Classification of Perceived Mental Stress Using A Commercially Available EEG Headband

Aamir Arsalan, Muhammad Majid, Amna Rauf Butt, and Syed Muhammad Anwar

***Abstract*—Human stress is a serious health concern, which must be addressed with appropriate actions for a healthy society. This paper presents an experimental study to ascertain the appropriate phase, when electroencephalography (EEG) based data should be recorded for classification of perceived mental stress. The process involves data acquisition, pre-processing, feature extraction and selection, and classification. The stress level of each subject is recorded by using a standard perceived stress scale questionnaire, which is then used to label the EEG data. The data are divided into two (stressed and non-stressed) and three (non-stressed, mildly stressed, and stressed) classes. The EEG data of twenty eight participants are recorded using a commercially available four channel Muse EEG headband in two phases i.e., pre-activity and post-activity. Five feature groups, which include power spectral density, correlation, differential asymmetry, rational asymmetry, and power spectrum are extracted from five bands of each EEG channel. We propose a new feature selection algorithm which selects features from appropriate EEG frequency band based on classification accuracy. Three classifiers i.e., support vector machine, the Naive Bayes and multi-layer perceptron are used to classify stress level of the participants. It is evident from our results that EEG recording during the pre-activity phase is better for classifying the perceived stress. An accuracy of 92.85% and 64.28% is achieved for two- and three-class stress classification respectively, while utilizing five groups of features from theta band. Our proposed feature selection algorithm is compared with existing algorithms and gives better classification results.**

***Index Terms*—perceived stress, electroencephalography, feature extraction, classification.**

I. INTRODUCTION

PSYCHOLOGICAL stress in humans is a serious health problem, which severely affects their working abilities. According to the world health organization, depression affects 350 million people around the globe and on an average one out of twenty people had an episode of depression [1]. Stress is a reaction of the human body to a challenging situation that disturbs the mental equilibrium and is induced by mental, physical, and emotional factors [2]. Human stress is generally categorized into two main categories i.e., perceived stress and acute stress. Perceived stress is a long term condition, which can occur due to social issues such as a bad career, an unhappy marriage, family issues, and poverty. Whereas, acute stress is an instantaneous condition, which could arise due to situations such as a nearly missed accident, a costly mistake at work, or an argument with a family member. Stress could increase the chances of serious health issues like stroke, depression, and heart attack [3], [4]. Therefore, an intelligent system is required that can assess and measure stress for ensuring a better human life style.

Human psychological stress is measured using subjective and objective measures [5], [6]. Subjectively, stress is either measured using questionnaires (developed by researchers) or interviews conducted by an expert psychologist. Perceived stress scale (PSS) [7], stress response inventory (SRI) [8], and state trait anxiety inventory (STAI) [9] are a few examples of questionnaires, which have been used for this task. Among these PSS is a commonly used questionnaire to observe perceived stress of an individual. Objectively, human stress is measured using physical as well as physiological methods. In physical measures, a change could be visible in the form of facial expressions [10], eye blink rate [11], and pupil dilation [12]. Whereas, physiological methods require sensors attached to the body of an individual to measure changes occurring in the body. In a stressful situation, the brain part (amygdala) associated with emotional processing sends a signal to the hypothalamus, which is responsible for preparing the body response. Hypothalamus activates the autonomic nervous system (ANS) of a human being. Stress and relaxation states are associated with sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) of ANS. Stress activates the SNS and suppresses the PNS within the human brain. ANS sends the signals to the adrenal glands, which responds by pumping adrenaline into the blood vessels. The circulation of adrenaline in human body results in a number of physiological changes. These changes are induced by stress in sympathetic and parasympathetic nervous system of ANS [13]. Various bio-markers which include heart rate (HR) and heart rate variability (HRV) [14], skin conductance [15], and electroencephalography (EEG) [16] have been used for measuring stress. Recent studies in neuroscience have established that the response of the human brain is affected by stress [4]. A situation is considered stressful or challenging depending on the perception of the human brain. Functional magnetic resonance imaging (fMRI), positron emission tomography (PET) and EEG are commonly used modalities to analyze brain activity in stressful conditions [17]. With recent advancements and commercial availability of EEG headsets at lower costs, EEG could be considered as a preferred modality for monitoring human stress [18].

In literature, different EEG based methods (which are non-invasive) are proposed to measure human stress. Most of these methods are for acute stress, which is induced in a laboratory

Manuscript received December 31, 2018.

Aamir Arsalan, Muhammad Majid, Amna Rauf Butt are with the Department of Computer Engineering, University of Engineering and Technology, Taxila, 47050 Pakistan (e-mail: m.majid@uettaxila.edu.pk).

Syed Muhammad Anwar is with the Department of Software Engineering, University of Engineering and Technology, Taxila, 47050 Pakistan.

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/JBHI.2019.2926407, IEEE Journal of Biomedical and Health Informatics

environment by using stressors such as stroop colour-word test, trier social stress test (TSST), cold pressor test, public speaking task, and mental arithmetic task [19], [20], [21], [22], [23]. Whereas, few studies are available to measure or classify perceived or chronic stress [24], [25], [26], [27], [28], [29], [30], [19]. EEG data in these methods is recorded either in an open eye or closed eye condition. In open eye condition, participants perform an activity or focus on a blank screen. On the other hand, in close eye condition participants are usually seated in a relaxed position with eyes closed. In [24], an EEG based study was presented to develop a relationship between the PSS score of the participants and their EEG signals. It was reported that people having a high level of perceived stress showed an increased beta activity as compared to non-stressed individuals. An EEG based study to quantify human stress using a single channel EEG headset in closed eye condition was presented in [25]. A regression analysis was performed by using beta waves to predict the PSS score of an individual with a confidence level of 94%, but no information was provided regarding the mental state of the subject while recording the EEG. Another EEG based perceived stress classification study based on correlation-based feature subset selection method was presented in [26]. The study found that among all EEG frequency bands, low beta, high beta, and low gamma are most highly correlated to perceived stress. In [27], it was reported that energy spectral density of alpha and beta activity within the left and right hemispheres of the brain in stressed and non-stressed subjects was significantly different. In [28], alpha asymmetry index was used as a measure of relationship between stress and EEG. Subjects with low chronic stress had dominant activity in the left hemisphere of the brain, whereas subjects with moderate and high stress showed dominant activity in the right hemisphere of the brain.

