Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Effects of Daily Stress in Mental State Classification

Soyeon Park, Student Member, IEEE, and Suh-Yeon Dong, Member, IEEE
Department of Information Technology Engineering, Sookmyung Women's University, Seoul 04310, Republic of Korea

Corresponding author: Suh-Yeon Dong (e-mail: sydong@sookmyung.ac.kr).

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (Nos. NRF-2019R1F1A1058011 and NRF-2018R1A4A1025559), and Sookmyung Women's University Research Grants (1-1803-1249)

ABSTRACT An external stimulus, event, or environment that stresses an individual is called a stressor. Many mental stress detection studies have been focused on the discrimination of the mental state with and without the experimental stressor. However, the mental state in the absence of experimental stressors may not represent accurately the nonstress (baseline) state because people inherently experience considerable stress in their daily lives. Therefore, we assumed that stress detection could be improved more accurately by considering the daily stress. In this study, functional near-infrared spectroscopy (fNIRS) was measured in 41 healthy participants to quantify their prefrontal cortical oxygenation during performing a cognitive task as an experimental stressor, considering individual daily stress level based on the self-report. We then extracted six signal features, including the slope, mean, standard deviation, peak, skewness, and kurtosis of the oxygenated hemoglobin concentration. Using various feature combinations and time windows, we successfully managed to classify daily stress (high/low) and the mental state (task/rest) with support vector machine classifiers. Specifically, individual daily stress level can be easily discriminated with signal features from fNIRS. Moreover, mental state classification performance improved significantly when the daily stress level was handled separately. The findings of this study show the feasibility of the fNIRS-based daily stress classification and can be used in the future to design a robust mental stress management system for the assessment of daily stress in individuals.

INDEX TERMS daily stress, functional near-infrared spectroscopy, stroop word color task, mental stress classification.

I. INTRODUCTION

Younger generations in their 20s are stressed every day owing to various academic, employment and office syndrome problems they face [1-2]. However, it is difficult for younger generations to be aware of the severity of mental stress on their health. Long-term exposure to stress can cause chronic mental or physical disorders, such as depression, heart disease, obesity, and diabetes [3–5]. Therefore, it is very important to evaluate stress at an early stage and enforce stress management programs as stress can cause problems even for healthy and younger people. Stress is a process of changes in the body or mind to be alert and counteract external or internal events, commonly referred to as stressors. Therefore, it is possible to quantify it by measuring changes in perceptual, behavioral, and physiological responses caused by stressors.

The most objective way is to measure specific hormones, such as cortisol and alpha amylase that are released in response to stress that are measurable in urine, blood, and saliva [6–8]. Another noninvasive and easier way is the monitoring of physiological responses, such as blood pressure, skin conductivity, electrocardiogram (ECG), electroencephalogram (EEG) [9], and others [10–11]. Among them, cortical physiological responses like EEG or functional near-infrared spectroscopy (fNIRS) have become extensively used in the estimation of stress with the development of wearable measurement systems [12–13]. A fNIRS is an optical imaging technique that uses near-infrared light to measure oxygenated and deoxygenated hemoglobin concentration changes in cortical brain. With recent advances in portable and multichannel fNIRS hardware, fNIRS is considered to be a cost-efficient and lightweight measurement system in comparison with other non-invasive brain imaging technique [14]. fNIRS-based brain–computer interface (BCI) studies have shown promising results in cognitive task [15] and depressive state classifications [16].

Most BCI classifications have utilized machine learning techniques, such as linear discriminant analysis (LDA) [17–19], support vector machine (SVM) [20–22], deep learning—such as convolutional neural networks [23]—and long short-term memory [24] —and cascade CNN-LSTM [25]. In recent fNIRS-based BCI studies, mental state classification performance varied depending on the type of mental task and classifier, but its accuracy was typically in the 60 to 90% range. SVM was preferably chosen owing to its reliable performance. For example, mental state classifications yielded outcomes equal to 70.64% [23], 84.44% [20], 83.56% [26], and 91% [21] . Cognitive tasks have been frequently used to induce mental stress. Typical examples include the mental arithmetic task [27], Trier Social Stress Test [28], and Stroop Color Word Task (SCWT) [29–30]. A SCWT is a reliable and valid task that causes mental stress, so we used this task as an experimental stress factor.

In many previous studies, cognitive tasks were considered as experimental stress factors [13], [31]. Conversely, in the resting state, an explicit task is not being performed and the participants were asked not to move and to remain relaxed to evaluate the condition that represented their stress-free states. However, we questioned whether it was possible to completely rule out the effects of daily stress. In some previous studies, salivary samples were obtained to quantify specific hormone levels as reference values to ensure that experimental stressors can induce mental stress. In [32], salivary cortisol levels were compared before, right after, and 20 min after the cognitive tasks, while in [13] salivary alpha amylases were measured through the baseline, task, and recovery phases. In another previous study, participants were required to follow a well-controlled psychophysiological protocol, such as an overnight stay at the research center, light breakfast with no caffeinated beverages, and cognitive tasks in the morning [33]. However, salivary cortisol samples can vary depending on the environment and time in which saliva is collected, and the measured value of cortisol is susceptible to oral conditions or diet. [34-35]. Moreover, even in the well-controlled protocol, it is not evident that the resting state can represent the stress-free condition. In other words, if the resting state was affected by the individual’s daily stress, it may have also adversely affected the performance of the mental state classification. We anticipate that if we could discriminate an individual's daily stress, we could obtain improved mental state classification performances.

Therefore, the specific goals of the present study are

- To discriminate self-reporting individual daily stress levels that are not induced by any experimental stress factor
- To prove that it is effective to consider daily stress in the mental state classification
- To investigate the features that mainly contributes in each classification process

To achieve the first goal, we designed an experimental paradigm to repeatedly report individual daily stress levels for two weeks and recorded fNIRS at the high- and low-stressful days selected in the 2nd week. In order to obtain fNIRS data on the days with the high and low stress, the range of each one’s responded stress indices during the 1st week was first checked. After that, based on the individual stress range, days with high- and low-stressful days were selected for each individual in the 2nd week. Therefore, fNIRS data were finally acquired for the selected two days, and a two-class classification (high/low) was performed with a variety of feature combinations and time windows. To achieve the second goal, we compared mental state classification (task/rest) performances with the use of models that considered daily stress versus those that did not. By the last goal, the features were identified that mainly contributed in each classification scheme.

II. MATERIALS AND METHODS

A. PARTICIPANTS

The study was conducted with 41 healthy female university students (mean (M) ± standard deviation (SD) age 21.93 ± 1.69). Patients with a past history of heart disease and neuropsychiatric disorders, those with current medications, or pregnant or prospective women, were excluded from this study. All experimental procedures involving human subjects were approved by the institutional review board at the Sookmyung Women’s University (SMWU-1902-HR-148-01). The entire experiment procedure was verbally introduced to the participants, and written informed consent was obtained.

B. EXPERIMENTAL PROCEDURE

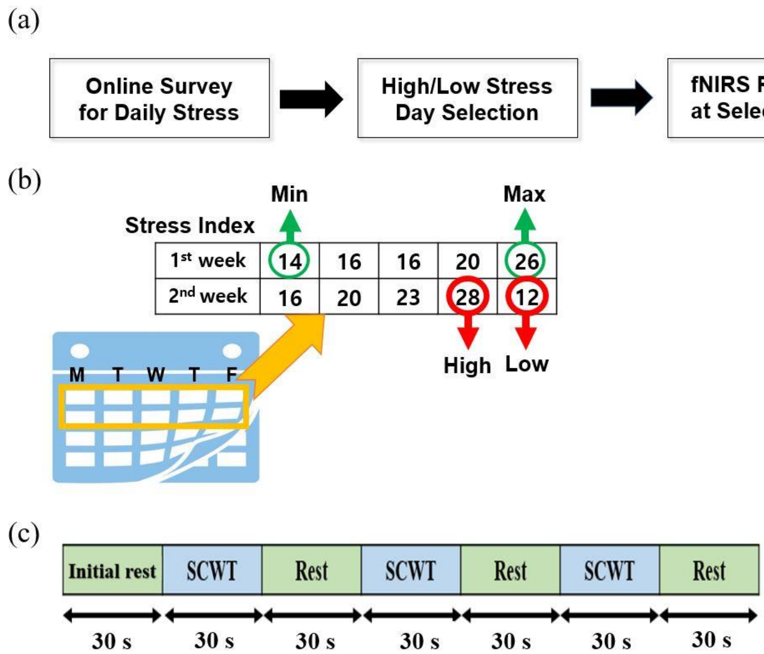

FIGURE 1. (a) Entire experiment procedure (b) Example stress indices used to explain how to select daily stress (high/low) days (c) Task protocol. In the fNIRS record, three repetitions of a Stroop Color Word Task (SCWT) with initial and intermediate rests are presented.

