

Evaluation of an integrated system of wearable physiological sensors for stress monitoring in working environments by using biological markers

Stefano Betti, Raffaele Molino Lova, Erika Rovini, Giorgia Acerbi, Luca Santarelli, Manuela Cabiati, Silvia Del Ry, and Filippo Cavallo*, *Member, IEEE*

Abstract—Objective: The objectives of this work is to develop and test the ability of a wearable physiological sensors system, based on ECG, EDA and EEG, to capture human stress and to assess whether the detected changes in physiological signals correlate with changes in salivary cortisol level, which is a reliable, objective biomarker of stress. **Methods:** 15 healthy participants, seven males and six females, mean age 40.8 ± 9.5 years, wore a set of three commercial sensors to record physiological signals during the Maastricht Acute Stress Test, an experimental protocol known to elicit robust physical and mental stress in humans. Salivary samples were collected throughout the different phases of the test. Statistical analysis was performed using a Support Vector Machine (SVM) classification algorithm. A correlation analysis between extracted physiological features and salivary cortisol levels was also performed. **Results:** 15 features extracted from heart rate variability, electrodermal and electroencephalography signals showed a high degree of significance in disentangling stress from a relaxed state. The classification algorithm, based on significant features, provided satisfactory outcomes with 86% accuracy. Furthermore, correlation analysis showed that the observed changes in physiological features were consistent with the trend of salivary cortisol levels ($R^2 = 0.714$). **Conclusion:** The tested set of wearable sensors was able to successfully capture human stress and quantify stress level. **Significance:** The results of this pilot study may be useful in designing portable and remote control systems, such as medical devices used to turn on interventions and prevent stress consequences.

Index Terms—Stress detection, Physiological sensors, Machine learning, Cortisol, Work environments.

Manuscript submitted December 1, 2016. This work was supported in part by European Community within the TransSafe Project (Sixth call of the Ambient Assisted Living Joint Programme (AAL JP) - AAL-6-2013) and by the Italian Ministry of Health within the Current Research Program performed at National Research Institutes (IRCCS).

S. Betti, E. Rovini, G. Acerbi, L. Santarelli and F. Cavallo* are with the BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera, Pisa, Italy (correspondence e-mail: filippo.cavallo@santannapisa.it). R. Molino Lova is with the Dept. of Cardiac Rehabilitation, IRCCS Fondazione Don Gnocchi, Florence, Italy. M. Cabiati and S. Del Ry are with the Institute of Clinical Physiology, National Research Council (CNR), Pisa, Italy.

I. INTRODUCTION

A. Motivation

ACCORDING to the modern definition, "Stress occurs where demands made on individuals do not match the resources available or meet the individual's needs and motivation. It will be the result if the workload is too large for strength and time available. Equally, a boring or repetitive task which does not use the potential skills and experience of some individuals will cause them stress" [1]. Many studies reveal that in recent years, stress and its consequences are playing an important role in modern society [2]. According to the American physiological association, 75% of the population affirm that in the last month, they have experienced a symptom related to stress. Furthermore, 25% of people think that a high level of stress has a huge impact on their physical health status. Similarly, the European Commission conducted a study [3] that indicates that more than 22% of employees in the EU think that their health or safety is at risk because of their work. This percentage increases considerably in the case of older people, and of jobs with high personal and public risk and irregular working hours. Generally, growing stress and anxiety, in addition to relapses in well-being of the individual, have a significant effect on the global economy and on society. For example, each year in the UK, work-related stress results in the loss of nearly 13 million working days at a cost of £12 billion [4]. This is especially true for physically and psychologically demanding working activities, such as in the transportation [5], medical [6], military [7], civil protection [8] and office [9] sectors. Another study showed that a significant percentage of people admit that they have never undertaken a strategy or activity to reduce and manage their stress [10]. There is therefore a need for efficient management of stress in workplaces, preferably with reliable methods for automated and real-time monitoring of a person's stress levels. Consequently, employers must find ways to keep their personnel motivated, fit and on duty while at the same time guaranteeing their personal safety. This scenario is more relevant considering that on-going demographic changes have resulted in a growing number of older people, thus increasing

the necessity of policies to maintain people at work longer, in a safe manner.

In this context, the general objective of this paper is to design, develop, test and evaluate an integrated system of wearable physiological sensors for stress monitoring in working environments by using biological markers.

B. Stress definition and assessment

From a biological point of view, life is characterized by the continuous interchange of energy and information with the environment within the narrow limits of a dynamic equilibrium known as “homeostasis” [11]. The term “stress” is used in biology to define a real or anticipated disruption of homeostasis, and, more generally, a real or anticipated threat to physiologic integrity [12]. In front of any event, limbic brain structures, such as the hippocampus, amygdala and prefrontal cortex, assess its destabilizing potential and, if the event does not match some cognitive representation based on previous subjective experience [11], then a general alarm reaction is triggered. Psychological changes, such as anxiety and fear, along with arousal, vigilance and alertness, are parts of the clinical picture of this alarm reaction, each of which being more or less represented according to individual susceptibility and resilience. Exposure to stressful events also triggers two main physiologic responses. The first and most immediate response is the activation of the sympatho-adrenal-medullary (SAM) axis due to a temporary suppression of the parasympathetic branch of the autonomous nervous system (ANS), with simultaneous activation of the sympathetic branch [13]. This results in an increased secretion of adrenalin and noradrenalin by the adrenal medulla, and an increased release of noradrenalin from sympathetic nervous terminations. This, in turn, results in a rapid increase in heart rate, heart inotropism, vascular tone, blood pressure, bronchial dilation and breathing frequency, thus predisposing the organism for a demanding task. The response of the SAM axis to a single stressful event lasts 1–2 minutes, owing to both the reflex parasympathetic branch re-activation [12, 13] and the exhaustion of the sympathetic activation effects related to the short plasmatic and inter-synaptic half-life of catecholamines. The second major response relates to the activation of the hypothalamus-pituitary-adrenal (HPA) axis, with the hypothalamus secreting corticotropin releasing hormone (CRH), which stimulates the secretion of adrenocorticotrophic hormone (ACTH) by the pituitary gland which, in turn, results in the secretion of cortisol by the adrenal cortex [11, 13]. The increased level of cortisol is responsible for the increase in blood glucose, which provides the organism with an efficient energetic substrate in view of a demanding task. Contrary to the stress response of the SAM axis, the response of the HPA axis to a single stressful event can last up to 60 minutes after the event [14]. These phylogenetically archaic responses to stress evolved under environmental selection pressure, as they provided living organisms with some evolutionary advantage, particularly the “fight or flight” response [15], which is intuitively aimed at increasing the chance of individual survival. These evolutionary advantageous responses may turn into disadvantageous health threatening

reactions [16], however, depending on individual susceptibility and resilience [17]. Sudden changes in catecholamine levels may, in fact, be responsible for the occurrence of severe cardiac arrhythmias, hypertension, stroke and coronary artery disease, whereas higher cortisol levels may be responsible for the occurrence of gastrointestinal diseases, such as gastric ulcers and ulcerative colitis, and cancer [18]. The physiologic responses to stress also affect cognitive-behavioural reactions in normally susceptible and resilient individuals, when the exposure to stressful events is repeated and without adequate periods of recovery between events. Imperfect perception, insufficient attention, inadequate or delayed information processing and errors of judgment frequently result from repeated exposure to stressful events [19, 20, 21] and may have serious consequences, particularly in working environment. With respect to these physiologic responses to stressful events, this study used electrodermal activity (EDA), heart rate variability (HRV) and electroencephalogram (EEG) to estimate the level of stress. This selection is also in accordance with the ranking of physiological and physical signals for measuring stress drawn up in a survey study conducted by Sharma [22]. In this chart, HRV, EDA and EEG take respectively first, second and third place. Whereas salivary cortisol was used to compare and evaluate the quality of the stress measurement.

II. RELATED WORKS

Traditionally, stress and anxiety have been detected using qualitative and subjective tools such as psychometric instruments and scales based on humans rating stress levels [23, 24]. Questionnaires can be self-compiled or performed by an expert [25, 26], but users often may not answer exactly how they are feeling and the psychometric responses could be dependent on participants’ personal wrong stress perception [27]. Moreover, this analysis is affected by a low level of objectivity and does not allow continuous monitoring over time.

