



A Novel Method for Classifying Driver Mental Workload Under Naturalistic Conditions With Information From Near-Infrared Spectroscopy

Anh Son Le^{1,2*}, Hirofumi Aoki¹, Fumihiko Murase³ and Kenji Ishida³

¹ Human Factors and Aging Laboratory, Institutes of Innovation for Future Society, Nagoya University, Nagoya, Japan,

² Department of Power Engineering, Faculty of Engineering, Vietnam National University of Agriculture, Hanoi, Vietnam,

³ Denso Corporation, Nisshin, Japan

OPEN ACCESS

Edited by:

Bruce Mehler,
Massachusetts Institute of
Technology, United States

Reviewed by:

Noman Naseer,
Air University, Pakistan
Sung-Phil Kim,
Ulsan National Institute of Science and
Technology, South Korea
Toshinori Kato,
KatoBrain Co., Ltd., Japan
Mauricio Muñoz,
Bosch Center for Artificial Intelligence,
Germany

*Correspondence:

Anh Son Le
leanhsonvn@gmail.com

Received: 12 April 2018

Accepted: 02 October 2018

Published: 26 October 2018

Citation:

Le AS, Aoki H, Murase F and Ishida K
(2018) A Novel Method for Classifying
Driver Mental Workload Under
Naturalistic Conditions With
Information From Near-Infrared
Spectroscopy.
Front. Hum. Neurosci. 12:431.
doi: 10.3389/fnhum.2018.00431

Driver cognitive distraction is a critical factor in road safety, and its evaluation, especially under real conditions, presents challenges to researchers and engineers. In this study, we considered mental workload from a secondary task as a potential source of cognitive distraction and aimed to estimate the increased cognitive load on the driver with a four-channel near-infrared spectroscopy (NIRS) device by introducing a machine-learning method for hemodynamic data. To produce added cognitive workload in a driver beyond just driving, two levels of an auditory presentation n-back task were used. A total of 60 experimental data sets from the NIRS device during two driving tasks were obtained and analyzed by machine-learning algorithms. We used two techniques to prevent overfitting of the classification models: (1) *k*-fold cross-validation and principal-component analysis, and (2) retaining 25% of the data (testing data) for testing of the model after classification. Six types of classifier were trained and tested: decision tree, discriminant analysis, logistic regression, the support vector machine, the nearest neighbor classifier, and the ensemble classifier. Cognitive workload levels were well classified from the NIRS data in the cases of subject-dependent classification (the accuracy of classification increased from 81.30 to 95.40%, and the accuracy of prediction of the testing data was 82.18 to 96.08%), subject 26 independent classification (the accuracy of classification increased from 84.90 to 89.50%, and the accuracy of prediction of the testing data increased from 84.08 to 89.91%), and channel-independent classification (classification 82.90%, prediction 82.74%). NIRS data in conjunction with an artificial intelligence method can therefore be used to classify mental workload as a source of potential cognitive distraction in real time under naturalistic conditions; this information may be utilized in driver assistance systems to prevent road accidents.

Keywords: near-infrared spectroscopy, cognitive distraction, classification, driver attention, mental workload, artificial intelligence

INTRODUCTION

Driver distraction is a major cause of traffic accidents (NHTSA, 2015). An analysis by the US Highway Traffic Safety Administration (NHTSA) showed that driver distraction can be categorized into three types: visual distraction, manual distraction, and cognitive distraction (NHTSA, 2012). Among these, cognitive distraction is the most difficult type to address, because it occurs within the driver's brain (Rizzo and Hurtig, 1987; Engström et al., 2005; Angell et al., 2006). Cognitive distraction is defined as the mental workload associated with a task that involves thinking about something other than driving. The detection of cognitive distraction imposed by a secondary task while driving might play an important role in creating a new driver-assistance system to reduce the incidence of traffic accidents.

Dong et al. (2011) categorized techniques for measuring mental workload while driving into five groups: (1) subjective metrics, (2) biological metrics, (3) physical metrics, (4) performance metrics, and (5) combinations of these metrics. Because the central goal of our research was to identify and improve a metric that might permit the detection of mental workload in real time and which could operate under real conditions in the presence of, for example, vibration from the vehicle, we examined only physical metrics in the present study.

One potential physical metric involves the use of eye-movement information. Many researchers have previously attempted to identify a relationship between mental workload and various items of information on the eye, such as blink (Tsai et al., 2007; Benedetto et al., 2011), pupil diameter (Backs and Walrath, 1992; Klingner et al., 2008; Schwalm et al., 2008; Klingner, 2010), saccades (Tsai et al., 2007; Pierce, 2009; Tokuda et al., 2009), gaze concentration (Wang et al., 2014), or eye fixation (Klingner, 2010). Each of these methods has its advantages and disadvantages. For example, pupil diameter has a strong relationship to the level of cognitive load but it is also highly sensitive to the frequent changes in light that occur while driving (Palinko and Kun, 2012). Another potential method is to use the involuntary eye movements based on the vestibulo-ocular reflex model that are simulated by head movements or by vibrations from the moving vehicle. In this method, differences between predicted and actual eye simulation are assessed as a measure of mental workload (Obinata et al., 2008, 2009, 2010; Aoki et al., 2015; Anh Son et al., 2016, 2017a,b,c,d,e, 2018; Le and Aoki, 2018; Son and Hirofumi, 2018; Son et al., 2018). However, the use of eye information to measure mental workload still has some limitations, such as oversensitivity to light, vibration, noise, and visual information.

In terms of a physical metric, monitoring of brain activity by electroencephalography or the use of the heart rate as an indicator of mental workload have been confirmed to be effective (Meshkati, 1988; Lee and Park, 1990; Jorna, 1992; Porges and Byrne, 1992; Veltman and Gaillard, 1996; Ryu and Myung, 2005; Henelius et al., 2009; Mehler et al., 2012; Cinaz et al., 2013; Angell and Perez, 2015). However, these techniques require physical attachment of the monitoring equipment and are highly sensitive to the driver age, body position, and muscle activity.

One method with a high potential for application is the use of information from near-infrared spectroscopy (NIRS) to classify mental workload (Kopton and Kenning, 2014). NIRS has been used in various fields; for example, in agriculture to check the quality of crops and in medicine to assess oxygenation and microvascular function. In terms of classifying mental workload, a relationship between mental workload and activity of the central nervous system has been confirmed by McBride and Schmorow (2005). Since their work, other researchers have attempted to classify levels of mental workload by applying artificial-intelligence analyses (Tsunashima and Yanagisawa, 2009; Herff et al., 2014; Ichikawa et al., 2014; Aghajani et al., 2017). All of these researchers showed that NIRS has considerable potential in quantifying mental workload while driving, especially in naturalistic cases (Kopton and Kenning, 2014; Liu et al., 2016).

Furthermore, Toshinori Kato and his group have done various investigations on NIRS data, especially how to filter the signals and map it (Kohri et al., 2002; Yoshino et al., 2013, 2015; Orino et al., 2015). In actual driving, his group pointed out that there was a relationship of the brain activity with the vehicle speed. Their research also confirmed that fNIRS data is one of a good solution for monitoring the driver status while driving especially in actual condition. Further, Liu et al. (2017) confirmed that the cognitive workload has a relationship with the hemodynamic activity level (Liu et al., 2017). His team also mentioned that the effective association can be weak in case of driving with subtasks. However, none of them investigated in applying machine learning with raw data to detect mental workload by secondary task while driving.

The central goal of our study was to examine whether or not it is possible to classify driver mental workload by using supervised learning with NIRS data obtained in a real vehicle. In this report, we initially point out the importance of detecting cognitive load while driving. We then review and summarize methods for evaluating cognitive workload that have been reported in the literature. We also discuss the differences in mental workload between doing one task and driving with secondary task. We then review the use of NIRS information to classify cognitive load, and we describe our experimental design and methods for analyzing NIRS data. The result of the classification are reported and, finally, we discuss our conclusions and any challenges that remain.

MATERIALS AND METHODS

Experimental Design

One female and four male subjects, who each held a driver's license (mean age: 38 ± 10 ; two professional drivers, and three newly qualified drivers), were recruited for this test. A total of 60 experiments were performed involving navigating a defined course alone (autocross) or following another vehicle (car-following) on a test course (**Figure 1**). The experiments were approved by the Ethical Review Board of Nagoya University's Institute of Innovation for Future Society. All subjects were provided with explanations regarding the

experimental procedure, and all gave their written informed consent.

The experiment procedure is shown in **Figure 2**. The subjects performed a series of tests involving a driving task (autocross or car following) during which drivers were asked to drive around 40 km/h. For portions of these drives, additional mental workload was introduced by asking subjects to engage in two levels of an n-back auditory digit recall task in which a number was verbally presented to the subject every 2 s. In the 1-back test, the subject was asked to press the “Yes” button when the number heard was the same as the previous one or the “No” button when it was different. In the 2-back test, the subject similarly had to remember the number preceding the previous one. The “Yes” and

“No” buttons were installed on the driving wheel so they could be easily pressed.

