Received June 2, 2017, accepted June 25, 2017, date of publication July 5, 2017, date of current version August 8, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2723622

Machine Learning Framework for the Detection of Mental Stress at Multiple Levels

AHMAD RAUF SUBHANI, WAJID MUMTAZ, MOHAMED NAUFAL BIN MOHAMED SAAD, NIDAL KAMEL, AND AAMIR SAEED MALIK

Centre for Intelligent Signal and Imaging Research, Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar 32610, Malaysia

Corresponding author: Aamir Saeed Malik (aamir_saeed@utp.edu.my)

This work was supported in part by the Ministry of Education Malaysia through the Higher Institution Centre of Excellence under Grant 0153CA-005 and in part by the Yayasan Universiti Teknologi PETRONAS Fundamental Research under Grant 0153AA-E54.

ABSTRACT Mental stress has become a social issue and could become a cause of functional disability during routine work. In addition, chronic stress could implicate several psychophysiological disorders. For example, stress increases the likelihood of depression, stroke, heart attack, and cardiac arrest. The latest neuroscience reveals that the human brain is the primary target of mental stress, because the perception of the human brain determines a situation that is threatening and stressful. In this context, an objective measure for identifying the levels of stress while considering the human brain could considerably improve the associated harmful effects. Therefore, in this paper, a machine learning (ML) framework involving electroencephalogram (EEG) signal analysis of stressed participants is proposed. In the experimental setting, stress was induced by adopting a well-known experimental paradigm based on the montreal imaging stress task. The induction of stress was validated by the task performance and subjective feedback. The proposed ML framework involved EEG feature extraction, feature selection (receiver operating characteristic curve, t-test and the Bhattacharya distance), classification (logistic regression, support vector machine and naïve Bayes classifiers) and tenfold cross validation. The results showed that the proposed framework produced 94.6% accuracy for two-level identification of stress and 83.4% accuracy for multiple level identification. In conclusion, the proposed EEG-based ML framework has the potential to quantify stress objectively into multiple levels. The proposed method could help in developing a computer-aided diagnostic tool for stress detection.

INDEX TERMS Absolute power, amplitude asymmetry, coherence, EEG, machine learning, mental stress levels, phase lag, relative power, support vector machine, t-test.

I. INTRODUCTION

Stress is commonly recognized as a state in which an individual is expected to perform too much under sheer pressure and in which he/she can only marginally contend with the demands. These demands can be psychological or social. It is known that psychosocial stress exists in daily life, which has resulted in poor quality of life by affecting people's emotional behavior, job performance, mental and physical health [1]. Psychosocial stress is a leading cause of several psychophysiological disorders. For example, it increases the likelihood of depression [2], stroke [3], heart attack and cardiac arrest [4]–[6].

The cure of stress requires it to be quantized into levels first. Clinically, stress has been evaluated using questionnaires and interviews, which are subjective methods. Alternatively, stress-related physical and physiological changes have also been utilized as objective indicators of stress [7]. For example, physically, stress changes the pupil dilation [8], blink rate [9] and facial gestures [10]. On the other hand, stress causes changes in the autonomic nervous system (ANS) [11]. Therefore, physiological biomarkers of stress from the ANS exist in the form of heart rate (HR) and heart rate variability (HRV) [12], respiration [13], and skin conductance [14]. According to the latest neuroscience, the human brain is the main target of mental stress [15] because the perceptions of the human brain determine whether a situation is threatening and stressful. To obtain the cortical response to stress, non-invasive neuroimaging modalities, such as electroencephalography (EEG), furnish the most suited modalities to measure functional changes in the brain. Importantly, EEGs have shown implication association with other stress indicators such as HR and HRV in general [16] and specifically in stress [17].

As far as the assessment of stress from EEG signals is concerned, various studies have extracted several electrophysiological features from EEG signals and employed classification algorithms with the highest accuracy of 96% for classification between two levels of stress [18]. These results validate that EEG is a potential assessment tool for stress. In recent studies, changes in the EEG absolute power and in connectivity measures such as coherence and mutual information have been shown to vary due to stress [19]. Similarly, asymmetry in EEG alpha power has been shown to be influenced by HRV biofeedback during stress therapy [20]. Another study discussed EEG alpha asymmetry and revealed stress-related disorders in a virtual reality environment [21]. EEG eigenvalue decomposition was utilized for stress level classification [22]. Another study proposed an EEG-based brainwave balancing index to assess the stress level of university students during their studies [23]. The time course of psychological stress was investigated through event-related potentials in a successfully designed stress-elicitation paradigm [24]. This study indicated that stress occurred in the early stages of cognitive processing. In the context of objectively identifying and differentiating stressful conditions from other conditions using EEG-based methods, various computational techniques have been used, such as support vector machine [25], [26], K-nearest neighbors (KNN) [22], [27], artificial neural networks (ANNs) [28], [29] and random forest [30]. Not only has stress been identified using EEG signals alone but also EEG signals have been fused with other modalities, such as skin conductance [30], functional near-infrared spectroscopy (fNIRS) [25] and electrocardiography (ECG) [31], in an aim to improve the identification of stress.

We hypothesize that EEG has the potential to objectively identify the levels of stress if proper analysis is conducted. We propose a machine learning-based objective framework for the identification of stress levels based on EEG signals from normal subjects during mental stress. In this paper, we present a novel methodology to detect mental stress levels by exploring quantitative differences between stress and control conditions as well as four levels of stress conditions. We have extracted five features from EEG signals: absolute power, relative power, coherence, amplitude asymmetry and phase lag. Our machine learning framework allowed for standardization of the extracted features followed by feature selection using the receiver operating characteristic (ROC) curve, the t-test and the Bhattacharya distance. The selected features were applied to three classifiers: logistic regression (LR), support vector machine (SVM) and naïve Bayes (NB) classifiers. Finally, the model validations were provided by cross-validation to avoid classifier over-fitting.

The structure of this paper is the following: Section II describes the detailed methodology, and the achievements and results are presented in section III. In section IV, a discussion of the results is provided, and the study's conclusions are presented in section V.

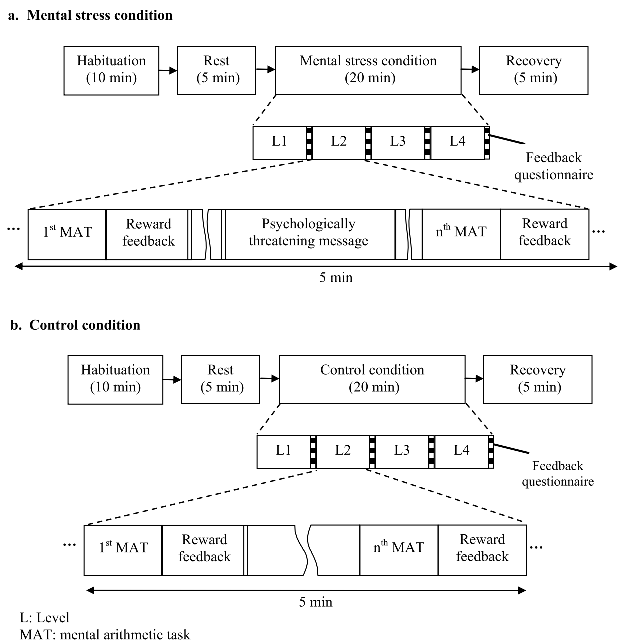


FIGURE 1. Experiment flow (a) mental stress condition and (b) control condition.