Recently, physiological signals have been used to measure stress with a higher level of objectiveness than that of psychometric instruments, and a lower level of intrusiveness and complexity than that of biological analysis. In fact, the development of new technologies and wearable sensors has allowed solving critical issues with common devices used in diagnostics: encumbrance, portability and high intrusiveness. Analysis of physiological signals seems to be a good trade-off between the previous techniques and, in particular, EDA, HRV and brain waves are commonly used in the literature to investigate levels of stress during different tasks [22]. Mental stress arouses an increase in sweat gland activity, resulting in a variation in skin conductance. EDA, as an indicator of sympathetic activation, has been used in previous works alone [33] or in combination with other physiological parameters to investigate stress level in computer users [34] and automotive drivers [35], as well as during other common tasks [36].

Additionally, several studies have shown that stress has an impact on the ANS [37] and, consequently, on cardiac activity. To study the effect of sympathetic and parasympathetic activities based on electrocardiogram (ECG) signal, it is necessary to analyse the HRV signal both in time and frequency

domains. HRV analysis has previously been used in different studies to detect stress during mental tasks [38], high workload [39], and car driving [40].

Brain activity is also subjected to changes due to mental and physical stress [41]. EEG signals are arranged in several frequency bands that represent different physical and emotional states. EEG analysis has been used to investigate the stress level of computer players [22] and to classify the execution of different stressful tasks [42].

Although the measurement of physiological parameters has been used in previous works to study the stress levels of subjects during different tasks, variations in physiological signals can be affected by multiple factors, such as environmental conditions, physical activities and body posture. So far, some research works have highlighted the necessity to use a combination of physiological and movement sensors in order to improve the robustness to detect stress. This sensor fusion approach, indeed, allows to investigate among the most efficient sensor feature selections and consequently to identify the most accurate data processing techniques. Typically, ECG, EDA and SpO2 are combined with some techniques of supervised learning [44, 45]; similarly, other examples of sensors include respiration [46], EMG and EEG [47]. Interestingly, also microphone, proximity and inertial sensors are used [43,45]. The overall accuracy in detecting stress with different supervised or unsupervised methods, to our best knowledge, ranges between 0.79 and 0.95; usually it is lower when data processing aims to detect at least 3 levels of stress (no stress, moderate stress, stress), rather than 2 levels (no stress, stress), which is the most common situation (Table I).

Most of the works [43-46] are based on a psychological / cognitive stress induction, which is often not adherent to real life, where also physical stress is typically included. Only, Xu et al. [47] proposed a stress induction with cognitive tests and a physical stress based on squat-stand exercise.

Another critical issue concerns the necessity to have a reference value of stress, that is used to validate the stress measured with sensors. The majority of works use self-reporting methods, such as the STAI questionnaire [43-47] (Table I). However, it is well known that these approaches suffer of high variability intra and among subjects and a low

level of objectivity. On the other hand, since stress levels cause a change in stress hormone production (e.g. high release of cortisol or catecholamine) [28, 29], the measurement of hormone concentrations is an objective method to detect stress. This approach requires invasive techniques (e.g. taking blood, saliva or urine samples) and lengthy expensive laboratory analysis procedures [22]. Some specific studies, however, use such monitoring techniques to evaluate stress in drivers [30], military field [31] and stress intervention techniques [32].

Based on the previous considerations, despite several previous studies concerning stress detection being conducted using different approaches, further investigations are required. Indeed, this paper aims to contribute with some specific points: (1) using the most relevant sensors used in stress detection, as investigated and discussed in [22], that guarantee an overall accuracy at least in line with state of the art; therefore the reliability of ECG, EDA and EEG features and the evaluation of their statistical significance is studied; (2) using a well defined and clinically certified stress induction protocol that includes alternated phases of relax and mental and physical stress induction, i.e. the Maastricht Acute Stress Test (MAST); (3) investigating the possibility to identify at least 3 levels of stress, overcoming the obvious situation of stressed or not stressed; (4) comparing the measured stress levels with objective biological methods based on cortisol [49].

Even if expensive experimental and laboratory procedures are required, this last point is fundamental, in order to ensure that changes in physiological signals are directly related to stress condition. Therefore the correlation between the physiological parameters and the concentration of hormones should be investigated. This critical point is not commonly studied [48] but it can be necessary for validating a stress detection algorithm obtained through the analysis of physiological signals by wearable sensors. This paper is structured as follows. In Section III, materials and methods are presented with a description of the system, the experimental methodology to collect data, the features extraction procedure, the signals analysis and classification algorithms. In Section IV, classification and correlation results are shown. Finally, discussion and conclusions are respectively Sections V and VI.

TABLE I:
RELATED WORKS WITH THE MAIN ATTRIBUTES.

Cit.	Sensors' Type	Subjects	Stress Induction	Processing / Classification	Comparison / Validation
[43]	IMU, EDA	28	Truck Driver Simulator	Support Vector Machine	Survey, Camera for facial expression
[44]	EDA, SpO2, ECG,	5	Cognitive task (Trier Social Stress Test (TSST))	Support Vector Machine	STAI questionnaire
[45]	ECG, EDA, SpO2, Microphone, Accelerometer, Proximity	18	Cognitive task (Trier Social Stress Test (TSST))	Support vector machine, AdaBoost, k-nearest neighbor	STAI questionnaire
[46]	ECG, Respiration	39	Psychological and cognitive tasks (Montreal Imaging Stress Task (MIST))	Support Vector Machine, Linear Discriminant Analysis, Adaboost, Nearest Neighbors	Survey
[47]	ECG, EDA, EMG, EEG, SpO2	44	Physical or cognitive task (squat-stand exercise, six computer-mediated tasks)	k-means, cluster-wise	STAI questionnaire



Fig. 1. Left and centre panels: Wearable sensors used to acquire physiological signals and their specific way to be worn; A) MindWave Mobile EEG headset, B) Zephyr BioHarness™3 (BH3) chest belt, C) Shimmer Sensor. Right panel: Picture depicting the execution of the cold pressure test.

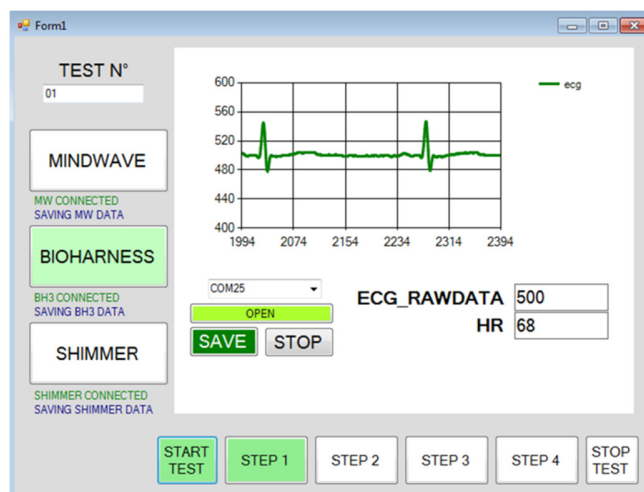


Fig. 2. Detail of the graphical user interface related to the BioHarness panel with dedicated commands for device connection management and acquisition of physiological signals.

III. MATERIALS AND METHODS

A. Participants

15 healthy participants, 8 males and 7 females, mean age 40.8 ± 9.5 years, were enrolled from among the employees of the Fondazione Don Gnocchi (FDG) of Florence, Italy. Eligibility was assessed by a physician based on medical history and visits to ascertain possible exclusion criteria, such as cardiac arrhythmias, coronary artery disease, chronic heart failure, uncontrolled hypertension or diabetes, acrocyanosis, Raynaud's disease, heavy smoking, substance abuse and/or medication known to affect stress response. Participants with oral lesions or salivary gland diseases were also excluded. The Ethical Committee of the FDG, where clinical experiments were conducted, approved the study protocol in a session held on December 16th, 2015. All participants signed an informed consent form before being included in the study. Unfortunately, of the 15 participants only 12 provided usable data for the analysis. In fact, one participant had carried out intensive physical training just before the test, compromising the usefulness of the collected data. Another participant complained of impending faint during the stress induction phase, so the test was immediately stopped. Finally, in one case the EDA data were not saved due to a malfunction of the acquisition system.

B. Instruments

In this section, we describe the components of the wireless sensor system to acquire physiological data during the experimental tests. To detect changes in physiological signals a designed multimodal sensor-compute infrastructure is required. The setup used involved three different wearable sensors and a personal computer provided with a dedicated interface.