A negative correlation between the PSS score and the alpha and beta activity ratio of the subject was reported in [29], hence alpha and beta ratio was negative for high and positive for low PSS scores. A study to correlate EEG, PSS score, and EEG temporal characteristics was presented in [30]. It was observed that subjects having high PSS score have an increased delta and theta band activity in the post-stimuli phase as compared to the pre-stimuli phase. It was also observed that the frontal channel theta activity increased in the post-stimuli phase. In [19], a study was presented to investigate whether features from single channel EEG signal can discriminate between relaxed, cognitive workload, and stressful states. Experimental results showed that EEG features could distinguish between these three states with an acceptable level of accuracy.

Most studies reported in literature have primarily focused on developing a relationship between EEG signals and PSS score and hence finding the brain activity patterns for subjects reporting high and low scores. Moreover, to the best of our knowledge, there are no studies reported in literature to find appropriate condition (pre- or post-activity) for observing EEG data to classify perceived stress. There are publicly available EEG datasets for acute stress, which is not the case for perceived stress. Also, perceived stress classification is taken as a two-class problem. To this end, our proposed study was

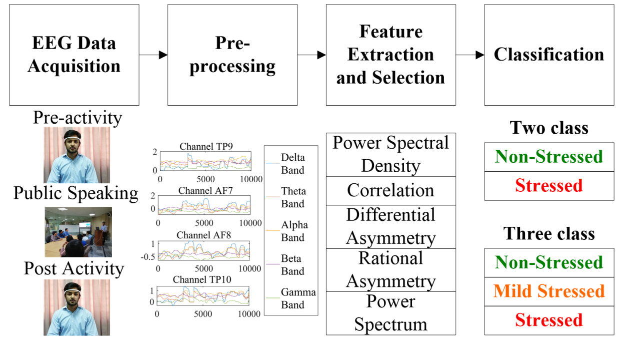


Fig. 1. Block diagram of the proposed methodology used to classify perceived stress using EEG.

focused on finding an appropriate time of EEG recording in perceived stress studies. Thereby we aim to determine whether pre- or post-activity phase would be better in classifying the perceived stress into three different classes. The participants involved in the study filled PSS questionnaire and the scores (after thresholding) were used to label subjects in multiple (two: non-stressed and stressed and three: stressed, mild- and non-stressed) classes. The EEG data were recorded during the pre- and post-activity phase using four channels Muse head-band. During the activity phase, subjects were asked to prepare a presentation and present it in front of an audience. Five groups of features were extracted from five EEG frequency bands of each channel. The most appropriate features (feature reduction) were selected using our proposed algorithm from different EEG bands based on accuracy. The selected features were used for classifying perceived stress into two- and three-classes. The major contributions of this study are,

1) A new EEG dataset for the classification of perceived stress was acquired, which is publicly available [1] for future research in the field of mental and stress related health issues.
2) We propose a new feature selection algorithm specifically for stress classification task, improving the classification accuracy when compared with existing feature selection algorithms.
3) We present multi-class classification of perceived stress based on pre- and post-activity phases of EEG recordings.
4) We found that pre-activity phase EEG recordings could be more discriminating for better classification of perceived stress.

The rest of the paper is organized as follows. Section II presents the materials and methods used in this study. Multi-class perceived stress classification results are presented in Section III. Section IV presents the comparison and discussion followed by conclusion in Section V.

II. Material and Methods

A block diagram of the proposed framework to classify perceived stress using EEG is presented in Fig. 1. This diagram

[1] https://sites.google.com/site/simplrgp/resources

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/JBHI.2019.2926407, IEEE Journal of Biomedical and Health Informatics

consists of four stages: EEG data acquisition, pre-processing, feature extraction and selection, and classification. The details of each stage is presented in the following subsections.

A. EEG Data Acquisition

1) Participants: A total of 28 subjects (13 males and 15 females) with an age ranging from 18 to 40 years participated in this study. All participants selected in this study had a minimum of twelve years of education and were either student or instructor at the university level. None of the participants reported any mental illness. The experiments for this study were designed in accordance to the Helsinki declaration. The board of advanced studies research and technological development at University of Engineering and Technology, Taxila has approved this study.

2) Apparatus: For recording EEG signals we used Interaxon Muse headband, which is a versatile and easy to use EEG recording system. It is a four channel headband with dry electrodes at positions $AF7$, $AF8$, $TP9$ and $TP10$. A reference electrode is placed on the forehead of the individual at position Fpz. The material used for the frontal electrodes is silver, whereas conductive silicon rubber is used for temporal electrodes. The Muse headband records EEG data at a sampling rate of 256 Hz. Muse headband was paired to a smart phone via a Bluetooth connection for data transmission. A mobile application (Muse Monitor) was used to record EEG data on smart phone, which was transferred to a PC for further offline processing.