II. MATERIALS AND METHODS

A. STUDY PARTICIPANTS

Forty-two healthy subjects, including eleven females (19-25 years of age), were selected for this study. They were selected based on having no previous medical record or head injury and not using any medication that might increase cardiac activation. The subjects were further scrutinized based on the results of the perceived stress scale (PSS) [32]. PSS is a ten-item inventory that grades the perception of an individual's stress into four levels based on the experiences of one past month. Four subjects who were placed in the fourth level of the PSS scale were excluded from participation because they already had stress. Data from six subjects were corrupted due to bad connections between the EEG cap and the scalp and had to be excluded. Ten subjects could not appear in both sessions. Finally, data from twenty-two subjects who participated in both experimental sessions were included in this paper. The subjects were asked to perform fasting for at least two hours before starting the experiment. Each subject signed an informed consent, agreeing to participate and was given an honorarium of RM 40 for his/her contribution. The experimental design had been approved by the ethics committee at Hospital Universiti Sains Malaysia (HUSM), Malaysia.

B. EXPERIMENT DESIGN

In this study, a computer-based mental arithmetic task (MAT) was employed to induce stress that was based on the paradigm of the Montreal Imaging Stress Task (MIST) [15]. The MIST was chosen because it has shown the capability of inducing reliable stress that involves the hypothalamic pituitary adrenal (HPA) axis [33], [34]. The tasks were presented using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) [35]. The pictorial elaboration of the experimental design is shown in Fig. 1. The experiment had three main components: stress, control and rest. The stress condition was induced by placing a deadline on solving a task (time pressure) accompanied with negative comments. In contrast, the control condition consisted of doing MAT without the extra challenge of time pressure or negative comments. The MIST paradigm has a unique feature of comparing the stress condition with a control condition of a similar nature. The rest condition included an eyes-open task while sitting comfortably.

Because each study participant has gone through both stress and control conditions, the conditions were performed on two separate days with a gap of at least seven days, to minimize the learning effect on the performance. In addition, to eliminate the expected effects of these sessions on the results, half of the subjects underwent the stress condition followed by the control condition while the other half of the subjects underwent those conditions in the opposite order. Both the stress and control condition had their own rest condition in the beginning. Moreover, a habituation phase and a recovery phase were also provided to the subjects. Before the stress condition, a practice session was provided to the subjects to observe their performance in solving the MAT. Hence, the time that a subject was given to solve the MAT during the stress condition was shorter than his/her average response time during the practice session.

Both the stress and control conditions were divided into four difficulty levels of 5 minutes of duration. The MAT comprised simple arithmetic calculations of up to two digit numbers (maximum 99) that involved four operands ($+, -, \times, \div$). The answer to every MAT was a single digit number. Level 1 included either addition or subtraction of two numbers (e.g., 5+3). Level 2 included paired multiplication with either addition or subtraction of three numbers (e.g., 4*6-16). Level 3 entailed multiplication with addition and subtraction among four numbers (e.g., 9-4*6-21). Level 4 included division along with previous operations to be performed among four numbers (e.g., 57/3-9-6).

In the *stress* condition, the arithmetic task was presented along with negative comments. Moreover, in the stress condition, due to having a limited time to solve the MAT, the performance accuracy remained below 50%. To induce stress on the subject, after every MAT, a feedback display message appeared on computer screen, such as ''correct'' or ''incorrect'' or ''no response'', according to the response to the task. In addition, the average performance of the subject at a given level as well as the response time to solve the MAT are displayed. To continue inducing stress, after every few MATs, the subjects were reminded to maintain the performance accuracy above an arbitrary threshold, which was described to him/her as the overall performance of other subjects. Additionally, in the stress condition, after a certain number of trials in every level, a stressful interrupt popped up that showed negative comments such as ''Don't guess answers'', ''Your performance is below average'', ''Don't panic. You have been given sufficient time to answer. Don't GUESS and don't give WRONG answers'', ''You are under observation from outside'', ''You have disappointed us with your performance''. In the *control* condition, there was no time limit to solve the MAT. The feedback display only revealed either ''correct'' or ''incorrect'' based on the response to the task.

After every interval of stress and control conditions, two minutes of idle time was granted to allow the subjects to perhaps feel comfortable while performing the task. During the break, the subjects were provided with a feedback questionnaire asking their insights about the preceding interval. In the experiment, the scores from the feedback questionnaires along with the task performance scores (percentage of correct responses) validated the induction of stress.

TABLE 1. Performance in every level of the stress and control conditions.

Level	Stress (mean ± standard deviation) x 100	Control (mean ± standard deviation) x 100
Level 1	56 ± 13 †‡	74 ± 7
Level 2	42 ± 10 †‡	64 ± 12
Level 3	32 ± 10 †‡	61 ± 12
Level 4	22 ± 11 †‡	49 ± 14

† indicates that the stress level has a mean that is significantly different from the control, and ‡ indicates that the level has a mean that is significantly different from all the other stress levels.

Table I shows the task performance in every level of the stress and control conditions. The task performance in the stress condition was significantly (based on a t-test with $P < 0.001$) different from the control condition in every level. Moreover, the task performance in every level of stress was also significantly different from the other stress levels (based on ANOVA and Tukey-Kramer post hoc test).

TABLE 2. Average response time of the tasks in every level of the stress and control conditions.

Level	Stress (mean ± standard deviation) (ms)	Control (mean ± standard deviation) (ms)
Level 1	828.6 ± 70.2 †‡	1358.1 ± 364.4
Level 2	1959.2 ± 210.1 †‡	3348.3 ± 863.4
Level 3	2680.0 ± 315.2 †‡	4762.86 ± 1284.9
Level 4	3076.2 ± 288.4 †‡	6690.1 ± 2070.1

† indicates that the stress level has a mean that is significantly different from the control, and ‡ indicates that the level has a mean that is significantly different from all the other stress levels.

Table II shows the response time of the tasks in every level of the stress and control conditions, as computed from E-Prime. Like the task performance, the response time of the tasks under the stress conditions was also significantly different from that of the control conditions in every level of stress as well as being different in every level of stress compared with all other levels of stress.

Table III shows the subjective rating of the allocated time to solve the task. Based on the response during the task,

TABLE 3. Average of the subjective response time of the tasks in every level of the stress and control conditions.

Level	Stress (mean ± standard deviation) x 100	Control (mean ± standard deviation) x 100
Level 1	5.7 ± 2.2	9.1 ± 1.5
Level 2	4.4 ± 2.1	8.6 ± 1.8
Level 3	3.8 ± 2.3	8.5 ± 1.7
Level 4	3.2 ± 2.1	7.9 ± 1.8

it can be concluded that our experiment successfully induced stress.

C. EEG DATA ACQUISITION

The EEGs were recorded from 128 channels using a Net Amps 300 amplifier (Electrical Geodesic Inc. (EGI), USA). The Ag/AgCl electrodes were mounted into an elastic net. All the electrodes were referenced to the Cz position. The impedance of all the electrodes was maintained below 50 KΩ throughout the recording. The signals were digitized at 500 Hz with a notch filter at 50 Hz. The EEG amplifier was placed inside the experiment room. The amplified and digitized EEG signal was transmitted to Net Station 4.43 recording software operating on the computer placed outside the experiment room via fiber optic cables.

The experiment was performed in an isolated room where the subject was sitting alone in front of a computer screen with the installed E-Prime. To control the experiment, the experimenter observed the acquired EEG signals as well as the task performance of the subject on a duplicate screen from outside the experiment room. The triggers from the E-Prime were sent to the Net Station to automatically start/stop the EEG recording as well as to mark the events through a TCP/IP link. Prior to the experimental set up, a timing test was performed to synchronize the clocks of the two computers that were running E-Prime and Net Station.

D. EEG ARTIFACT REDUCTION

EEG signals were pre-processed offline by employing a 0.1-Hz filter to remove the DC artifacts and a 50-Hz notch filter to remove the line noise. Nineteen EEG channels according to the 10-20 system were selected against the average mastoid reference. Pre-processing of EEG signals was performed in Net Station 4.43. For further processing and feature extraction, EEG data were exported to Neuroguide [36]. Further processing was performed at a sampling rate of 128 Hz. The eye-blink and muscle artifacts were manually removed by discarding that portion of the recording from the data. Both eye-blink and muscle artifacts were detectable by visual inspection. For example, the eye-blinks created a peak at approximately 10 Hz, while the muscle artifacts appeared at a higher frequency in the power spectral density graph of the EEG signals. The internal consistency and reliability of the cleaned EEG data were measured by computing the split-half reliability, and test-retest reliability measures that were above 90% for every EEG channel. To perform the analysis, sixty seconds of cleaned EEG data were selected from each level of the stress and control conditions.

