A wearable EEG instrument for real-time frontal asymmetry monitoring in worker stress analysis

(Article begins on next page)

A wearable EEG instrument for real-time frontal asymmetry monitoring in worker stress analysis

Pasquale Arpaia[1], *Senior Member, IEEE,* Nicola Moccaldi[1], Roberto Prevete[1], Isabella Sannino[2], Annarita Tedesco[3]

***Abstract*—A highly-wearable single–channel instrument, conceived with off-the-shelf components and dry electrodes, is proposed for detecting human stress in real time by electroencephalography (EEG). The instrument exploits EEG robustness to movement artifacts with respect to other biosignals for stress assessment. The single-channel differential measurement aims at analyzing the frontal asymmetry, a well-claimed EEG feature for stress assessment. The instrument was characterized metrologically on human subjects. As triple metrological references, standardized stress tests, observational questionnaires given by psychologists, and performance measurements were exploited. Four standard machine learning classifiers (SVM, k-NN, Random Forest, and ANN), trained on 50% of the data set, reached more than 90% accuracy in classifying each 2-s epoch of EEG acquired from stressed subjects.**

***Index Terms*—EEG, Stress, Brain-Computer Interface, Cobot, Industry 4.0, Smart Manufacturing.**

I. Introduction

Stress is a psycho-physical pathological response to emotional, cognitive, or social tasks, perceived as excessive by an individual. Many stimuli of different nature (physical, toxic, emotional), external to individuals, could disturb their homeostasis and psychological well-being, bringing to an adaptive or non-adaptive response [1]. In industrial work, stress has negative impact on safety, on the quality of the outcome and, thus, on the cost of the production process as a whole [2]. Technological innovation, indeed, has introduced new sources of stress (*stress 4.0*). Intelligent automated systems in their various configurations, robots or cobots in collaborative meaning [3], interact continuously with individuals in a constant relationship of cooperation and, at the same time, of unconscious competition.

In literature, different indicators of stress status are proposed, arising from products of neuroendocrine reactions affecting sympathetic and parasympathetic nervous systems [4]. Some biochemical and biophysical markers are measured usually by invasive methods: (i) Cortisol Concentration in blood or saliva; (ii) Galvanic Skin Response; (iii) Heart Rate; and (iv) Brain Activity.

*This work was supported by the the project Advanced Virtual Adaptive Technologies e-hEAlth (AVATEA, CUP. B13D18000130007) POR FESR CAMPANIA 2014/2020 and by the Project "Excellence Department" of Italian Ministry of Instruction and University Research.

[1]Laboratory of Augmented Reality for Health Monitoring (ARHeMLab), Department of Electrical Engineering and Information Technology, University of Naples Federico II

[2]Department of Electronics and Telecommunications (DET), Politecnico di Torino, Turin, Italy

[3]Laboratoire de l 'Intégration du Matériau au Système, University of Bordeaux, Talence cedex, France.

Cortisol is a hormone produced by the adrenal glands with the aim of preserving homeostasis in all conditions tending to alter the normal body balance. Cortisol concentration in blood has been used as the first index of the individual's response to stress. It is measured through repeated blood samples, or through saliva samples, by means of less invasive methods but with less significance [2].

Skin conductance is a further parameter associated to the activation of the sympathetic nervous system and, therefore, to stress. Stress induces an increase in the epidermis moisture and, therefore, a reduction in skin resistance.

Furthermore, stress generates peripheral vasoconstriction that causes a decrease in wave amplitudes of electrocardiogram (ECG) and an increase in the heart rate [4].

Brain activity produces electrical signals as a response to all kind of internal and external stimuli. The signals are recorded either through functional Magnetic Resonance Imaging, Positron Emission Tomography, or electroencephalography (EEG). All these techniques detect brain activity changes in the limbic system and frontal regions.

EEG is the most widely used because it is easy to implement and little intrusive; moreover, EEG signals can be classified effectively through a frequency analysis. Some use cases of stress recognition based on EEG are given in the literature[5], [6]. Significant is a wearable EEG device for construction workers [7]. High vulnerability characterizes the activities on-site of the workers during a construction process; so, they suffer from load stress. By including an EEG device into their protective helmet, brain waves are monitored and analyzed along the activity, by highlighting possible emotional states and, therefore, actual attention levels [8]. However, state-of-the-art solutions exhibit at least one of the following weaknesses: (i) limitations for daily on-field use, e.g. due to a large number of wet electrodes and use of wired systems; (ii) accuracy less than 90%, even in case of simultaneous ECG and EEG measurements [9]; and (iii) high cost, up to thousands of dollars [10].

In this paper, a highly-wearable single-channel instrument, conceived with off-the-shelf components and dry electrodes, for human stress real-time detection based on EEG, is presented. The single-channel differential measurement aims at analyzing the Frontal Asymmetry, a well-claimed EEG feature for stress assessment. The differential channel, in fact, acquires two signal from symmetric regions of the scalp and calculates the difference [9]. The instrument exploit EEG robustness to movement artifacts compared to other stress assessment biosignals. In particular, in Section II, concepts related to biosignals and stress, EEG for stress assessment, and frontal EEG asymmetry are recalled. Instrument architecture and operation are illustrated in Section III. Hardware, firmware, and software are described in Section IV. In Section V, the experimental psychological validation of the stressors effectiveness, a Principal Component Analysis, and a noise robustness assessment are reported.

II. Background

A. EEG for Stress Assessment

Several methods for stress assessment, like self-assessment scale, or questionnaires, follow a psychological approach [11]. As an example, in human-robot interaction, questionnaires to analyze the psychological effect of cycle time on operators [12] highlighted frustration, effort, and a dissatisfaction feeling about own performance.

As more direct and objective tools for stress detection, biosignals have been proposed in several studies [13]. Physiological parameters, as EEG signals, blood volume pulse (BVP), electro-oculogram (EOG), salivary cortisol level (SCL) [14], heart rate variability (HRV) [15], galvanic skin response (GSR), or electromyography (EMG) are assessed [4].

