Journal of Ambient Intelligence and Humanized Computing (2021) 12:4409–4418
https://doi.org/10.1007/s12652-019-01571-0

ORIGINAL RESEARCH

Multilevel assessment of mental stress via network physiology paradigm using consumer wearable devices

Matteo Zanetti[1] · Teruhiro Mizumoto[2] · Luca Faes[3] · Alberto Fornaser[1] · Mariolino De Cecco[1] · Luca Maule[1] · Martina Valente[4] · Giandomenico Nollo[1]

Received: 7 November 2018 / Accepted: 6 November 2019 / Published online: 14 November 2019

Abstract

Mental stress is a physiological condition that has a strong negative impact on the quality of life, affecting both the physical and the mental health. For such a reason, accurate measurements of stress level can be helpful to provide mechanisms for prevention and treatment. This paper proposes a procedure for the classification of different mental stress levels by using physiological signals provided by low invasive wearable devices. 17 healthy volunteers participated in this study. Three different mental states were elicited in them: a resting condition, a stressful cognitive state, and a sustained attention task. The acquired physiological signals were: a one lead electrocardiogram (ECG), a respiratory signal, a blood volume pulse (BVP), and 14 channels of a 10–20 electroencephalogram (EEG). For all subjects, 59 time series of 300 samples each were structured by including the RR series, the respiratory series, the pulse arrival time (PAT) series, and the delta, theta, alpha, beta power series of the 14 EEG channels. Different classifiers were implemented to assess the mental stress level starting from a pool of 3481 features computed from the aforementioned physiological quantities, using the Network Physiology paradigm. The highest achieved accuracy was 84.6%, from logistic regression and random forest classifiers, cross validated by mean of leave-one-person-out analysis. A further analysis was carried out to evaluate the classification accuracy using only cardio-respiratory signals, since the latter are more suitable to be used in real-life scenarios. In this case, the highest achieved accuracy was 76.5% obtained by the random forest classifier.

Keywords Stress assessment · Wearable devices · Network Physiology · Measurements · Classification · Machine learning

1 Introduction

Stress is a key resource that can make the difference between life and death (Selye 1974). However, dangerous situations for health can happen, if stress mechanisms activate in an useless manner and for a long time. Indeed, an unhealthy level of stress is a direct cause of diseases and disorders (Cohen et al. 2007; Kemeny 2003), such as sleep disorder, difficulty in concentration and decision, short-term memory loss, altered mood, depression and anxiety, inflammation and cardiovascular problems. Stress is considered one of the most serious social problem in today's society for its high social cost (Hassard et al. 2017). For the aforementioned reasons, accurate measurements of stress levels are necessary to apply mechanism for prevention and treatment.

Different areas are interested in mental stress assessment, such as the ones related to cardiovascular risk (Vaccarino et al. 2018; Esler 2017; Curtis and O'Keefe 2002), exercises to reduce stress level (Eda et al. 2017; Gauche et al. 2017),

✉ Matteo Zanetti
matteo.zanetti@unitn.it

Mariolino De Cecco
mariolino.dececco@unitn.it

Giandomenico Nollo
giandomenico.nollo@unitn.it

1 Department of Industrial Enginnerig, University of Trento, via Sommarive, 9, 38123 Povo, TN, Italy

2 Graduate School of Information Science, Nara Institute of Science and Technology (NAIST), 8916-5, Takayama, Ikoma, Nara Prefecture, Japan

3 Department of Engineering, University of Palermo, 90133 Palermo, Italy

4 Center for Mind/Brain Sciences (CIMeC), University of Trento, Corso Bettini, 31, 38068 Rovereto, TN, Italy

work-related stress (Kopyt et al. 2017; Seoane et al. 2014; Zheng et al. 2015), student stress (Vanitha and Krishnan 2016), and environmental stress (Steinheuser et al. 2014). Accurate measurement of stress and effort can be also helpful in ambient assisted living (Calvaresi et al. 2017) scenarios. Indeed, also in this context interesting applications are emerging (Pisoni et al. 2016; Kikhia et al. 2018). This kind of measurements can be helpful for therapists, providing information not directly perceivable by mean of observation.

Stress detection is usually performed acquiring physiological signals. Among the many physiological signals available, the most relevant for stress detection are (Sharma and Gedeon 2012):

- the cortisol levels, usually measured in saliva. The drawback of these kind of measurements is that they are invasive and it is very difficult to obtain a continuous monitoring of such levels;
- the cardiovascular system activity, usually monitored through electrocardiography (ECG), blood volume pulse (BVP) and arterial blood pressure (ABP);
- the respiratory system activity, that is strongly related to cardiovascular system activity;
- the electrodermal activity (EDA), i.e. the electrical conductivity of the skin surface;
- the muscle activity, measured through electromyography (EMG), which measures the level of discharge of the motor nerve fibers that innervate the muscle;
- the brain activity, measured through electroencephalography (EEG).