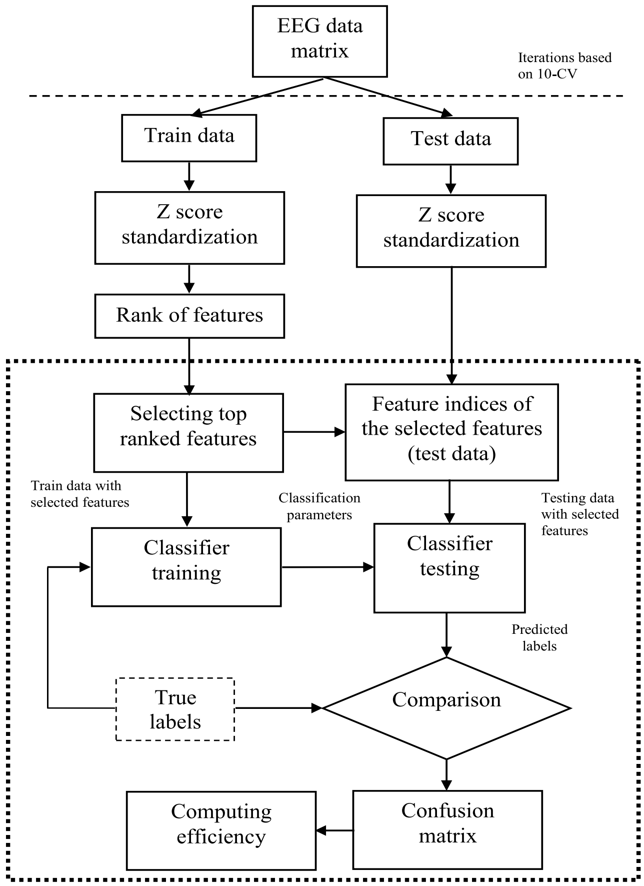


FIGURE 2. Proposed ML framework.

E. PROPOSED ML FRAMEWORK

For the identification of stress levels, three analytical cases were performed. In *case one*, each of the four levels of stress was compared with the initial level of control (a binary classification), and in *case two*, every level of stress was compared with its respective level of control (a binary classification), and in *case three*, each level of stress was compared with all the other levels of stress (one vs. all classification). For every case, the proposed framework, as shown in Fig. 2, was applied.

Figure 2 shows an overview of the proposed ML method, which involves a description of the EEG feature extraction, selection, classification and validation. For the feature extraction, one minute of artifact-free EEG epochs were selected from the stress and control conditions per level per subject. The feature extraction has been implicated into many features such as absolute and relative power, coherence, amplitude asymmetry and phase lag. The features were arranged column-wise in a matrix, and each column was denoted as x_i, where $i = 1 \ldots Nc$. The rows of the matrix represent stress conditions with 2 physiological conditions per patient; the matrix was termed the EEG data matrix. The matrix was denoted by $L = [(xi,\ yi),\ i{=}1\ \ldots Nc]$ and included both the feature space matrix and the corresponding output class labels, y = [Stress, Controls]. A detailed description for each sub-process is provided in the respective subsections.

1) EEG FEATURE EXTRACTION

a: ABSOLUTE POWER (AP)

In this paper, the EEG absolute power was estimated by first converting the EEG signal to frequency domain using fast Fourier Transform (FFT). The FFT was applied using a tapered cosine window of 256 samples with 75% overlapping. The cosine window is defined in (1)

$$w(x) = \begin{cases} \frac{1}{2}\left\{1+\cos\left(\frac{2}{r}\left[x-\frac{r}{2}\right]\right)\right\}, & 0 \le x \le \frac{r}{2} \\ 1, & \frac{r}{2} \le x \le -\frac{r}{2} \\ \frac{1}{2}\left\{1+\cos\left(\frac{2}{r}\left[x-1+\frac{r}{2}\right]\right)\right\}, & 1-\frac{r}{2} \le x \le 1 \end{cases} \quad (1)$$

In this study, the EEG absolute power was computed for each channel, which included the frontal (Fp1, Fp2, F3, F4, F7, F8, Fpz), temporal (T3, T4, T5, T6), parietal (P3, P4, P7, P8), occipital (O1, O2), and central (C3, C4) channels. Moreover, the values were computed for frequency bands between 1 to 45 Hz, such as the delta (1 to 4 Hz), theta (4 to 8 Hz), alpha 1 (8 to 10 Hz), alpha 2 (10 to 12 Hz), beta 1 (12 to 15 Hz), beta 2 (15 to 18 Hz), beta 3 (18 to 25 Hz), gamma 1 (30 to 35 Hz), gamma 2 (35 to 40 Hz) and gamma3 (40 to 45 Hz) bands. The EEG power in different frequency bands and scalp locations were features during the proposed machine learning process.

b: RELATIVE POWER (RP)

The relative power finds the rhythmicity in the EEG signals. The relative power was derived from the absolute power of the frequency bands as the power in a specific frequency band divided by the total power [37], as shown in (2).

$$Relative\ power = \frac{Power\ in\ band}{Total\ power} \times 100\% \quad (2)$$

In this paper, the features of RP were computed for all frequency bands of absolute power across 19 electrodes. These features were computed for every subject and in every level of the experiment.

c: COHERENCE

Coherence is a brain connectivity measure that reports the degree of association between two brain locations. The idea in measuring the coherence in this study was to indicate impurities in the coherences between the stress and control conditions. Mathematically, it can be represented as follows [37]:

$$Coherence = \frac{|H_{uv}^2|}{|H_u|\,|H_v|} \quad (3)$$

In this representation, the numerator is the cross-spectrum between the two signals, and the two terms in the denominator represent the auto-spectra of the individual signals. This representation interprets the Pearson correlation coefficient for the variables in the frequency domain. For further knowledge about coherence, please refer to [38]. The coherence was computed between 171 electrode pairs for each of the frequency bands for every subject in every level of the experiment.

d: PHASE LAG

The phase difference is also a measure of connectivity that describes the lead or lag between two EEG signals from different locations. The phase is the arctangent of the ratio of quadrature components derived from the FFT. The phase of a particular signal is generally defined as follows [37]:

$$Phase = Arc\tan\left(\frac{b}{a}\right) \quad (4)$$

where b represents the "imaginary" or "out-of-phase" component, and a represents the "real" or "in-phase" component of the signal. The real and imaginary components of the signals were computed from the FFT. Then, the phase difference between the signals from the two locations is computed by subtracting their individual phases, as shown in (5):

$$Phase\ difference = Arc\tan\left(\frac{b2}{a2}\right) - Arc\tan\left(\frac{b1}{a1}\right) \quad (5)$$

The phase difference was computed in radians and converted to degrees. The absolute phase delay was computed by squaring and then taking the square root of the squared difference.

The phase difference was also computed for 171 location pairs for each of the frequency bands for every subject in every level of the experiment.

e: AMPLITUDE ASYMMETRY

The asymmetry is also a measure of the connectivity, and it reflects the relative stimulation between two brain locations. The asymmetry was found by taking the difference between the signals' amplitudes and, then, normalizing it to the sum of their amplitudes, as shown in (6), where M and N are the instantaneous amplitudes of the given signals.

$$Asymmetry = \frac{M-N}{M+N} \quad (6)$$

The amplitude of the asymmetry was computed for 171 pairs of locations of each of the frequency bands for every subject in every level of the experiment.