3) Experimental Procedure: Each participant was escorted to a quiet and temperature controlled room with consistent lightening conditions, where they were introduced to the experimental procedure, and asked to sign the consent form. The subjects were also asked to fill the demographic questionnaire (bio data sheet) for recording the age, gender, and information about mental illnesses. Subjects were then directed to fill in the PSS questionnaire, which was designed by expert psychologists to measure the level of perceived stress. It is a 10 item questionnaire that measures how much stress an individual had in the last thirty days. Subjects can answer each question on a scale of 0 to 4, where 0 means a particular event never occurred to the individual, and 4 means an event occurred frequently in the past thirty days. The score of each question in the questionnaire sums up to the total perceived stress score of the subject, which is between $0-40$. The PSS scores were used to label each subject in either of the two (non-stressed and stressed) or three (non-stressed, mildly stressed, and stressed) classes.

Once the PSS questionnaire was filled, the pre-activity EEG data of subjects were recorded for a duration of 3 minutes in an open eye condition while sitting in a relaxed position. The subject were given a relaxation period (after pre-activity recording) and then asked to prepare a talk on an unknown topic. The slides (with content on the topic) were given to participants for helping with the preparation. Once the participants indicated that they are ready, they were given another relaxation period, and then taken to another room for presentation in front of a real audience. The size of this audience was between 15 to 25 people including students and faculty members. The EEG data were recorded during the presentation (activity phase) but was not used in this study. After the presentation, the subjects were taken back to the recording room and the post-activity EEG data were recorded for a duration of 3 minutes.

B. Pre-processing

The recorded EEG signal from participants were pre-processed before feature extraction and classification. An on-board driven right leg (DRL) feedback circuit was used for noise cancellation within the recorded EEG signals. DRL circuits make sure that EEG electrodes have proper skin contact. A clean signal can be identified by thresholding statistical properties such as mean power, standard deviation of power, maximum amplitude, standard deviation of amplitude, kurtosis of amplitude and skewness of amplitude of the EEG signal [31]. We used the on-board noise cancellation mechanism of muse headband, which considered the EEG signal as clean if the incoming signal had variance, amplitude, and kurtosis values less than a pre-determined threshold. The EEG frequency bands were obtained by using the on-board digital signal processing module, which took fast Fourier transform for raw EEG signals with an overlap of 90% on window size of 256. Delta (0-4 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-50 Hz) frequency bands were obtained from the muse headband.

C. Feature Extraction and Selection

For analysis of the recorded EEG data, five groups of features were extracted from each frequency band of each channel. These included power spectral density (PSD), correlation (C), divisional asymmetry (DASM), rational asymmetry (RASM), and power spectrum (PS). PSD describes the power distribution of the signal over a particular frequency. PSD was calculated using Welch method [32] with an overlap of 50%. The mean and variance of the PSD of each band from each channel were used as features from this group. A total of forty values were obtained from four channels and five bands for this feature group, represented as $F_{G_1} = [f_\delta(1,8)\ f_\theta(1,8)\ f_\alpha(1,8)\ f_\beta(1,8)\ f_\gamma(1,8)]$. Correlation is a statistical measure that indicates the extent to which two values vary together. In this study, the correlation between asymmetric channels for the left and right hemisphere of the brain was calculated. A correlation between two pairs of electrodes i.e., ($TP9$, $TP10$), and ($AF7$, $AF8$) was calculated. A total of ten values i.e., five for each pair from each frequency band was calculated for this feature group represented as $F_{G_2} = [f_\delta(1,2)\ f_\theta(1,2)\ f_\alpha(1,2)\ f_\beta(1,2)\ f_\gamma(1,2)]$. DASM is the difference between the absolute power of asymmetric channels of the left and right hemisphere of the brain. A total of ten features for DASM i.e., five for each pair from each band were used for this feature group and is represented as $F_{G_3} = [f_\delta(1,2)\ f_\theta(1,2)\ f_\alpha(1,2)\ f_\beta(1,2)\ f_\gamma(1,2)]$. RASM is the ratio between the absolute power of asymmetric channels of the left and right hemisphere of the brain. A total of ten features for RASM i.e., five for each pair from each band were used for this feature group and is represented as $F_{G_4} = [f_\delta(1,2)\ f_\theta(1,2)\ f_\alpha(1,2)\ f_\beta(1,2)\ f_\gamma(1,2)]$. Power spectrum consists of average absolute power of four scalp electrodes in the five frequency bands of the EEG signal. A total of twenty features for power spectrum i.e., five for each channel were used for this feature group and is represented as $F_{G_5} = [f_\delta(1,4)\ f_\theta(1,4)\ f_\alpha(1,4)\ f_\beta(1,4)\ f_\gamma(1,4)]$.

In order to train the classification model, we propose a feature selection algorithm specifically designed for stress classification task based on accuracy. Generally accuracy is used to measure the fraction of correct predictions. A value close to 1 suggests superior performance of the classification model. Therefore we selected features from those EEG frequency bands which enhance the classification accuracy. All possible combinations (form five frequency bands) were utilized to select the optimum frequency band. The classification model was trained by using 90% of the data, and the remaining 10% of the data was used for testing. For the selection of features, 1000 iterations were considered for each combination. Features from those frequency bands were selected, which gave the maximum accuracy. The details of our proposed features selection method are given in Algorithm 1.

Algorithm 1: Feature Selection based on EEG Frequency Bands

Input :
1) Extracted feature groups $FV = [F_{G_1}\ F_{G_2}\ F_{G_3}\ F_{G_4}\ F_{G_5}]$.
2) Class Labels = $[NS, S]$ or $[NS, MS, S]$
3) Classifier = SVM, NB, MLP.
4) EEG frequency bands = δ, θ, α, β, γ.
5) Number of EEG bands = N.

Output: Set of selected features F_S.

ind = 1;
for *i*=1 *to N* **do**
  Compute all possible combinations $T = {}^{N}C_i$
  for *j*=1 *to length(T)* **do**
    $F_{temp}(1, ind)$ = Select subset of features from FV based on combination $T(j)$.
    C = Array of 1000 elements.
    for *k = 1 to 1000* **do**
      Randomly select
      90% data for training
      10% data for testing
      C(1,k) = Compute classification accuracy using input classifier and class labels.
    end
    $A(1, ind) = sum(C)/1000$.
    ind = ind + 1;
  end
end
$indf$ = Find index that gives maximum accuracy.
$F_S = F_{temp}(1, indf)$
return F_S.