Compared to other biosignals, EEG proved better latency and robustness to artifacts due to physical activity [6] [16]. In Industry I4.0 scenarios, EEG has been widely applied to assess individuals' stress in workplace in order to improve workers' safety, health, well-being, and productivity [7][17] [18]. Thanks to the ease of application and removal, dry electrodes are increasingly used to reliably search human cognitive states in real-life conditions. They guarantee the quality of the EEG signal which approaches the wet sensors, as demonstrated in [19], [20]. Beside that, EEG is regarded as one of the most reliable and effective techniques for identifying fatigue and monitoring stress level in drivers [21], [5], [22]. Different classification methods try to face the main problems of EEG signals, including the low signal-to-noise ratio, their non stationarity over time within or between users, and the limited amount of training data typically available to calibrate the classifiers[23].

A large number of informative and measurable properties (features) of EEG signals, can be used both in time and frequency domains. Their accurate selection is crucial for the accuracy and the computational cost of classification [24]. Both types of features reap the benefit from being extracted after spatial filtering. Independent Component Analysis and Canonical Correlation Analysis are useful methods for muscle artifact removal in EEG data [25], [26]. Several supervised learning algorithms could be exploited to assess workers stress by using subjects' EEG signals. The classification can be assisted by: linear classifiers, neural networks, non-linear Bayesian classifiers, nearest neighbour classifiers, random forest, naive bayes, and decision tree [27][28]. Linear classifiers are the most popular algorithms for Brain Computer Interface (BCI) applications, such as, Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM)The LDA is used to assess mental fatigue in [29], it divides the data into hyperplanes representing the different classes, with very-low computational burden. A discrimination based on hyperplanes was also used in SVM, with recognition rate of 75.2% to identify three different level of stress out of four, using EEG features and six statistical features in [30]. Meanwhile a better prediction accuracy of 90.5% and 92%, combining different acquired signals, were reported in [21], [22] for drivers. Various supervised machine learning algorithms, using sliding and fixed windowing procedures, were tested in [7]: k-Nearest Neighbors, Gaussian Discriminant Analysis, SVM with different similarity functions (linear, Gaussian, cubic, and quadratic). Among the state-of-the-art classifiers, the SVM yielded the highest classification accuracy of 90.1% [31], using a single-channel EEG. As well as, the highest accuracy of 88.0% was reached by SVM in [32], where individuals' stress was recognized by exploiting only EEG signal as input of the classifier. Tab I summarizes the reported accuracy of different classifiers, including the numbers and type of EEG electrodes, without reference electrodes, the numbers of different classes according to the acquired bio-signals used as model input.

B. Frontal EEG asymmetry

Systematic alterations in frontal EEG asymmetry, in response to specific emotional stimuli, can be exploited to analyze emotional response [35]. In particular, EEG asymmetry proved to be capable of predicting state-related emotional changes and responses. For example, a greater self-reported happiness or positively-valued stimuli might be expected to be associated with greater relative left frontal activity. Therefore, greater relative right frontal activity would be expected in response to negative stimuli [36],[37]. However, fear or happiness response to stimuli may either be attenuated or amplified according to any given individual's trait pattern of frontal EEG asymmetry [36].

Different models were presented: Baron and Kenney linear model may predict individual's response to fear relevant stimuli. According to the relative difference between the left and right hemisphere, the EEG asymmetry may serve both to amplify and attenuate the effect of the fear relevant stimuli. Some individuals show the increase of relative right versus left sided activity in response to negative cues and the increase of left versus right sided activity to positive cues [38]. Coan and Allen [39] presented another linear model to predict emotional experience using emotion type and trait frontal EEG asymmetry. The frontal EEG asymmetry may serve as a useful liability marker also for depression and anxiety [40]. Many works, using EEG caps with a limited number of electrodes, demonstrated that stress causes changes in regions of prefrontal and frontal areas [14] [6] [36].

III. Design

In this Section, (A) the *Basic Ideas*, (B) the *Architecture*, (C) the *Operation*, and (D) the *Feature Extraction and Classification* of the instrument are presented.

A. Basic Ideas

The concept design of the real-time stress monitoring instrument was based on the following main basic ideas.

Table I
STATE OF ART OF STRESS CLASSIFICATION

Classifier	Reference	Reported Accuracy%	Acquired Signals	Classes	n° Electrodes	n° Subjects
Artificial Neural Network (ANN)	[21]	76.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
	[28]	79.2%	EEG,SCL,BVP,PPG	2 levels of stress	5 Wet	15
Cellular Neural Network (CNN)	[21]	92.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
Decision Tree	[21]	84.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
Fisher linear discriminant analysis (FLDA)	[22]	90.5%	EEG,EOG	2 alert and fatigue states	32 Wet	8
Gaussian Discriminant Analysis (GDA)	[7]	74.9%	EEG,GSR	2 high or low stress level	14 Wet	11
K-Nearest Neighbors (k-NN)	[7]	65.8%	EEG,GSR	2 high or low stress level	14 Wet	11
	[30]	76.7%	EEG	2 levels of stress	14 Wet	9
Linear discriminant analysis (LDA)	[29]	77.5%	EEG	3 low, medium, high mental fatigue	16 Wet	10
	[6]	86.0%	EEG,ECG,EMG,GSR	3 stress,relax,and neutral	4 Wet	10
Naive Bayes (NB)	[21]	77.0%	EEG,ECG,GSR	2 no-stress/stress	14 Wet	22
	[32]	69.7%	EEG,ECG,GSR	2 mental workload and stress	2 Wet	9
Random Forest (RF)	[33]	79.6%	EEG,EMG,ECG,GSR	4 cognitive states	8 Wet	12
	[27]	84.3%	EEG,ECG,BVP	3 mental stress states	14 Wet	17
Support vector machine (SVM)	[30]	75.2%	EEG	3 levels of stress	14 Wet	9
	[7]	80.3%	EEG, SCL	2 high or low stress level	14 Wet	11
	[34]	85.4%	EEG	2 positive or negative emotion	14 Wet	11
	[9]	87.5%	EEG,ECG,HRV	2 stress and rest	2 Wet	7
	[31]	88.0%	EEG	2 levels of stress	14 Wet	10
	[32]	90.1%	EEG,ECG,GSR	2 mental workload and stress	2 Wet	9

- *High wearability*: a single differential channel allows the use of only two frontal electrodes in the area FP1 e FP2 according to the EEG International 10–20 system for placement of EEG electrodes on the scalp. The reference electrode is applied to the earlobe. These positions have been used in reports of successful studies on stress [32]. Active dry electrodes avoid the inconvenient of electrolytic gel. A wireless module allows the user to carry out work activity during the EEG acquisition.
- *High accuracy and low latency*: Despite the use of a single differential acquisition channel, a time-domain based machine learning algorithm brings to an accuracy of 98.3 ± 0.4 in stress detection. A time window of 512 samples guarantees a latency of 2 s.
- *Off-the-shelf components*: The measurement of frontal asymmetry by EEG at very-low density (single channel) allows high wearability, maximum accuracy, and low latency by exploiting the lowest cost hardware on the market (< 200 $) [10].