2) EEG DATA MATRIX AND Z-SCORE STANDARDIZATION

The feature extraction implicated in the EEG data matrix involved the number of rows (*data points = 44*). The matrix might not be centered and could be unequally distributed. Hence, the data standardization was performed by involving the z-score standardization. The standardization was performed by computing the values column-wise by subtracting each element value with its column-wise mean and dividing by the corresponding standard deviation.

3) EEG FEATURE SELECTION

The extracted features might be either irrelevant (due to low feature-class correlations) or redundant (due to high feature-feature correlations). For this paper, feature selection was achieved by first selecting those features that had high feature-class correlations. Second, we sorted the selected features using a rank-based method that assigns a weight value to each feature, as shown in (7) [39]–[41]. For this purpose, three feature selection criteria, the ROC, t-test and Bhattacharya distance [40], were used.

$$r = z \times (1 - \alpha \times \rho) \tag{7}$$

The z-value for each feature corresponds to the absolute value of the feature selection criterion. For the ROC, z corresponds to the area between the empirical ROC curve and the random classifier slope and could vary from 0 and 0.5, which indicates a bad to good classification ability, accordingly. A high z-value (equal to or close to 0.5) corresponded to the ability of a feature to discriminate from other features within classes. For the t-test, z corresponds to the absolute value two-sample T-test with a pooled variance estimate. For Bhattacharya, z is the minimum achievable classification error or the Chernoff bound. In (7), α is the scalar weight (from 0 to 1) for *rho,* which is the average value of the cross-correlation coefficient between the candidate feature and all the already picked features. A large value of *rho* overshadows the significance statistic. This arrangement means that the higher the value of rho is, the higher the correlation of the feature with previously selected features and hence the likely it is to be excluded from the output list.

In this way, the features were arranged in descending order, i.e., the top-ranked features were listed at the top of the list. Furthermore, only the top-ranked features were selected for training and testing the classifier models. To find the minimum number of features that would be sufficient to train the classifier models without over-fitting, an empirical process was adopted. In this iterative procedure, the classification performances of the classification models for each of the feature subsets based on the top 1, 2, 3, 4, 5, 10, 15, and 20 features were observed. Finally, the highest classification performances were reported.

4) CLASSIFICATION MODELS

In this study, the LR classifier was used to model the relationship between the reduced set of features and the corresponding treatment outcomes (stress and control), *y =[stress, Control]*, according to (8) [42]. For the LR classifier, the coefficient estimations were based on the maximum likelihood method. The LR classifier resulted in a likelihood value $l(x)$, where $0 \leq l(x) \leq 1$, which was an indication of the condition, associated with either stress or control. If $l(x)$ was greater than the threshold = 0.5, then the condition was declared to be stress, and otherwise, it was associated with the control group.

$$F(z) = E(Y/x) = \frac{1}{1 + e^{-z}} \tag{8}$$

where Y indicates the class labels, which are assigned the value of either 'Stress' or 'Controls'. In addition, x represents a combination of different EEG features. To obtain the LR model from the logistic function, we used (9):

$$z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \tag{9}$$

where z is a linear combination of α plus β_1 multiplied by X_1, plus β_2 multiplied by X_2, and so on until plus β_k multiplied with X_k, where the X_k are the independent variables, and α, and β_i are constant terms that represent unknown parameters. Furthermore, by replacing the value of z from (9) to (8), the following (10) represents the logistic function:

$$F(z) = E(Y/x) = \frac{1}{1 + e^{-(\alpha + \sum \beta_i X_i)}} \tag{10}$$

In terms of a response and non-response, the risk of a person to be a non-responder or a responder is estimated and is represented by Y or $l(x)$. The LR classifier resulted in a likelihood value of $l(x)$, where $0 < l(x) \leq 1$, which was an indication of the subjects' association with either the stress or controls category. If $l(x)$ was greater than the *threshold* = *0.5*, then the condition was declared as 'Stressed' and, otherwise, as 'control'.

The second classification model that was employed was the SVM classifier with a linear kernel [43]. It can classify the feature space based on a 'hyperplane' that separated the stress and control conditions according to the class labels. The SVM is a high efficiency classifier model and is used here for comparison purposes. According to the SVM, a linear decision boundary can be found based on this high-dimensional space. The use of a linear kernel instead of a nonlinear kernel reduced the risk of over-fitting the data and improved the performance for our data and significantly reduced the overall model complexity. In summary, the LR classifier generated probability values to categorize stress or controls, and the SVM developed a hyperplane to achieve the maximum classification accuracy.

The third classification model was the NB classification [44], which is based on generating the conditional posterior probabilities for each sample while involving the target condition, i.e., stress vs. control. The classifier was formed by assigning the sample to the class for which the sample had higher posterior probabilities.

In this study, all the three classifier models were implemented in Matlab. Table IV shows the specific values of the parameter assigned for each classifier during training and testing. Regarding the LR classifier, the link function showed the relationship between the EEG features and the clinical outcomes. The value was set as the 'logit' since the classifier that was used was logistic regression. A two-class classification assumes a binomial distribution; this aspect was set as binomial because the data were supposed to originate from 2 classes, i.e., stress and control. The offset value was set to 1, whereas the mathematical model of the LR classifier included a constant term. Regarding the SVM classifier, the C values were assigned as 0.787 for the stress class and

TABLE 4. Values of the parameter assigned for each classifier during training and testing while discriminating the stress and control conditions.

S.No.	Classifier	Parameters	Value
1	Logistic Regression	Link Function	Logit
		Distribution	Binomial
		Offset	1
		Constant term	A constant term is added in the model
2	Support Vector Machine (SVM)	C for class 1 (N/2xN1)	0.787
		C for class 2 (N/2xN2)	1.3684
		Degree of polynomial	1
		No. of classes	2
		Kernal function	Linear
3	Naïve Baysian Classification	Distribution	Normal
		Prior	Uniform distribution for all classes

1.3684 for the control class. The values were computed with formula (*N/2×N1*) and (N/2×N2), respectively. The variable 'N' denoted the total number of study participants; *N1* indicated the number of stressed subjects, and *N2* indicated the number of controls. Other parameters, such as 'Degree of polynomial', 'No. of classes', and 'Kernel function', were assigned as *1, 2,* and *'Linear'*, respectively. The parameters for the Naïve Bayesian were assigned with a normal uniform distribution for the stress and control classes.

5) VALIDATION OF CLASSIFICATION MODELS

After the classifier design, a fair evaluation requires an assessment of its performance over a range of selected features and classifier designs (with suitable coefficient values until convergence), which corresponds to many subjects. To address this consideration, we evaluated the classification performance based on 10-fold cross validation by dividing the data sample points (Study participants) into 10 equal segments. During each round, 9 of the segments were utilized as the training subset, and the remaining 1 was the test subset. Ten-fold cross validation provides a fair test of validation in cases where the data points are limited while utilizing features for both testing and training the classifier models.

For each feature subset, 100-times runs of the simulations were performed involving 10-fold cross validation to achieve the box plot representations of the accuracies, sensitivities and specificities. Since the individual iterations resulted in 100 different values of the performance metrics (the accuracy, sensitivity and specificity), the final confusion matrix was computed by averaging over 100 times. The performance metrics computed from the confusion matrix were presented by (11-13). The sensitivity of a classification model corresponds to the percentage of true cases (TP) that are correctly classified as cases defined by (11). The specificity of a classification model refers to the percentage of true non-cases (TN) that are correctly classified as non-cases, as described by (12). The accuracy of a classification model illustrates the percentage of correctly classified cases and non-cases among all the example points, as depicted in (13).

$$Sensitivity = \frac{TP}{TP + FN} \tag{11}$$

$$Specificity = \frac{TN}{TN + FP} \tag{12}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

III. RESULTS

The best performances of the feature sets for the identification of stress in *case one* are shown in Table V. For the identification of levels 1-3, the relative power produced the best performance with the t-test and NB in level 1 (accuracy 94.0%, sensitivity 95.0% and specificity 91.7%), with the t-test and SVM in level 2 (accuracy 93.9%, sensitivity 96.7%, and specificity 92.5%) and with the t-test and NB Level 3 (accuracy 94.6%, sensitivity 98.3% and specificity 93.3%). For the identification of level 4, the amplitude asymmetry produced the best performance with the t-test and NB (accuracy 91.7%, sensitivity 94.2% and specificity 90.0%).