D. Classification

In this study, three different classifiers were used to classify perceived stress using EEG signals into two and three classes. The details of the classifiers are as follows.

1) Support Vector Machine (SVM): In SVM, data items are placed in an n-dimensional space. The classification of data points is performed by finding a hyper plane, which separates the data point of the classes in the best possible manner. To find the optimum hyper plane, SVM performs an iterative training algorithm, which tries to minimize an error function. Points near the hyper plane are termed as support vectors. SVM can classify the data in a linear or non-linear manner. We used the polynomial (degree: $d = 2$) kernel function with complexity constant $C = 1$.

2) The Naive Bayes (NB): The Naive Bayes classification algorithm is based on the generation of conditional probabilities of the classes in which the data is to be classified. The conditional probabilities of each input value based on the class probabilities are calculated. Training of the Naive Bayes algorithm is comparatively faster than other classification algorithms, as no complex optimization parameters are required. Test sample is assigned to the class for which the classifier has the highest probability.

3) Multilayer Perceptron (MLP): Multilayer perceptron is a class of feed forward neural network, which is used for the classification purpose. Simplest kind of MLP consists of three layers, i.e., input, hidden, and the output layer. A transfer function is used for mapping the input of the neurons of each layer to the output. In MLP, input data is transformed into linearly separable points by using a non-linear transformation. The transformation function is applied in the hidden layer. We used a network with 4 hidden layers, a learning rate of 0.3 with momentum rate of 0.2.

III. Experimental Results

A. Data Labeling

Subjects were labeled into non-stressed and stressed group for two-class problem, and non-stressed, mildly stressed and stressed for three-class problem using their PSS scores. The PSS score of each participant is shown in Fig. 2. The mean and variance of the recorded PSS scores of 28 subjects participated in this study was $\mu = 22.2$ and $\sigma = 6.46$, respectively. For the two-class problem, a threshold value of 20 was selected. Subjects having PSS score between 0-20 were labeled as non-stressed, whereas participants having score between 21-40 were labeled as stressed. For the three-class problem, subjects

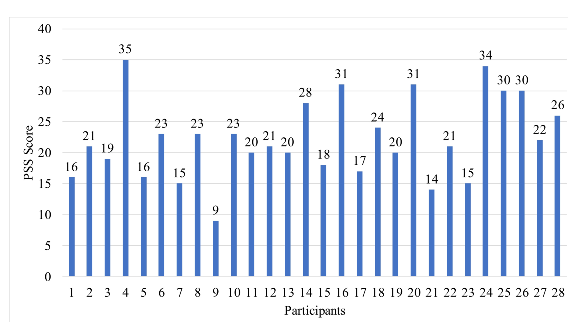
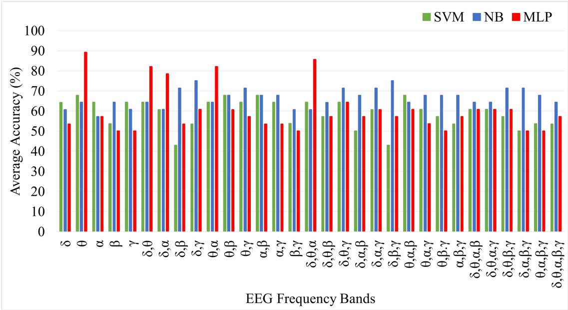
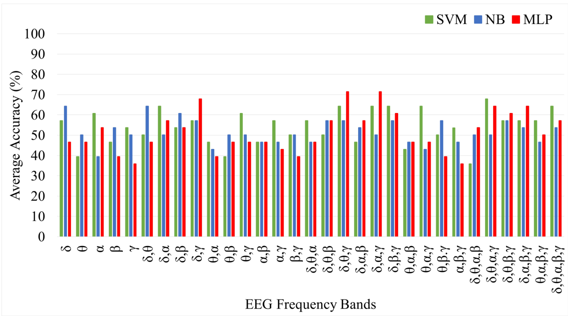


Fig. 2. Distribution of PSS scores of the subjects participated in this study.



(a)



(b)

Fig. 3. Classification accuracy of SVM, NB and MLP classifiers based on features extracted from different combinations of EEG frequency bands for the two class classification problem based on (a) pre-activity (b) post-activity phase recording.

TABLE I
PERFORMANCE COMPARISON OF THE PROPOSED FEATURE SELECTION ALGORITHM IN TERMS OF CLASSIFIER USED, EEG BANDS FOR FEATURE EXTRACTION, ACCURACY, F-MEASURE AND KAPPA STATISTICS FOR TWO CLASS AND THREE CLASS PERCEIVED STRESS CLASSIFICATION USING PRE-ACTIVITY EEG RECORDING.