B. Architecture

The architecture of the proposed instrument is highlighted in Fig. 1, in an example of interaction with a cobot. Prefrontal asymmetry is measured by two *Electrodes* as the difference of brainwaves from position FP1 and FP2, according to 10/20 system. The differential signal is referred to the earlobe. Analog signal is digitized by the *Acquisition Unit* and is sent, via wires, to the *Wi-Fi Transmission Unit*. Digital data arrives at the *Processing Unit* through wireless communication for real-time elaboration. Suitable *features* are extracted from each EEG record to compress data and increase significance. A *Classifier* receives the feature arrays, detects the stress condition, and assess its level. The measured stress is sent to the Cobot.

C. Operation

The instrument allows to detect the onset and to assess the level of the stress arising from the concurrence of high

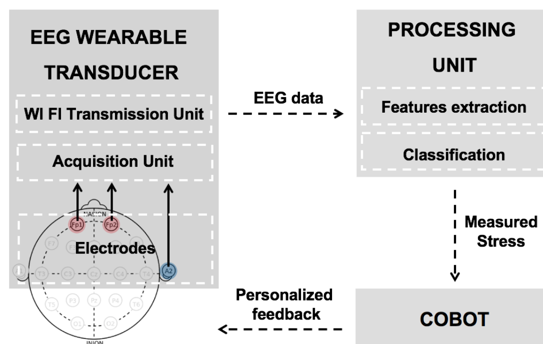

Figure 1. Architecture of the real-time stress monitoring instrument in Cobot interaction.

mental load and negative emotional conditions, during the interaction with a Cobot. Once the worker fixes the electrodes on the forehead and on the earlobe, the Processing Unit interface allows to check signal quality both in time and in frequency domain. Subsequently, the stress measurement starts and the acquired data are sent in real time to the Processing Unit, by updating the user condition assessment every 2 s. Measurement results are sent to the Cobot in order to adapt its behavior to the worker stress conditions.

D. Feature extraction and classification

Preliminary experiments in frequency domain highlighted poor accuracy results. Therefore, data analysis was carried out in the time domain. According to the state of the art [38], a EEG time window of 2 s was chosen as the optimal solution considering the trade-off accuracy vs latency. In time domain, EEG tracks are divided into 2–s records of 512 samples. In this way, raw data are composed of 512 features, i.e. each feature corresponds to just one sample. Feature Extraction was carried out by a standard machine learning technique, the Principal Component Analysis (PCA). This allows to compress data [41] and to approximate signals as a linear combination of a restricted number of orthogonal components. Therefore, data variance is most efficiently explained. Accordingly, a multi-variable signal can be represented as a smaller number of coefficients of the linear combination of the components. PCA also performs a filter function, because it highlights the components with maximum variance (information) of the data. Therefore, selecting only the components with the greatest variability improves signal-to-noise ratio.

For the classifier design, a linear separability test of the data was carried out by an euclidean distance-based K-means algorithm with low computational burden [42]. If a problem is linearly separable, a nonlinear classifier complicates the model unnecessarily and makes the correct learning of the classifier parameters less effective [43]. K-means algorithm estimates k means (centroids) in order to partition data into k clusters where each observation belongs to the cluster with the nearest mean. Then, in case of few outliers, a linear classifier is justified. Therefore, a preliminary analysis was realized.

1) Preliminary Analysis: Ten subjects were divided into two classes with different stress level: (i) *control group*, only cognitive load, and (ii) *experimental group*, cognitive load but with negative emotions. Data were recorded during all the tests with a differential single-channel digitizer, sampling at 256 sample/s. The signal was elaborated in time domain and without artifact filtering according to [34]. For each volunteer, two EEG tracks of 20 s were processed and divided into 2–s records of 512 samples. The resulting matrix 200x512 was divided in two clusters using the standard K-means algorithm with $K = 2$. In Fig. 2, the result of the clustering algorithm is reported. The two experimental groups were separated by

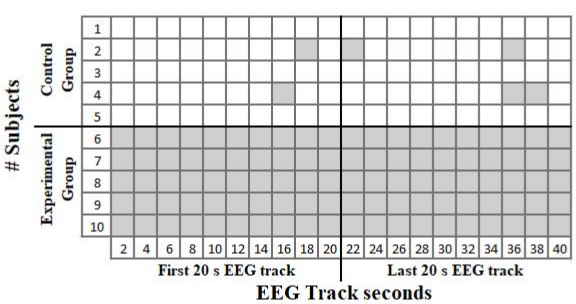

Figure 2. K-means classification (white: class 1; black: class 2) among the 2 different time phases according to subjects belonging group.

K-means almost cleanly: on the first five rows, the arrays of the experimental group records, and on the other rows, the ones of the control group. These results suggested that also a linear classifier can be used to discriminate the points of the two groups adequately.

IV. Realization

A. Hardware

1) Data Acquisition Unit: It is based on the differential single-channel 10-bit digitizer EEG-SMT by Olimex, with maximum sampling rate of 256 sample/s, an EEG amplifier, and an Atmel ATmega16 Alf and Vegard Reduced Instruction Set Computer processor microcontroller. The gain of the analog-to-digital converter (ADC) of the transducer was set to be 6427 V/V. A right-leg driver [driven rightled (circuit) (DRL)] signal increases the common-mode noise rejection. Universal serial bus is used for both data communication and powering. Moreover, the EEG-SMT has an analog three stages pass-band filter from 0.16 to 59.00 Hz. A previous work proved its suitability for wearable, low-cost, and non-invasive brain activity monitoring, by means of a single differential channel [10].