As shown in Fig. 1(a), daily stress was assessed with online subjective questionnaires. Participants were asked to respond to questionnaires that comprised 15 items based on their moods and behaviors during the 24 h periods that preceded the test, and repeat it everyday for two weeks. The stress questionnaire was adapted to university students regarding stress-relevant questionnaire items [36–37]. Questionnaire items were selected and adapted from the stress-response scale, stress-awareness scale, and included university-student-related questions, such as for example, relationships, academics, etc. Answers were based on a 5-point scale: very unlikely (4), somewhat unlikely (3), neutral (2), somewhat likely (1), and very likely (0). The total score from all 15 answers was determined as a stress index for that day. Because reported stress indices varied from person to person, we thought it was inappropriate to set a threshold and apply it equally to everyone. Therefore, we identified the individual criterion through the range of individual stress indices from the 1st week's responses. In the mornings of the days during the 2nd week, according to the personal criteria of the 1st week, the high- or low-daily-stressful days were selected. The approach used to determine high- and low-daily-stressful days with example indices is shown in Fig. 1(b). For example, if 14 is the minimum and 26 is the maximum value among the responded stress indices in the 1st week, we assume that this person has a stress index close to 14 for low stress and 26 for high stress. In the 2nd week, when the stress index of the self-questionnaire on the day is 16 or 20, it will be skipped. If the index is above 26 (for example, 28 in fig. 1(b)), this day is selected as the day with high stress. Similarly, the day with the index of 12 is selected as the day with low stress since it is less than the minimum value. In other words, the individual stress index of the 1st week is used as a standard, and the stress index of the day of the 2nd week determines whether to select a day with high or low stress. If any index in the 2nd week did not meet the criterion for high and low stress, stress report continued for few more days. If any individual's stress indices did not change significantly for two weeks, that participant was dropped out. Once the day was selected as a high- or low-daily-stress day, the experimenter asked the participant to visit the laboratory to record fNIRS measurements in the afternoon of that day. For consistency, the survey responses were collected at the same time in the morning for two weeks. Additionally, for the same reason, the fNIRS recording time was always fixed in the afternoon.

When a subject visited the laboratory, the entire experiment procedure, including the fNIRS recording, was verbally introduced again. After the participant wore the fNIRS device on the head, the cognitive task was performed. After the initial rest period that spanned 30 s, the 30 s SCWT and the 30 s rest periods alternated and were repeated three times. Therefore, the entire experiment lasted 210 s in Fig. 1(c). SCWT was presented using E-Prime software (E-Prime 3.0, Psychology Software Tools Inc., Sharpsburg, MD, USA). Participants were required to press a keyboard (left or right) to respond to each SCWT trial. To familiarize them with the experiment and the task, all the recruited participants conducted a practice experiment in the 1st week. Data recorded in the practice experiment were not used for further analysis. In summary, each participant participated in three fNIRS recordings including the training recording, but only analyzed two main recordings: one on a high daily stress day and the other on a low daily stress day.

C. DATA RECORDING AND PREPROCESSING

Cortical hemodynamic variations in the prefrontal cortex (PFC) region were recorded using a high-density NIRS device (NIRSIT, OBELAB, Seoul, Korea). Fig. 2 shows the setup of wearable fNIRS device and the sources and detector array of the device. The sensor was composed of 24 dual-wavelength laser diodes (780/850 nm) and 32 detectors separated by a 1.5 cm unit distance [38]. A 3 cm distance separated the laser and detector pairs at 48 sensing areas. Thus, we analyzed 48 channels. The optical signal variation of each channel was sampled at 8.138 Hz, and the threshold signal-to-noise ratio (SNR) was 30 dB. Detected light signals in each wavelength were filtered by low-pass filtering discrete cosine transform (DCT) 0.1 Hz and high-pass filtering DCT 0.005 Hz to remove physiological and environmental noise. And then, converted into hemodynamic parameters (oxygenated hemoglobin concentration) using the differential pathlength factor (DPF) method. The DPF values of 780nm and 850nm are 5.075 and 4.64 in respectively [39-40]. Relative hemodynamic changes in each channel during each trial of each task were extracted using the modified Beer–Lambert law (MBLL) [41].

As shown in Fig. 2(b) and 2(c), 48 channels at a source–detector distance of 3 cm from the predefined array were

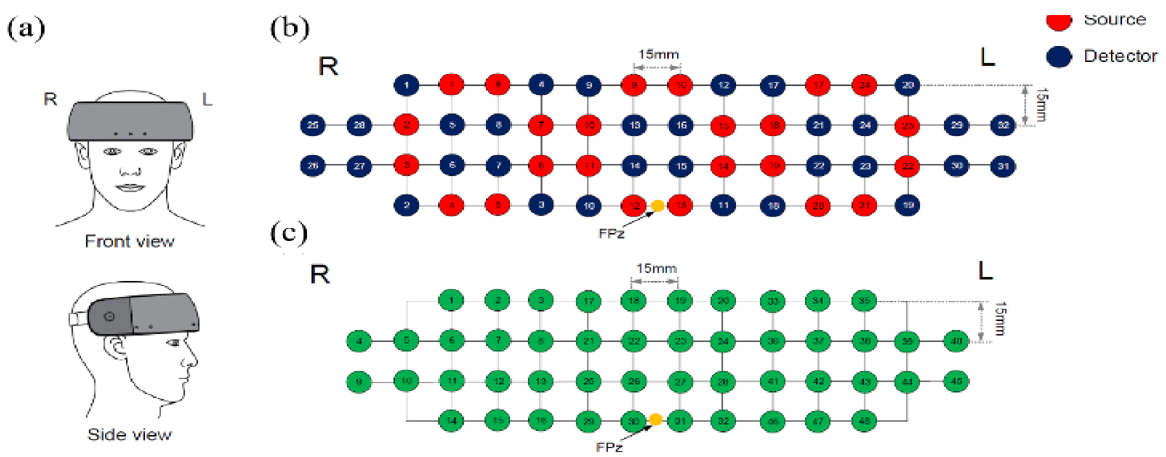

FIGURE 2. Data acquisition configuration, (a) device setup, (b) optode placement, and (c) channel position.

acquired. This was adequate for an in-depth penetration of 20 mm from the scalp to measure the microvasculature of the cerebral cortex. The center of the lowermost probes was aligned at the FPz of the international 10–20 EEG system to remove positional uncertainty among subjects. After the rejection of four channels with poor average SNRs, 44 channel values were used for our analysis.

D. FEATURE EXTRACTION AND CLASSIFICATION

For the classification, we extracted signal features from the oxygenated hemoglobin (HbO_2) concentration during the task and rest state, respectively. Signal features of the rest state were extracted from the responses recorded during the initial rest, i.e., the first 30 s in Fig. 1(c). Therefore, the term "rest state" from now on indicates only the initial rest. Conversely, signal features of the task state were extracted from the average response of three tasks, also known as a block-averaged response. Moreover, feature extraction was performed for three different timeframes during a period of 10 s. Because both the rest and task states were 30 s, each state was divided into three windows, namely, window 1 (0–10 s), window 2 (10–20 s), and window 3 (20–30 s). For the signal feature, signal slope (SS), signal mean (SM), signal variance (SV), signal peak (SP), signal kurtosis (SK), and signal skewness (SSK), were calculated using built-in functions in R (R Software, version 3.6, R Foundation for Statistical Computing, Vienna, Austria) such as the *mean, var, max, kurto and skew*. The SS value was computed based on the slope of the linear regression in the respective time window. All feature values were scaled between 0 and 1 using min–max normalization to improve classification performance (1).

$$\hat{z} = \frac{z - z_{min}}{z_{max} - z_{min}} \quad (1)$$

In the above equation, where $z \in R^n$, n is a number of samples, $\hat{z}$ indicated the re-scaled z with a range between 0 and 1, z_{max} is the largest value in z, and z_{min} is the smallest value in z. In summary, six signal features (SS, SM, SV, SP, SK, and SSK) were obtained in three timeframes (window 1 to 3) for each mental state (rest and task), and for each daily stress level (high and low).