The best performance of every feature set in *case two* is shown in Table VI. It has been observed that the maximum performance was shown by the relative power, in levels 1, 3 and 4, and absolute power, in level 1. Moreover, the maximum performance was achieved with the t-test across every level and with the NB classifier across levels 1-3 and LR at level 4. The overall accuracy for the case two identification of stress was 94.0% in level 1.

IV. DISCUSSION

This paper provides a framework for the identification of stress at multiple levels using EEGs. For this purpose, an experimental paradigm, based on MIST, was designed that induced four levels of stress based on time pressure, distraction and evaluative pressure, to mimic social pressure. For the comparison of the four levels of stress, four identical control conditions were available with the same task difficulty as the stress conditions. By the same task difficulty, we mean that the nature of the arithmetic task was the same. The findings on the task performance and response time validate the experiment paradigm to induce stress, as shown in Tables I-III.

For the identification of stress, the analysis was conducted for three cases. In case one, four levels of stress were individually identified in comparison with the initial control condition, as a two-class problem. Case two was also a two-class problem, in which every individual level of stress was identified using the similar level of stress. Case three, however, was a multiclass identification of stress, in which every level of stress was identified from the other levels of stress.

The extracted features from the EEG signals were the absolute power, relative power, coherence, amplitude asymmetry

TABLE 5. Highest performance of every feature in combination with the classifier and feature selection method in every level of stress with respect to the first level of control.

Feature	Highest performance	Classifier/Feature selection
Level 1		
Absolute power	Acc. 91.75 Sen. 95.00 Spec. 90.00	NB/ t-test
Relative power	Acc. 94.00 Sen. 95.00 Spec. 91.67	NB/ t-test
Coherence	Acc. 90.75 Sen. 95.00 Spec. 87.50	NB/ t-test
Amplitude asymmetry	Acc. 87.2 Sen. 85.8 Spec. 86.4	SVM/ ROC
Phase lag	Acc. 81.83 Sen. 85.00 Spec. 81.67	SVM/ ROC
Level 2		
Absolute power	Acc. 92.25 Sen. 95.83 Spec. 90.00	LR/ t-test
Relative power	Acc. 93.88 Sen. 96.67 Spec. 92.50	SVM/ t-test
Coherence	Acc. 83.85 Sen. 90.83 Spec. 77.71	SVM/ Bhat
Amplitude asymmetry	Acc. 91.25 Sen. 95.83 Spec. 90.00	NB/ t-test
Phase lag	Acc. 86.25 Sen. 90.83 Spec. 80.83	NB/ t-test
Level 3		
Absolute power	Acc. 91.00 Sen. 95.00 Spec. 90.83	NB/ t-test
Relative power	Acc. 94.58 Sen. 98.33 Spec. 93.33	NB/ t-test
Coherence	Acc. 83.00 Sen. 87.50 Spec. 82.50	NB/ t-test
Amplitude asymmetry	Acc. 91.25 Sen. 95.00 Spec. 89.17	NB/ t-test
Phase lag	Acc. 81.00 Sen. 84.17 Spec. 80.83	NB/ t-test
Level 4		
Absolute power	Acc. 89.38 Sen. 92.00 Spec. 88.33	SVM/ ROC
Relative power	Acc. 90.58 Sen. 90.83 Spec. 90.83	LR/ ROC
Coherence	Acc. 83.75 Sen. 86.67 Spec. 85.83	NB/ t-test
Amplitude asymmetry	Acc. 91.75 Sen. 94.17 Spec. 90.00	NB/ t-test
Phase lag	Acc. 82.17 Sen. 86.67 Spec. 85.83	NB/ t-test

Acc.: accuracy, Bhat: Bhattacharyya distance, LR: linear regression, NB: naïve Bayes, ROC: receiver operating characteristic, Sen.: sensitivity, Spec.: specificity and SVM: support vector machine.

TABLE 6. Highest performance of every feature in combination with the classifier and feature selection method in every level of stress with respect to the similar control condition.

Feature	Highest performance	Classifier/Feature selection
Level 1		
Absolute power	Acc. 91.8 Sen. 95.0 Spec. 90.0	NB/ t-test
Relative power	Acc. 94.0 Sen. 95.0 Spec. 91.7	NB/ t-test
Coherence	Acc. 90.8 Sen. 95.0 Spec. 87.5	NB/ t-test
Amplitude asymmetry	Acc. 87.2 Sen. 85.8 Spec. 86.4	SVM/ ROC
Phase lag	Acc. 81.8 Sen. 85.0 Spec. 81.7	SVM/ ROC
Level 2		
Absolute power	Acc. 93.1 Sen. 95.8 Spec. 91.7	NB/ t-test
Relative power	Acc. 91.4 Sen. 97.5 Spec. 86.7	NB/ ROC
Coherence	Acc. 86.0 Sen. 87.5 Spec. 91.7	LR/ ROC
Amplitude asymmetry	Acc. 87.8 Sen. 90.0 Spec. 85.0	LR/ Bhatt
Phase lag	Acc. 85.8 Sen. 90.0 Spec. 87.5	NB/ ROC
Level 3		
Absolute power	Acc. 89.9 Sen. 95 Spec. 84.2	SVM/ t-test
Relative power	Acc. 92.5 Sen. 95 Spec. 93.3	NB/ t-test
Coherence	Acc. 87.0 Sen. 95.0 Spec. 79.2	SVM/ ROC
Amplitude asymmetry	Acc. 91.5 Sen. 95.0 Spec. 95.5	LR/ Bhatt
Phase lag	Acc. 82.8 Sen. 85 Spec. 83.3	NB/ t-test
Level 4		
Absolute power	Acc. 90.2 Sen. 95.0 Spec. 84.2	SVM/ t-test
Relative power	Acc. 92.0 Sen. 95.0 Spec. 90.8	LR/ t-test
Coherence	Acc. 86.7 Sen. 95 Spec. 80.4	SVM/ Bhat
Amplitude asymmetry	Acc. 87.8 Sen. 91.7 Spec. 88.3	LR/ Bhatt
Phase lag	Acc. 77.5 Sen. 73.3 Spec. 81.7	LR/ t-test

Acc.: accuracy, Bhat: Bhattacharyya distance, LR: linear regression, NB: naïve Bayes, ROC: receiver operating characteristic, Sen.: sensitivity, Spec.: specificity and SVM: support vector machine.

and phase-lag. Before presenting them to the classifier, these features were standardized using the z-score followed by a rank-based feature selection approach for which three techniques, ROC, t-test and Bhattacharya distance, were used. Finally, for the classification, LR, SVM and naïve Bayes classifiers were used. The results showed that in identifying the four level of stress in *case one,* the maximum accuracies were found to be 94.0%, 93.9%, 94.6% and 90.6%, respectively (Table V). In *case two*, the maximum accuracies in identifying the levels of stress were 94%, 93.1%, 92.5% and 92.0%, respectively (Table VI). To identify the level of stress from the other levels of stress in *case three*, the maximum accuracies were 83.43%, 77.28% 75.61% and 75.47%, respectively (Table VII).

Despite the rapid development of physical and physiological biomarkers of stress, limited reports discuss the application of EEG signals for the assessment of stress. In the earlier review about the psychophysiological biomarkers of stressors, EEG was not included as a biomarker of stress [45], even though EEG signals can more effectively illustrate the stress levels. As discussed in an earlier study, EEG signals illustrated relaxation (contrary to stress) levels, which heart rate and blood pressure failed to represent [46]. Exposure to physiological biomarkers other than EEG is probably observed because stress, after originating in the amygdala, ultimately initiates responses in the ANS [47]. As summarized in Table VIII, the employment of EEGs for the assessment of stress started as late as 2010.