2) Dry Electrodes: Brain signals are acquired by two dry active electrodes (Olimex EEG-AE), coated with a thin layer of silver chloride to guarantee the best contact impedance. The contact surface is extended by pins of conductive material. In this way, the quality of the acquired signal is preserved even with the electrode on a thick layer of hair. The reference passive dry electrode (Olimex EEG-PE) was applied to the earlobe. The electrodes on the user's forehead are fixed with a tight headband. The electrode on the earlobe is fixed with a clip, to ensure electrical connection.

3) Transmission unit: A Wi-Fi communication channel was implemented to enhance wearability, throughout a Raspberry Pi 3 single-board computer, used as server, connected via UART to the EEG-SMT. The Raspberry Pi 3 uses a BCM43438 wireless chip and operates at ISM frequency bands (2.4 GHz).

B. Signal processing and classification

In time domain, EEG tracks are divided into 2–s records of 512 samples. In this way, raw data are composed of 512 features, i.e. each feature corresponds to just one sample. Then, a feature reduction process is realized by PCA. The first four Principal Components are considered as input in the successive classification step. As emerged in a previous exploratory experimental campaign, the first 4 components guarantee at the same time high accuracy as well as low uncertainty and computational burden. The mean of accuracy obtained by the first four Principal Components resulted significantly greater than that of the first three Principal Components, with a confidence level of 93.0% (one-tailed t test). machine learning classifier distinguishes records of a stressed or no stressed subject. The results of the preliminary experiments (Fig. 2) show that the two groups are separated from the K-means quite clearly. These results suggested that also a linear classifier can be used to discriminate data of the two groups adequately. The length of records determines the latency of 2 s.

C. Software

1) Raspberry: The EEG signals, digitized by the EEG-SMT Olimex, are acquired by the Raspberry via UART by means of a dedicate software in C and installed on the Raspberry Pi 3. The baud rate is set to 57600 bit/s, with packet size 8, without parity bit. The Raspberry Pi 3 acts also as a Wi-Fi server, receiving from the EEG-SMT the command of start of the acquisition, and sending to the computer station the acquired data. This allows the users to freely move during real life. In view of a stand-alone device release, the computational power guaranteed by the raspberry allows processing to be carried out directly on board.

2) Processing Unit: A specifically designed Matlab graphical user interface (GUI) allows easy interaction with Olimex EEG-SMT, through graphical icons and visual indicators. Moreover, by observing the display windows, EEG signal can be monitored both in time and frequency domain. Meanwhile, Matlab scripts implement the machine learning classifiers.

V. Metrological Characterization

A. Experimental Setup

Seventeen volunteers underwent an initial screening test administered by the psychologist. Seven participants were excluded from the experiment owing to excess in smoke, high score in anxiety and depression at questionnaires, and low performance at short memory tests. Therefore, ten healthy young volunteers (average age 25 years) of whom five women and five men, participated in the study. The informed consent, containing all the information about the experiment, was provided and signed by the subjects. The protocol was explained by the psychologist. Participants were divided equally into control and experimental groups, to complete a task, which induces mental load, together with (experimental group) or without (control group) negative social feedback. In particular, the Stroop Color and Word Test (SCWT) [44], a neuropsychological test extensively used for both experimental and clinical purposes, aimed to challenge subject using a complex cognitive task. In this test, subjects are required to read as fast as possible color-words printed in an inconsistent color ink, and to name the color of the ink instead of reading the word. This is to be done in a limited time punctuated by the psychologist who also gave information about errors during the performance. Environment was specifically designed in order to stress participants, by means of an attractive prize and an extremely out of range performance. Before and after the Stroop Test, subjects were required to complete two questionnaires: (i) STAI State form [45], to evaluate current anxiety state, and (ii) Rosemberg inventory [46], to assess participants' self-esteem. In them, they had to reflect their emotions in the specific moments during their exercise. Moreover, at the end of experimental tests, participants filled a rating of the experience in the Likert scale. The two groups, experimental and control, were subjected to the same protocols, but only the experimental group was stressed emotionally. During the experiment, the device did not annoy or distract the subject. After each trial, the psychologist asked for feedbacks in order to ensure the safety of participants. They did not experience any discomfort related to the electrode band; after a few minutes, they no longer noticed the device. The most significant 40 s were extracted from each individual test of 180 s. The initial and concluding stages are potentially the most inhomogeneous among them, that is, the most challenging in order to find a regularity, intra individual and even more intra group. The first 10 s of the test, regarded as cognitive warm up, were excluded. Therefore, only the later 20 s were deemed. The final 10 s were discarded, due to observations of the psychologist. The specialist noticed that some subjects showed a renouncing attitude, once realized the impossibility to complete the task. Hence, the previous 20 s were considered. Subsequently, for each subject, the two 20-s EEG tracks were divided in 2-s records. Each record is characterized by 512 time domain features, i.e. the 512 samples contained in 2 s. The total number of records were 200, namely 20 records for 10 subjects of 2 s each. Five subjects were taken from control group and five from experimental (stressed) group. In this way, the total EEG-samples from each group were 51,200. A matrix with 200 records on the rows was obtained by placing the first 100 records referred to the initial and final 20 s stress of control group and subsequently 100 records related to initial and final 20 s of experimental group.

B. Psychological validation

A unique stress index was estimated as sum of normalized indexes to assess the general stress induced to participants. The indexes of performance, anxiety, self-esteem, perceived stress, and motivation, were obtained from parametric STAI and Rosemberg tests, as well as from task performance. One-way ANOVA was used to evaluate stress and motivation indexes on groups with a significance level α=0.05. The experimental group was more stressed compared to the control group, as evidenced by the One-way ANOVA (F=7.49; p=0.026). Instead, any significant difference between gender was noticed. A relevant difference in motivation between groups (F=14.52; p=0.005) showed that control group was more motivated than experimental group at the end of the experiment. Tab. II shows that, once arranged the stress index in decreasing order, the experimental group is more stressed than control one.