Using the extracted features, we first classified the daily stress levels as high or low. The daily stress classification performance was evaluated based on signal features from each mental state. Second, we performed mental state classification that classified rest and task into two models. The former combined the signal features of the two daily stress levels (so doubled the number of samples), and the latter processed them separately. The former handles the two daily stress levels separately, and the latter combines the signal features of the two. In the latter case, the model was divided into a model that doubled the number of samples and a model that matched the number of samples to the former. This classification process can be summarized as a Fig. 3. A classification scheme was implemented using a support vector machine (SVM) and linear discriminant analysis (LDA) using the Python Scikit-learn 0.23.1 library [23]. We tested every possible combination of six features by selecting combinations that included from one up to six features. The number of selected features in feature combinations was denoted as an *m*-feature-combination, whereby *m* indicates the number of selected features. For example, if only one feature was used as an input of the classifier, it becomes a one-feature-combination, and a total of six combinations could exist (SS, SM, SP, SSK, SK, and SV respectively). Additionally, each feature value was obtained for all 44 channels. Thus, the size of input matrix became an n-by-($m\times 44$), wherein n indicates the number of samples (participant) and m indicates the number of selected features in range of [1, 2, …, 6]. The total number of tested combinations was 63, reflecting a sum of 6 for one-, 15 for two-, 20 for three-, 15 for four-, 6 for five-, and 1 for all six-feature-combination. Finally, these feature combinations were obtained for three timeframes, and were tested with a machine learning classifier with a 5x5 cross-validation. The k-fold cross-validation can be used both when optimizing the hyperparameters of a model on a dataset, and when comparing and selecting models on a dataset. One approach to run both procedures at the same time without leading to a biased evaluation of the model performance is to nest the hyperparameter optimization procedure under the model selection procedure called nested cross-validation. Therefore, in our study, we made 5 folds divided into train set and test set in the outer loop of nested cross-validation. For train set of each fold, a grid search using another 5-fold cross-validation is performed to find the optimal parameters. Then, the scores of the test set of the fold divided from the outside are measured using the optimal parameter settings, and averaged.

III. RESULTS

A. PERCEIVED DAILY STRESS LEVELS

We investigated whether there were statistically significant differences between the selected two days reported as high and low stress days. A paired-samples *t*-test showed a significant difference between the high and low stress levels ($t(40) = 15.91$, $p < 0.001$). As expected, the average of reported stress level high stress day (M = 30.83, SD = 9.38) was statistically significantly higher than that of the low stress day (M = 18.52, SD = 9.26). The selected two days were significant, although the reported daily stress levels widely among individuals. Therefore, we confirmed that the selected two days based on fNIRS recordings

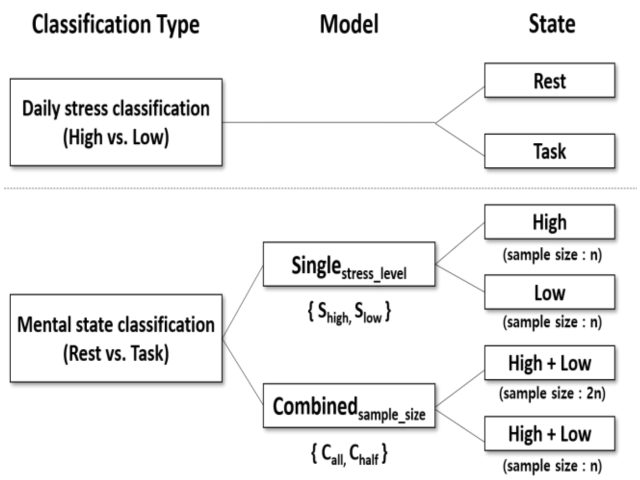

FIGURE 3. **The classification processes.**

could represent high- and low-daily-stressful days for tested individuals.

B. DAILY STRESS CLASSIFICATION

First, we performed daily stress classification using binary stress labels based on individual perceived daily stress indices. Classification accuracies were obtained by testing every possible feature combination that was extracted from three different time windows. We used SVM and LDA. However, in almost all cases SVM outperformed LDA, so the classification accuracies reported here were all obtained from SVM.

Table 1 summarizes the classification performance with standard deviation for one-feature and two-feature-combinations in all three windows. The highest accuracy in each column is shown in bold. In one-feature-combination, SS, SM, and SP reported relatively higher accuracies in both states and all three windows. In two-feature-combination, feature combinations, including these three features (SS, SM and SP), yielded relatively higher accuracies. In terms of windows, window 2 and 3 outperformed window 1 for both feature combinations. To enable the understanding of feature characteristics for better classification, we counted the number of features that reported the accuracy above 80% in each combination. Fig. 4 shows the histograms that indicate the number of features with daily stress classification accuracies of 80% or higher in all windows (1, 2, and 3) and in all states (rest and task) in one-, two-, and three-feature-combination. For an example of a one-feature-combination in window 2 and a task state, the daily stress classification accuracy above 80% was obtained with the use of four features (SS, SM, SP and SSK) among the six used, as indicated in Table 1. In another example of a two-feature combination in window 2 and the rest state, accuracies > 80% were obtained with 15 feature pairs (with the exception of the pair (SV, SK)) among 16 possible pairs. Since window 2 and 3 are relatively better than window 1, we focused on window 2 and 3 for the further analysis.

Moreover, we identified the existence of a trend toward a better performance in some specific features associated with the one- and the two-feature-combination. In our effort to identify the features with a classification accuracy of 80% or higher in window 2 and 3, SS, SM, and SP features appeared in both the rest and task states, and SSK only appeared in the task state. To evaluate the statistical characteristics of each feature, we tested each feature in high- and low-daily stress level states with paired t-tests, while we ignored the effects of the channels. In the rest state, the SM and SP features for all windows showed statistically significant differences according to the daily stress levels (at window 1, SP: $t(1803) = -3.978$, $p < 0.001$, SM: $t(1803) = -4.619$, $p < 0.001$, at window 2, SP: $t(1803) = -6.853$, $p < 0.001$, SM: $t(1803) = -6.318$, $p < 0.001$, at window 3, SP: $t(1803) = -8.277$, $p < 0.001$, SM: $t(1803) = -7.367$, $p < 0.001$). Conversely, in the task state, SSK was found in both windows 2 and 3 ($t(1803) = 2.268$, $p < 0.05$ at window 2, and $t(1803) = 2.972$, $p < 0.01$ at window 3).

In comparison with the daily stress classification performance according to the mental state (rest and task), the averaged classification accuracies were obtained as 85.5, 95.53, 98.25, 99.53, 99.83, and 100%, according to the number of feature combinations in the rest state of window 2, while for the task state of window 2 we obtained 82.67, 93.13, 97.25, 99, 99.67, and 100%. In both states, the average accuracy improved by 100% as the number of features increased. However, the accuracy of the rest state was slightly higher than that of the task state.

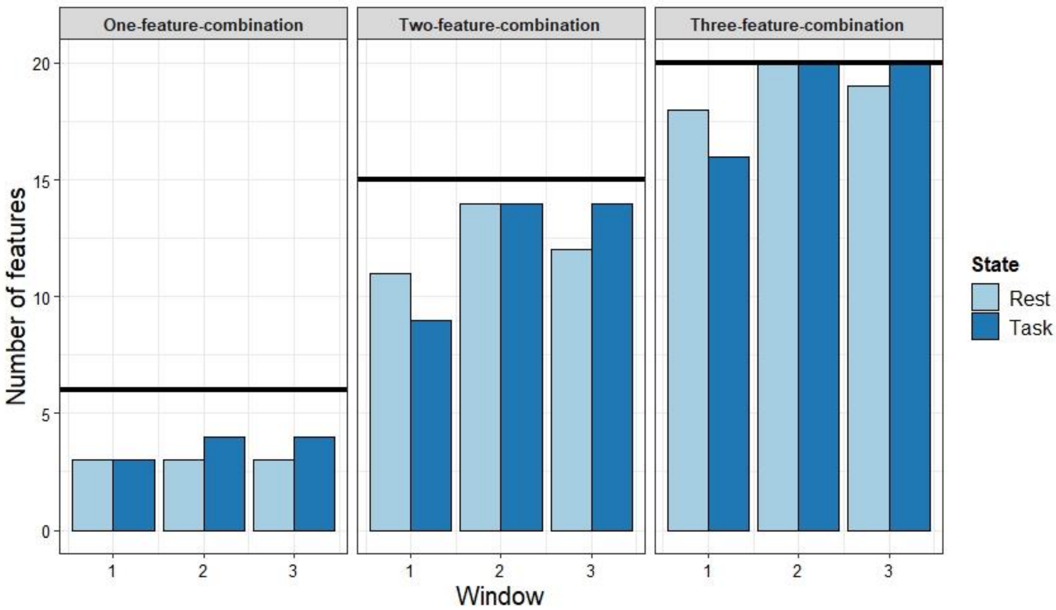

FIGURE 4. Number of features with daily stress classification accuracy of 80% or higher in all windows (1, 2, and 3) and all states (rest and task) in single, two-, and three-feature-combination. The thick black horizontal line indicates the total number of features that can be obtained from each combination.