As our main aim is to create an algorithm for an advanced driver-assistance system to help prevent traffic accidents by identifying driver cognitive distraction, the classification needed to be reliable, quick, and easy. To achieve this, we used a commercial four-channel NIRS system (Astem Corp., ~~Fukuoka~~, Japan) which was placed on the forehead of the subject, where the signals from the four channels are almost the same (**Figure 3**). This device can measure blood oxyhemoglobin (oxy-Hb) and deoxyhemoglobin (deoxy-Hb) levels at wavelengths of 770 nm (probe distance 35 mm) and 830 nm (probe distance is 40 mm), and oxygen saturation at 35 mm.

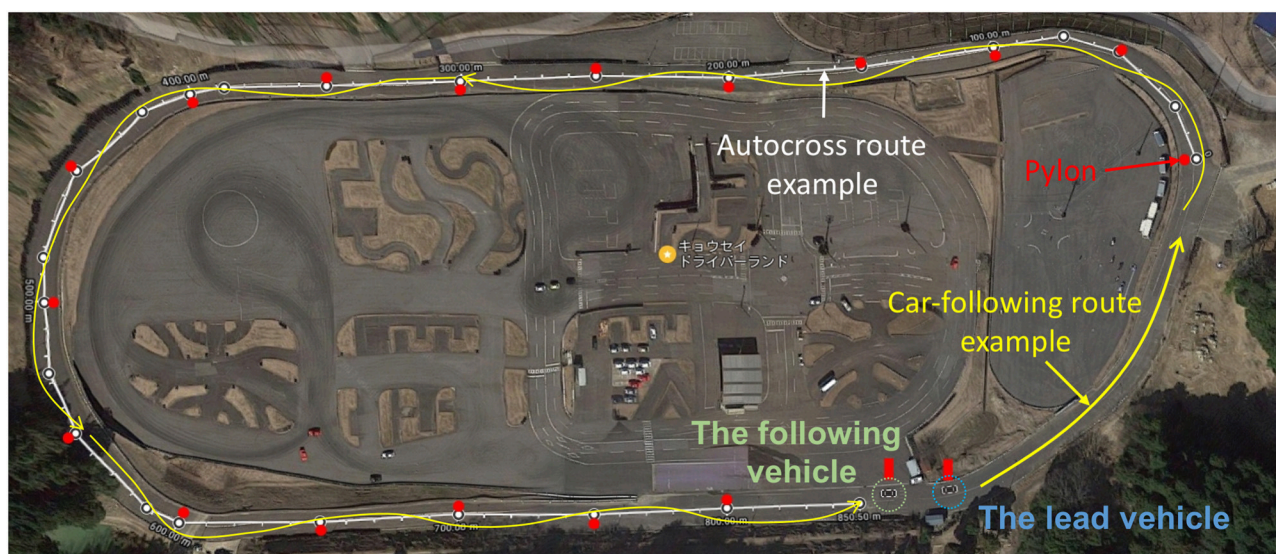


FIGURE 1 | Test course.

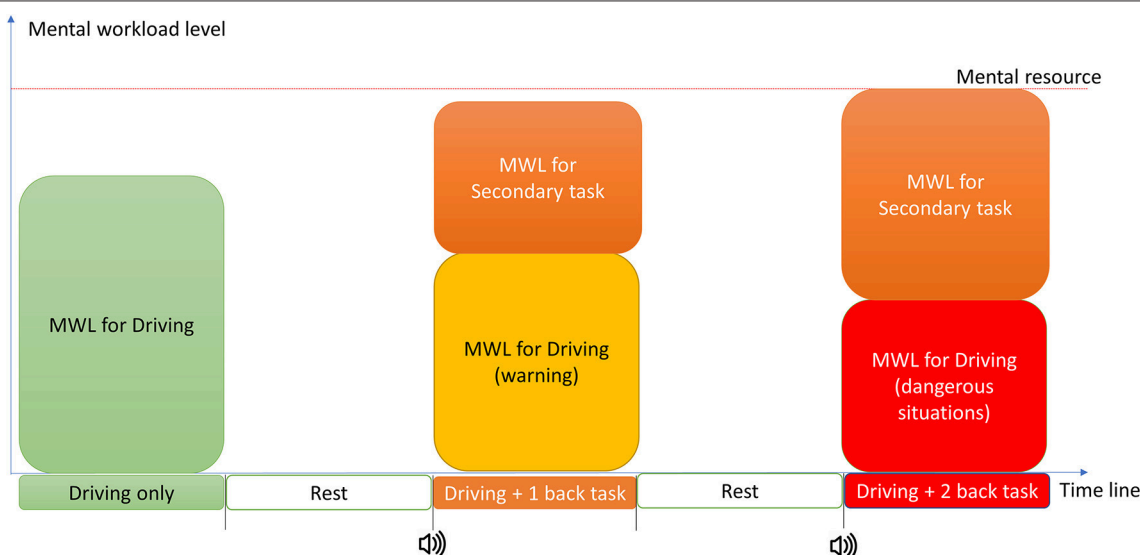
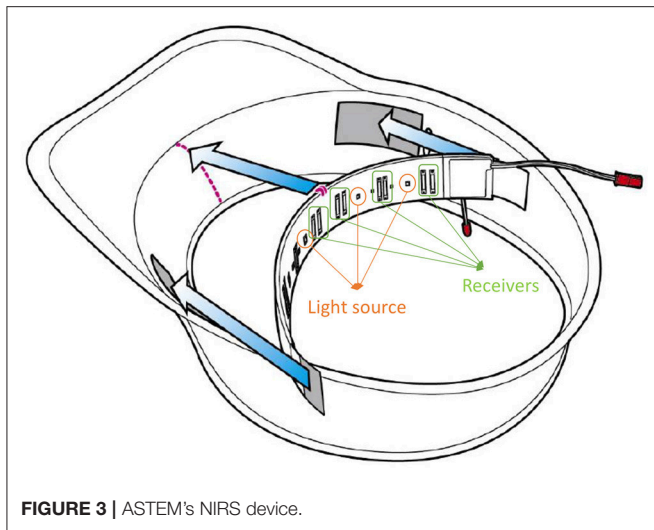


FIGURE 2 | Experiment procedure and hypothetical image of MWL.



Furthermore, to keep the class-balance data set for machine learning, the subjects were asked to drive around the course for each task (driving task only, driving plus 1-back task, and driving plus 2-back task), and repeat it twice (average time for each trial was $1 \text{ min } 54 \pm 16 \text{ s}$ depending upon the speed actually driven). Therefore, the sample data input to machine learning step was class-balance.

Data Processing

Figure 4 provides an overview of the method we used to preprocess the NIRS data. Because of noise arising from movement artifacts (Cooper et al., 2012; Kirilina et al., 2013), the raw NIRS data from each channel were preprocessed by using a modified form of the Beer–Lambert Law (Huppert et al., 2016) with removal of lost data (time shift) (Kirilina et al., 2013), bandpass filtration (0.02–1 Hz) (Ichikawa et al., 2014), and Kalman filtration (Abdelnour and Huppert, 2009), before finally being transformed into features.

In our experiment, the raw data of each channel includes (**Figure 5**): oxyhemoglobin at 35 mm (OxHb35), deoxyhemoglobin at 35 mm (DoxHb35), total oxyhemoglobin (ToxHb35), absolute tissue saturation (StO₂), oxyhemoglobin at 40 mm (OxHb40), deoxyhemoglobin at 40 mm (DoxHb40). After filtering, all of the information from NIRS will be transform to 6 features namely OxHb35, DoxHb35, ToxHb35, StO₂, OxyHb40, and DoxHb40 for preparing to input for machine learning step.

These features were then processed to create training data and testing data for subject-dependent, subject-independent, channel-independent, and subject-independent plus channel-independent cases. After taking all of the data, they were divided into 75% (training data) and 25% (testing data, which is not used to improve the model, but to measure its predictive performance; **Figure 6**). Because we used four channels for the forehead, the data combinations were obtained merely by combining the data together. To prevent overfitting during machine learning, a fivefold cross-validation and a principal-component analysis were applied before the data were used to train the system. The fivefold cross-validation was conducted in the following three

steps. First, the training data (75% of all data) was split into 5-fold. Second, a model for each fold using all the data outside the fold (75% of the training data) was trained and validated. After that, the features were transformed with PCA to reduce the dimensionality of the predictor space (we applied 5 principal components).

Definition:

- Subject-dependent (Subject-dependent + channel-dependent; **Figure 7**): data of each channel for each subject in all task was combined as the input of the machine learning step. Total 20 datasets were prepared for running machine learning.
- Subject-independent (Subject-independent + channel-dependent; **Figure 8**): data of each channel for all subjects in all task was combined as the input of the machine learning step. Total 4 datasets were prepared for running machine learning.
- Channel-independent (Channel-independent + Subject-dependent; **Figure 7**): data of all channel for each subject in all task were combined as the input of machine learning step. Total 5 datasets were prepared for running machine learning.
- Subject-independent + Channel-independent (**Figure 8**): data of all channel for all subject in all task were combined as the input of machine learning step. Only one data set was prepared for running machine learning.