For the induction of stress in the experimental conditions, various experimental tasks have been used in studies, for example, an arithmetic task or a Stroop task. However, whether these tasks induced stress was neither validated nor discussed in many of these studies. For example, the arithmetic task and the Stroop task were utilized to induce high and low levels of stress besides the rest condition as no stress [18]. The tasks that basically produce a cognitive load can be used to induce stress by following certain paradigms. For example, the arithmetic task was presented under social evaluative pressure in the Trier social stress task [48] as well as under time pressure in the Montreal imaging stress task (MIST) [15] and the Stroop task, in which Gaussian noise causes visual fluctuations [49] to induce stress. In this scenario, the validation that the achieved results in the presented studies were due to the induction of stress was unanswered.

The outcome of a classifier strongly depends on the number of samples used for training and testing. The previous studies achieved accuracy based on a small number of samples, and usually, the reported outcomes had the highest accuracy, i.e., the best outcome of only one subject. For example, the reported accuracy of 96% in [18] was achieved from one out of ten subjects. In our study, the achieved results are based on twenty-two subjects. This number of subjects is higher than the number of subjects in the previous studies. Moreover, the reported outcome of the classifier

TABLE 7. Highest performance of every feature in combination with the classifier and feature selection method in every level of stress with respect to the other levels of stress.

Feature	Highest performance	Classifier/Feature selection
Level 1		
Absolute power	Acc. 75.64 Sen. 100 Spec. 41.67	LR/Bhat
Relative power	Acc. 75.89 Sen. 100 Spec. 45.83	LR/t-test
Coherence	Acc. 83.43 Sen. 93.69 Spec. 53.33	LR/ROC
Amplitude asymmetry	Acc. 74.26 Sen. 97.02 Spec. 22.50	NB/Bhat
Phase lag	Acc. 73.03 Sen. 98.57 Spec. 39.17	LR/Bhat
Level 2		
Absolute power	Acc. 75.47 Sen. 100 Spec. 37.5	LR/Bhat
Relative power	Acc. 75.36 Sen. 100 Spec. 37.5	LR/ROC
Coherence	Acc. 77.28 Sen. 97.74 Spec. 28.33	LR/ROC
Amplitude asymmetry	Acc. 74.89 Sen. 98.57 Spec. 12.5	NB/Bhat
Phase lag	Acc. 73.74 Sen. 98.57 Spec. 27.50	LR/Bhat
Level 3		
Absolute power	Acc. 75.28 Sen. 100 Spec. 40.0	LR/Bhat
Relative power	Acc. 75.36 Sen. 100 Spec. 40.83	LR/ROC
Coherence	Acc. 75.61 Sen. 100 Spec. 27.5	LR/Bhat
Amplitude asymmetry	Acc. 75.03 Sen. 99.29 Spec. 16.67	NB/Bhat
Phase lag	Acc. 73.25 Sen. 98.33 Spec. 31.67	LR/Bhat
Level 4		
Absolute power	Acc. 75.11 Sen. 100 Spec. 36.67	LR/Bhat
Relative power	Acc. 75.31 Sen. 100 Spec. 38.33	LR/ROC
Coherence	Acc.75.47 Sen. 99.29 Spec. 29.17	LR/Bhat
Amplitude asymmetry	Acc. 74.78 Sen. 99.29 Spec. 14.17	NB/Bhat
Phase lag	Acc. 73.9 Sen. 98.57 Spec. 30.0	LR/Bhat

Acc.: accuracy, Bhat: Bhattacharyya distance, LR: linear regression, NB: naïve Bayes, ROC: receiver operating characteristic, Sen.: sensitivity, Spec.: specificity and SVM: support vector machine.

TABLE 8. Comparison between our results and the results of recent EEG based methods.

Authors	Objective	EEG feature	Classifier	Performance
2 Level Stress				
Al-Shargie et al. [50]	To fuse EEG with fNIRS for an improved assessment of stress	Average normalized alpha power at PFC	SVM	89.8
Jun, G et al. [18]	To identify low and high levels of stress	Relative difference of alpha and beta power	SVM	96
Vanitha et al. [51]	To detect stress	Hilbert-Huang transform	SVM	89.07
Hou et al. [52]	To recognize the stress level	Fractal dimension + statistical features	SVM KNN	85.71
Al-Shargie et al. [25]	To fuse EEG with fNIRS for an improved assessment of stress	Normalized alpha and beta powers at PFC	SVM	91.7
Al-Shargie et al. [53]	To quantize stress into 3 levels vs control	Alpha power at PFC	SVM	94
Norizam et al. [54]	To determine a stress index that identifies stress levels	Asymmetry ratio, relative energy ratio, Spectral centroids, Spectral entropy	KNN FCM and FKM	88.89
Sani et al. [55]	To identify EEGs from stress and non-stress	Energy spectral density	SVM	83.33
Proposed framework	To identify stress and control	AP, RP, Co, AA, PL	LR, SVM, NB	94.58
Multiple Level Stress				
Jun, G et al. [18]	To identify low and high levels of stress	Relative difference of alpha and beta power	SVM	75
Hou et al. [52]	To recognize the stress level	Fractal dimension + statistical features	SVM KNN	75.22
Proposed framework	To identify stress and control	AP, RP, Co, AA, PL	LR, SVM, NB	83.43

AA: amplitude asymmetry, AP: absolute power, C: coherence, KNN: k-nearest neighbors, LR: linear regression, NB: naïve Bayes, Perf.: performance, RP: relative power, PFC: prefrontal cortex, PL: phase lag, SVM: support vector machine.

(accuracy, sensitivity and specificity) was the median value of all the subjects/repetition of a classifier. Moreover, amongst the studies that compared more than two levels of stress, the performance of our proposed framework (accuracy of 83.43%) outperformed other preceding findings (e.g. accuracy of 75.22% and 67.06% in [52] for recognizing three and four levels of stress, respectively).

An EEG signal is highly sensitive to noise and artifacts. Although there exist several artifact removal techniques, such as independent component analysis (ICA), which separates the artifact space from the signal space and reconstructs the EEG signal, still the risk of artifacts cannot be completely removed. Moreover, to compute the coherence and phase lag from an EEG signal, ICA is not recommended because the reconstruction of an EEG signal in turns harms the raw digital samples and ultimately distorts the computations of the coherence and phase lag [56]. Therefore, for our dataset, we adopted the manual cleaning of the EEG signal from artifacts through visual inspection, i.e., discarding the portion of EEG signal affected by artifacts and selecting only the correct EEG portion for analysis. Split-half reliability and test-retest reliability measures were followed to provide internal consistency and reliability of the signals.

Al-Shargie *et al.* [50], [25] discussed the fusion of EEG and fNIRS for the detection of stress with higher accuracy. Their accuracy with prefrontal EEG signals was 91.7% [25] and 89.9% [1], which increased to 95.1% and 97%, respectively, when fNIRS data were fused with EEG. Our analysis showed a maximum accuracy of 94.58%, which is slightly lower than the result that fusion of the EEG and fNIRS produced. Our experimental condition to induce stress was like MIST. Based on our results, it can be concluded that the EEG signal alone is capable of classifying stress from controls if a thorough analysis is conducted. Hence, the cumbersome routine of fusing two modalities as well as the expensive cost of fNIRS can be avoided.

Possibly, our proposed ML model is confounded with some outliers' other than the relevant patterns extracted from the brain's activity. We have ruled out this concern by (1) properly adopting artifact removal techniques as mentioned before, (2) standardizing the extracted features based on z-scores, (3) randomly selecting each data point such that each data point in the feature space can be used for training and testing of the classification. Considering these precautions, we can conclude that the results shown here are unbiased and a true representation of the information from the recorded EEG data in the stress condition.