Cross Validation	Classes	Classifier	EEG band	*Acc*	F_m	*K*
10-fold	Two	MLP	All	57.14	0.57	0.14
			θ	**89.28**	**0.89**	**0.77**
		SVM	All	53.50	0.53	0.04
			θ	67.85	0.61	0.27
		NB	All	64.28	0.61	0.22
			θ	64.28	0.61	0.22
	Three	MLP	All	32.14	0.31	-0.03
			θ	**60.71**	**0.60**	**0.40**
		SVM	All	35.71	0.31	0.007
			θ	42.85	0.38	0.10
		NB	All	39.28	0.33	0.04
			θ	42.85	0.36	0.10
Leave -one-out	Two	MLP	All	71.42	0.67	0.36
			θ	**92.85**	**0.91**	**0.80**
		SVM	All	57.14	0.57	0.14
			θ	57.14	0.46	0.02
		NB	All	71.42	0.68	0.37
			θ	75.00	0.72	0.44
	Three	MLP	All	32.78	0.32	-0.04
			θ	**64.28**	**0.64**	**0.27**
		SVM	All	39.28	0.39	-0.22
			θ	46.42	0.46	-0.08
		NB	All	42.85	0.41	-0.19
			θ	50.00	0.50	-0.02

with PSS score between 0 to $\lceil \mu - \frac{\sigma}{2} \rceil$ (0-19) were labelled as non-stressed, those with PSS score between $\lceil \mu - \frac{\sigma}{2} \rceil + 1$ to $\lceil \mu + \frac{\sigma}{2}) \rceil - 1$ (20-25) were labelled as mildly stressed, and those with PSS score between $\lceil \mu + \frac{\sigma}{2} \rceil$ to 40 (26-40) were labelled as stressed. Based on two class labelling, 12 participants were labeled as non-stressed and 16 as stressed. For three class problem, 9 participants were labeled as non-stressed, 11 as mildly stressed, and 8 as stressed.

B. Performance of Feature Selection

Here we present the performance of our proposed feature selection algorithm in classifying perceived stress into two- and three-classes using EEG data recorded at two different phases for the same subject. We used both 10-fold and leave-one-out cross validation to measure the classifier's performance. The parameters used to evaluate the performance of our proposed system are accuracy (A_{cc}), F-measure (F_m), and kappa statistics (K). Accuracy is defined as the percentage of correctly classified instances from the total number of instances available. F-measure is the harmonic mean of precision and recall, and kappa statistics is a parameter used to measure the reliability among two human ratings.

For classification of perceived stress, three different classifiers were used i.e., SVM, NB, and MLP. Features were extracted from all frequency bands of EEG data i.e., δ, θ, α, β, and γ resulting in a feature vector length of 90. The number of features for each feature group PSD, C, DASM, RASM, and PS were 40, 10, 10, 10, and 20 respectively. For feature selection, we tested all possible combinations of frequency bands from each feature group. Features from only those frequency bands were selected which enhanced the classification accuracy. Fig. 3 (a) and Fig. 3 (b) shows the average classification accuracy of two class problem based on 10-fold cross validation for all three classifiers using selected features from different bands in pre- and post-activity phase of EEG recordings respectively. It is evident that the highest accuracy of 89.28% was given by MLP classifier in the pre-activity phase, when all feature groups of θ band (feature vector length of 18) were used. Whereas in case of post-activity phase the highest classification accuracy of 71.42% was achieved by MLP classifier when all feature groups from δ, θ, and γ band (feature vector length of 54) were used. Pre-activity phase has a higher classification accuracy as compared to post-activity phase with a minimum feature vector length of 18 using only θ band. It can be observed that pre-activity phase EEG recording was better in classifying perceived stress as compared to post-activity phase EEG recordings. A t-test was applied on the absolute power of EEG bands from pre- and post-activity phases where only θ band was found significantly different having p-value of 0.04. This could be due to the fact that presentation induces instantaneous stress in some users (which affects the θ band activity) and reduces the classification accuracy of perceived stress using post-activity EEG data.

Based on these findings we only used pre-activity phase EEG recordings for three class classification problem and features from θ band having feature vector length of 18. Our proposed feature selection algorithm gave features from the theta band including mean and variance of power spectral

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/JBHI.2019.2926407, IEEE Journal of Biomedical and Health Informatics

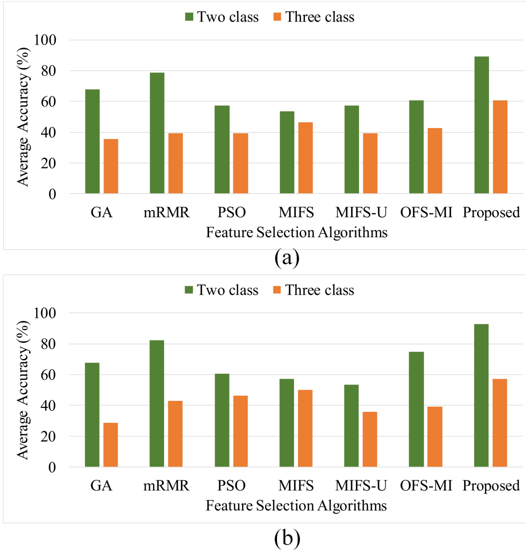


Fig. 4. Performance comparison of the proposed feature selection algorithm with GA, mRMR, PSO, MIFS, MIFS-U, OFS-MI feature selection algorithms in terms of average accuracy for two- and three- class classification of perceived stress using (a) 10-fold cross validation (b) leave-one-out cross validation.

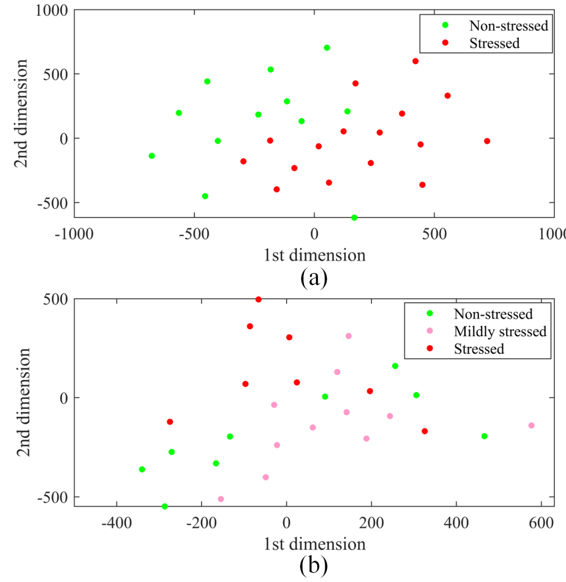


Fig. 5. Graphical visualization of selected features from the proposed feature selection algorithm for perceived stress classification (a) two-class (b) three-class.