TABLE 1
DAILY STRESS CLASSIFICATION (HIGH VS. LOW) ACCURACY (%) FOR ONE- AND TWO-FEATURE-COMBINATION. NUMBERS IN PARENTHESES INDICATE STANDARD DEVIATIONS.

	Window 1 (0–10 s)		Window 2 (10–20 s)		Window 3 (20–30 s)	
State	Rest	Task	Rest	Task	Rest	Task
One-feature-combination						
SS	**98.0 (0.04)**	**94.0 (0.07)**	98.0 (0.04)	**98.0 (0.03)**	95.0 (0.05)	98.0 (0.03)
SM	94.0 (0.04)	**94.0 (0.07)**	99.0 (0.02)	96.0 (0.05)	**100.0 (0.00)**	98.0 (0.030
SV	69.0 (0.09)	63.0 (0.08)	79.0 (0.05)	62.0 (0.09)	68.0 (0.12)	75.0 (0.08)
SP	61.0 (0.11)	66.0 (0.07)	**100.0 (0.00)**	89.0 (0.08)	99.0 (0.02)	**99.0 (0.02)**
SK	65.0 (0.08)	55.0 (0.09)	65.0 (0.12)	65.0 (0.03)	66.0 (0.05)	67.0 (0.06)
SSK	83.0 (0.08)	82.0 (0.10)	72.0 (0.12)	86.0 (0.11)	58.0 (0.10)	88.0 (0.15)
Two-feature-combination						
(SS, SM)	98.0 (0.04)	94.0 (0.07)	99.0 (0.02)	98.0 (0.03)	99.0 (0.02)	99.0 (0.02)
(SS, SV)	94.0 (0.05)	**95.0 (0.05)**	95.0 (0.04)	98.0 (0.03)	94.0 (0.04)	99.0 (0.02)
(SS, SP)	96.0 (0.05)	93.0 (0.09)	**100.0 (0.00)**	96.0 (0.05)	99.0 (0.02)	99.0 (0.02)
(SS, SK)	99.0 (0.02)	**95.0 (0.05)**	99.0 (0.02)	98.0 (0.03)	98.0 (0.03)	98.0 (0.03)
(SS, SSK)	**100.0 (0.00)**	94.0 (0.06)	98.0 (0.03)	99.0 (0.02)	98.0 (0.03)	99.0 (0.02)
(SM, SV)	94.0 (0.05)	92.0 (0.12)	**100.0 (0.00)**	98.0 (0.03)	99.0 (0.02)	**100.0 (0.00)**
(SM, SP)	95.0 (0.05)	**95.0 (0.07)**	**100.0 (0.00)**	95.0 (0.05)	**100.0 (0.00)**	99.0 (0.02)
(SM, SK)	94.0 (0.05)	**95.0 (0.07)**	**100.0 (0.00)**	98.0 (0.05)	99.0 (0.02)	99.0 (0.02)
(SM, SSK)	98.0 (0.05)	**95.0 (0.07)**	**100.0 (0.00)**	99.0 (0.02)	98.0 (0.03)	**100.0 (0.00)**
(SV, SP)	71.0 (0.07)	73.0 (0.07)	95.0 (0.05)	98.0 (0.03)	**100.0 (0.00)**	98.0 (0.03)
(SV, SK)	78.0 (0.07)	66.0 (0.07)	78.0 (0.12)	66.0 (0.06)	58.0 (0.06)	70.0 (0.11)
(SV, SSK)	86.0 (0.07)	72.0 (0.14)	87.0 (0.05)	81.0 (0.09)	59.0 (0.12)	90.0 (0.13)
(SP, SK)	66.0 (0.07)	65.0 (0.05)	**100.0 (0.00)**	95.0 (0.05)	95.0 (0.07)	99.0 (0.02)
(SP, SSK)	74.0 (0.10)	73.0 (0.11)	**100.0 (0.00)**	95.0 (0.06)	96.0 (0.05)	98.0 (0.04)
(SK, SSK)	85.0 (0.05)	76.0 (0.07)	82.0 (0.09)	83.0 (0.12)	61.0 (0.11)	86.0 (0.09)

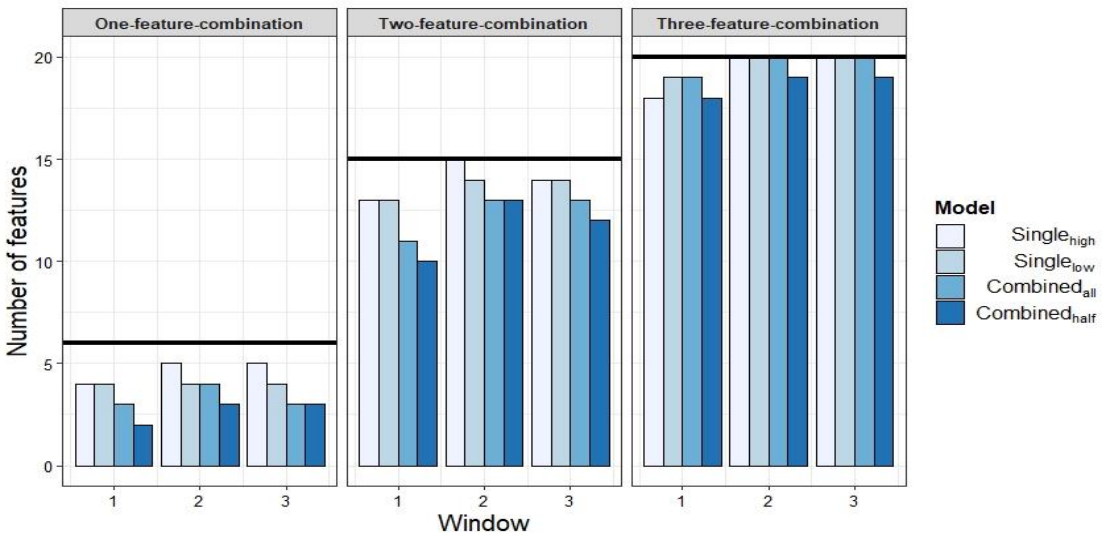

FIGURE 5. Number of features with mental state classification accuracy of 70% or higher in all windows with one- to three-feature-combination. The horizontal black bold line indicates the total number of features that can be obtained from each combination.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035799, IEEE Access

TABLE 2
MENTAL STATE CLASSIFICATION (REST VS. TASK) ACCURACY (%) FOR ONE- AND TWO-FEATURE-COMBINATION. NUMBERS IN PARENTHESES INDICATE STANDARD DEVIATIONS.