The choice of the wearable sensor devices to be included in the study was based on three different criteria: the nature of the measured signals (HRV, EDA, EEG), the accuracy of measurements and the unobtrusiveness of the sensors. Considering the above criteria, the chosen devices had to be able to measure cardiac, electrodermal and brain activities in an unobtrusive way, in order to guarantee a minimum level of acceptability from a user's perspective. For this purpose, the BioHarness™3 (Zephyr, Maryland USA) (Fig. 1A), Shimmer

Sensor (Shimmer, Ireland) (Fig. 1B) and MindWave Mobile EEG headset (Neurosky, California USA) (Fig. 1C) were selected.

The Zephyr BioHarness™3 (BH3) is a Bluetooth chest belt able of retrieving signals derived from ECG, such as heart rate (HR) and inter-beat-interval (IBI), at a sampling frequency of 250 Hz. The Shimmer is an EDA-monitoring wearable sensor composed of two special finger electrodes and a main unit that is able to send data to a personal computer with a sample frequency equal to 51.2 Hz by means of Bluetooth connection. The MindWave EEG headset is a Bluetooth device able to capture single channel EEG raw data at a sample rate of 512 Hz, and to provide an index of attention and meditation of the user using the frequency power spectrum density (sample rate 1 Hz). The three wearable sensors used are shown in Fig. 1.

Each sensor transmitted through a Bluetooth link to a USB Bluetooth 2.1+EDR Dongle connected with a personal computer. The BH3 sensor transmitted (i) the IBI message payload of 40 byte and (ii) the ECG message payload of 136 byte. The IBI message rate depends on the HR frequency of the individual (approx. 1 message every 18 seconds) while the ECG messages was sent at 4Hz. The MindWave sensor transmitted two type of messages: (i) the EEG wave power EEG index with 35 byte at 1Hz and (ii) the EEG raw data with 1 byte payload at 500Hz. Shimmer sensor sent only one type of messages, including the GSR signal with a 10 byte payload at 51.2 Hz. The entire system transmitted about 1600 byte/s of data payloads, that is about one half of the theoretical Bluetooth 2.1+EDR maximum data rate.

Sensors data were processed on PC instrumented with an Intel i7 processor and 4GB RAM. A dedicated multithread application was developed to the saving of data for the post processing (Fig. 2). The same application also displayed in runtime the acquired data for the experimenter's reference. For such purpose, a dedicated graphical user interface was designed. The sensing and communication software was implemented in C# language in Visual Studio (Microsoft Corporation, Washington, USA) environment to handle the required tasks during data acquisition and managing control. It

was implemented starting from the Software Development Kit (SDK) released from each sensors company and allowed to manage wearable sensors connection, data communication and saving of different comma-separated value files. Time synchronization is essential for any communications networks, especially for wireless sensor networks. In the described system a global timestamp provided from the internal clock of personal computer represented a foundation for merging individual sensor readings.

For the collection of saliva samples, 10 Salivette Cortisol (Salivette® device, Sarstedt, Germany) were used for each subject (double swab for five samplings). At a specific sample time, the user had to chew two swabs for 1 minute in order to produce and acquire a sufficient quantity of saliva for the analysis.

C. Experimental Protocol

Participants underwent the MAST [13], a quick and non-invasive test able to induce physical and mental stress, and elicit robust sympathetic and glucocorticoid responses. The MAST was conceived and validated at the Maastricht University, Maastricht, The Netherlands, and has been described in detail by Smeets [13]. In this experimental loop test, wearable sensors and physiological signal acquisition were incorporated into the original MAST protocol. The test was arranged in four different phases: Relax, Stress, Recovery1 and Recovery2 (Fig. 3). During the Relax phase, after having obtained the basal saliva sample, participants wore the set of sensors resting on a chair in a room for 10 minutes, without using a mobile phone, and without music, external sounds or any other stimulus. A short presentation lasting 5 minutes was then shown to instruct participants about the upcoming tasks.

The explanation was followed by the Stress phase, consisting of a rotation of two different kinds of stressors: five periods of cold pressure stress alternated with four periods of mental arithmetic tasks. Each cold pressure period consisted of the immersion of the non-dominant hand in ice-cold water (5°C) (Fig. 1). The mental stress test consisted of backwards counting from 2043 in steps of 17 as quickly and as accurately as possible. A temporised presentation drove the subject during the stress induction phase with visual and vocal commands.

The Recovery phase from the stress induction was separated into two different sub-phases. During Recovery1, the participant continued wearing the sensors. During the last 30 minutes of the protocol (Recovery2), however, the sensors were removed and the physiological signals were not monitored to ensure to replicate the initial conditions in which the participant collected the first saliva sample (tPre). In fact, the presence of the wearable sensors on the participant's body could *per se* represent a cause of stress. Furthermore, four saliva samples were taken: the first one immediately after the induction of stress (t0) and the other ones after 10 (t10), 30 (t30) and 40 minutes (t40). Saliva samples were immediately frozen at -20°C and sent to the laboratory of the Consiglio Nazionale delle Ricerche (CNR) in Pisa, Italy. Additionally, blood pressure was recorded at the beginning and at the end of the test. All tests were performed in the early afternoon, between 2 and 4 p.m., to avoid the physiologic peak cortisol levels, which typically occur in the early morning and characterise the circadian cortisol rhythm. For safety reasons, a cardiologist was present

during all experimental sessions and an emergency trolley was always kept within reach.

D. Physiological Data Analysis

The physiological data acquired during the whole experimentation were analysed offline using Matlab® R2012a (MathWorks, Massachusetts USA). The data were examined in relation to three different phases: Relax (10 minutes of recording), Stress (15 minutes), Recovery1 (10 minutes).

Three different datasets containing HRV, EDA and EEG data for each phase were obtained. All these data were then appropriately filtered and processed in order to obtain a set of salient features. The relationship between the sequence of phases, cortisol trends and the variation in physiological parameters was then investigated. 16 features were extracted from the EDA signal in the time domain; 16 features were extracted from HRV through time and frequency domains analysis and 6 features related to frequency power spectral density were captured from MindWave output. The final set was composed of 38 different parameters.

The Shimmer sensor provides as output galvanic resistance that has been converted into galvanic skin conductance (SC). The SC is characterised by a slow varying tonic activity and a relatively fast varying phasic activity [50]. The latter is characterised by local peaks, startles, of 1 to 5 seconds of duration [51]. Importantly, the GSR is obtained by evaluating such peaks features. Contrariwise, the tonic phase is related to other uncorrelated sweating activities taking place over period of time longer than the one characterising the startle events. Therefore, the SC signal frequency content is entirely located within 2 Hz [52]. The Shimmer frequency sampling is set to 51.2 Hz, and thus, we applied a fourth-order Butterworth low-pass filter with a cut-off frequency off 2 Hz, so as to erase all components not linked with the SC signal. A further filtration aimed to extract the phasic signal from the acquired SC signal is obtained by means of a moving average filter (moving window set to 5 seconds). This filter removes the phasic signal from the tonic signal that is subsequently recovered and treated by means of an ad-hoc algorithm of startle detection and features extraction. Startles, detected according a trough-to-peak method, are classified measuring the rise time, the fall time, the amplitude, the duration, the area and the root mean square (Table II).

The BH3 device provides HRV data that specific the temporal distance between one heart beat and the following one. The analysis of cardiac signal was structured to investigate both the time domain and the frequency domain. An algorithm to extract the main features, based on IBI data, was developed. Ectopic rhythm, which is an irregular heart rhythm due to a premature heartbeat, was identified and corrected. Intervals that changed by more than 20% from the previous were replaced with the mean value of five neighbouring IBI intervals centred on the ectopic one. By correcting ectopic rhythm, a Normal-to-Normal (NN) interval sequence, appropriate for HRV analysis, was obtained starting from inter-beat (RR) interval sequence [53]. The resulting signals were used to compute the time domain features listed in Table III. NN interval sequence is, by

its nature, an irregularly sampled time sequence. Fourier-based power spectrum estimates require time series that are regularly sampled in time. Spectrum estimates taken from irregularly time-sampled signals can introduce additional harmonics into the power spectrum. For this reason, after a signal smooth detrending, the NN interval sequence was resampled at 4Hz, as suggested by Mali et al. [54]. Power Spectral Density (PSD) estimates can give information about the amount of power in which three principal frequency bands contribute to the time series: high frequency (HF), low frequency (LF) and very low frequency (VLF). The 10 HRV parameters, extracted from the frequency-domain are reported in Table IV. These features were calculated with a parametric autoregressive (AR) model of the order 16 with coefficients determined based on the Burg method [55].