density (8 feature values), correlation (2 feature values), rational asymmetry (2 feature values), divisional asymmetry (2 feature values), power spectrum (4 feature values). Table I presents the performance comparison of our proposed feature selection algorithm in terms of EEG band used, classification accuracy, F-measure, and kappa statistics for two- and three-class perceived stress classification problem using pre-activity EEG recording. We observed that MLP outperforms both SVM and NB classifiers and gives the highest accuracy, F-measure and kappa statistics for both two- and three-class classification of perceived stress. The highest accuracy of 89.28% and 60.71% was achieved by MLP using 10-fold cross validation. Similarly a highest accuracy of 92.85% and 64.28% was achieved by MLP using leave-one-out cross validation. The proposed feature selection algorithm was compared with genetic algorithm (GA) [33], minimum redundancy maximum relevance (mRMR) [34], particle swam optimization (PSO) [35], mutual information feature selection (MIFS) [36], uniform mutual information feature selection (MIFS-U) [37] and optimal feature selection using mutual information (OFS-MI) [38] based feature selection algorithms for two- and three-class perceived stress classification. The comparison in terms of average accuracy for two- and three-class perceived stress classification using 10-fold and leave-one-out cross validation is shown in Fig. 4. In case of 10-fold cross validation, the proposed feature selection algorithm achieves 21.43%, 10.71%, 32.14%, 35.71%, 32.14%, and 28.57% better average accuracy for two class and 25%, 21.43%, 21.43%, 14.29%, 21.43%, and 17.86% better accuracy for three class perceived stress classification than GA, mRMR, PSO, MIFS, MIFS-U, and OFS-MI, respectively. Similarly in case of leave-one-out cross validation, the proposed algorithm gives 35%, 10.71%, 32.14%, 35.71%, 39.28%, and 17.85% better accuracy for two class and 28.57%, 14.29%, 10.7%, 7.14%, 21.43%, and 17.86% better accuracy for three class perceived stress classification than GA, mRMR, PSO, MIFS, MIFS-U, and OFS-MI respectively. This improvement is due to the fact that the proposed feature selection algorithm was designed specifically for stress classification and selects features from all feature groups of θ band. Whereas, existing feature selection algorithms selects features from different bands and different feature groups. The selected features by existing algorithm from those bands, which are not significantly discriminating in terms of stress group can degrade the classification performance. The features selected using our proposed algorithm for perceived stress classification are visualized in two dimensions (Fig. 5) by using the t-Distributed Stochastic Neighbor Embedding (t-SNE) scheme [39] which is a dimensionality reduction technique to visualize high dimension data. We observed that two class (stressed and non-stressed) are more visually distinct as compared to three class (non-, mildly-stressed and stressed). This showed that the selected features effectively distinguished between two- and three-class perceived stress groups with accuracy of 92.85% and 64.28%.

IV. Discussion

In this paper, we have presented an experimental study to find an appropriate phase for recording the EEG data to classify perceived human stress into two- and three-classes. We showed that the pre-activity phase EEG recording could be better for stress classification, which was further improved by using our proposed feature selection algorithm. Table II present the performance comparison of our proposed scheme with recent studies conducted for the classification of human stress. The methods selected for comparison either classifies

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/JBHI.2019.2926407, IEEE Journal of Biomedical and Health Informatics

TABLE II
PERFORMANCE COMPARISON OF THE PROPOSED SCHEME WITH STATE-OF-THE-ART METHODS FOR HUMAN STRESS CLASSIFICATION USING EEG.

Method, Year	Number of Participants	Number of EEG Channels	Significant EEG band	Accuracy (Classes)	Classifier	Stress Type
[19], 2017	09	01	Alpha, Beta	83.33% (2)	SVM	Acute
[20], 2017	22	128	—	94.60% (2) 83.40% (multi-level)	Naive Bayes	Acute
[21], 2016	10	14	Beta	96.00% (2)	SVM	Acute
[22], 2016	22	07	Alpha, Beta	91.70% (2)	SVM	Acute
[23], 2015	09	14	—	85.71% (2)	SVM	Acute
[26], 2018	28	01	Low Beta, High Beta, Low Gamma	78.57% (2)	SVM	Perceived
[25], 2017	28	01	Beta	71.40% (2)	Naive Bayes	Perceived
[24], 2015	28	01	Beta	71.42% (2)	SVM	Perceived
Proposed, 10-fold	**28**	**04**	**Theta**	**89.30% (2) 60.71% (3)**	**MLP**	**Perceived**
Proposed, leave-one-out	**28**	**04**	**Theta**	**92.85% (2) 64.28% (3)**	**MLP**	**Perceived**

acute stress [19], [20], [40], [21], [22], [23], or perceived stress using EEG [25], [24], [26]. In [19], a single channel EEG study with 9 participants was presented to assess different levels of mental stress. Stress was induced in participants by mental workload and a public speaking task. Classification accuracy of 83.33% for two class classification was reported with alpha and beta as significant bands. A machine learning based multiple level human stress detection system using 128 EEG electrodes was proposed in [20]. Montreal imaging stress task was used as a stressor. An accuracy of 94.6% and 83.4% was reported for two level and multi-level stress classification respectively. In [23], acute stress level classification based on EEG was presented where stroop color-word test was used to induce stress in 9 subjects. Classification accuracy of 85.71%, 75.22% and 67.06% was reported for two, three and four level stress identification respectively. In [21], mental arithmetic task was used to induce stress in participants and EEG signals were acquired using 14 channel Emotiv EEG headset. Stress was classified into two- and three-levels with a classification accuracy of 96% and 75% respectively. A multi-modal human stress measurement was proposed in [22], where EEG and functional near infrared spectroscopy (fNIRS) were used to classify acute stress. A classification accuracy of 96.6% was achieved with EEG and fNIRS data.