	Window 1 (0–10 s)				Window 2 (10–20 s)				Window 3 (20–30 s)			
	S_{high}	S_{low}	C_{all}	C_{half}	S_{high}	S_{low}	C_{all}	C_{half}	S_{high}	S_{low}	C_{all}	C_{half}
One-feature-combination												
SS	95.0 (0.07)	**99.0 (0.02)**	93.0 (0.02)	90.0 (0.03)	96.0 (0.05)	**98.0 (0.03)**	**94.0 (0.06)**	92.0 (0.03)	96.0 (0.05)	96.0 (0.03)	**94.0 (0.02)**	95.0 (0.04)
SM	**99.0 (0.02)**	93.0 (0.03)	**94.0 (0.03)**	**99.0 (0.02)**	**98.0 (0.03)**	**98.0 (0.03)**	92.0 (0.06)	**95.0 (0.02)**	97.0 (0.05)	**98.0 (0.03)**	**94.0 (0.04)**	**99.0 (0.02)**
SV	68.0 (0.10)	63.0 (0.12)	58.0 (0.07)	64.0 (0.08)	78.0 (0.08)	73.0 (0.05)	78.0 (0.02)	67.0 (0.09)	74.0 (0.03)	64.0 (0.14)	62.0 (0.04)	57.0 (0.08)
SP	72.0 (0.17)	73.0 (0.07)	56.0 (0.07)	66.0 (0.12)	**98.0 (0.03)**	**98.0 (0.03)**	93.0 (0.03)	93.0 (0.07)	**99.0 (0.02)**	**98.0 (0.03)**	93.0 (0.06)	96.0 (0.05)
SK	63.0 (0.17)	57.0 (0.11)	57.0 (0.05)	51.0 (0.04)	66.0 (0.03)	56.0 (0.08)	52.0 (0.06)	54.0 (0.07)	67.0 (0.12)	44.0 (0.10)	65.0 (0.09)	63.0 (0.09)
SSK	72.0 (0.12)	87.0 (0.04)	71.0 (0.08)	53.0 (0.13)	73.0 (0.09)	67.0 (0.17)	65.0 (0.04)	52.0 (0.06)	76.0 (0.11)	76.0 (0.03)	65.0 (0.08)	58.0 (0.08)
Two-feature-combination												
(SS, SM)	94.0 (0.07)	98.0 (0.03)	**98.0 (0.02)**	**96.0 (0.08)**	98.0 (0.05)	**100.0 (0.00)**	97.0 (0.03)	98.0 (0.05)	99.0 (0.02)	98.0 (0.03)	96.0 (0.03)	98.0 (0.03)
(SS, SV)	94.0 (0.07)	98.0 (0.03)	95.0 (0.02)	93.0 (0.05)	98.0 (0.03)	98.0 (0.03)	**98.0 (0.02)**	84 (0.08)	98.0 (0.03)	96.0 (0.05)	95.0 (0.01)	93.0 (0.05)
(SS, SP)	90.0 (0.10)	**99.0 (0.02)**	91.0 (0.04)	92.0 (0.06)	99.0 (0.02)	**100.0 (0.00)**	97.0 (0.03)	96 (0.05)	**100.0 (0.00)**	99.0 (0.02)	**98.0 (0.02)**	98.0 (0.03)
(SS, SK)	94.0 (0.07)	98.0 (0.03)	92.0 (0.04)	89.0 (0.08)	95.0 (0.06)	95.0 (0.05)	93.0 (0.03)	93 (0.06)	99.0 (0.02)	95.0 (0.05)	96.0 (0.02)	93.0 (0.07)
(SS, SSK)	98.0 (0.03)	**99.0 (0.02)**	95.0 (0.04)	88.0 (0.05)	98.0 (0.05)	99.0 (0.02)	95.0 (0.03)	85.0 (0.05)	**100.0 (0.00)**	98.0 (0.03)	92.0 (0.02)	97.0 (0.07)
(SM, SV)	93.0 (0.07)	96.0 (0.03)	96.0 (0.01)	**96.0 (0.08)**	99.0 (0.02)	95.0 (0.05)	94.0 (0.05)	94.0 (0.04)	96.0 (0.05)	98.0 (0.03)	92.0 (0.05)	95.0 (0.06)
(SM, SP)	**99.0 (0.02)**	95.0 (0.02)	96.0 (0.02)	95.0 (0.05)	99.0 (0.02)	**100.0 (0.00)**	96.0 (0.03)	**99.0 (0.02)**	99.0 (0.02)	98.0 (0.05)	95.0 (0.03)	**100.0 (0.00)**
(SM, SK)	98.0 (0.05)	95.0 (0.05)	93.0 (0.02)	95.0 (0.02)	98.0 (0.05)	95.0 (0.06)	94.0 (0.04)	95.0 (0.07)	96.0 (0.05)	96.0 (0.05)	92.0 (0.03)	94.0 (0.04)
(SM, SSK)	94.0 (0.10)	98.0 (0.03)	**98.0 (0.02)**	94.0 (0.00)	**100.0 (0.00)**	98.0 (0.03)	90.0 (0.07)	96.0 (0.03)	**100.0 (0.00)**	99.0 (0.02)	95.0 (0.03)	93.0 (0.06)
(SV, SP)	79.0 (0.10)	78.0 (0.05)	62.0 (0.03)	72.0 (0.06)	99.0 (0.02)	98.0 (0.03)	94.0 (0.04)	95.0 (0.05)	99.0 (0.02)	96.0 (0.03)	93.0 (0.03)	93.0 (0.05)
(SV, SK)	60.0 (0.12)	57.0 (0.09)	65.0 (0.08)	65.0 (0.12)	79.0 (0.03)	63.0 (0.11)	66.0 (0.08)	72.0 (0.08)	67.0 0.10)	46.0 (0.14)	64.0 (0.08)	56.0 (0.01)
(SV, SSK)	82.0 (0.13)	89.0 (0.06)	70.0 (0.09)	59.0 (0.09)	84.0 (0.06)	78.0 90.13)	74.0 (0.08)	66.0 (0.12)	81.0 (0.14)	74.0 (0.05)	73.0 (0.07)	54.0 (0.05)
(SP, SK)	62.0 (0.11)	65.0 (0.12)	58.0 (0.06)	60.0 (0.04)	**100.0 (0.00)**	95.0 (0.05)	94.0 (0.07)	93.0 (0.04)	98.0 (0.03)	96.0 (0.05)	94.0 (0.04)	94.0 (0.04)
(SP, SSK)	76.0 (0.12)	85.0 (0.09)	68.0 (0.07)	62.0 (0.07)	99.0 (0.02)	99.0 (0.02)	92.0 (0.06)	95.0 (0.05)	**100.0 (0.00)**	**100.0 (0.00)**	94.0 (0.07)	98.0 (0.03)
(SK, SSK)	72.0 (0.03)	85.0 (0.09)	76.0 (0.09)	50.0 (0.09)	84.0 (0.06)	71.0 (0.17)	68.0 (0.04)	62.0 (0.03)	79.0 (0.08)	77.0 (0.05)	69.0 (0.09)	60.0 (0.06)

The variables S_{high} , S_{low}, C_{all} and C_{half} specified in the table above indicate $Single_{high}$, $Single_{low}$, $Combined_{all}$ and $Combined_{half}$ models, respectively.

C. MENTAL STATE CASSIFICATION

Similar to the conventional fNIRS-based BCI approaches in mental stress classificaion, we classified the mental state between task and rest that we experimentally defined. There were two different daily stress levels, so we used the following four classification models to deal with daily stress level:

(1) $Single_{stress\_level}$ model: classification only with samples obtained from either daily stress level. Specifically, there were two models, namely, $single_{high}$ and $single_{low}$. The former was a classification with samples obtained when the reported daily stress level was high, and the latter was the classification with low-daily stress level samples

(2) $Combined_{sample\_size}$ model: classification with samples with the use of both daily stress levels indicates that the "sample_size" is "all" (which is equal to twice the value of the $Single_{stress\_level}$ model), while classifications with randomly selected samples with the same length of the $Single_{stress\_level}$ model indicated that the "sample_size" was "half."

We counted the number of features that reported accuracies > 70% in each model. As shown in Fig. 5, we could see that the number of features that came out was more in window 2 and 3 than in window 1. So we focused on window 2 and 3 as in the daily stress classification.

Table 2 lists the overall mental state classification results for one- and two-feature-combination in all three windows and in all four models. In window 2, the averaged classification accuracies were 84.83, 95.27, 98.30, 98.67, 99 and 99% as the number of features increased in the case of $Single_{high}$, while the corresponding values were 81.67, 92.27, 96.75, 98.27, 99.67, and 100% in the cases of $Single_{low}$. For the case of the combined model, the averaged accuracies were 79, 89.4, 90.7, 96.67, 97.67, and 98% in $Combined_{all}$, while the corresponding values were 75.5, 88.2, 86.15, 97, 98.67 and 99% in the $Combined_{half}$. Comparison of the model types indicated that $Single_{stress\_level}$

model outperformed the Combined$_{sample_size}$ models. This indicated that consideration of the daily stress levels helped improve the mental state classification performance. Similar with the daily stress classification, consideration of the number of features used in the model indicated that the accuracy improved as the number of features increased. For example, in the cases of one-feature-combination, the lowest accuracies were in the range between 50 to 60% for all models in window 2, whereas the same values increased to a range of 98 to 100% in six-feature-combination.