Table II
Stress index distribution (descending sort)

Subject	Stress Index	Group
1	1,68	Experimental
2	1,66	Experimental
3	1,52	Experimental
4	1,38	Control
5	1,21	Experimental
6	0,77	Experimental
7	0,69	Control
8	0,54	Control
9	0,14	Control
10	-0,06	Control

C. Stress Classification

Four different machine learning classifiers were used for validating the proposed method, by distinguishing stressed subject signals from no-stressed subject signals: (i) SVM (linear Kernel), (ii) k-nearest neighbors (n_neighbors = 9), (iii) Random Forest (criterion = 'gini', max_depth = 118, min_samples_split = 49) , and (iv) ANN (one hidden layer, activation function for hidden node = hyperbolic tangent, loss function = cross entropy cost, post processing = soft max, training algorithm = Resilient Propagation). In tab III the optimized iperparameters for each classifier are reported.

The behavior of each classifier was also evaluated when the input was pre-processed by PCA.

Importantly, a subject-wise leave-two-out cross-validation evaluation was uniformly conducted in all the experiments

Table III
CLASSIFIER OPTIMIZED IPERPARAMETERS AND RANGE OF VARIATION

Classifier	Iperparameter	Variation range
SVM	Cost parameter (C)	[0.1, 10.1] step = 1.0
Random Forest	n_estimators	{90, 180, 270, 360, 450}
k-NN	n_neighbors	[5, 15] step = 2
ANN	number of internal node	{25, 50, 100, 200}

in order to build a model capable of generalizing to new subjects. In case of small dataset according to [47], the Leave-p-out cross-validation (LPOCV) guarantees better statistical significance with respect to Leave-one-out cross-validation (LOOCV). Applying LOOCV to our dataset, the cross-validation process is repeated for k = 10 times, i.e. k = n (the number of subjects in the original sample). Instead, LPOCV requires training and validating the model C_p^n times, where C_p^n is the binomial coefficient, n the number of subjects in the original sample, and p is the number of subjects reserved only for the test. In our case (leave-two-out) the two subjects always belong to different groups (experimental vs control). In this way, a higher statistical significance was obtained (k = 25), by keeping training and test datasets balanced concerning the two classes. Therefore, for each iteration, one subject for group was left out from training set and used in the test set.

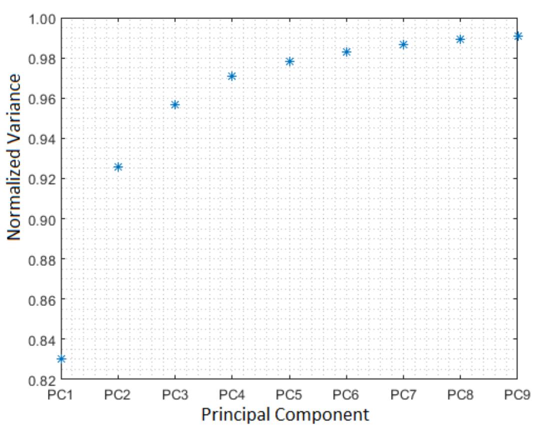

Figure 3. Cumulative Explained Variance in the PCA.

1) PCA Analysis: Each classifier was fed with both raw data (2-s EEG epoch) and PCA pre-processed data. In particular, for each iteration of the LPOCV method [47] the first p principal components were computed on the training set. Then both training and test set were projected on them. Finally, the reduced representations of both data sets were input to the classifiers. The number of principal components p was varied: $p \in \{0, 1, 2, \dots, 9\}$, where $p = 0$ corresponds to consider original data without PCA. The cumulative explained variance by the first nine components is greater than 99%, when PCA is applied on the dataset as a whole (Fig. 3). Therefore, this result highlights an intrinsic dimensionality of the data actually equal to no more than 9 (with respect to 512) and, in this case, the use of PCA for features extraction is validated. The results of the cross-validation strategy are shown in Tab. IV, as mean and uncertainty, with and without PCA. The lowest average

Table IV
CLASSIFIERS ACCURACY (MEAN AND UNCERTAINTY PERCENTAGE) IN ORIGINAL DATA (O.D.) AND PRINCIPAL COMPONENTS HYPERPLANES

	SVM	Random Forest	k-NN	ANN
O.D.	97.5 ± 0.6	98.5 ± 0.3	98.5 ± 0.4	99.2± 3.1
PC1	90.5 ± 5.3	98,6 ± 0.3	98,9 ± 0.3	98.5± 3.8
PC2	78.5 ± 7.1	98,8 ± 0.2	98,0 ± 0.5	98.8± 3.9
PC3	93.2 ± 3.4	98,4 ± 0.5	98,5 ± 0.4	98.7± 4.3
PC4	98.3 ± 0.4	98,9 ± 0.3	98,5 ± 0.4	99.1± 2.4
PC5	97.8 ± 0.4	98,8 ± 0.5	98,5 ± 0.4	99.2± 2.8
PC6	97.4 ± 0.6	98,4 ± 0.5	98,5 ± 0.4	98.9± 3.3
PC7	97.8 ± 0.5	99.0 ± 0.4	98,5 ± 0.4	98.9± 3.6
PC8	97.4 ± 0.6	98,6 ± 0.5	98,5 ± 0.4	99.0± 3.5
PC9	97.9 ± 0.5	98,9 ± 0.5	98,5 ± 0.4	98.9± 4.1

Table V
F-MEASURE TEST RESULTS FOR SVM (MEAN AND UNCERTAINTY PERCENTAGE)

		Precision (%)	Recall (%)
O.D. Hyperplane		96,5 ± 1,0	98,4 ± 0,7
	PC1	89,2 ± 5,1	92,2 ± 5,3
	PC2	81,1 ± 6,2	81,1 ± 6,7
	PC3	96,4 ± 1,5	93,6 ± 3,1
	PC4	98,2 ± 0,5	98,5 ± 0,7
P.C. Hyperplanes	PC5	97,2 ± 0,5	98,5 ± 0,7
	PC6	96,4 ± 1,1	98,5 ± 0,7
	PC7	97,2 ± 0,8	98,5 ± 0,7
	PC8	96,4 ± 1,1	98,5 ± 0,7
	PC9	97,2 ± 0,8	98,7 ± 0,6

accuracy for data without PCA is obtained by SVM and is equal to 97.5%. An F-measure test was carried out to assess the classification performance of the worst classifier (SVM). Results are reported in Tab. V.