The Classification Method

Previous studies on the classification of mental workload from NIRS data have used an SVM (Devos et al., 2009; Ichikawa et al., 2014; Aghajani et al., 2017), linear discriminant analysis (Luu and Chau, 2009), the hidden Markov model (Sitaram et al., 2007; Zimmermann et al., 2013), or artificial neural networks (Chan et al., 2012; Thanh Hai et al., 2013). However, most of these studies involved complicated multichannel NIRS systems. In our study, because of the large number of samples and the low number of channels, we applied supervised learning in MATLAB 2017b (MathWorks Inc., Natick, MA, United States) (**Figure 5**). We used the 75% of the data to train several well-known models, including the decision-tree model, the discriminant-analysis model, the logistic-regression model, SVMs, nearest-neighbor classifiers, and ensemble classifiers. The performance of these classifiers was determined from the accuracy, as calculated by using the equation shown below:

$$\text{Accuracy } (A_{cc}) = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

In addition, in case of applying SVMs to perform multiclass classification (just driving vs. 1-back vs. 2-back), the transformation technique was applied to reduce the multiclass classification problem to a set of binary classification subproblems, with one SVM learner for each subproblem. One-vs.-All trains one learner for each class. It learns to distinguish one class from all others will be applied in our case.

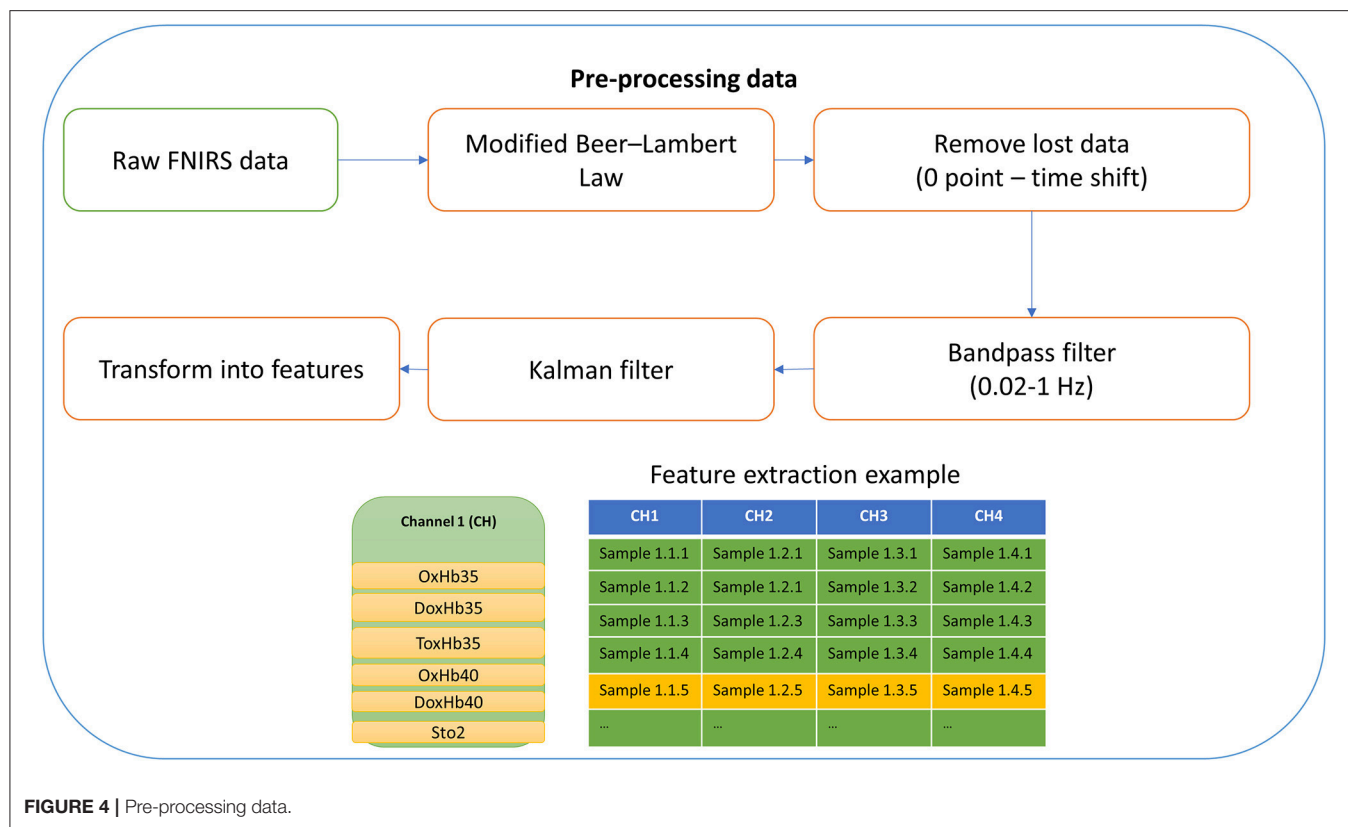


FIGURE 4 | Pre-processing data.