To observe the feature characteristics associated with the performance improvement according to considerations of the daily stress levels, we computed the differences of Single$_{stress_level}$ and Combined$_{sample_size}$ models for each feature combination. We averaged two Single$_{stress_level}$ models and computed the differences between the averaged outcome and each Combined$_{sample_size}$ model. As a result of sorting the feature combinations in the order of the greatest difference, SSK values with the greatest differences for all frames were included in the feature combinations. Fig. 6 shows the average mental state classification accuracy in the case of the averaged Single$_{stress_level}$ models and the two Combined$_{sample_size}$ models from one- to six-feature-combination in window 3. Considering that the sample size of combined$_{all}$ is doubled compared with that of the Single$_{stress_level}$ model, the difference in performance can be observed to be remarkable. The difference is even more pronounced when comparing the performance with a Combined$_{sample_size}$ model with the same sample size, referred to as Combined$_{half}$. This difference in performance tended to decrease as the number of features used in the feature combination increased. These tendencies were the same in window 2.

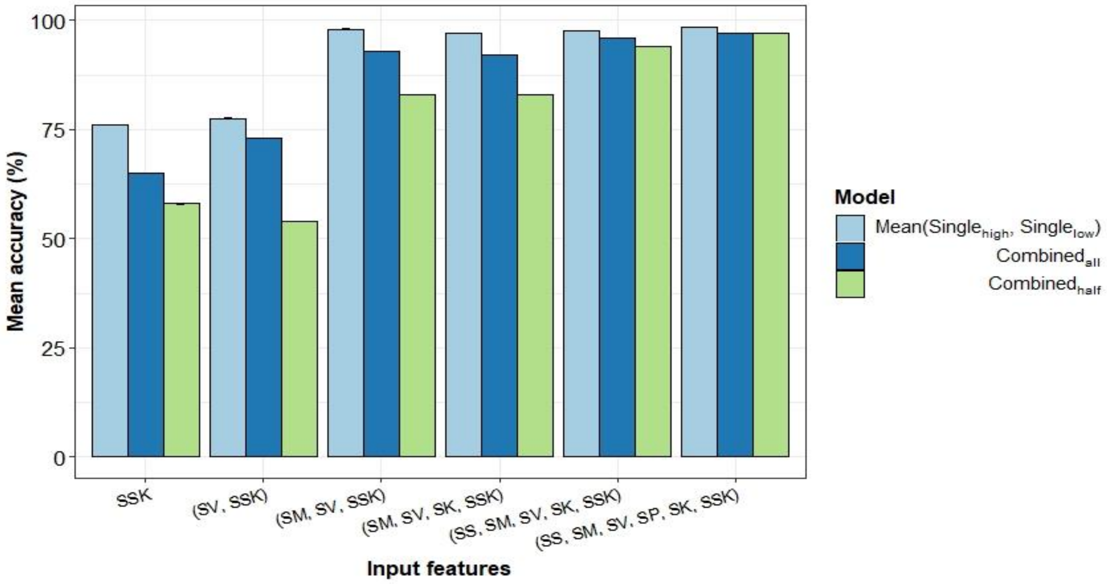

FIGURE 6. Average mental state classification accuracy for selected feature combinations in averaged Single$_{stress_level}$ models and the two Combined$_{sample_size}$ models (window 3).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035799, IEEE Access

IV. DISCUSSION

During the last few decades, researchers have expressed growing interests in measuring mental stress and its effects on our physical health [42–44]. Because the stress concept has been argued based on different theories regarding its nature, development of universal measures of mental stress have been challenging [45]. One of the currently used global measures of perceived stress is the perceived stress scale (PSS) that measures the degree to which situations in an individual’s life are appraised as stressful [46]. The items in PSS consist of general questions aimed at exploring the extent to which one perceives his or her life as unpredictable, uncontrollable, and overloaded compared with his/her life status the previous month. Therefore, the concept we defined in this study is not suitable for daily stress because it can change considerably on a daily basis, but not over a month period. Thus, we modified the questionnaires to evaluate daily stress levels for our subject group, i.e., university students. While we assessed their daily stress levels, we found that perceived daily stress levels for individuals had surprisingly different distributions shown in Fig. 7. In other words, it is not appropriate to define one specific value to determine each response as high or low for all subjects. Thus, the proposed method was employed to define individual reference values. However, there are obvious limitations in that outcomes relied solely on the stress questionnaire. Even though we have verified that there were statistically significant differences between high- and low-daily-stress levels, the objectivity of daily stress levels can be further enhanced by adding other measures, such as the salivary cortisol hormone levels [7].

The effects of daily stress have not yet been investigated extensively. Our previous study examined the effect of daily stress on heart rate variability (HRV) and found significant differences in values and trends for specific HRV parameters during Stroop tasks between high- and low-daily-stressful days [47]. Similar to these results, the new findings in this study are that the daily stress levels also differ in prefrontal oxygenation measured by fNIRS, and that the fNIRS signal features can be used to classify daily stress levels with stable performance. Therefore, daily stress levels can affect physiological responses that can be evaluated by ECG and fNIRS, even in the absence of experimental stressors. Identifying differences in prefrontal oxygenation caused by daily stress could lead to improved performance in mental state classification. Mental state classification generally aims to separate mental states between rest and task, and is one of frequently attempted problems in brain–computer interface research [21], [23], [27]. However, it may have a negative impact on state classification performance if resting states can vary from person-to-person depending on the daily stress levels of tested individuals. As hypothesized earlier, the effect of daily stress was not negligible in mental state classification. When n was the same, the averaged performances of the $\text{Single}_{\text{stress\_level}}$ models ($\text{single}_{\text{high}}$ and $\text{single}_{\text{low}}$) that considered daily stress, were much higher than the $\text{Combined}_{\text{sample\_size}}$ model ($\text{combined}_{\text{half}}$), and yielded a difference of at most 35% in the cases of three-feature combination with SM, SV, and SSK at window 3.

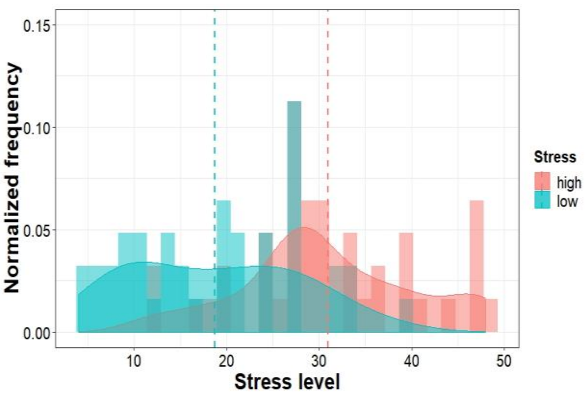

FIGURE 7. Distribution of stress indices. The high and low stress indices of participants are depicted. The Y-axis indicates the normalized frequencies from responses obtained from 41 subjects (showing the highest and lowest responses). A density curve was added to the histogram based on the probability density to make it easier to understand the distribution. The vertical dashed lines represent the means of the groups.

Comparing the classification performance in terms of the types of input features, we found that there were some specific features that had contributed significantly to the daily stress and mental state classification performance. The characteristics of these features were evaluated based on statistical tests, and the SP and SM at rest and the task SSK yielded statistically significant differences according to the daily stress levels. Especially, SP and SM features have been frequently used in several BCI studies with a good performance [15], [48–49]. The signal feature that was included in the feature combination that showed the greatest difference in mental state classification performance was signal kurtosis, namely, SSK. In other words, it can be said that the SSK feature was most affected by daily stress in the mental state classification. This is not a feature that was used frequently like SM or SP, but considering that there were only a few studies on the effects of daily stress, the remarkable aspect of this feature is an important fact that was newly discovered in this study. The SSK also exhibited statistically significant differences according to daily stress, especially during the execution of specific tasks. Therefore, given that it is a signal characteristic that reflects the effects of daily stress, it can be observed that it affects the mental state classification performance considerably.