SVM classification output when p=2 is shown in Fig. 4 in PCs space. The PCs plot shows vectors distribution with respect to Support Vector. In that, the diamonds are associated to the control group, while the circles represent the experimental group.

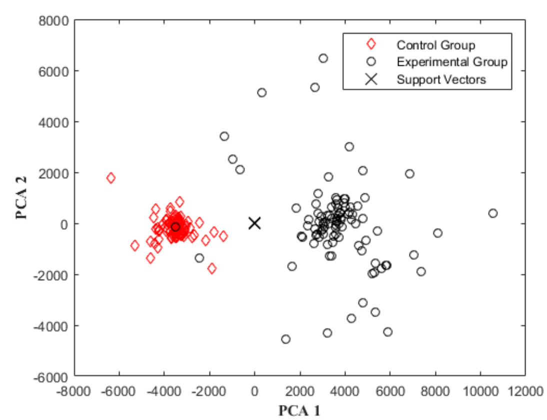

Figure 4. SVM data distribution in PCs space, p=2, 92,6% of explained variance.

Even with a temporal resolution of 2 s, satisfying results can be obtained in discriminating stress conditions. Generally PCA allows to obtain comparable or better average accuracy

Table VI

ACCURACY (MEAN AND UNCERTAINTY) IN ORIGINAL DATA (O.D.) AND PRINCIPAL COMPONENTS HYPERPLANES AT VARYING AMPLITUDE OF RANDOM GAUSSIAN NOISE

	Noise σ percentage value				
	4	8	12	16	20
O. D.	97.9 ± 0.5	97.0 ± 0.7	96.1 ± 0.8	95.2 ± 0.9	92.2 ± 1.2
PC1	90.1 ± 5.3	89.7 ± 5.2	88.2 ± 5.2	86.4 ± 5.1	84.1 ± 4.8
PC2	79.0 ± 7.0	77.9 ± 7.1	77.5 ± 6.7	74.7 ± 6.6	74.5 ± 6.5
PC3	93.2 ± 3.4	92.4 ± 3.5	88.7 ± 3.4	85.7 ± 3.3	84.4 ± 3.4
PC4	98.2 ± 0.4	97.1 ± 0.7	95.9 ± 0.7	92.7 ± 0.9	90.8 ± 1.1
PC5	97.6 ± 0.4	97.2 ± 0.5	94.6 ± 0.7	91.6 ± 0.1	89.9 ± 0.1
PC6	97.7 ± 0.6	96.6 ± 0.7	95.1 ± 1.0	93.3 ± 1.0	90.4 ± 1.1
PC7	97.9 ± 0.5	96.8 ± 0.7	96.2 ± 0.8	93.7 ± 0.9	91.6 ± 0.1
PC8	97.5 ± 0.6	96.8 ± 0.6	96.3 ± 0.7	93.5 ± 0.1	90.5 ± 0.1
PC9	98.1 ± 0.5	97.2 ± 0.6	96.6 ± 0.6	93.7 ± 0.9	91.9 ± 0.1

Table VII

ACCURACY (MEAN AND UNCERTAINTY) IN ORIGINAL DATA (O.D.) AND PRINCIPAL COMPONENTS HYPERPLANES AT VARYING AMPLITUDE OF HOMOGENEOUS NOISE %

	Noise Percentage value				
	4	8	12	16	20
O. D.	97.4 ± 0.6	97.4 ± 0.6	97.1 ± 0.7	92.6 ± 2.2	88.7 ± 3.3
PC1	90.7 ± 5.3	90.8 ± 5.2	90.5 ± 5.3	89.9 ± 5.4	88.0 ± 5.6
PC2	78.0 ± 7.2	77.4 ± 7.2	76.0 ± 7.4	73.4 ± 7.2	68.9 ± 7.0
PC3	93.0 ± 3.6	90.8 ± 3.8	85.1 ± 4.5	81.8 ± 4.5	72.0 ± 4.5
PC4	98.2 ± 0.4	97.4 ± 0.7	94.2 ± 1.8	89.7 ± 3.2	85.2 ± 3.2
PC5	97.5 ± 0.4	97.6 ± 0.4	93.5 ± 1.8	88.8 ± 3.1	85.3 ± 3.3
PC6	97.4± 0.6	97.8 ± 0.5	95.6 ± 0.1	91.1 ± 3.2	87.6 ± 3.4
PC7	97.8 ± 0.5	97.8 ± 0.5	95.4 ± 1.5	91.1 ± 3.1	87.6 ± 3.4
PC8	97.5 ± 0.6	97.4 ± 0.6	94.9 ± 1.5	89.9 ± 3.0	86.7 ± 3.4
PC9	97.8 ± 0.5	97.5 ± 0.5	94.9 ± 1.54	90.9 ± 2.9	88.1 ±3.3

when $p > 3$ and, correspondingly, a lower uncertainty. This last result suggests a better noise robustness with PCA.

In bi-dimensional case, PCA highlights that the variance of the control group is lower. Among the two groups, a significant difference in dispersion around the mean value as well as amplitudes comes out. Results of Fig. 2 confirm the data separability, even using only the first principal component, capable of explaining almost the 90% of variance. The good correlation between psychometric data and the exposure to different experimental set up (emotionally stressful and not) founds the experimental set up reliability in conditioning participants with regard to the study variable. The high accuracy level suggests that the signal acquired through a single channel preserves the information concerning the frontal asymmetry elicited from an emotional stress condition.

2) Noise and Bias Robustness: Noise robustness was tested on the worst classifier (SVM) in order to assess the robustness of the proposed method. The subject-wise leave-two-out cross-validation strategy was repeated but with a further noise parameter, both (i) to make more generic the proposed method, and (ii) to verify the occurrence of possible bias during acquisition. The second evaluation is aimed to test the noise robustness of classification accuracy after PCA. In particular, two different kinds of noise were considered. In the first test, aimed at generalization, a random Gaussian noise with zero-mean and $\sigma \in \{0.04, 0.08, 0.12, 0.16, 0.20\}$, multiplied by the absolute value of the data maximum, was added. Results are reported in Tab. VI.