The above mentioned methods were compared with our proposed scheme on the basis of the number of participants in the experiment, number of EEG electrodes, significant EEG band, and accuracy achieved. The number of participants involved in our proposed study (28) was highest as compared to other studies related to stress classification using EEG signals. Apart from [25], [24], [26], [19], [22], all other experiments have used a minimum of 14 electrodes. The studies using a single electrode achieves a maximum classification accuracy of 83.33% and 71.4% in case of acute and perceived stress classification respectively, whereas with an addition of 3 electrodes a classification accuracy of 89.28% (10-fold) and 92.85% (leave-one-out) was achieved in our proposed scheme for two class classification of perceived stress. For multi-level acute stress, a classification accuracy of 83.4% was achieved with 128 electrodes [20]. Whereas our proposed methodology achieved an accuracy of 60.71% (10-fold) and 64.28% (leave-one-out) for multi-class perceived stress (which is relatively difficult to classify) classification with 4 electrodes. In previous EEG based perceived stress classification studies, single (frontal) channel headset was used and beta band was reported as significant for perceived stress classification based on t-test. We applied a t-test on absolute power of different EEG bands of stressed and non-stressed groups and only theta band was found significantly different with a p-value of 0.005. This significance was also supported by the classification results achieved by our proposed features selection method from theta band which has also been reported in literature for stress monitoring task [41]. The proposed method was compared with HRV and accelerometer based perceived stress monitoring methods [42]. The heart and motion data of 8 participants were acquired for duration of two weeks. A classification accuracy of 85.7% was achieved with bagging classifier for the two-class problem. Our proposed EEG based method gave 7.1% better classification accuracy when compared with HRV based method [42]. Our proposed study used a commercially available easy to wear EEG headband and EEG data were acquired in open eye condition (which allows mobility while recording) which suggest that this can be used in an out of laboratory environment. Moreover, we identified that pre-activity EEG recording gives best classification results for perceived stress, which makes this study distinct from existing studies.

V. CONCLUSION

In this paper, an experimental study is presented with an aim to find a suitable phase for EEG recording to classify perceived mental stress into multiple classes. Five groups of features: power spectral density, correlation, differential asymmetry, rational asymmetry, and power spectrum were extracted from delta, theta, alpha, beta, and gamma bands of EEG data during pre- and post-activity phase in an open eye condition. Features were selected from EEG frequency bands, which contribute in enhancing the classification accuracy. Three classification algorithms were used with the selected features, and MLP classifier gave the best results for two- and three-class stress classification. The highest accuracy of 89.28% and 60.71% (10-fold cross validation) and 92.85% and 64.28% (leave-one-out cross validation) was achieved by using MLP classifier for two- and three-class classification respectively. In the future, we intend to classify acute mental stress using EEG signals in response to a public speaking task.

REFERENCES

[1] W. H. Organization *et al.*, "Depression fact sheet. 2012," *URL: http://www. who. int/mediacentre/factsheets/fs369/en/index. html [accessed 2014-08-22][WebCite Cache]*, 2013.

[2] S. Reisman, "Measurement of physiological stress," in *Proceedings of the IEEE 1997 23rd Northeast Bioengineering Conference,*. IEEE, 1997, pp. 21–23.

[3] C. Espinosa-Garcia, I. Sayeed, S. Yousuf, F. Atif, E. G. Sergeeva, G. N. Neigh, and D. G. Stein, "Stress primes microglial polarization after global ischemia: therapeutic potential of progesterone," *Brain, behavior, and immunity*, vol. 66, pp. 177–192, 2017.

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

[5] C. Gross and K. Seebaß, "The standard stress scale (sss): Measuring stress in the life course," in *Methodological Issues of Longitudinal Surveys*. Springer, 2016, pp. 233–249.

[6] T. Föhr, A. Tolvanen, T. Myllymäki, E. Järvelä-Reijonen, S. Rantala, R. Korpela, K. Peuhkuri, M. Kolehmainen, S. Puttonen, R. Lappalainen *et al.*, "Subjective stress, objective heart rate variability-based stress, and recovery on workdays among overweight and psychologically distressed individuals: a cross-sectional study," *Journal of Occupational Medicine and Toxicology*, vol. 10, no. 1, p. 39, 2015.

[7] S. Cohen, T. Kamarck, R. Mermelstein *et al.*, "Perceived stress scale," *Measuring stress: A guide for health and social scientists*, 1994.

[8] K. B. Koh, J. K. Park, and C. H. Kim, "Development of the stress response inventory," *Journal of Korean Neuropsychiatric Association*, vol. 39, no. 4, pp. 707–719, 2000.

[9] C. D Spielberger, R. Gorsuch, R. E Lushene, P. Vagg, and G. A Jacobs, *Manual for the State-Trait Anxiety Inventory (Form Y1 – Y2)*, 01 1983, vol. IV.

[10] A. Deschênes, H. Forget, C. Daudelin-Peltier, D. Fiset, and C. Blais, "Facial expression recognition impairment following acute social stress." *Journal of vision*, vol. 15, no. 12, pp. 1383–1383, 2015.

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

[12] J. T.-y. Wang, "Pupil dilation and eye tracking," *A handbook of process tracing methods for decision research: A critical review and user's guide*, pp. 185–204, 2011.

[13] G. C. Dieleman, A. C. Huizink, J. H. Tulen, E. M. Utens, H. E. Creemers, J. van der Ende, and F. C. Verhulst, "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, 2015.