The MindWave EEG headset mobile device provides a raw EEG signal, a power spectrum in different frequency ranges (alpha1: 8–9Hz, alpha2: 10–12Hz, beta1: 13–17Hz, beta2: 18–30Hz, delta: 1–3Hz, gamma1: 31–40Hz, gamma2: 41–50 Hz, theta: 4–7Hz), attention level and mediation level [56]. The device automatically provides a number of parameters as output (Table V).

Relax	Stress	Recovery 1	Recovery 2
10'	15'	10'	20'

SALIVARY CORTISOL SAMPLING

Fig. 3. Sequence, duration (minutes) and description of the different phases of the experimental protocol.

TABLE II:
FEATURES EXTRACTED FROM SKIN CONDUCTANCE (SC) SIGNAL
(M: MEAN SD: STANDARD DEVIATION)

Feature Name	Description
# Startle	Number of Startle detected
Startle Amp	Value (Mean±Std.) of a startle amplitude
Sum Rise Time	Sum of all detected startles time duration
Sum Fall Time	Sum of all detected startles fall time
Rise Rate	Value (Mean±Std.) of a startle rise time
Decay Rate	Value (Mean±Std.) of a startle fall time
Phasic Value	Value (Mean±Std.) of the phasic signal
Startle Time	Value (Mean±Std.) of a startle duration
Startle RMS	Value (Mean ± Std.) of the Root Mean Square of the curve identifying a startle
RMS overall	Value of the Root Mean Square of the phasic signal referred to an entire phase

TABLE III
FEATURES EXTRACTED FROM HRV SIGNAL IN TEMPORAL DOMAIN.

Feature Name	Description
IBI Mean	Mean of inter-beat interval corresponding to R-to-R interval
SDNN	Standard deviation of all Normal RR intervals (NN intervals)
HR Mean	Mean of heart rate
SDHR	Standard deviation of the heart rate
RMSSD	Square root of the mean of the squared differences between adjacent normal RR intervals
pNN50	Percentage of differences between adjacent normal RR intervals exceeding 50 ms

TABLE IV
FEATURES EXTRACTED FROM HRV SIGNAL IN FREQUENCY DOMAIN.

Feature Name	Description
VLF Peak	Frequency peak in very low frequency (VLF) range (0.00–0.04 Hz)
VLF Power	Signal power by Power Spectral Density (PSD) in VLF
%VLF	Percentage of signal power in the VLF with respect to the total signal power
LF Peak	Frequency peak in low frequency (LF) range (0.04–0.15 Hz)
LF Power	Signal power by PSD in LF
%LF	Percentage of signal power in the LF with respect to the total signal power
HF Peak	Frequency peak in high frequency (HF) range (0.15–0.40 Hz)
HF Power	Signal power by PSD in HF
%HF	Percentage of signal power in the HF with respect to the total signal power
Alpha Ratio (LF/HF)	Ratio between LF and HF powers

TABLE V
FEATURES EXTRACTED FROM EEG SIGNAL

Feature Name	Description
Attention	NeuroSky index for user's level of mental "focus" or "concentration" [0–100] (n.u.)
Meditation	NeuroSky index for user's level of mental "calmness" or "relaxation" [0–100] (n.u.)
EEG Alpha 1	NeuroSky index of EEG signal power in frequency range 8–9Hz (n.u.)
EEG Alpha 2	NeuroSky index of EEG signal power in frequency range 10–12Hz (n.u.)
EEG Beta 1	NeuroSky index of EEG signal power in frequency range 13–17Hz (n.u.)
EEG Beta 2	NeuroSky index of EEG signal power in frequency range 18–30Hz (n.u.)

In addition, all the parameters were scaled in amplitude for each subject between 0 and 1, in order to eliminate the interpersonal variability of physiological signals. For each person a baseline was acquired before the test execution and a set of baseline features was extracted and memorized. Therefore, all the features calculated during the test protocol have been normalized using the corresponding baseline features.

The Kolmogorov-Smirnov (KS) test was applied to the dataset described above to verify the normal distribution of features achieved from physiological signals. The KS test showed that data were not normally distributed. Thus, Kruskal-Wallis (KW), a non-parametric statistical analysis test, was used for comparing data acquired in different phases, in order to verify a significant disagreement (p-value < 0.05) on the basis of the extracted parameters.

Furthermore, the linear correlation between the significant parameters was calculated using the Spearman's coefficient. If the value of correlation between two features was at least rho = 0.8, the less significant feature was deleted. The remaining features were then used for Principal Component Analysis (PCA) to identify how the investigated groups, in relation to the relax, stress and recovery phases, could be visualised and separated in the space of the Principal Components (PCs). Finally, the most important PCs, those which included more than 80% of the overall variance of data, were taken into

account in order to train and test a Support Vector Machine (SVM) classifier, which had to be able to correctly classify a subject as stressed or not-stressed.

E. Salivary Cortisol Data Analysis

Saliva samples, collected in the Salivette, were thawed, mixed and centrifuged at 3000 g for 10 minutes in order to remove particulate material. Only samples with clear and clean saliva were used. After centrifugation, a total volume of 600–800 μl was obtained for each sample.

Cortisol concentrations were measured in 50 μl of saliva, in duplicate, via a commercially available immunoassay (Cortisol ELISA, RE 52611, IBL International Hamburg, Germany). The methodological performance of the assay was checked by dedicated experiments performed before the beginning of the study. The Cortisol ELISA has a detection range of 0.015–4 $\mu\text{g/dL}$ for saliva samples; within-assay variability was evaluated using two saliva samples and resulted in < 10% variation: CK1 = 1.64 ± 0.03 $\mu\text{g/dL}$ (n = 5 duplicate assays, CV = 4%) and CK2 = 0.53 ± 0.01 $\mu\text{g/dL}$ (n = 5 duplicate assays, CV = 5%). Between-assay variability, measured on a single repeated sample, resulted < 5% variation (1.45 ± 0.1 $\mu\text{g/dL}$, CV = 3.2%). Accuracy was evaluated using dilution tests providing a linear response, between 50 and 6.125 μL of saliva progressively diluted as required by manufacturer. Two control samples were assayed in each run for quality control.

Optical densities (OD) of saliva samples were read at 450 nm and were calibrated on a standard curve to measure the cortisol concentration of each sample. The interpolation of the dose–response curves was computed using a five-parameter logistic function (log scale) and calculated by Milliplex analyst integrated with Luminex xPONENT software (MILLIPLEX™Analyst, Merck Millipore Corporation, D, and Luminex Corpor. USA). Differences between more than two independent groups were analysed by Fisher’s test after ANOVA. The results are expressed as mean \pm standard error of the mean (SEM) and a p-value < 0.05 was considered significant. The statistical analysis was carried out using the Stat-View 5.0.1 (SAS Institute Inc., SAS Campus Drive, Cary, NC, USA).

F. Physiological and Cortisol Correlation

The correlation between the salivary cortisol concentration curve and the selected set of significant features was investigated to verify the capability of the physiological monitoring system to capture different stress conditions. The cortisol concentration curve provided an objective contribution for establishing the stress response in subjects, so it was used as reference for the comparison.

There is a time lag in hormonal response, typically 20 minutes after application of the stimulus [30], whereas the physiological response could be considered instantaneous. On this basis, the salivary cortisol concentration values obtained from samples collected during the experimental test were interpolated and resampled with a polynomial fitting of degree four for each subject. The new samples matched the centres of physiological monitoring windows (Relax, Stress, Recovery1),

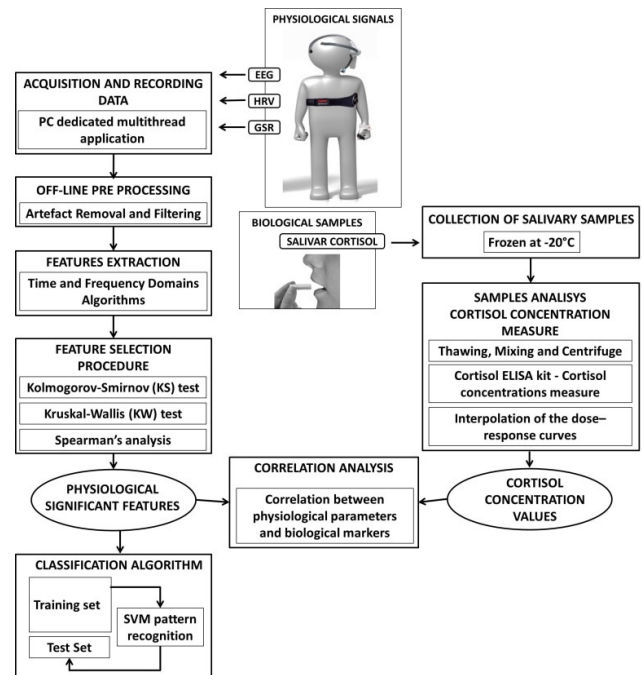


Fig. 4. Iterative scheme about the entire process of acquisition, processing and analysis of physiological and biological data

considering 20 minutes of delay (Fig. 5). In this way, the investigation of the actual correspondence between physiological and biological data became possible. A Multiple Linear Regression using Least Squares Fit (confidence significant level of 0.05) of cortisol values on a physiological parameters matrix was performed.