In the second test, aimed to verify bias, a constant value was added to each subject signal of the test sets. In this way, the signal of each subject was treated with a different random bias. Bias levels were chosen randomly within intervals of increasing amplitude ($\sigma \in [0.04 - 0.20]$, step = 0.04). For this reason, the global effect on the entire data set is noise. Results are reported in Tab. VII.

In both the cases, the instrument shows good noise robustness. The classifier with PCA performs better if the noise level is less than 12% of absolute value of the maximum of data. Performance degrades in any case with higher noise levels. In this study, the differential channel and the PCA are exploited to face the problem of artifacts. A differential channel intrinsically rejects the common mode noise. PCA on the differential channel EEG acts like a pass band filter owing to its intrinsic reduction of the signal dimensionality in the PC domain. The combined effect of this two filtering effects improves the signal-to-noise ratio significantly. The experimental analysis of noise robustness validated, ex post, the proposed method.

VI. CONCLUSIONS

A method to assess stress condition in real time trough a high-wearable EEG-based device has been proposed. EEG signal amplitudes variations between prefrontal right and left zone were acquired through a single differential channel. The induced stress status was verified by a psychologist through (i) questionnaires administered before and after the stress test, and (ii) performance assessment. Time domain features were used in the classification procedure. Four standard machine learning classifiers (SVM, k-NN, Random Forest, and ANN) reached more than 90% accuracy in distinguishing each 2-s epoch of EEG. Generally, PCA allows to obtain a better noise robustness. The results show the adequacy of the proposed solution based on a single-acquisition channel and time domain-based feature selection. In the worst case, the SVM Linear -Kernel classifier succeeded in discriminating stress conditions with an accuracy of 97.5 ± 0.6% and a latency of 2 s. For latency above 4 s the accuracy reaches 100%. Noise robustness was tested in order to exclude the impact of bias during signal acquisition and to empower generality to the results. The proposed method gives a new way to detect prefrontal asymmetry traditionally associated to emotional stress condition. Further future experimental activity with a larger number of subjects are necessary to consolidate the statistical significance of these preliminary results.). Electrode scalp locations used in this study, FP1 and FP2, are considered as sensitive to ocular artifacts. However, our experiments did not highlight this problem. In any case, further experimental campaigns will be carried out on new areas of the scalp. In this way, the impact on the classifier of the information produced by both the EEG signals and the eye movements will be deepen. Alternative classifiers and strategies for the feature extraction such as sparse dictionary learning will be implemented.

Acknowledgment

This work is dedicated to the late profs. Felice Cennamo, Massimo D'Apuzzo, and Domenico Grimaldi. Authors thank Andrea Grassi and Leda Bilo for useful suggestions, as well as Carmela Laezza for support in data elaboration

References

[1] M. Poole and M. Warner, *The IEBM handbook of human resource management*. International Thomson Business, 1998.

[2] J. Fischer, A. Calame, A. Dettling, H. Zeier, and S. Fanconi, "Objectifying psychomental stress in the workplace–an example," *International archives of occupational and environmental health*, vol. 73, no. 1, pp. S46–S52, 2000.

[3] B. Siciliano and O. Khatib, *Springer handbook of robotics*. Springer, 2016.

[4] N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey," *Computer methods and programs in biomedicine*, vol. 108, no. 3, pp. 1287–1301, 2012.

[5] M. Choi, G. Koo, M. Seo, and S. W. Kim, "Wearable device-based system to monitor a driver's stress, fatigue, and drowsiness," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 3, pp. 634–645, 2017.

[6] J. Minguillon, E. Perez, M. Lopez-Gordo, F. Pelayo, and M. Sanchez-Carrion, "Portable System for Real-Time Detection of Stress Level," *Sensors*, vol. 18, no. 8, p. 2504, 2018.

[7] H. Jebelli, S. Hwang, and S. Lee, "EEG-based workers' stress recognition at construction sites," *Automation in Construction*, vol. 93, pp. 315–324, 2018.

[8] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device," *IEEE Transactions on information technology in biomedicine*, vol. 14, no. 2, pp. 410–417, 2009.

[9] J. W. Ahn, Y. Ku, and H. C. Kim, "A Novel Wearable EEG and ECG Recording System for Stress Assessment," *Sensors*, vol. 19, no. 9, p. 1991, 2019.

[10] L. Angrisani, P. Arpaia, D. Casinelli, and N. Moccaldi, "A Single-Channel SSVEP-Based Instrument With Off-the-Shelf Components for Trainingless Brain-Computer Interfaces," *IEEE Transactions on Instrumentation and Measurement*, 2018.

[11] G. Matthews, L. Joyner, K. Gilliland, S. Campbell, S. Falconer, and J. Huggins, "Validation of a comprehensive stress state questionnaire: Towards a state big three," *Personality psychology in Europe*, vol. 7, pp. 335–350, 1999.

[12] S. L. Mïller-Abdelrazeq, S. Stiehm, M. Haberstroh, and F. Hees, "Perceived Effects of Cycle Time in Human-Robot-Interaction," in *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*. IEEE, 2018, pp. 25–30.

[13] J. Wijsman, B. Grundlehner, H. Liu, H. Hermens, and J. Penders, "Towards mental stress detection using wearable physiological sensors," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 1798–1801.

[14] A.-M. Brouwer, M. A. Neerincx, V. Kallen, L. van der Leer, and M. ten Brinke, "EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress," *Journal of Cybertherapy & Rehabilitation*, vol. 4, no. 1, pp. 21–34, 2011.

[15] S. GR *et al.*, "Development of four stress levels in group stroop colour word test using HRV analysis." *Biomedical Research (0970-938X)*, vol. 28, no. 1, 2017.

[16] A. Mohammed and L. Wang, "Brainwaves driven human-robot collaborative assembly," *CIRP annals*, vol. 67, no. 1, pp. 13–16, 2018.