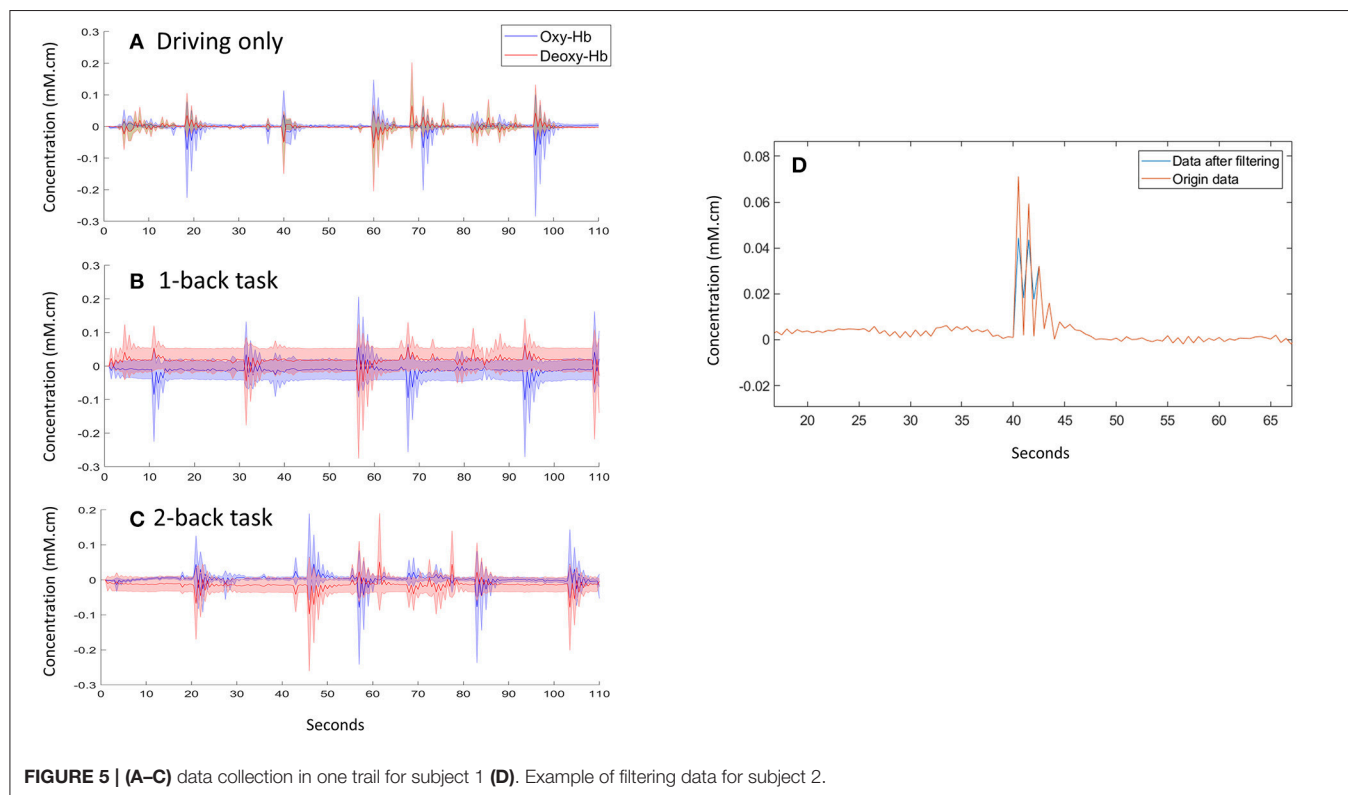
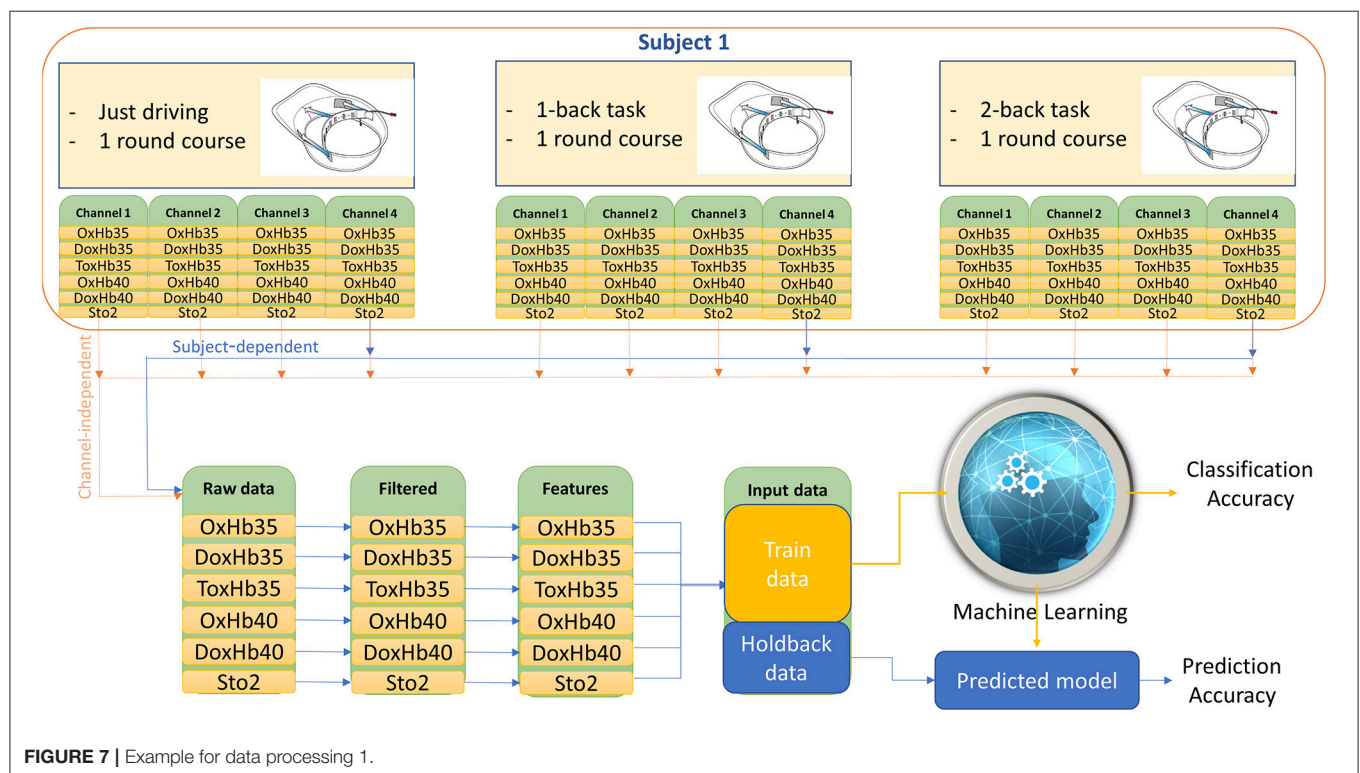
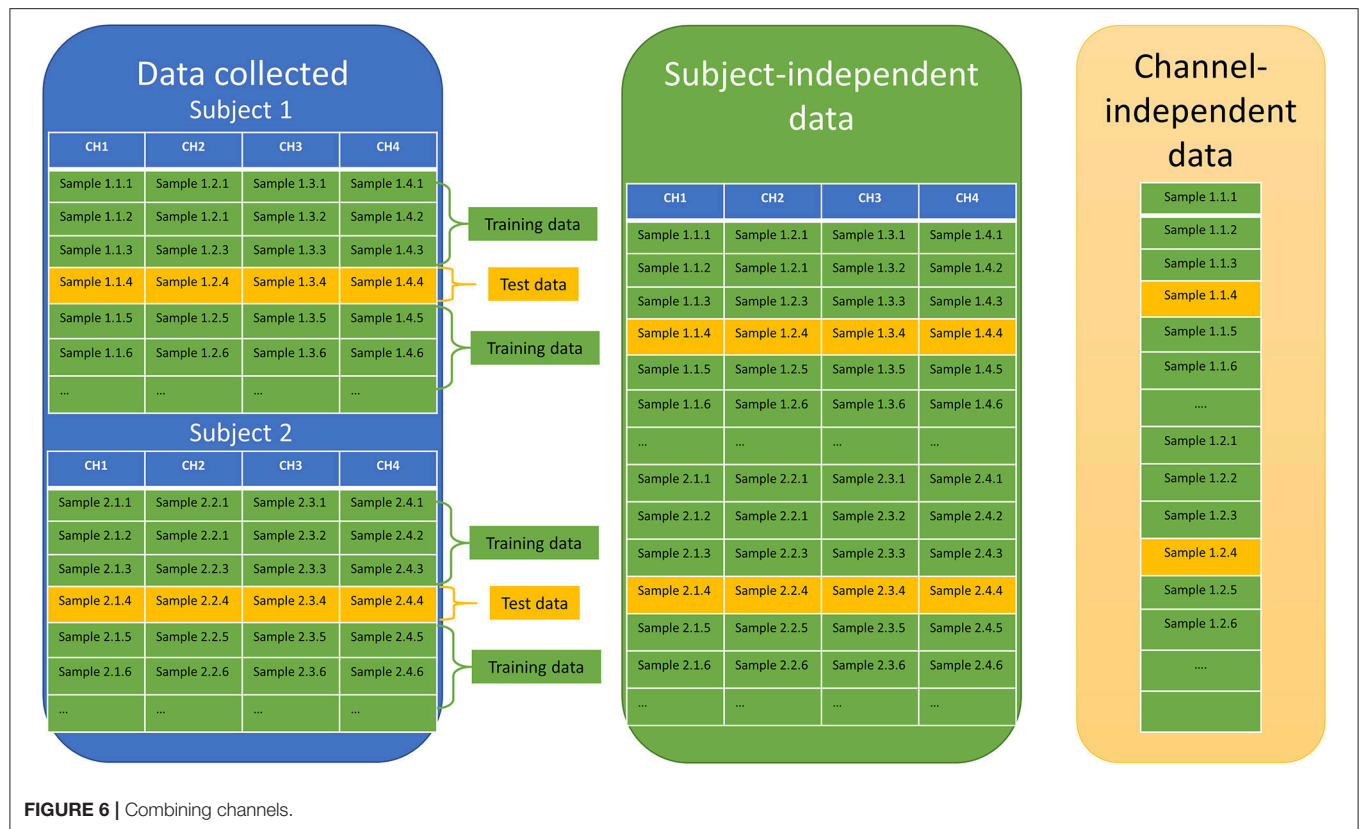
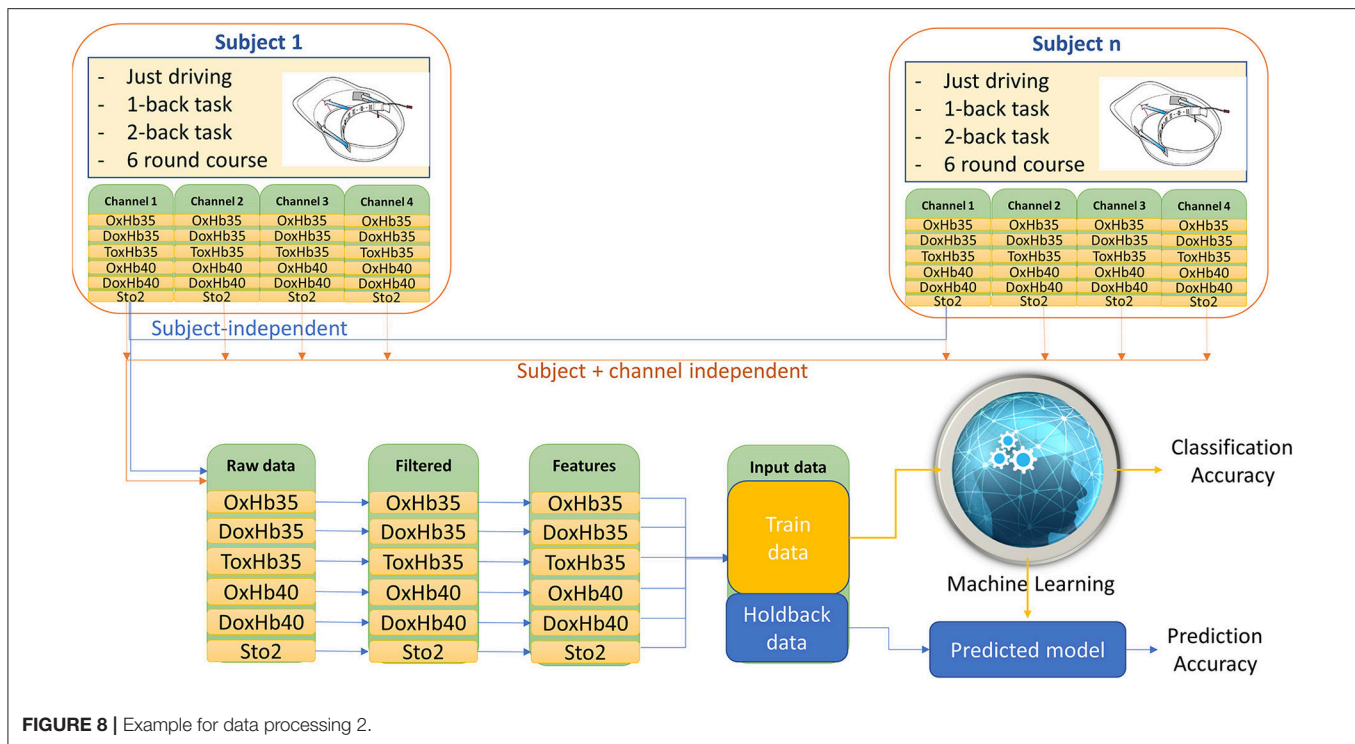


FIGURE 5 | (A–C) data collection in one trail for subject 1 (D). Example of filtering data for subject 2.