V. CONCLUSION

Accurate and reliable identification of stress is essential and requires a valid experimental methodology and analysis framework. The main contribution of this paper lies in developing an experimental paradigm for successfully inducing stress at multiple levels and providing a framework involving

EEG data analysis for the identification of stress at multiple levels. The proposed framework identified stress with a maximum accuracy of 94.6% between 2 levels of stress and the control and 83.43% between stress and the other levels of stress. Our results suggest that EEG signals have the potential to reliably identify stress levels. However, multiple levels of stress require further analysis and validation. The compactness of the EEG system makes it a strong modality for clinical use for the diagnosis of mental stress.

REFERENCES

[1] A. Abbott, "Stress and the city: Urban decay," *Nature* vol. 490, no. 7419, pp. 162–164, 2012."

[2] R. S. Duman, "Neurobiology of stress, depression, and rapid acting antidepressants: Remodeling synaptic connections," *Depression Anxiety*, vol. 31, no. 4, pp. 291–296, Apr. 2014.

[3] C. Espinosa-Garcia *et al.*, "Stress exacerbates global ischemia-induced inflammatory response: Intervention by progesterone," *Stroke*, vol. 48, no. 1, p. ATP83, 2017.

[4] F. Lederbogen *et al.*, "City living and urban upbringing affect neural social stress processing in humans," *Nature*, vol. 474, pp. 498–501, Jun. 2011.

[5] A. P. Allen, P. J. Kennedy, J. F. Cryan, T. G. Dinan, and G. Clarke, "Biological and psychological markers of stress in humans: Focus on the trier social stress test," *Neurosci. Biobehavioral Rev.*, vol. 38, pp. 94–124, Jan. 2014.

[6] H. Ursin and H. Eriksen, "The cognitive activation theory of stress," *Psychoneuroendocrinology*, vol. 29, no. 5, pp. 567–592, Jun. 2004.

[7] N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey," *Comput. Methods Programs Biomed.*, vol. 108, no. 3, pp. 1287–1301, Dec. 2012.

[8] J. T.-Y. Wang, "Pupil dilation and eye tracking," in *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*, A. K. M. Schulte-Mecklenbeck and J. G. Johnson, Eds. New York, NY, USA: Psychology Press, 2011, pp. 185–204.

[9] S. Gowrisankaran, N. K. Nahar, J. R. Hayes, and J. E. Sheedy, "Asthenopia and blink rate under visual and cognitive loads," *Optometry Vis. Sci.*, vol. 89, no. 1, pp. 97–104, 2012.

[10] A. Deschênes, H. Forget, C. Daudelin-Peltier, D. Fiset, and C. Blais, "Facial expression recognition impairment following acute social stress," *J. Vis.*, vol. 15, no. 15, p. 1383, 2015.

[11] G. C. Dieleman *et al.*, "Alterations in HPA-axis and autonomic nervous system functioning in childhood anxiety disorders point to a chronic stress hypothesis," *Psychoneuroendocrinology*, vol. 51, pp. 135–150, Jan. 2015.

[12] A. R. Subahni, L. Xia, and A. S. Malik, "Association of mental stress with video games," in *Proc. 4th Int. Conf. Intell. Adv. Syst. (ICIAS)*, Jun. 2012, pp. 82–85.

[13] J. Wielgosz, B. S. Schuyler, A. Lutz, and R. J. Davidson, "Long-term mindfulness training is associated with reliable differences in resting respiration rate," *Sci. Rep.*, vol. 6, pp. 1–6, 2016.

[14] A. Liapis, C. Katsanos, D. Sotiropoulos, M. Xenos, and N. Karousos, "Recognizing emotions in human computer interaction: Studying stress using skin conductance," in *Human-Computer Interaction* (Lecture Notes in Computer Science), vol. 9296, J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, and M. Winckler, Eds. Cham, Switzerland: Springer, 2015, pp. 255–262.

[15] K. Dedovic, R. Renwick, N. K. Mahani, V. Engert, S. J. Lupien, and J. C. Pruessner, "The montreal imaging stress task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain," *J. Psychiatry Neurosci.*, vol. 30, no. 5, pp. 319–325, 2005.

[16] A. R. Subhani, L. Xia, and A. S. Malik, "Association of electroencephalogram scalp potential with autonomic nervous system," *Adv. Sci. Lett.*, vol. 19, no. 5, pp. 1340–1344, 2013.

[17] A. R. Subhani, L. Xia, A. S. Malik, and Z. Othman, "Quantification of physiological disparities and task performance in stress and control conditions," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 2060–2063.

[18] G. Jun and K. G. Smitha, "EEG based stress level identification," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 3270–3274.

[19] J. F. Alonso, S. Romero, M. R. Ballester, R. M. Antonijoan, and M. A. Mañanas, "Stress assessment based on EEG univariate features and functional connectivity measures," *Physiol. Meas.*, vol. 36, no. 7, p. 1351, 2015.

[20] I. Dziembowska, P. Izdebski, A. Rasmus, J. Brudny, M. Grzelczak, and P. Cysewski, "Effects of heart rate variability biofeedback on EEG alpha asymmetry and anxiety symptoms in male athletes: A pilot study," *Appl. Psychophysiol. Biofeedback*, vol. 41, no. 2, pp. 141–150, Jun. 2016.

[21] P. Cipresso *et al.*, "EEG alpha asymmetry in virtual environments for the assessment of stress-related disorders," Studies in Health Technology and Informatics, Newport Beach, CA, USA, 2012, pp. 102–104.

[22] P. M. Pandiyan and S. Yaacob, "Mental stress level classification using eigenvector features and principal component analysis," *Commun. Inf. Sci. Manage. Eng.*, vol. 3, no. 5, p. 254, 2013.

[23] N. Adnan, Z. Hj Murat, R. S. S. A. Kadir, and N. Hj Mohamad Yunos, "University students stress level and brainwave balancing index: Comparison between early and end of study semester," in *Proc. IEEE Student Conf. Res. Develop. (SCOReD)*, Dec. 2012, pp. 42–47.

[24] J. Yang, M. Qi, L. Guan, Y. Hou, and Y. Yang, "The time course of psychological stress as revealed by event-related potentials," *Neurosci. Lett.*, vol. 530, no. 1, pp. 1–6, Nov. 2012.

[25] 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.

[26] Y. Jung and Y. I. Yoon, "Multi-level assessment model for wellness service based on human mental stress level," *Multimedia Tools Appl.*, vol. 76, no. 9, pp. 11305–11317, May 2016.

[27] R. Khosrowabadi, C. Quek, K. K. Ang, S. W. Tung, and M. Heijnen, "A brain-computer interface for classifying EEG correlates of chronic mental stress," in *Proc. Int. Joint Conf. Neural Netw.*, San Jose, CA, USA, Jul. 2011, pp. 757–762.

[28] A. Gaggioli *et al.*, "A decision support system for real-time stress detection during virtual reality exposure," in *Proc. 21st Med. Meets Virtual Reality Conf. (NextMed/MMVR)*, Manhattan Beach, CA, USA, 2014, pp. 114–120.

[29] K. S. Rahnuma, A. Wahab, N. Kamaruddin, and H. Majid, "EEG analysis for understanding stress based on affective model basis function," in *Proc. 15th IEEE Int. Symp. Consum. Electron. (ISCE)*, Singapore, Jun. 2011, pp. 592–597.

[30] K. Kalimeri and C. Saitis, "Exploring multimodal biosignal features for stress detection during indoor mobility," in *Proc. 18th ACM Int. Conf. Multimodal Interact.*, 2016, pp. 53–60.

[31] A. Secerbegovic, S. Ibric, J. Nisic, N. Suljanovic, and A. Mujcic, "Mental workload vs. stress differentiation using single-channel EEG," in *International Conference on Medical and Biological Engineering*. Singapore: Springer, 2017, pp. 511–515.