The whole process of data acquisition, recording, processing and analysis is shown in Fig. 4.

IV. RESULTS

In this section, the results obtained from the analysis of physiological data and the biological analysis of salivary samples are reported. Additionally, the correlation between the features’ extracted signals and the concentration of salivary cortisol is investigated.

A. Physiological and Cortisol Correlation

The 38 features extracted from HRV, EDA and EEG signals are reported as mean values and standard deviations in Table VI. At the end of each row, the p-values, calculated with KW test for non-parametric data, are also shown. A p-values index, calculated using the features extracted from signals recorded during relax and stress phases, provided information on the presence of potential significant differences between the two investigated conditions. 22 parameters marked with a * in Table VI represent concrete variations in physiological responses to MAST.

A further reduction of the database was deemed necessary in order to proceed with the data analysis. The redundancy of the information could be an obstacle for the accuracy of stress detection and some parameters were highly correlated with others. In particular, five features (HR Mean, SDNN, VLF

> REPLACE THIS LINE WITH YOUR PAPER IDENTIFICATION NUMBER (DOUBLE-CLICK HERE TO EDIT) <

Power, LF Power, %HF), close to cardiac activity, showed a high correlation with at least one other parameter. For these reasons the final set of parameters, able to distinguish between relax and stress status, was reduced from 22 to 15 features. HRV, EDA and EEG signals contributed to the final dataset with 5, 5 and 5 features, respectively. IBI Mean, SDNN, VLF Power, %VLF, %HF represent the cardiac activity in time and frequency domain. Startle Amp, Rise Time, #Startle, SD StartleRMS and Phasic Value are parameters related to skin conductance. Lastly, Attention, Meditation, Alpha1, Alpha2, Beta1 are linked to brain activity. The PCA related to the Relax and Stress phases is presented in Fig. 6. The two different groups were visualised and separated in the PC space. PCA analysis showed that the 7 most important PCs contained more than 80% of the overall variance in the data. This is why, on the first 7 linearly uncorrelated variables, a SVM algorithm was trained and tested. As the aim was to classify the subjects as stressed or not stressed, and to assess the quality of the system in terms of sensitivity, specificity and overall accuracy, a k-fold cross-validation technique ($k = 5$) was implemented to randomly select and assess training and testing datasets. A third order polynomial kernel for the SVM algorithm was chosen as the most suitable to recognise the performances of the subjects as belonging to relax or stress phases. To reduce variability, multiple rounds of cross-validation were performed using different random partitions, and the validation results were averaged over the rounds. The SVM classifier proved to be able to distinguish whether or not a person was stressed with performances summed up in 84.0% Sensitivity, 90.0% Specificity and 86.0% Overall Accuracy (Fig. 6). All of the features extracted by electro-cardiac, electro-dermal or brain activity were used in PCA and SVM, taking into account only the Relax and Stress phases. The same dataset of features was calculated also for the Recovery1 stage. Using the same significant parameters of the previous analysis, the dataset was expanded with data regarding Recovery1 and the PCA was performed again. The new group (green) was not comparable to either Relax or Stress, but was located in an intermediate position between the two clusters (Fig. 7).

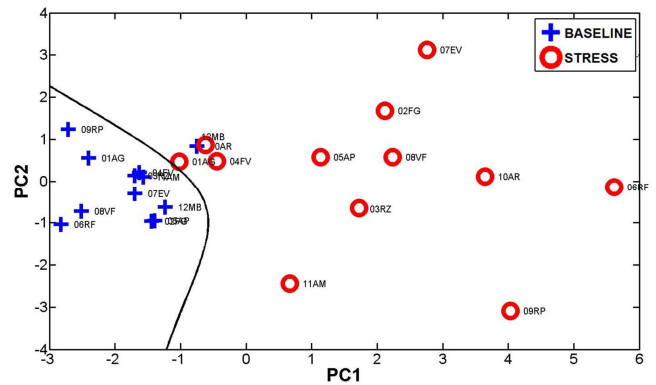


Fig. 6. PCA related to the Relax and Stress phases. The arrangement of data is shown on a plan consisting of the two main PCs to ensure a greater understanding and ease of viewing. A particular iteration of the SVM k-fold cross-validation technique is shown.

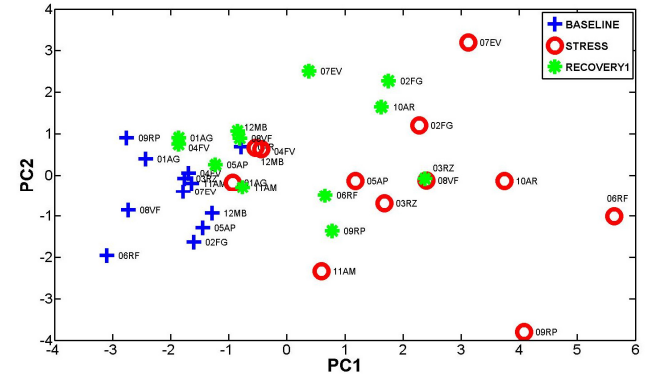


Fig. 7. PCA related to the Relax, Stress and Recovery1 phases. The two main PCs were used to ensure a greater understanding and ease of viewing. Recovery1 phase resulted in an intermediate position between Relax and Stress phases probably because the effects of the MAST dilate the actual time in which the subject returns to the initial conditions.

B. Salivary Cortisol Results

The assessment of salivary cortisol levels was aimed both at verifying the real capability of our MAST protocol and setup reproduction to induce stress, and at obtaining a biological stress marker with which to compare the physiological data. Results are shown in Table VII.

C. Correlation Results

HRV, EDA and EEG features that composed the final dataset used for the SVM classification were also involved in the correlation study with the cortisol. In particular, the correlation was studied between the vector containing salivary cortisol levels and a dataset composed of 15 physiological features representing ECG, EDA and EEG (IBI Mean, SDNN, VLF Power, %VLF, %HF for the cardiac activity in time and frequency domain; Startle Amp, Rise Time, #Startle, SD StartleRMS, Phasic Value related to skin conductance; Attention, Meditation, Alpha1, Alpha2, Beta1 linked to brain activity). Both physiological features and cortisol levels were normalised with respect to baseline values. There was a high level of correlation between the physiological features and the cortisol trend (Table VIII).

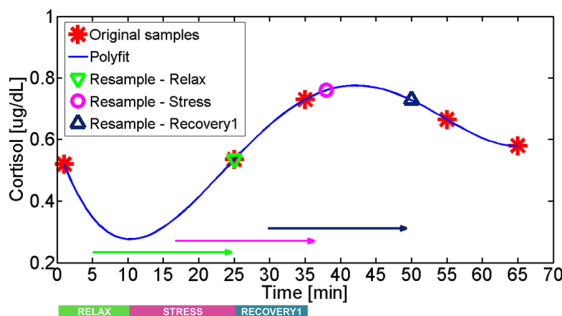


Fig. 5. Change in cortisol levels from before to after a stress response. Red: Value of the original samples collected at the beginning of the relax phase, at the end of the stress phase, at the end of Recovery1, and during and at the end of Recovery2. Blue: Polynomial fitting (degree four) of the original samples. Green: resampled value related to relax phase considering the 20 minutes of delay (response of hormones to a stressor). Violet: resampled value related to stress phase considering the 20 minutes of delay (response of hormones to a stressor). Grey: resampled value related to Recovery1 phase considering the 20 minutes of delay (response of hormones to a stressor).