Finally, this study was conducted by 41 young healthy subjects, but was limited to female university students. Because if subjects’ occupation varied, their daily stress levels originated from different daily stressors. Accordingly, the levels of daily stress levels can also vary [50]. Moreover, we controlled the study’s outcomes by focusing on the study of females only. This is attributed to the findings of a prior report that indicated that the stress response was significantly affected by the subject’s gender [51]. We anticipate to extend our study to male subjects to ascertain if the gender has effects on daily stress on prefrontal oxygenation and on mental state classification.

V. CONCLUSION

This study investigated the effects of daily stress with a machine learning algorithm and a fNIRS passive brain–computer interface. After we selected six feature values associated with prefrontal cortical oxygenation for two different levels of daily stress and two mental states, we found differences according to daily stress levels and experimental conditions based on classification. In daily stress classification, it was shown that it was possible to classify daily stress, not mental tasks, based on the high-classification performance for various feature combinations and all the frames. In mental state classification, we found that the existing mental stress classification performance can vary depending on the effects of daily stress by evaluating the differences between the average of the single models and the combined model. In addition, the specific signal features such as SM, SP, and SSK played an important role to reflect these effects. In particular, the SSK was highly associated with the influence of daily stress in mental state classifications. The fNIRS signals investigated in this study were frequently mentioned in the literature on mental stress and were found to be experimentally significant. From these results, it was concluded that an individual's daily stress could influence the fNIRS response, and that the mental state classification outcomes could change according to this influence. Future research investigating the influence of individual differences on perceived stress in fNIRS will help elucidate individual characteristics, and would thus allow studies on the cognitive abilities of individuals to daily stress, and on personalized fNIRS research.

REFERENCES

[1] S. Baek, E. Ha, E. Ryo, and M. Song, “Effects of life stress and depression in undergraduates on suicidal ideation,” J. Korean *Acad. Child Heal. Nurs*., vol. 18, no. 4, pp. 157–163, Dec. 2012.

[2] P. Reanaree, P. Tananchana, W. Narongwongwathana, and C. Pintavirooj, “Stress and office-syndrome detection using EEG, HRV and hand movement,” BMEiCON 2016 - 9th Biomed. Eng. Int. Conf., 2017, doi: 10.1109/BMEiCON.2016.7859624.

[3] C. Hammen, “Stress and depression,” *Annu. Rev. Clin. Psychol.*, vol. 1, no. 1, pp. 293–319, 2005, doi: 10.1146/annurev.clinpsy.1.102803.143938.

[4] B. S. McEwen, “Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators,” *Eur. J. Pharmacol.*, vol. 583, no. 2–3. pp. 174–185, 2008.

[5] R. A. Ajjan and P. J. Grant, “Cardiovascular disease prevention in patients with type 2 diabetes: The role of oral anti-diabetic agents,” *Diabetes Vasc. Dis. Res.*, Vol. 3, No. 3, pp. 147-158, 2006.

[6] F. Figueroa-Fankhanel, “Measurement of stress,” *Psychiatr. Clin. North Am.*, vol. 37, no. 4, pp. 455–487, 2014, doi: 10.1016/j.psc.2014.08.001.

[7] D. H. Hellhammer, S. Wüst, and B. M. Kudielka, “Salivary cortisol as a biomarker in stress research,” *Psychoneuroendocrinology*, vol. 34, no. 2, pp. 163–171, 2009, doi: 10.1016/j.psyneuen.2008.10.026.

[8] S. Cohen, W. J. Doyle, and A. Baum, “Socioeconomic status is associated with stress hormones,” *Psychosom. Med.*, vol. 68, no. 3, pp. 414–420, 2006, doi: 10.1097/01.psy.0000221236.37158.b9.

[9] P. Sawangjai, S. Hompoonsup, P. Leelaarporn, S. Kongwudhikunakorn, and T. Wilaiprasitporn, “Consumer Grade EEG Measuring Sensors as Research Tools: A Review,” IEEE Sens. J., vol. 20, no. 8, pp. 3996–4024, 2020, doi: 10.1109/JSEN.2019.2962874.

[10] B. Czéh et al., “Stress-induced changes in cerebral metabolites, hippocampal volume, and cell proliferation are prevented by antidepressant treatment with tianeptine.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 22, pp. 12796–801, Oct. 2001.

[11] C. M. Vander Weele, C. Saenz, J. Yao, S. S. Correia, and K. A. Goosens, “Restoration of hippocampal growth hormone reverses stress-induced hippocampal impairment,” Front. *Behav. Neurosci.*, vol. 7, p. 66, Jun. 2013.

[12] G. Jun and K. G. Smitha, “EEG based stress level identification,” 2016 IEEE Int. Conf. Syst. Man, Cybern. *SMC 2016*, Budapest, Hungary Oct. 9-12, 2016, doi: 10.1109/SMC.2016.7844738.

[13] F. Al-Shargie, M. Kiguchi, N. Badruddin, S. C. Dass, A. F. M. Hani, and T. B. Tang, “Mental stress assessment using simultaneous measurement of EEG and fNIRS,” *Biomed. Opt. Exp.*, Vol. 7, no. 10, pp. 3882-3898, 2016.

[14] T. Liu, M. Pelowski, C. Pang, Y. Zhou, and J. Cai, “Near-infrared spectroscopy as a tool for driving research,” Ergonomics, vol. 59, no. 3, pp. 368–379, 2016, doi: 10.1080/00140139.2015.1076057.

[15] K. S. Hong, N. Naseer, and Y. H. Kim, “Classification of prefrontal and motor cortex signals for three-class fNIRS-BCI,” *Neurosci. Lett.*, vol. 587, pp. 87–92, Feb. 2015.

[16] R. Takizawa *et al.*, “Neuroimaging-aided differential diagnosis of the depressive state,” *Neuroimage*, vol. 85, pp. 498–507, 2014, doi: 10.1016/j.neuroimage.2013.05.126.

[17] N. Naseer, F. M. Noori, N. K. Qureshi, and K. S. Hong, “Determining optimal feature-combination for LDA classification of functional near-infrared spectroscopy signals in brain-computer interface application,” *Front. Hum. Neurosci.*, vol. 10, no. MAY2016, pp. 1–10, 2016, doi: 10.3389/fnhum.2016.00237.

[18] N. K. Qureshi, N. Naseer, F. M. Noori, H. Nazeer, R. A. Khan, and S. Saleem, “Enhancing classification performance of functional near-infrared spectroscopy-brain-computer interface using adaptive estimation of general linear model coefficients,” *Front. Neurorobot.*, vol. 11, no. Jul. 2017, doi: 10.3389/fnbot.2017.00033.

[19] J. Shin, J. Kwon, and C. H. Im, “A ternary hybrid EEG-NIRS brain-computer interface for the classification of brain activation patterns during mental arithmetic, motor imagery, and idle state,” *Front. Neuroinform*., vol. 12, no. February, pp. 1–9, 2018, doi: 10.3389/fninf.2018.00005.

[20] R. A. Khan, N. Naseer, S. Saleem, N. K. Qureshi, F. M. Noori, and M. J. Khan, “Cortical tasks-based optimal filter selection: An fNIRS study,” *J. Healthc. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/9152369.

[21] F. Al-Shargie, “Fusion of fNIRS and EEG signals: mental stress study”, Apr. 2019. [Online]. Available: engrxiv.org/kaqew.

[22] U. Chaudhary, B. Xia, S. Silvoni, L. G. Cohen, and N. Birbaumer, “Brain–Computer Interface–based communication in the completely locked-in state,” *PLoS Biol.*, vol. 15, no. 1, pp. 1–25, 2017, doi: 10.1371/journal.pbio.1002593.

[23] U. Asgher, K. Khalil, Y. Ayaz, R. Ahmad, and M. J. Khan, “Classification of mental workload (MWL) using support vector machines (SVM) and convolutional neural networks (CNN),” *2020 3rd Int. Conf. Comput. Math. Eng. Technol. Idea to Innov. Build. Knowl. Econ. iCoMET 2020*, Sukkur, Pakistan, Jan. 29-30, 2020, doi: 10.1109/iCoMET48670.2020.9073799.

[24] U. Asgher et al., “Enhanced accuracy for multiclass mental workload detection using long short-term memory for Brain–Computer Interface,” *Front. Neurosci.*, vol. 14, no. June, pp. 1–19, 2020, doi: 10.3389/fnins.2020.00584.