[17] B. Wallace, F. Knoefel, R. Goubran, R. A. L. Zunini, Z. Ren, and A. Maccosham, "EEG/ERP: within episodic assessment framework for cognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 10, pp. 2525–2534, 2017.

[18] A. Lay-Ekuakille, P. Vergallo, G. Griffo, F. Conversano, S. Casciaro, S. Urooj, V. Bhateja, and A. Trabacca, "Entropy index in quantitative EEG measurement for diagnosis accuracy," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 6, pp. 1440–1450, 2013.

[19] L.-D. Liao, S.-L. Wu, C.-H. Liou, S.-W. Lu, S.-A. Chen, S.-F. Chen, L.-W. Ko, and C.-T. Lin, "A novel 16-channel wireless system for electroencephalography measurements with dry spring-loaded sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 6, pp. 1545–1555, 2014.

[20] Y.-C. Chen, B.-S. Lin, and J.-S. Pan, "Novel noncontact dry electrode with adaptive mechanical design for measuring EEG in a hairy site," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 12, pp. 3361–3368, 2015.

[21] S. Mühlbacher-Karrer, A. H. Mosa, L.-M. Faller, M. Ali, R. Hamid, H. Zangl, and K. Kyamakya, "A driver state detection system—combining a capacitive hand detection sensor with physiological sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 4, pp. 624–636, 2017.

[22] Z. Gao, S. Li, Q. Cai, W. Dang, Y. Yang, C. Mu, and P. Hui, "Relative wavelet entropy complex network for improving EEG-based fatigue driving classification," *IEEE Transactions on Instrumentation and Measurement*, no. 99, pp. 1–7, 2018.

[23] D. J. Krusienski, M. Grosse-Wentrup, F. Galán, D. Coyle, K. J. Miller, E. Forney, and C. W. Anderson, "Critical issues in state-of-the-art brain–computer interface signal processing," *Journal of neural engineering*, vol. 8, no. 2, p. 025002, 2011.

[24] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, p. 031005, 2018.

[25] X. Chen, H. Peng, F. Yu, and K. Wang, "Independent vector analysis applied to remove muscle artifacts in EEG data," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1770–1779, 2017.

[26] X. Chen, X. Xu, A. Liu, M. J. McKeown, and Z. J. Wang, "The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings," *IEEE transactions on instrumentation and measurement*, vol. 67, no. 2, pp. 359–370, 2017.

[27] M. Zanetti, T. Mizumoto, L. Faes, A. Fornaser, M. De Cecco, L. Maule, M. Valente, and G. Nollo, "Multilevel assessment of mental stress via network physiology paradigm using consumer wearable devices," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–10, 2019.

[28] S. A. Hosseini and M. A. Khalilzadeh, "Emotional stress recognition system using EEG and psychophysiological signals: Using new labelling process of EEG signals in emotional stress state," in *2010 international conference on biomedical engineering and computer science*. IEEE, 2010, pp. 1–6.

[29] P. Zhang, X. Wang, X. Li, and P. Dai, "EEG feature selection based on weighted-normalized mutual information for mental fatigue classification," in *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*. IEEE, 2016, pp. 1–6.

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

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

[32] 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.

[33] D. Bian, J. W. Wade, A. Swanson, Z. Warren, and N. Sarkar, "Physiology-based affect recognition during driving in virtual environment for autism intervention." in *PhyCS*, 2015, pp. 137–145.

[34] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, "Emotion classification using minimal EEG channels and frequency bands," in *The 2013 10th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE, 2013, pp. 21–24.

[35] J. A. Coan and J. J. Allen, "Frontal EEG asymmetry as a moderator and mediator of emotion," *Biological psychology*, vol. 67, no. 1-2, pp. 7–50, 2004.

[36] I. Papousek, E. M. Weiss, G. Schulter, A. Fink, E. M. Reiser, and H. K. Lackner, "Prefrontal EEG alpha asymmetry changes while observing disaster happening to other people: cardiac correlates and prediction of emotional impact," *Biological psychology*, vol. 103, pp. 184–194, 2014.

[37] R. N. Goodman, J. C. Rietschel, L.-C. Lo, M. E. Costanzo, and B. D. Hatfield, "Stress, emotion regulation and cognitive performance: The predictive contributions of trait and state relative frontal EEG alpha asymmetry," *International Journal of Psychophysiology*, vol. 87, no. 2, pp. 115–123, 2013.

[38] J. Minguillon, M. A. Lopez-Gordo, and F. Pelayo, "Stress assessment by prefrontal relative gamma," *Frontiers in computational neuroscience*, vol. 10, p. 101, 2016.

[39] J. A. Coan and J. J. Allen, "Frontal EEG asymmetry and the behavioral activation and inhibition systems," *Psychophysiology*, vol. 40, no. 1, pp. 106–114, 2003.

[40] R. J. Davidson, "Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums," *Psychophysiology*, vol. 35, no. 5, pp. 607–614, 1998.
[41] I. Jolliffe, *Principal component analysis*. Springer, 2011.
[42] E.-C. Chang, S.-C. Huang, and H.-H. Wu, "Using k-means method and spectral clustering technique in an outfitter's value analysis," *Quality & Quantity*, vol. 44, no. 4, pp. 807–815, 2010.
[43] C. M. Bishop *et al.*, *Neural networks for pattern recognition*. Oxford university press, 1995.
[44] P. Renaud and J.-P. Blondin, "The stress of stroop performance: Physiological and emotional responses to color–word interference, task pacing, and pacing speed," *International Journal of Psychophysiology*, vol. 27, no. 2, pp. 87–97, 1997.
[45] C. D. Spielberger, "Manual for the State-Trait Anxiety Inventory STAI (form Y)(" self-evaluation questionnaire")," 1983.
[46] J. Blascovich and J. Tomaka, "Measures of self-esteem," *Measures of personality and social psychological attitudes*, vol. 1, pp. 115–160, 1991.
[47] A. Celisse and S. Robin, "Nonparametric density estimation by exact leave-p-out cross-validation," *Computational Statistics & Data Analysis*, vol. 52, no. 5, pp. 2350–2368, 2008.