TABLE VI
FEATURES EXTRACTED FROM HRV, EDA AND EEG SIGNALS: NON-NORMALISED MEAN VALUES \pm STANDARD DEVIATIONS AND SIGNIFICANCE IN DISTINGUISHING RELAX FROM STRESS CONDITIONS.

Features	Source	Relax	Stress	Recovery	P-value
IBI Mean (s)	HRV	820.5 \pm 95.4	739.7 \pm 94.8	824.8 \pm 87.6	0.037*
SDNN (s)	HRV	50.3 \pm 17.3	63.3 \pm 21.4	58.8 \pm 19.9	0.022*
HR Mean (bpm)	HRV	74.4 \pm 8.0	83.3 \pm 9.8	73.9 \pm 6.9	0.013*
SDHR (bpm)	HRV	4.6 \pm 1.3	6.9 \pm 1.8	5.3 \pm 1.2	0.000*
RMSSD (s)	HRV	34.0 \pm 15.3	33.7 \pm 17.3	37.6 \pm 18.6	0.355
pNN50 (%)	HRV	15.1 \pm 14.8	12.9 \pm 11.7	16.0 \pm 16.1	0.563
Alpha Ratio	HRV	2.5 \pm 1.5	3.3 \pm 1.4	3.3 \pm 1.6	0.073
VLF Power (s ² ·10 ⁸)	HRV	2.1 \pm 1.8	7.8 \pm 7.0·E8	4.0 \pm 1.9·E8	0.000*
LF Power (s ² ·10 ⁸)	HRV	4.5 \pm 4.7·E8	10.7 \pm 8.6·E8	7.1 \pm 4.6·E8	0.000*
HF Power (s ² ·10 ⁸)	HRV	2.4 \pm 2.5·E8	3.4 \pm 5.0·E8	3.0 \pm 4.2·E8	0.225
%VLF	HRV	23.5 \pm 15.3	35.4 \pm 15.2	39.1 \pm 15.7	0.000*
%LF	HRV	50.4 \pm 17.6	51.4 \pm 9.9	49.9 \pm 10.6	0.908
%HF	HRV	26.1 \pm 18.9	13.2 \pm 9.8	11.0 \pm 9.1	0.000*
VLF Peak (Hz)	HRV	0.024 \pm 0.007	0.020 \pm 0.007	0.021 \pm 0.005	0.204
LF Peak (Hz)	HRV	0.084 \pm 0.017	0.079 \pm 0.014	0.083 \pm 0.008	0.686
HF Peak (Hz)	HRV	0.261 \pm 0.043	0.224 \pm 0.028	0.241 \pm 0.033	0.133
Startle Amp (μ S)	EDA	0.11 \pm 0.10	0.44 \pm 0.67	0.36 \pm 0.75	0.001*
SD Startle Amp (μ S)	EDA	0.084 \pm 0.076	0.43 \pm 0.70	0.43 \pm 0.98	0.018*
Rise Rate (s)	EDA	1.00 \pm 0.39	1.30 \pm 0.33	1.23 \pm 0.36	0.018*
SD Rise Rate (s)	EDA	0.40 \pm 0.29	0.54 \pm 0.22	0.57 \pm 0.26	0.563
Decay Rate (s)	EDA	1.52 \pm 0.80	1.41 \pm 0.40	1.30 \pm 0.50	0.686
SD Decay Rate (s)	EDA	0.68 \pm 0.41	0.73 \pm 0.28	0.81 \pm 0.33	0.686
Startle Time (s)	EDA	2.52 \pm 0.95	2.71 \pm 0.66	2.53 \pm 0.79	0.563
SD Startle Time (s)	EDA	0.82 \pm 0.52	0.93 \pm 0.30	1.10 \pm 0.41	0.682
Phasic Value (μ S)	EDA	0.011 \pm 0.024	-0.012 \pm 0.028	0.002 \pm 0.004	0.024*
SD Phasic Value (μ S)	EDA	0.369 \pm 0.535	0.286 \pm 0.353	0.28 \pm 0.45	1.000
Sum Fall Time (s)	EDA	13.6 \pm 11.3	19.0 \pm 12.7	13.1 \pm 7.9	0.028*
Sum Rise Time (s)	EDA	11.4 \pm 10.2	17.8 \pm 11.8	12.3 \pm 7.7	0.003*
# Startles	EDA	10.7 \pm 10.1	13.3 \pm 8.3	9.8 \pm 5.7	0.018*
RMS overall (μ S)	EDA	0.371 \pm 0.534	0.288 \pm 0.354	0.278 \pm 0.449	1.000
Startle RMS (μ S)	EDA	0.116 \pm 0.213	0.108 \pm 0.167	0.080 \pm 0.099	0.105
SD Startle RMS (μ S)	EDA	0.022 \pm 0.025	0.103 \pm 0.241	0.077 \pm 0.097	0.021*
Attention	EEG	50.27 \pm 6.92	45.41 \pm 6.94	48.19 \pm 6.83	0.033*
Meditation	EEG	61.12 \pm 8.41	53.06 \pm 6.30	53.97 \pm 4.95	0.006*
Alpha 1	EEG	34.2 \pm 24.9	36.4 \pm 7.5	28.9 \pm 11.4	0.024*
Alpha 2	EEG	23.0 \pm 10.6	26.5 \pm 5.7	22.5 \pm 9.8	0.004*
Beta 1	EEG	14.6 \pm 5.2	21.3 \pm 4.9	18.3 \pm 7.7	0.000*
Beta 2	EEG	23.8 \pm 24.0	18.2 \pm 3.8	17.1 \pm 7.9	0.015*

* Significant difference between groups (p < 0.05)

V. DISCUSSION

The aim of the current paper was to develop and investigate the use of a wearable sensor system for stress detection in response to physical and mental stress induction. During the experimental test, subjects were monitored using three wearable devices that measured HRV, EDA and EEG. The analysis of these physiological signals allowed the extraction of parameters whose variation was highly correlated with changes in different mental/physical conditions. The data acquired have led to satisfactory results, and thus this work can be considered a promising pilot study in which the feasibility of the system is demonstrated. In fact, this study was able to qualitatively (PCA) and objectively (SVM) distinguish the different stages of the test carried out by the subjects. In accordance with the literature

[22], the accuracy of wearable sensors and the choice of the analysed signals allowed the identification of significant changes in physiological parameters between the Relax and Stress phases. The main innovation of this study is the verification of how these changes are actually linked to changes in stress level. The high degree of correlation between cortisol, as a validated biomarker of stress, and the selected physiological parameters confirms the efficacy of the stress detection system.

Additionally, this work also took into account the Recovery1 phase, which does not overlap with the Relax phase, as shown by the relative cluster in the PCA analysis. The presence of a third cluster, visually located between Relax and Stress, opens the possibility to obtain a scale of different levels of stress. The ability to detect transient stress levels will allow the user to be

warned before the level exceeds an alarm threshold. The significant difference observed here between the values of Relax and Recovery could arise from the double nature of the MAST. In fact, the combination of physical and cognitive components dilates the actual time in which the subject returns to the initial conditions.

The MAST protocol's ability to induce stress was previously demonstrated in a study conducted by Smeets [13]. The cortisol curve trend achieved while analysing our 12 subjects was comparable with that reported by Smeets. The combination of cognitive and physical tasks proposed in this study included several modifications, such as an increase of the water temperature from 2°C to 5°C. The temperature was modified due to the Italian population having a different sensitivity to the cold than the northern European population involved in the original study. Despite these variations, the salivary cortisol concentration curve is comparable to the one obtained at the University of Maastricht with the original protocol. This similarity confirms the ability of the test to induce stress in the users even if a low number of subjects were involved in the trial due to its complexity and the long analysis time dictated by biological samples [22].

Despite the data processing of this work was offline performed, some considerations about the concrete use in real situations, particularly with a real time approach, could be already discussed. Due to nature of the used signals, specific requirements should be included in the implementation of algorithms. For example, recent studies demonstrated that reducing the window duration for HRV analysis below 4 min negatively affects the outcomes [57]. Based on this assumption, a 5-minute window with a 50% of overlapping was chosen to extrapolate stress parameters, thus automatically providing an output value in range 1-3 (1: relaxed, 2: normal, 3: stressed) every 2.5 minutes. On the other hand, the computational load was not a problem, since the time required to get an output from the algorithm is limited to 0.7 seconds.