[25] A. Ditthapron, N. Banluesombatkul, S. Ketrat, E. Chuangsuwanich, and T. Wilaiprasitporn, “Universal Joint Feature Extraction for P300 EEG Classification Using Multi-Task Autoencoder,” IEEE Access, vol. 7, pp. 68415–68428, 2019, doi: 10.1109/ACCESS.2019.2919143.

[26] E. A. Aydin, “Subject-Specific feature selection for near infrared spectroscopy based brain-computer interfaces,” *Comput. Methods Programs Biomed.*, vol. 195, p. 105535, 2020, doi: 10.1016/j.cmpb.2020.105535.

[27] F. Al-Shargie, T. B. Tang, and M. Kiguchi, “Stress assessment based on decision fusion of EEG and fNIRS signals,” *IEEE Access*, vol. 5, pp. 19889–19896, 2017, doi: 10.1109/ACCESS.2017.2754325.

[28] D. Rosenbaum et al., “Cortical hemodynamic changes during the Trier Social Stress Test: An fNIRS study,” *Neuroimage*, vol. 171, no. 1, Oct. 2017, pp. 107–115, 2018, doi: 10.1016/j.neuroimage.2017.12.061.

[29] J. León-Carrion et al., “The hemodynamics of cognitive control: The level of concentration of oxygenated hemoglobin in the superior prefrontal cortex varies as a function of performance in a modified Stroop task”, *Behav. Brain Res.*, vol. 193, no. 2, pp. 248–256, 2008.

[30] O. Kofman, N. Meiran, E. Greenberg, M. Balas, and H. Cohen, “Enhanced performance on executive functions

associated with examination stress: Evidence from task-switching and Stroop paradigms," *Cogn. Emot.*, vol. 20, no. 5, pp. 577–595, 2006, doi: 10.1080/02699930500270913.

[31] N. Z. Gurel, H. Jung, S. Hersek, and O. T. Inan, "Fusing Near-Infrared Spectroscopy with Wearable Hemodynamic Measurements Improves Classification of Mental Stress," IEEE Sens. J., vol. 19, no. 19, pp. 8522–8531, 2019, doi: 10.1109/JSEN.2018.2872651..

[32] S. Lotfan, S. Shahyad, R. Khosrowabadi, A. Mohammadi, and B. Hatef, "Support vector machine classification of brain states exposed to social stress test using EEG-based brain network measures," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 199–213, 2019, doi: 10.1016/j.bbe.2018.10.008.

[33] A. M. L. Sheng, M. Christopher, "Linking daily stress processes and laboratory-based heart rate variability in a national sample of midlife and older adults," *Physiol. Behav.*, vol. 176, no. 1, pp. 100–106, 2016, doi: 10.1016/j.gde.2016.03.011.

[34] M. K. Gordon, E. Peloso, A. Auker, and M. Dozier, "Effect of flavored beverage crystals on salivary cortisol enzyme-immunoreactive assay measurements," Dev. Psychobiol., vol. 47, no. 2, pp. 189–195, 2005, doi: 10.1002/dev.20081.

[35] D. Zava, "Saliva Laboratory Testing: Advantages and Disadvantages," Townsend Lett. Dr. Patients, pp. 120–124, 2004.

[36] H. C. Shin, "Measuring stress with questionnaires," *J. Korean Med. Assoc.*, vol. 56, no. 6, p. 485, 2013, doi: 10.5124/jkma.2013.56.6.485.

[37] J. O. Park and Y. S. Seo, "Validation of the perceived stress scale (PSS) on samples of Korean University students". *Korean J. Psychol.*, vol. 29, pp.611-629, 2010.

[38] J. K. Choi, J. M. Kim, G. Hwang, J. Yang, M. G. Choi, and H. M. Bae, "Time-divided spread-spectrum code-based 400 fW-detectable multichannel fNIRS IC for portable functional brain imaging," *IEEE J. Solid-State Circuits*, vol. 51, no. 2, pp. 484–495, 2016, doi: 10.1109/JSSC.2015.2504412.

[39] J. K. Choi, M. G. Choi, J. M. Kim, and H. M. Bae, "Efficient data extraction method for near-infrared spectroscopy (NIRS) systems with high spatial and temporal resolution," IEEE Trans. Biomed. Circuits Syst., vol. 7, no. 2, pp. 169–177, 2013, doi: 10.1109/TBCAS.2013.2255052.

[40] D. A. Boas, T. Gaudette, G. Strangman, X. Cheng, J. J. A. Marota, and J. B. Mandeville, "The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics," Neuroimage, vol. 13, no. 1, pp. 76–90, 2001, doi: 10.1006/nimg.2000.0674.

[41] D. T. Delpy, M. Cope, P. van der Zee, S. Arridge, S. Wray, and J. Wyatt, "Estimation of optical pathlength through tissue from direct time of flight measurement," *Phys. Med. Biol.*, vol. 33, no. 12, pp. 1433–1442, 1988.

[42] A. Bhide, R. Durgaprasad, L. Kasala, V. Velam, and N. Hulikal, "Electrocardiographic changes during acute mental stress," *Int. J. Med. Sci. Public Heal.*, vol. 5, no. 5, p. 835, 2016, doi: 10.5455/ijmsph.2016.19082015137.

[43] H. G. Kim, E. J. Cheon, D. S. Bai, Y. H. Lee, and B. H. Koo, "Stress and heart rate variability: A meta-analysis and review of the literature," *Psychiatry Investig*., vol. 15, no. 3, pp. 235–245, 2018, doi: 10.30773/pi.2017.08.17.

[44] R. K. Henderson, H. R. Snyder, T. Gupta, and M. T. Banich, "When does stress help or harm? The effects of stress controllability and subjective stress response on Stroop performance," *Front. Psychol.,* vol. 3, no. 179, pp. 1-15, Jun. 2012, doi: 10.3389/fpsyg.2012.00179.

[45] S P. Martin and P. Brantley, "Stress, coping, and social support in health and behavior.", Handbook of clinical health psychology: Volume 2. Disorders of behavior and health, pp. 233–267, 2004.

[46] S. Cohen, T. Kamarck and R. Mermelstein, "A global measure of perceived stress", *J. Health Soc. Behav.*, vol. 24, no. 4, p. 385, 1983.

[47] S. Y. Park, J. W. Yu, S. H. Kang, H. J. Lee, and S. Y. Dong, "Effect of daily stress on heart-rate variability during Stroop color word task", *2020 42th Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC2020)*, Montreal, Canada, July. 20-24, 2020 (in press).

[48] N. Naseer and K. S. Hong, "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.*, vol. 553, pp. 84–89, 2013, doi: 10.1016/j.neulet.2013.08.021.

[49] M. J. Khan and K.-S. Hong, "Passive BCI based on drowsiness detection: an fNIRS study," *Biomed. Opt. Express*, vol. 6, no. 10, p. 4063, 2015, doi: 10.1364/boe.6.004063.

[50] L. Narayanan, S. Menon, and P. E. Spector, "Stress in the workplace: A comparison of gender and occupations," *J. Organ. Behav.*, vol. 20, no. 1, pp. 63–73, 1999, doi: 10.1002/(SICI)1099-1379(199901)20:1<63::AID-JOB873>3.0.CO;2-J.

[51] P. Punita, K. Saranya, M. Chandrasekar, and S. S. Kumar, "Gender difference in heart rate variability in medical students and association with the level of stress," *Natl. J. Physiol. Pharm. Pharmacol.*, vol. 6, no. 5, pp. 431–437, 2016, doi: 10.5455/njppp.2016.6.0102325042016.

SOYEON PARK is currently pursuing the B.S./M.S. integrated degree program at the Department of Information Technology Engineering (double B.S. major at the Department of Statistics), Sookmyung Women's University, Seoul, South Korea. She is currently a Research Assistant at the HCI Laboratory under the supervision of Professor Suh-Yeon Dong. Her research interests include human–computer interaction, signal processing, and machine learning.

SUH-YEON DONG received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2010, 2011, and 2016, respectively. She is currently an Assistant Professor at the Department of Information Technology Engineering, Sookmyung Women's University, Seoul, Korea. Her research interests include machine-learning-based biosignal processing and cognitive neuroscience.