[32] S. Cohen, T. Kamarck, and R. Mermelstein, "A global measure of perceived stress," *J. Health Social Behavior*, vol. 24, no. 4, pp. 385–396, Dec. 1983.

[33] K. Dedovic, A. Duchesne, J. Andrews, V. Engert, and J. C. Pruessner, "The brain and the stress axis: The neural correlates of cortisol regulation in response to stress," *NeuroImage*, vol. 47, no. 3, pp. 864–871, Sep. 2009.

[34] J. C. Pruessner *et al.*, "Deactivation of the limbic system during acute psychosocial stress: Evidence from positron emission tomography and functional magnetic resonance imaging studies," *Biol. Psychiatry*, vol. 63, no. 2, pp. 234–240, Jan. 2008.

[35] *E-Prime User's Guide*, Psychol. Softw. Tools, Pittsburgh, PA, USA, 2012.

[36] R. Thatcher. (2008). NeuroGuide manual and tutorial. Applied Neuroscience, St. Petersburg, FL, USA. [Online]. Available: http://www.AppliedNeuroscience.com/NeuroGuide_Deluxe.pdf

[37] T. F. Collura, "Towards a coherent view of brain connectivity," *J. Neurotherapy*, vol. 12, nos. 2–3, pp. 99–110, 2009.

[38] A. R. Subhani, A. S. Malik, N. Kamil, M. Naufal, and M. Saad, "Using resting state coherence to distinguish between low and high stress groups," in *Proc. 6th Int. Conf. Intell. Adv. Syst. (ICIAS)*, Aug. 2016, pp. 1–4.

[39] J. R. Carter and C. A. Ray, "Sympathetic neural responses to mental stress: Responders, nonresponders and sex differences," *Amer. J. Physiol. Heart Circulatory Physiol.*, vol. 296, no. 3, pp. H847–H853, 2009.

[40] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. San Diego, CA, USA: Academic, 2008.

[41] H. Mamitsuka, "Selecting features in microarray classification using ROC curves," *Pattern Recognit.*, vol. 39, no. 12, pp. 2393–2404, Dec. 2006.
[42] A. Chiesa and A. Serretti, "A systematic review of neurobiological and clinical features of mindfulness meditations," *Psychol. Med.*, vol. 40, no. 8, pp. 1239–1252, Aug. 2010.
[43] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
[44] H. D. Critchley, J. Tang, D. Glaser, B. Butterworth, and R. J. Dolan, "Anterior cingulate activity during error and autonomic response," *NeuroImage*, vol. 27, no. 4, pp. 885–895, Oct. 2005.
[45] T. Chandola, A. Heraclides, and M. Kumari, "Psychophysiological biomarkers of workplace stressors," *Neurosci. Biobehavioral Rev.*, vol. 35, no. 1, pp. 51–57, Sep. 2010.
[46] T. A. Lin and L. R. John, "Quantifying mental relaxation with EEG for use in computer games," in *Proc. Int. Conf. Internet Comput.*, 2006, pp. 409–415.
[47] B. L. Seaward, *Managing Stress: Principles and Strategies for Health and Wellbeing*, 7th ed. Boston, MA, USA: Jones & Bartlett, 2011.
[48] C. Kirschbaum, K. M. Pirke, and D. H. Hellhammer, "The 'Trier social stress testâĂŹ—A tool for investigating psychobiological stress responses in a laboratory setting," *Neuropsychobiology*, vol. 28, nos. 1–2, pp. 76–81, 1993.
[49] H. Peng, Y. Gao, and X. Mao, "The roles of sensory function and cognitive load in age differences in inhibition: Evidence from the Stroop task," *Psychol. Aging*, vol. 32, no. 1, pp. 42–50, 2017.
[50] F. Al-Shargie, T. B. Tang, and M. Kiguchi, "Assessment of mental stress effects on prefrontal cortical activities using canonical correlation analysis: An fNIRS-EEG study," *Biomed. Opt. Exp.*, vol. 8, no. 5, pp. 2583–2598, 2017.
[51] V. Vanitha and P. Krishnan, "Real time stress detection system based on EEG signals," *Biomed. Res.*, vol. 27, pp. 271–275, Aug. 2016.
[52] X. Hou, Y. Liu, O. Sourina, Y. R. E. Tan, L. Wang, and W. Mueller-Wittig, "EEG based stress monitoring," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2015, pp. 3110–3115.
[53] F. Al-Shargie, T. B. Tang, N. Badruddin, and M. Kiguchi, "Mental stress quantification using EEG signals," in *Proc. Int. Conf. Innov. Biomed. Eng. Life Sci.*, 2015, pp. 15–19.
[54] S. Norizam, "Determination and classification of human stress index using nonparametric analysis of EEG signals," Ph.D. dissertation, Faculty Elect. Electron. Eng., Univ. Teknol. Mara, Shah Alam, Malaysia, 2015. [Online]. Available: http://umpir.ump.edu.my/16490/
[55] M. M. Sani, H. Norhazman, H. A. Omar, N. Zaini, and S. A. Ghani, "Support vector machine for classification of stress subjects using EEG signals," in *Proc. IEEE Conf. Syst., Process Control (ICSPC)*, Dec. 2014, pp. 127–131.
[56] T. H. Budzynski, H. K. Budzynski, J. R. Evans, and A. Abarbanel, *Introduction to Quantitative EEG and Neurofeedback: Advanced Theory and Applications*, 2nd ed. Cambridge, MA, USA: Academic, 2008.

AHMAD RAUF SUBHANI received the B.E. degree in electrical engineering from Air University, Islamabad, Pakistan, in 2009, and the M.S. degree from Universiti Teknologi PETRONAS, Malaysia, in 2013, where he is currently pursuing the Ph.D. degree in electrical engineering.

From 2010 to 2011, he was a Research Assistant with the Center for Advanced Studies in Telecommunication. His research interests include biomedical signal processing and machine learning for rehabilitation.

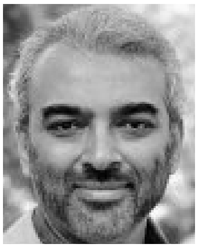
WAJID MUMTAZ received the Ph.D. degree in electrical engineering from Universiti Teknologi PETRONAS (UTP), Malaysia, in 2017.

He is currently a Research Scientist with the Center for Intelligent Signal and Imaging, UTP, Malaysia. He has authored over ten articles. His research interest includes neuro-signal processing and machine learning for diagnosis.

MOHAMED NAUFAL BIN MOHAMED SAAD received the master's degree from the Ecole Nationale Supérieure d'Ingénieurs de Limoges, France, and the Ph.D. degree in telecommunication from the Université de Limoges, France, in 2005. He is currently an Associate Professor with the Electrical and Electronic Engineering Department, Universiti Teknologi Petronas (UTP), Malaysia. He is a Core Research Member with the Center for Intelligent Signal and Imaging Research, UTP. He is also the Director of Mission Oriented Research in Biomedical Technology. His research interests include neuro signal processing, medical imaging, and communication.

NIDAL KAMEL received the Ph.D. degree in the scope of Statistical Signal Processing in telecommunication from the Technical University of Gdansk, Poland.

His is currently an Associate Professor with the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, where he is also the Head of Neuro-Signal Processing Group. He is the author of a book *EEG/ERP Analysis* and over 50 articles. His research interests include statistical signal processing and brain signal processing.



AAMIR SAEED MALIK received the Ph.D. degree in information and mechatronics from the Gwangju Institute of Science and Technology, South Korea, in 2008.

He was the Ex-Director of the Mission-Oriented Research for Biomedical Technology with Universiti Teknologi PETRONAS (UTP). He is attached with CISIR. He is currently an Associate Professor with the Department of Electrical and Electronic Engineering, UTP. He has authored three books and more than 100 journal articles. His research interests include brain sciences and neuroimaging.