computation through fuzzy systems and machine learning systems," in *2006 IEEE International Conference on Fuzzy Systems*, 2006, pp. 678-684: IEEE.
[36] L. Breiman, "Bagging predictors," *Machine learning,* vol. 24, no. 2, pp. 123-140, 1996.
[37] F. Shic, B. Scassellati, and K. Chawarska, "The incomplete fixation measure," in *Proceedings of the 2008 symposium on Eye tracking research & applications*, 2008, pp. 111-114: ACM.
[38] P. Blignaut, "Fixation identification: The optimum threshold for a dispersion algorithm," *Attention, Perception, & Psychophysics,* vol. 71, no. 4, pp. 881-895, 2009.
[39] S. W. Samn and L. P. Perelli, "Estimating aircrew fatigue: a technique with application to airlift operations," School of Aerospace Medicine Brooks Afb tx1982.
[40] V. Renata, F. Li, C.-H. Lee, and C.-H. Chen, "Investigation on the Correlation between Eye Movement and Reaction Time under Mental Fatigue Influence," in *2018 International Conference on Cyberworlds (CW),* 2018, pp. 207-213: IEEE.