The most accurate model for classification was selected and used in the prediction step with the testing data. All parameters of any model were kept the same in both classification and prediction step.

Definition (**Figure 9**):

- Classification accuracy: the accuracy of the classification using 75% of the data for train step. The model of highest accuracy was selected to use in the prediction step.
- Accuracy on the testing data: the accuracy of prediction using 25% of the data for the prediction step. In this step, the prediction model was exported from the selected model in the classification step and apply it for the new data (25% testing data).

RESULTS

Subject-Dependent Classification Analysis

Data for each channel for each subject were separated for the purposes of training and prediction. The classification between driving only, driving with a 1-back task, and driving with a 2-back task showed a good performance and a high accuracy (the classification accuracy increased from 81.30 to 95.40%, and the accuracy for the testing data increased from 82.18 to 96.08%) (**Video 1**). Details of the accuracy are shown in **Table 1**.

Subject-Independent Classification Analysis

With the main aim of comparing the effects of individual characteristics on the accuracy of classification, we also performed a classification with the data for each channel for all subjects. The results are shown in **Figure 10**. The accuracies

in classifying the driver's mental workload from each channel were found to be in the range 84.9 to 89.5%, and the accuracy in predicting testing data increased from 84.08 to 89.91%. These results indicated that individual characteristics affected the accuracy of classification of the mental workload.

Channel-Independent Classification Analysis

Before combining the data from all channels from the NIRS together, we performed a multiple comparison of oxy-HB and deoxy-HB levels for all the channel data from each subject in the same task by means of a Bonferroni test. The results showed that there was no significant difference between the data from the various channels in the same task (the p -value in all cases was >0.05). According to the result, we decided to combine the data from all the channels together when checking the accuracy of classification.

First, the data from the four channels for each subject were combined to test the accuracy of classification. The results of this classification are shown in **Figure 11**. The accuracy of classification increased from 80.8 to 88.6%, and the accuracy on the testing data increased from 83.4 to 88.2%. This shows that acceptably accurate results of classification can be obtained simply by combining the data for the various channels.

Subject-Independent + Channel-Independent Classification Analysis

Finally, to examine the potential for creating a system real-time classification of driver cognitive load to prevent accidents, we

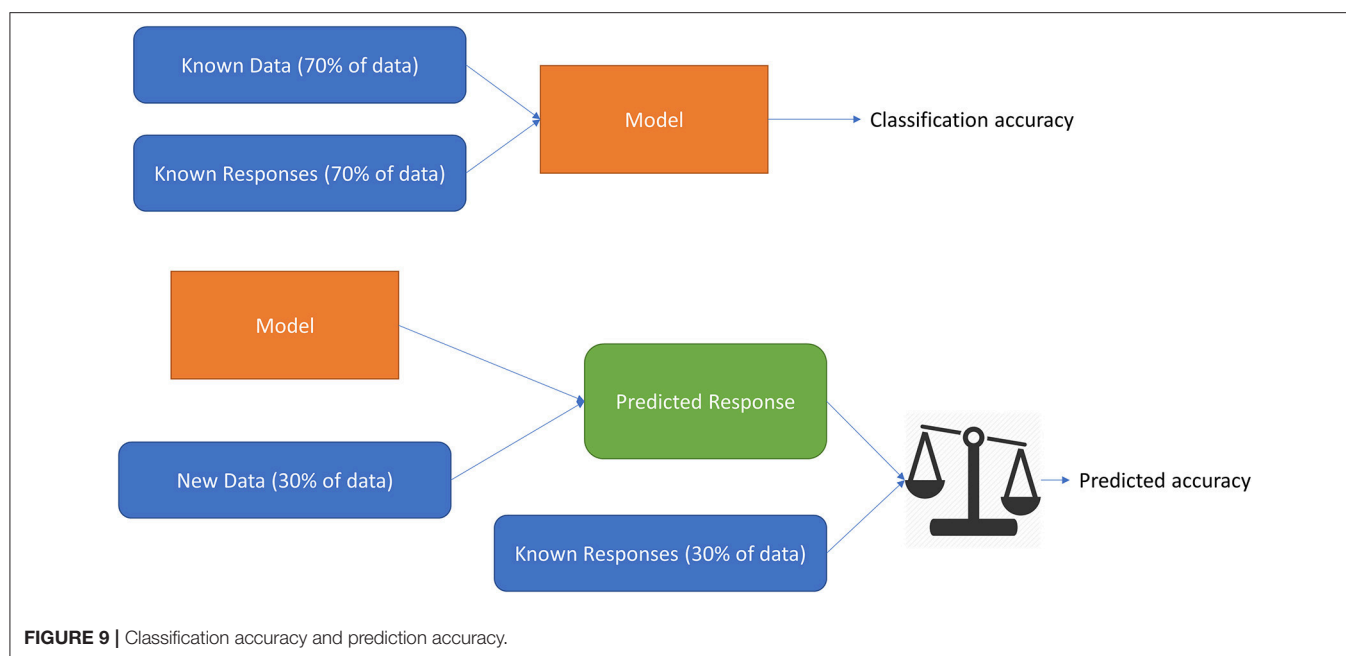


TABLE 1 | The classification accuracy (the accuracy on the testing data) (%).

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
Channel 1	94.0 (93.3)	95.4 (96.1)	91.4 (91.3)	92.1 (91.3)	89.7 (91.1)
Channel 2	88.9 (87.9)	92.1 (88.3)	89.7 (89.9)	92.6 (92.3)	87.3 (91.0)
Channel 3	92.9 (90.4)	89.8 (90.6)	91.9 (88.7)	85.2 (82.3)	83.0 (86.4)
Channel 4	94.4 (92.5)	89.9 (90.0)	87.8 (89.5)	85.7 (85.4)	81.3 (82.2)

combined the data from all the channels and all the subjects and we used the combined data to train the classification. We then examined the effect of this combination on the accuracy of classification and on the accuracy of prediction. As expected, the accuracy of classification was 82.9% and the accuracy of prediction was 82.71%, which were similar values to those previously obtained. This showed that the position of the channel on the forehead did not have a significant effect on the accuracy, and it confirmed that a compact NIRS device can capture the cognitive distraction of a driver, even under naturalistic conditions.