[14] A. Hernando, J. Lázaro, E. Gil, A. Arza, J. M. Garzón, R. López-Antón, C. de la Cámara, P. Laguna, J. Aguiló, and R. Bailón, "Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment," *IEEE journal of biomedical and health informatics*, vol. 20, no. 4, pp. 1016–1025, 2016.

[15] Q. Xu, T. L. Nwe, and C. Guan, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE journal of biomedical and health informatics*, vol. 19, no. 1, pp. 275–281, 2015.

[16] H. Jebelli, M. M. Khalili, and S. Lee, "A continuously updated, computationally efficient stress recognition framework using electroencephalogram (eeg) by applying online multi-task learning algorithms (omtl)," *IEEE journal of biomedical and health informatics*, 2018.

[17] J. M. Soares, A. Sampaio, L. M. Ferreira, N. C. Santos, P. Marques, F. Marques, J. A. Palha, J. J. Cerqueira, and N. Sousa, "Stress impact on resting state brain networks," *PLoS One*, vol. 8, no. 6, p. e66500, 2013.

[18] A. Asif, M. Majid, and S. M. Anwar, "Human stress classification using eeg signals in response to music tracks," *Computers in biology and medicine*, vol. 107, pp. 182–196, 2019.

[19] A. Secerbegovic, S. Ibric, J. Nisic, N. Suljanovic, and A. Mujcic, "Mental workload vs. stress differentiation using single-channel eeg," in *CMBEBIH 2017*. Springer, 2017, pp. 511–515.

[20] A. R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel, and A. S. Malik, "Machine learning framework for the detection of mental stress at multiple levels," *IEEE Access*, vol. 5, pp. 13 545–13 556, 2017.

[21] G. Jun and K. Smitha, "Eeg based stress level identification," in *Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 003 270–003 274.

[22] 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," *Biomedical optics express*, vol. 7, no. 10, pp. 3882–3898, 2016.

[23] X. Hou, Y. Liu, O. Sourina, Y. R. E. Tan, L. Wang, and W. Mueller-Wittig, "Eeg based stress monitoring," in *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*. IEEE, 2015, pp. 3110–3115.

[24] S. M. U. Saeed, S. M. Anwar, M. Majid, and A. M. Bhatti, "Psychological stress measurement using low cost single channel eeg headset," in *Signal Processing and Information Technology (ISSPIT), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 581–585.

[25] S. M. U. SAEED, S. M. ANWAR, and M. Majid, "Quantification of human stress using commercially available single channel eeg headset," *IEICE Transactions on Information and Systems*, vol. 100, no. 9, pp. 2241–2244, 2017.

[26] S. M. U. Saeed, S. M. Anwar, M. Majid, M. Awais, and M. Alnowami, "Selection of neural oscillatory features for human stress classification with single channel eeg headset," *BioMed Research International*, 2018.

[27] N. H. A. Hamid, N. Sulaiman, Z. H. Murat, and M. N. Taib, "Brainwaves stress pattern based on perceived stress scale test," in *Control and System Graduate Research Colloquium (ICSGRC), 2015 IEEE 6th*. IEEE, 2015, pp. 135–140.

[28] N. Sulaiman, M. N. Taib, S. Lias, Z. H. Murat, S. A. M. Aris, M. Mustafa, N. A. Rashid, and N. H. A. Hamid, "Intelligent system for assessing human stress using eeg signals and psychoanalysis tests," in *Computational Intelligence, Communication Systems and Networks (CICSyN), 2011 Third International Conference on*. IEEE, 2011, pp. 363–367.

[29] N. H. A. Hamid, N. Sulaiman, S. A. M. Aris, Z. H. Murat, and M. N. Taib, "Evaluation of human stress using eeg power spectrum," in *Signal Processing and Its Applications (CSPA), 2010 6th International Colloquium on*. IEEE, 2010, pp. 1–4.

[30] R. Luijcks, C. J. Vossen, H. J. Hermens, J. van Os, and R. Lousberg, "The influence of perceived stress on cortical reactivity: a proof-of-principle study," *PloS one*, vol. 10, no. 6, p. e0129220, 2015.

[31] I. Daly, F. Pichiorri, J. Faller, V. Kaiser, A. Kreilinger, R. Scherer, and G. Müller-Putz, "What does clean eeg look like?" in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2012, pp. 3963–3966.

[32] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[33] J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," in *Feature extraction, construction and selection*. Springer, 1998, pp. 117–136.

[34] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 8, pp. 1226–1238, 2005.

[35] J. Kennedy, "Particle swarm optimization," *Encyclopedia of machine learning*, pp. 760–766, 2010.

[36] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Transactions on neural networks*, vol. 5, no. 4, pp. 537–550, 1994.

[37] N. Kwak and C.-H. Choi, "Input feature selection for classification problems," *IEEE transactions on neural networks*, vol. 13, no. 1, pp. 143–159, 2002.

[38] T. W. Chow and D. Huang, "Estimating optimal feature subsets using efficient estimation of high-dimensional mutual information," *IEEE Transactions on Neural networks*, vol. 16, no. 1, pp. 213–224, 2005.

[39] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.

[40] K. Smitha, N. Y. Xin, S. S. Lian, and N. Robinson, "Classifying subjective emotional stress response evoked by multitasking using eeg," in *Systems, Man, and Cybernetics (SMC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 3036–3041.

[41] S. A. Awang, P. M. Pandiyan, S. Yaacob, Y. M. Ali, F. Ramidi, and F. Mat, "Spectral density analysis: theta wave as mental stress indicator," in *International Conference on Signal Processing, Image Processing, and Pattern Recognition*. Springer, 2011, pp. 103–112.

[42] M. Wu, H. Cao, H.-L. Nguyen, K. Surmacz, and C. Hargrove, "Modeling perceived stress via hrv and accelerometer sensor streams," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 1625–1628.