VI. CONCLUSION

We proposed this study to demonstrate that a system, developed by integrating commercial and wearable devices, was effectively able to detect a stress reaction in subjects that underwent the MAST, a quick and non-invasive test shown to elicit robust sympathetic and glucocorticoid stress responses. This study involved both the use of physiological wearable sensors and salivary cortisol analysis, which is considered a reliable biological marker of stress. The data processing performed in this study allowed to identify an optimized set of physiological features able to detect stress. The comparison of the information gained from sensors and salivary samples has been established, producing very interesting results.

The MAST protocol provoked a realistic stress reaction in the subjects, which was detected and evidenced by the variation of several salient physiological features. This variation was also effectively reflected in the cortisol trend, considered as reliable marker, and was usable as a reference for the induction of stress. The correlation calculated between physiological responses and cortisol values was significant and positive.

TABLE VII
MEAN AND STANDARD DEVIATION OF CORTISOL CONCENTRATION IN SALIVARY SAMPLES

	Mean [nmoli/l]	Std Dev [nmoli/l]
tPre	18.08	9.29
t0	20.45	6.68
t10	31.18	14.28
t30	24.88	7.85
t40	18.87	7.27

TABLE VIII
CORRELATION ANALYSIS BETWEEN CORTISOL LEVELS AND PHYSIOLOGICAL PARAMETERS

R ²	0.714
Multiple R	0.845
p-value	0.021

Furthermore, the wearable sensors and the physiological features also reflected the recovery status of the subjects, demonstrating that the variation shown was effectively due to the cognitive and physical stress status.

Since the correlation with a reliable biological marker was satisfying and the capability of the system for stress detection has been demonstrated, future works will include improvements in features extraction and classification algorithms and additional experiments for testing more accurate stress models in real scenarios. In order to obtain a system able to detect stress levels in real time for immediate action to be taken to prevent chronic stress and health consequences, the algorithms will be also translated into an appropriate language to ensure its use on portable devices, thus proving a feedback related to stress in real environments and applications.

ACKNOWLEDGMENT

The authors are deeply grateful to Dr. Tom Smeets, from the Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands, who kindly provided them with PowerPoint presentations with specific instructions for the MAST. This study has been partially supported by the European Community within the TransSafe Project (Sixth call of the Ambient Assisted Living Joint Programme (AAL JP) - AAL-6-2013-64) and by the Italian Ministry of Health within the Current Research Program performed at National Research Institutes (IRCCS).

REFERENCES

- [1]. C. Raymond.(2000). Stress the real millennium bug. *Stress News*, 12(4). <http://www.isma.org.uk/stressnw/millbug.htm#> (accessed on 30 October 2016).
- [2]. Statistic Brain (2015). *Stress Statistics*. Available online: <http://www.statisticbrain.com/stress-statistics> (accessed on 30 October 2016).
- [3]. European Foundation for the Improvement of Living and Working ConditionsEuropean (2010). *Working Condition Survey*, <http://www.eurofound.europa.eu/surveys/ewcs/2010/datatables> (accessed on 30 October 2016).
- [4]. R. Blaug Kenyon & R. Lekhit, (2007). *Stress at work: A report prepared for the work foundation's principal partners*.

> REPLACE THIS LINE WITH YOUR PAPER IDENTIFICATION NUMBER (DOUBLE-CLICK HERE TO EDIT) < 11

- [5]. J. G. Rodrigues et al, "A mobile sensing approach to stress detection and memory activation for public bus drivers". *IEEE Trans. Intell. Transp. Syst.*, vol. 16(6), pp. 3294-3303, 2015.
- [6]. M. Zecca et al, "Using the Waseda Bioinstrumentation System WB-1R to analyze Surgeon's performance during laparoscopy-towards the development of a global performance index". In *2007 IEEE/RSJ IROS*, pp. 1272-1277, Oct. 2007.
- [7]. H. T. Peng et al, "Performance evaluation of a salivary amylase biosensor for stress assessment in military field research". *J Clin Lab Anal.*, vol. 30(3), pp. 223-230, 2016.
- [8]. P. Li et al, "A smart safety helmet using IMU and EEG sensors for worker fatigue detection". *International Symposium on Robotic and Sensors Environments (ROSE)*, pp. 55-60, Oct. 2014.
- [9]. A. Alberdi et al, "Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review.". *J Biomed Inform.*, vol. 59, pp. 49-75, 2016.
- [10]. American Psychological Association (2015). *Stress in America: Paying With Our Health*. Available: <https://www.apa.org/news/press/releases/stress/2014/stress-report.pdf> (accessed on 30 October 2016).
- [11]. E. R. De Kloet et al, "Stress and the brain: From adaptation to disease". *Nat Rev Neurosci.*, vol. 6(6), pp. 463-475, 2005.
- [12]. Y. M. Ulrich-Lai et al, "Neural regulation of endocrine and autonomic stress responses". *Nat Rev Neurosci.*, vol. 10(6), pp. 397-409, 2009.
- [13]. T. Smeets et al, "Introducing the Maastricht Acute Stress Test (MAST): A quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses". *Psychoneuroendocrinology*, vol. 37(12), pp. 1998-2008, 2012.
- [14]. S. K. Drost, et al. "Corticosterone levels in the brain show a distinct ultradian rhythm but a delayed response to forced swim stress". *Endocrinology*, vol. 149(7), pp. 3244-3253, 2008.
- [15]. W. B. Cannon, "The emergency function of the adrenal medulla in pain and the major emotions". *Am J Physiol Cell Physiol*, vol. 33(2), pp. 356-372, 2014.
- [16]. B. S. McEwen, "Stress, adaptation, and disease: Allostasis and allostatic load". *Ann N Y Acad Sci.*, vol. 840(1), pp. 33-44, 1988.
- [17]. T. Smeets, "Autonomic and hypothalamic-pituitary-adrenal stress resilience: Impact of cardiac vagal tone". *Biol Psychol*, vol. 84(2), pp. 290-295, 2010.
- [18]. B. S. McEwen, "Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators". *Eur J Pharmacol*, vol. 583(2), pp. 174-185, 2008.
- [19]. Acerbi, Giorgia, et al. "A Wearable System for Stress Detection Through Physiological Data Analysis." *Italian Forum of Ambient Assisted Living*. Springer, Cham, 2016.
- [20]. K.A Brookhuis & D. deWaard, "Monitoring drivers' mental workload in driving simulators using physiological measures". *Accid Anal Prev*; vol. 42: pp. 898-903, 2010.
- [21]. A. Di Domenico & M.A. Nussbaum, "Effects of different physical workload parameters on mental workload and performance". *Int J Ind Ergonom.* vol. 41, pp. 255-260, 2011
- [22]. N. Sharma & T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey". *Comput Methods Programs Biomed.* vol. 108(3), pp. 1287-1301, 2012.
- [23]. T. M. Marteau & H. Bekker, "The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI)". *Br J Clin Psychol.* vol. 31(3), pp. 301-306, 1992.
- [24]. W. S. Helton. "Validation of a short stress state questionnaire". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications, 2004, pp. 1238-1242.
- [25]. S. Cohen et al, "A global measure of perceived stress". *J Health Soc Behav.* pp. 385-396, 1983.
- [26]. I. Ulstein et al, "High score on the Relative Stress Scale, a marker of possible psychiatric disorder in family carers of patients with dementia". *Int J Geriatr Psychiatry*, vol. 22(3), pp. 195-202, 2007.
- [27]. D. G. Elmes & E. B. Zechmeister, "Research methods in psychology". Nelson Education, 2011. *Research methods in psychology*
- [28]. H. M. Burke et al, "Depression and cortisol responses to psychological stress: A meta-analysis". *Psychoneuroendocrinology*, vol. 30(9), pp. 846-856, 2005.
- [29]. D. A. Padgett & R. Glaser, "How stress influences the immune response". *Trends Immunol.*, vol. 24(8), pp. 444-448, 2003.
- [30]. M. Yamaguchi & J. Sakakima, "Evaluation of driver stress in a motor-vehicle driving simulator using a biochemical marker". *J Int Med Res.* vol. 35(1), pp. 91-100, 2007.
- [31]. H. T. Peng, et al, "Performance evaluation of a salivary amylase biosensor for stress assessment in military field research". *J Clin Lab Anal.*, vol. 30(3), pp. 223-230, 2016.
- [32]. S. Khalfa et al, "Effects of relaxing music on salivary cortisol level after psychological stress". *Ann N Y Acad Sci.*, vol. 999(1), pp. 374-376, 2003.
- [33]. J. Bakker et al, "What's your current stress level? Detection of stress patterns from GSR sensor data". In: *2011 IEEE 11th ICDMW*. IEEE, p. 573-580, 2011.
- [34]. J. Zhai & A. Barreto, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables" In: *Conf Proc IEEE Eng Med Biol Soc.* pp. 1355-1358, 2006.
- [35]. R. R. Singh et al, "A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals". *Biomed Signal Process Control*, vol. 8(6), pp. 740-754, 2013.
- [36]. Y. Shi et al. "Galvanic skin response (GSR) as an index of cognitive load". In: *CHI'07 extended abstracts on Human factors in computing systems*. ACM, pp. 2651-2656, 2007.
- [37]. L. L. Watkins et al, "Anxiety reduces baroreflex cardiac control in older adults with major depression". *Psychosom Med.* vol. 61(3), pp. 334-340, 1999.
- [38]. J. Taelman et al, "Influence of mental stress on heart rate and heart rate variability". In: *4th European conference of the IFMBE*. Springer Berlin Heidelberg, pp. 1366-1369, 2009.
- [39]. R. Orsila et al, "Perceived mental stress and reactions in heart rate variability—a pilot study among employees of an electronics company". *Int J Occup Saf Ergon.* vol. 14(3), pp. 275-283, 2008.
- [40]. J. A. Healey & R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors". *IEEE Trans. Intell. Transp. Syst.* vol. 6(2), pp. 156-166, 2005.
- [41]. E. M. Holczerberger et al, "Electroencephalographic coherences during emotion identification task", 2012.
- [42]. N. H. A. Hamid et al, "Evaluation of human stress using EEG power spectrum". In: *Signal Processing and Its Applications (CSPA), 2010 6th International Colloquium on*. IEEE, pp. 1-4, 2010.
- [43]. D. S. Lee, et al. "Stress Events Detection of Driver by Wearable Glove System." *IEEE Sensors Journal* 17.1 (2017): 194-204.
- [44]. V. Sandulescu et al. "Stress detection using wearable physiological sensors." *International Work-Conference on the Interplay Between Natural and Artificial Computation*. Springer, Cham, 2015.
- [45]. Mozos, Oscar Martinez, et al. "Stress detection using wearable physiological and sociometric sensors." *International journal of neural systems* 27.02 (2017): 1650041.
- [46]. Han, Lu, et al. "Detecting work-related stress with a wearable device." *Computers in Industry* 90 (2017): 42-49.
- [47]. Qianli Xu et al. "Cluster-based analysis for personalized stress evaluation using physiological signals." *IEEE journal of biomedical and health informatics* 19.1 (2015): 275-281.
- [48]. W. S. Liew et al, "Classifying stress from heart rate variability using salivary biomarkers as reference", 2015
- [49]. A. Levine et al, "Measuring cortisol in human psychobiological studies". *Physiol. Behav.* vol. 90(1), pp. 43-53, 2007.
- [50]. M. Benedek, & C. Kaernbach, "A continuous measure of phasic electrodermal activity". *J. Neurosci. Methods*, Vol. 190, pp. 80-91, 2010.
- [51]. W. Boucsein, "Electrodermal activity". Springer Science & Business Media, 2012.
- [52]. J Schumm et al. "Effect of movements on the electrodermal response after a startle event". In: *2008 Int Conf Pervasive Comput Technol Healthc*. IEEE, pp. 315-318, 2008.
- [53]. N. E. A. Lippman et al, "Comparison of methods for removal of ectopy in measurement of heart rate variability". *Am J Physiol Heart Circ Physiol*, vol. 267(1), pp. H411-H418, 1994.
- [54]. B. Mali et al, "Matlab-based tool for ECG and HRV analysis". *Biomed Signal Process Control*, vol. 10, pp. 108-116, 2014.
- [55]. U. R. Acharya et al "Heart rate variability: A review". *Med Biol Eng Comput.* vol. 44(12), pp. 1031-1051, 2006.
- [56]. MindWave User Guide, Available: http://developer.neurosky.com/docs/lib/exe/fetch.php?media=mindwave_user_guide_en.pdf, p.14. (accessed on 30 October 2016).
- [57]. N. Bourdillon et al. "Minimal Window Duration for Accurate HRV Recording in Athletes", *Frontiers in neuroscience* 11 (2017).