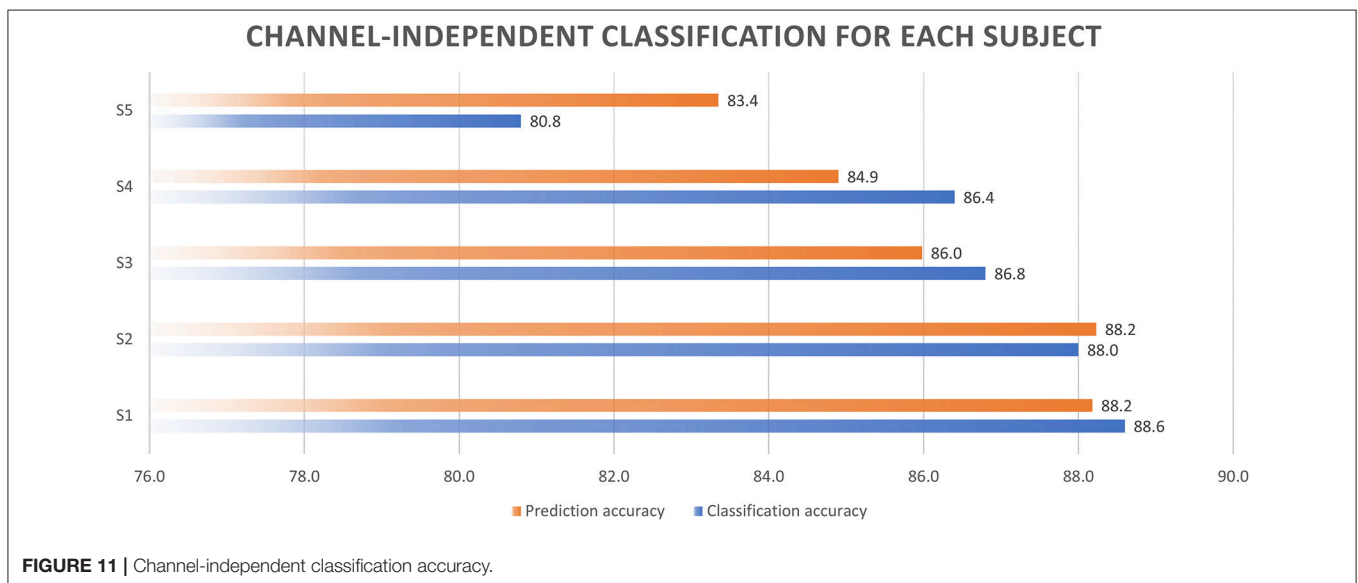
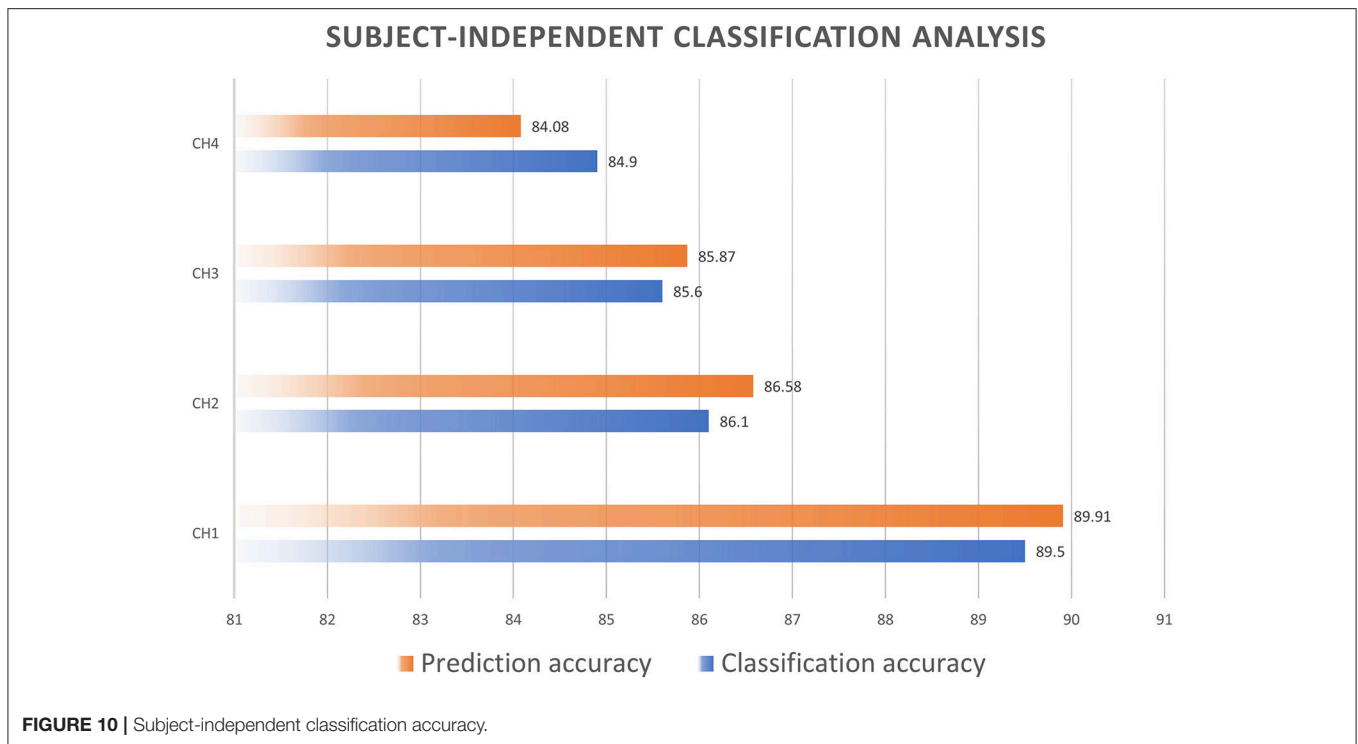
Compare with the result done by Naseer and Hong (2015) and Hong et al. (2015), which was used fNIRS signal, and then show the possibility of the hybrid feature extraction to classification with motor imagery tasks. The highest classification accuracy was around 77.5% with multi-class LDA. Furthermore, the classification of the right—and left—wrist motor imageries also done by Naseer and Hong (2013) using fNIRS information. By reducing the time span within the task period to 2–7 s, the accuracy for classification was increased to 77.56 and 87.28%. Here, we performed the classification with a machine learning algorithm and get better results with accuracy around 82.9% (in the case of subject-independent + channel-independent). The difference of the accuracy may depend on the difference in mental workload level, different experiment condition, and so on.

DISCUSSION

Model Selection for Classification of NIRS Data

As we have mentioned above, most previous researchers have used an SVM (Devos et al., 2009; Ichikawa et al., 2014; Aghajani et al., 2017), linear discriminant analysis (Luu and Chau, 2009), the hidden Markov model (Sitaram et al., 2007; Zimmermann et al., 2013), or artificial neural networks (Chan et al., 2012; Thanh Hai et al., 2013) to classify mental workload. In this study, we trained the data with some new models, such as the k-nearest neighbors model (k-NN) and the bagged tree (random forests) model, depending on the number of samples. Our results showed that the random forests model provided the highest accuracy, even with large numbers of samples, whereas the cubic SVM showed the worst performance (The average accuracies of each model in all previous analyses are shown in Figure 12). In addition, the k-NN model is also suitable for classifying mental workloads by using NIRS data because of its ability to maintain similar levels of accuracy even when the sample size changes markedly.

The SVM, a well-known method that has been previously applied in various classifications of NIRS data, showed very good performance with small numbers of samples. However, when there were large numbers of samples, the SVM was very slow and its accuracy was low compared with other methods. For example, in the case of a channel-independent test for Subject 1, where the sample size was over 15,000 samples, the accuracy of the SVM was 68.2%, compared with 87.3% for the random forests method and 88.6% for the k-NN classifier. In addition, the SVM took 1,341 s to perform the classification, whereas the k-NN classifier required only 88.6 s. Similar effects were observed in subject-dependent and subject-independent classifications. The sudden reduction in the accuracy of the SVM might arise from



differences in the data after data combination, as well as from effects of individual characteristics.

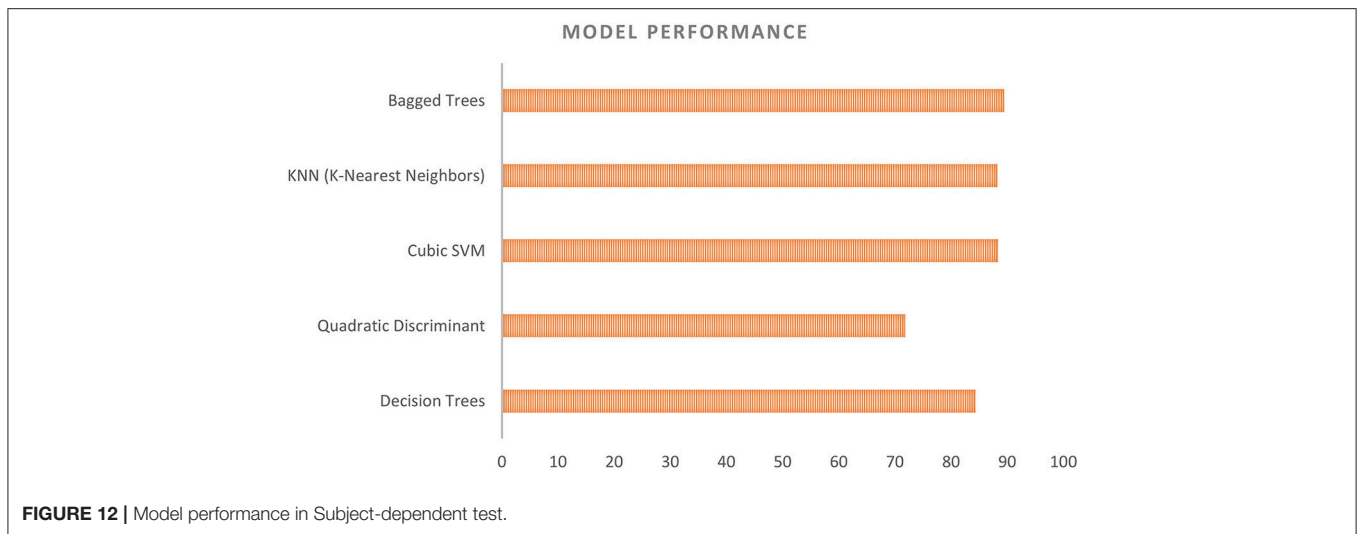
On the other hand, the way to select testing sample also plays a very importance role in the classification. In our case, we selected the testing sample following $X \rightarrow X \rightarrow X \rightarrow Y \rightarrow X \rightarrow X \rightarrow Y \rightarrow X \dots$, where X is a sample taken for the 75% and Y is a testing sample. It may make the nearest neighbor classifier that will perform very well, probably because the variation from Y to its neighbors in time (the X before and after Y) will be very low.

We also believe that for higher numbers of subjects, a lower accuracy is attained for subject-independent classification.

Consequently, for large numbers of subjects, the machine-learning algorithm should be changed to a deep learning or convolution neural-network algorithm, which can still show good performance with large quantities of data.

The Potential For Using NIRS Data to Evaluate Levels of Driver Mental Workload

This study is one example of the application of machine learning in classifying driver mental workload from data obtained with a simple commercial NIRS device, which has a high potential



for routine use with drivers because of its acceptable price. However, we believe that the use of a combination of channels for NIRS is necessary because signal losses tended to occur often under naturalistic conditions. In some cases, however, one or two channels provided sufficient signals due to the activities of the driver.

The ability to unobtrusively detect changes in mental workload is relevant since high levels of cognitive load can reduce a driver's ability to anticipate and respond to emergent dangers in the driving environment. Broadly considered, these findings suggest various lines of potential research related to the development of advanced driver assistance systems (e.g., a new method to prevent accidents by detecting levels of mental workload that may lead to cognitive distraction), basic human factors insight (exploring the relationship between individual characteristics and objective indicators of mental workload), and mathematical modeling (combining channel, improve accuracy by applying different technical).

In conclusion, as previously suggested (Kopton and Kenning, 2014; Unni et al., 2017), simple NIRS has considerable potential for capturing driver mental workload, especially under naturalistic conditions.

LIMITATIONS

The relatively small sample size used in this study (a total of 5 subjects including one female and four males) could be considered a limitation. While we believe that the NIRS signals were found to be predictive for this small sample under our specific set of conditions, it would be worthwhile to repeat the experiment with a larger sample and a wider range of conditions (e.g., driving track, time of day, gender balance, driver skill level, age, etc.).

CONCLUSIONS

Our study suggested that it is possible to use NIRS data to classify levels of driver mental workload, even in a naturalistic situation. Furthermore, a simple combination of forehead channels was shown to provide acceptably high accuracies of classification. While the fNIRS sensors employed in this study required contact with the participants' skin, the lightweight ball cap configuration was much less intrusive than more traditional electrophysiological measures used in related work. We also confirmed the potential of using machine learning (channel- and subject-independent) to predict possible driver cognitive distraction, a critical factor in road safety.

AUTHOR CONTRIBUTIONS

AL, HA, FM, and KI generated the idea. AL and HA planned the experiments, collected and analyzed the data, and prepared the manuscript for this study.

ACKNOWLEDGMENTS

This study was partially supported by DENSO CORPORATION and by the Center of Innovation Program (Nagoya University COI: Mobility Innovation Center) from Japan Science and Technology Agency.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2018.00431/full#supplementary-material>

Video 1 | NIRS toolbox for classifying driver mental workload.

REFERENCES

- Abdelnour, A. F., and Huppert, T. (2009). Real-time imaging of human brain function by near-infrared spectroscopy using an adaptive general linear model. *Neuroimage* 46, 133–143. doi: 10.1016/j.neuroimage.2009.01.033
- Aghajani, H., Garbey, M., and Omurtag, A. (2017). Measuring mental workload with EEG+fNIRS. *Front. Hum. Neurosci.* 11:359. doi: 10.3389/fnhum.2017.00359
- Angell, L., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W., et al. (2006). *Driver Workload Metrics Project*. Task 2 final Report. Security, 460.
- Angell, L., and Perez, M. (2015). *Identification of Cognitive Load in Naturalistic Driving*. Blacksburg, VA: Virginia Tech Transportation Institute.
- Anh Son, L., Aoki, H., and Suzuki, T. (2017a). Evaluation of driver distraction with changes in gaze direction based on a vestibulo-ocular reflex model. *J. Transp. Technol.* 7, 336–350. doi: 10.4236/jtts.2017.73022
- Anh Son, L., Hamada, H., Inagami, M., Suzuki, T., and Aoki, H. (2017d). “Effect of mental workload and aging on driver distraction based on the involuntary eye movement,” in *Advances in Human Aspects of Transportation, Vol. 484, Advances in Intelligent Systems and Computing*, eds N. Stanton, S. Landry, G. Di Bucchianico, and A. Vallicelli (Cham: Springer), 349–359. doi: 10.1007/978-3-319-41682-3_30
- Anh Son, L., Hiroto, H., Tatsuya, S., and Hirofumi, A. (2017b). “Driver distraction evaluation using reflex eye movement simulation,” in *Proceedings of the Future Active Safety Technology Towards Zero Traffic Accidents Conference* (Nara).
- Anh Son, L., Inagami, M., Hiroto, H., Tatsuya, S., and Hirofumi, A. (2018). Towards online detection of driver distraction: Eye-movement simulation based on a combination of vestibulo-ocular reflex and optokinetic reflex models. *Transp. Res. Part F Traffic Psychol. Behav.* doi: 10.1016/j.trf.2018.02.019
- Anh Son, L., Inagami, M., Suzuki, T., Aoki, H., and Obinata, G. (2016). “Driver’s internal state estimation based on the eye movement,” in *Proceedings of the 2016 International Conference University Village* (Nagoya).
- Anh Son, L., Makoto, I., Hiroto, H., Tatsuya, S., and Hirofumi, A. (2017c). “Evaluation of cognitive distraction in a real vehicle based on the reflex eye movement,” in *Proceedings of the 2017 Japan Society Automotive Engineers Annual Congress* (Yokohama).
- Anh Son, L., Makoto, I., Hiroto, H., Tatsuya, S., and Hirofumi, A. (2017e). The effect of visual stimulus on voluntary eye movement based on a VOR / OKR Model. *Int. J. Automot. Eng.* 8, 37–44. doi: 10.20485/jsaeijae.8.2_37
- Aoki, H., Anh Son, L., Hamada, H., Suzuki, T., and Obinata, G. (2015). “Quantitative evaluation of mental workload with voluntary eye movements by means of the vestibulo-ocular reflex,” in *Proceedings of the 2015 Japan Society Automotive Engineers Annual Congress* (Sapporo), 780–785.
- Backs, R. W., and Walrath, L. C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Appl. Ergon.* 23, 243–254.
- Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., and Montanari, R. (2011). Driver workload and eye blink duration. *Transp. Res. Part F Traffic Psychol. Behav.* 14, 199–208. doi: 10.1016/j.trf.2010.12.001
- Chan, J., Power, S., and Chau, T. (2012). Investigating the need for modelling temporal dependencies in a brain-computer interface with real-time feedback based on near infrared spectra. *J. Near Infrared Spectrosc.* 20, 107–116. doi: 10.1255/jnirs.971
- Cinaz, B., Arnrich, B., La Marca, R., and Tröster, G. (2013). Monitoring of mental workload levels during an everyday life office-work scenario. *Pers. Ubiquitous Comput.* 17, 229–239. doi: 10.1007/s00779-011-0466-1
- Cooper, R. J., Selb, J., Gagnon, L., Phillip, D., Schytz, H. W., Iversen, H. K., et al. (2012). A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. *Front. Neurosci.* 6:147. doi: 10.3389/fnins.2012.00147
- Devos, O., Ruckebusch, C., Durand, A., Duponchel, L., and Huvenne, J. P. (2009). Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. *Chemom. Intell. Lab. Syst.* 96, 27–33. doi: 10.1016/j.chemolab.2008.11.005
- Dong, Y., Hu, Z., Uchimura, K., and Murayama, N. (2011). Driver inattention monitoring system for intelligent vehicles: a review. *IEEE Trans. Intell. Transp. Syst.* 12, 596–614. doi: 10.1109/TITS.2010.2092770
- Engström, J., Johansson, E., and Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transp. Res. Part F Traffic Psychol. Behav.* 8, 97–120. doi: 10.1016/j.trf.2005.04.012
- Henelius, A., Hirvonen, K., Holm, A., Korpela, J., and Muller, K. (2009). Mental workload classification using heart rate metrics. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2009, 1836–1839. doi: 10.1109/IEMBS.2009.5332602
- Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., and Schultz, T. (2014). Mental workload during n-back task—quantified in the prefrontal cortex using fNIRS. *Front. Hum. Neurosci.* 7:935. doi: 10.3389/fnhum.2013.00935
- Hong, K. S., Naseer, N., and Kim, Y. H. (2015). Classification of prefrontal and motor cortex signals for three-class fNIRS-BCI. *Neurosci. Lett.* 587, 87–92. doi: 10.1016/j.neulet.2014.12.029
- Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2016). HomER: a review of time-series analysis methods for nearinfrared spectroscopy of the brain. *Appl. Opt.* 48, 280–298. doi: 10.1016/j.drugalcdep.2008.02.002.A
- Ichikawa, H., Kitazono, J., Nagata, K., Manda, A., Shimamura, K., Sakuta, R., et al. (2014). Novel method to classify hemodynamic response obtained using multi-channel fNIRS measurements into two groups: exploring the combinations of channels. *Front. Hum. Neurosci.* 8:480. doi: 10.3389/fnhum.2014.00480
- Jorna, P. G. (1992). Spectral analysis of heart rate and psychological state: a review of its validity as a workload index. *Biol. Psychol.* 34, 237–257. doi: 10.1016/0301-0511(92)90017-O
- Kirilina, E., Yu, N., Jelzow, A., Wabnitz, H., Jacobs, A. M., and Tachtsidis, I. (2013). Identifying and quantifying main components of physiological noise in functional near infrared spectroscopy on the prefrontal cortex. *Front. Hum. Neurosci.* 7:864. doi: 10.3389/fnhum.2013.00864
- Klingner, J. (2010). “Fixation-aligned pupillary response averaging,” in *Proceedings of the 2010 Symposium Eye-Tracking Research & Applications* (Austin, TX), 275–282.
- Klingner, J., Kumar, R., and Hanrahan, P. (2008). “Measuring the task-evoked pupillary response with a remote eye tracker,” in *Proceedings of the 2008 Symposium Eye Tracking Research & Applications-ETRA’08* 1, 69 (Savannah, GA).
- Kohri, S., Hoshi, Y., Tamura, M., Kato, C., Kuge, Y., and Tamaki, N. (2002). Quantitative evaluation of the relative contribution ratio of cerebral tissue to near-infrared signals in the adult human head: a preliminary study. *Physiol. Meas.* 23, 301–312. doi: 10.1088/0967-3334/23/2/306
- Kopton, I. M., and Kenning, P. (2014). Near-infrared spectroscopy (NIRS) as a new tool for neuroeconomic research. *Front. Hum. Neurosci.* 8:549. doi: 10.3389/fnhum.2014.00549
- Le, A. S., and Aoki, H. (2018). Smart eye and polhemus data for vestibulo-ocular reflex and optokinetic reflex model. *Data Br.* 18, 882–885. doi: 10.1016/j.dib.2018.03.097
- Lee, D. H., and Park, K. S. (1990). Multivariate analysis of mental and physical load components in sinus arrhythmia scores. *Ergonomics* 33, 35–47. doi: 10.1080/00140139008927092
- Liu, T., Pelowski, M., Pang, C., Zhou, Y., and Cai, J. (2016). Near-infrared spectroscopy as a tool for driving research. *Ergonomics* 59, 368–379. doi: 10.1080/00140139.2015.1076057
- Liu, Z., Zhang, M., Xu, G., Huo, C., Tan, Q., Li, Z., et al. (2017). Effective connectivity analysis of the brain network in drivers during actual driving using near-infrared spectroscopy. *Front. Behav. Neurosci.* 11:11. doi: 10.3389/fnbeh.2017.00211
- Luu, S., and Chau, T. (2009). Decoding subjective preference from single-trial near-infrared spectroscopy signals. *J. Neural Eng.* 6:016003. doi: 10.1088/1741-2560/6/1/016003
- McBride, D. K., and Schmorow, D. (eds.). (2005). *Quantifying Human Information Processing*. Lanham, MA: Lexington Books.
- Mehler, B., Reimer, B., and Coughlin, J. F. (2012). Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups. *Hum. Factors* 54, 396–412. doi: 10.1177/0018720812442086
- Meshkati, N. (1988). Heart rate variability and mental workload assessment. *Hum. Ment. Workload* 52, 101–115. doi: 10.1016/S0166-4115(08)62384-5
- Naseer, N., and Hong, K.-S. (2015). fNIRS-based brain-computer interfaces: a review. *Front. Hum. Neurosci.* 9:3. doi: 10.3389/fnhum.2015.00003