Stefano Betti received MSc in Biomedical Engineering in 2013, University of Pisa, Italy. Currently he is research assistant at the BioRobotics Institute of Scuola Superiore Sant'Anna, Italy. His main research interests are in the fields of assistive technology, wearable sensors,

physiological signals analysis, stress and emotion detection, assistive robotics, home automation and smart appliance.



Raffaele Molino Lova is a Cardiologist and Cardiovascular Surgeon and has been involved in the clinical management of the Cardiac Rehabilitation Unit of the Fondazione Don Gnocchi, Centre of Florence, where are admitted older persons recently undergone cardiac surgery. At

the moment he is Senior Researcher and his research interest is in the field of exercise physiology in older persons, in gas exchange kinetics and in the energy cost of walking. He is co-author of more than 70 ISI peer-reviewed papers and has been acknowledged by the Italian Ministry of Education as Full Professor in Rehabilitation Medicine on a national basis.



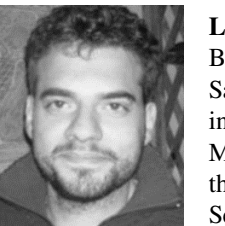
Erika Rovini received MSc in Biomedical Engineering in 2011, University of Pisa, Italy. From 2011 to 2014, she was Assistant Researcher and currently she is PhD student at the BioRobotics Institute, Scuola Superiore Sant'Anna, Italy. Her research interests include biomedical signal processing,

pattern recognition, ambient assisted living and wearable sensors.



Giorgia Acerbi received MSc in Biomedical Engineering in 2011 at University of Pisa, Italy. Currently she is PhD student at the BioRobotics Institute of Scuola Superiore Sant'Anna, Italy, focusing on novel models and approached of acceptability of technology and robotic services,

integrating physiological, psychological and sociological factors.



Luca Santarelli, is research assistant at the BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy. He received a BSc in Biomedical Engineering in 2007 and MSc in Electronics Engineering in 2013 at the University of Pisa, Italy. From

September 2013 to August 2016 he has been PhD student in the department of Physics at UCL, London (UK) and fellow of the EU FP7 project "Initial Training

Network CONTEST (Collaborative Network for Training in Electronic Skin Technology)" as a Marie Curie Early Stage researcher (ESr). During this period he has been mainly developing conjugated polymer-based flexible and stretchable electronic devices. As research assistant, his main research interest is the development of wearable movement and physiological sensors.



Manuela Cabiati is a researcher of Biochemistry and Molecular Biology Laboratory at Institute of Clinical Physiology-Italian National Research Council (CNR). She is devoted to biomedical technologies and experimental strategies applied to cardiometabolic fields. In the evaluation

of biomarkers, the molecular medicine approach is used: expression, production, release and function mechanism are studied by application of transcriptomic approach in addition to classical immunometric determinations. She has published over 50 peer-review publications (H-index 12).



Silvia Del Ry is the research director of the Biochemistry and Molecular Biology Laboratory at Institute of Clinical Physiology-Italian National Research Council (CNR). Her research activity is devoted to biomedical technologies and experimental strategies applied to

cardiovascular, cardiometabolic and oncology fields. She has been working in discovery and validation of serum and tissue biomarkers by high-throughput analysis for peripheral and tissue level determination and by advanced Real-Time PCR analysis for evaluation of mRNA expression producing more than 150 peer-review publications and an H-index of 27. In the evaluation of biomarkers, the molecular medicine approach is used: expression, production, release and function mechanism are studied by application of transcriptomic approach in addition to classical immunometric determinations.



Filippo Cavallo, MScEE, Phd in Bioengineering, is Assistant Professor at BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy, focusing on cloud and social robotics, wireless and wearable sensor systems, biomedical processing, acceptability and AAL roadmapping. He participated

in various National and European projects. He was visiting researcher at the the EndoCAS Center of Excellence, Pisa; at the Takanishi Lab, Waseda University, Tokyo; at Tecnalia Research Center, Spain. He was granted from the ISRR Committee as Fellowship Winner for best PhD thesis in Robotics; from the ACCESS-IT 2009 for the Good Practice Label in Alzheimer Project; from the Well-Tech Award for Quality of Life with the Robot-Era Project. He is author of various papers on conferences and ISI journals.