- Naseer, N., and Hong, K. S. (2013). Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface. *Neurosci. Lett.* 553, 84–89. doi: 10.1016/j.neulet.2013.08.021
- NHTSA (2012). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. *Dep. Transp.* 77, 1–34.
- NHTSA (2015). *Research Note Distracted Driving 2013*. Dot Hs 812 132, 1–8. Available online at: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812381>.
- Obinata, G., Fukuda, K., Moriyama, A., Tokuda, S., and Kim, Y. W. (2010). Evaluating the influence of distractions to drivers based on reflex eye movement model. *IFAC* doi: 10.3182/20100831-4-FR-2021.00045
- Obinata, G., Tokuda, S., Fukuda, K., and Hamada, H. (2009). “Quantitative evaluation of mental workload by using model of involuntary eye movement,” in *Engineering Psychology and Cognitive Ergonomics SE - 24 Lecture Notes in Computer Science*, ed D. Harris (Berlin; Heidelberg: Springer), 223–232. doi: 10.1007/978-3-642-02728-4_24
- Obinata, G., Usui, T., and Shibata, N. (2008). On-line method for evaluating driver distraction of memory-decision workload based on dynamics of vestibulo-ocular reflex. *Rev. Automot. Eng.* 29, 627–632. doi: 10.11351/jsaereview.29.627
- Orino, Y., Yoshino, K., Oka, N., Yamamoto, K., Takahashi, H., and Kato, T. (2015). Brain activity involved in vehicle velocity changes in a sag vertical curve on an expressway: vector-based functional near-infrared spectroscopy study. *J. Transport. Res. Board* 2518, 18–26. doi: 10.3141/2518-03
- Palinko, O., and Kun, A. (2012). “Exploring the influence of light and cognitive load on pupil diameter in driving simulator studies,” in *Proceedings Symposium Eye Tracking Research & Application*, 413–416.
- Pierce, E. T. (2009). *Mental Workload Measurement Using the Intersaccadic Interval*. ProQuest Diss. Theses, 160. Available online at: [http://ezproxy.net.ucf.edu/login?url=http://search.proquest.com/docview/858993124?accountid=10003%5Chttp://sfx.fcla.edu/ucf?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+\\$&\\$&\\$theses&sid=ProQ:ProQuest+\\$Dissertations+\\$&\\$&\\$T](http://ezproxy.net.ucf.edu/login?url=http://search.proquest.com/docview/858993124?accountid=10003%5Chttp://sfx.fcla.edu/ucf?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre=dissertations+$&$&$theses&sid=ProQ:ProQuest+$Dissertations+$&$&$T).
- Porges, S. W., and Byrne, E. A. (1992). Research methods for measurement of heart rate and respiration. *Biol. Psychol.* 34, 93–130. doi: 10.1016/0301-0511(92)90012-J
- Rizzo, M., and Hurtig, R. (1987). Looking but not seeing: attention, perception, and eye movements in simultanagnosia. *Neurology* 37, 1642–1648.
- Ryu, K., and Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int. J. Ind. Ergon.* 35, 991–1009. doi: 10.1016/j.ergon.2005.04.005
- Schwalm, M., Keinath, A., and Zimmer, H. D. (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. *Hum. Factors Assist. Autom.* 1–13.
- Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., et al. (2007). Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *Neuroimage* 34, 1416–1427. doi: 10.1016/j.neuroimage.2006.11.005
- Son, L. A., and Hirofumi, A. (2018). *Evaluating Driver Mental Workload in a Realistic Condition Using Eye Movement Sensor*. Nagoya: Tokai Engineering Complex. doi: 10.1299/jsmetokai.2018.67.408
- Son, L. A., Suzuki, T., and Aoki, H. (2018). Evaluation of Cognitive Distraction in a Real Vehicle Based on the Reflex Eye Movement. *Int. J. Automot. Eng.* 9, 1–8. doi: 10.20485/jsaiejae.9.1_1
- Thanh Hai, N., Cuong, N. Q., Dang Khoa, T. Q., and Van Toi, V. (2013). Temporal hemodynamic classification of two hands tapping using functional near-infrared spectroscopy. *Front. Hum. Neurosci.* 7:516. doi: 10.3389/fnhum.2013.00516
- Tokuda, S., Palmer, E., Merkle, E., and Chaparro, A. (2009). Using saccadic intrusions to quantify mental workload. *Hum. Factors Ergon. Soc. Annu. Meet. Proc.* 53, 809–813. doi: 10.1518/107118109X12524442637660
- Tsai, Y. F., Viirre, E., Strychacz, C., Chase, B., and Jung, T.-P. (2007). Task performance and eye activity: predicting behavior relating to cognitive workload. *Aviat. Space. Environ. Med.* 78, 176–185.
- Tsunashima, H., and Yanagisawa, K. (2009). Measurement of brain function of car driver using functional near-infrared spectroscopy (fNIRS). *Comput. Intell. Neurosci.* 2009:164958. doi: 10.1155/2009/164958
- Unni, A., Ihme, K., Jipp, M., and Rieger, J. W. (2017). Assessing the driver's current level of working memory load with high density functional near-infrared spectroscopy: a realistic driving simulator study. *Front. Hum. Neurosci.* 11:167. doi: 10.3389/fnhum.2017.00167
- Veltman, J. A., and Gaillard, A. W. (1996). Physiological indices of workload in a simulated flight task. *Biol. Psychol.* 42, 323–342. doi: 10.1016/0301-0511(95)05165-1
- Wang, Y., Reimer, B., Dobres, J., and Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transport. Res. Part F* 26, 227–237. doi: 10.1016/j.trf.2014.08.003
- Yoshino, K., Oka, N., Yamamoto, K., Takahashi, H., and Kato, T. (2013). Functional brain imaging using near-infrared spectroscopy during actual driving on an expressway. *Front. Hum. Neurosci.* 7:882. doi: 10.3389/fnhum.2013.00882
- Yoshino, K., Oka, N., Yamamoto, K., Takahashi, H., and Kato, T. (2015). Correlation of prefrontal cortical activation with changing vehicle speeds in actual driving: a vector-based functional near-infrared spectroscopy study. *Front. Hum. Neurosci.* 9:665. doi: 10.3389/fnhum.2015.00665
- Zimmermann, R., Marchal-Crespo, L., Edelmann, J., Lamberg, O., Fluet, M. C., Riener, R., et al. (2013). Detection of motor execution using a hybrid fNIRS-biosignal BCI: A feasibility study. *J. Neuroeng. Rehabil.* 10:4. doi: 10.1186/1743-0003-10-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Le, Aoki, Murase and Ishida. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.