For what concerns the use of the aforementioned parameters for the detection of mental stress, large differences arise in the literature. These differences are mainly due to aspects such as different protocols and equipment for signal monitoring, in addition to the data analysis performed (Smets et al. 2015). Subhani et al. (2017) considered features extracted from a professional 128 channel EEG to distinguish among 4 different levels of stress. The reached accuracy was 83.4%. Hou et al. (2015) reached an accuracy of 67.1%, 75.2%, and 85.7% in distinguishing respectively among 4, 3, and 2 different levels of stress. They used features extracted from EEG signals obtained from a wireless headset to train a support vector machine (SVM) classifier. In Smets et al. (2015), an analysis of different classification algorithms was performed to distinguish between stressful and non-stressful situations. The acquired physiological signals were ECG, respiration, EDA and temperature. The acquisition of such signals was performed using wireless devices, even if quite invasive, since the recording of the ECG was performed applying electrodes on the skin. A similar setup was used in Huysmans et al. (2018), where unsupervised learning was used to distinguish between relax or stress phases. The authors obtained an accuracy of 84.6% using personalized dynamic Bayesian networks and an accuracy of 82.7% using generalized support vector machines (SVM). Sandulescu et al. (2015) used wearable devices to monitor EDA and pulse plethysmograph (PPG) signals to detect stressful situations in five participants. The maximum accuracy was 83.08% using an SVM algorithm. Mohino-Herranz et al. (2015) used ECG and thoracic electrical bioimpedance (TEB) signals provided by wearable devices to distinguish between low mental load and mental overload, reaching an accuracy of 67.7%.

Other works in the literature made use of deep learning (LeCun et al. 2015) techniques for detecting mental stress. In Masood and AlGhamdi (2019) a convolutional neural network (CNN) framework was employed to assess the improvement in the classification accuracy adding neural signals to the traditional physiological signals used for stress detection, i.e. heart rate variability (HRV) and EDA. The authors reached an accuracy of 90% in distinguish between stress and non-stress situations. Vuppalapati et al. (2018) used EEG features to distinguish between 4 different levels of stress, reaching an accuracy of 83.43%. As the authors claimed, their accuracy was dependant on the accuracy of the machine learning model used and its datasets. In Jaques et al. (2017), deep learning techniques were used to implement a mood prediction system. In particular, the authors demonstrate how personalized models can provide substantial performance enhancements. Finally, a survey on machine learning techniques for stress detection can be found in Panicker and Gayathri (2019).

In this paper, we try to develop a model capable to distinguish among 3 different mental stress levels among 17 different subjects. With respect to the previous presented works, the novelty of our approach consists in using the new paradigm of Network Physiology (Bashan et al. 2012) to perform stress detection. In this approach, each organ system is seen as a node of a complex network of physiological dynamical interactions. Using the Network Physiology, we overcome the traditional, reductionist approach, in which the function of a single organ is studied in isolation. The considered systems are studied by looking at the coupling among their output signals. We try to quantify such physiological interactions, using information theory quantities, in order to distinct among different mental stress levels. We start from the framework described in Zanetti et al. (2018), where the Network Physiology paradigm was used to distinguish between stressful and non-stressful situations in one single subject. The novelty of this work with respect to Zanetti et al. (2018) consists in the increasing number of states to be distinguished, i.e. 3 vs 2, and the development of an inter-subject model. Indeed, in Zanetti et al. (2018) the procedure was only tested with one single participant.

2 System and hardware configuration

The acquisition of the physiological signals was performed using low invasive and consumer wearable devices. A sensorized t-shirt, by Smartex[1], provides the ECG and the respiratory signal at a sampling frequency of respectively 250 Hz and 25 Hz. The respiratory signal is acquired through a piezoresistive sensor situated at the level of the ribcage. A wristband, by Empatica[2], provides the BVP signals at a sampling rate of 64 Hz. The EEG signals were acquired using the 14 channels Emotiv[3] EPOC PLUS wireless headset (international 10–20 locations), which has a sampling frequency of 256 Hz for every channel.

In order to obtain accurate vital signs acquisition, it is important to wear these devices correctly. In particular, the Smartex t-shirt must be of the right size to provide a good contact of the skin with the ECG electrodes and not to have the piezoresistive sensor too much stretched or loose. The Empatica wristband must be wear not uncomfortably tight, but snugly enough to prevent bad illumination conditions caused the dispersion of the light from the PPG sensor on wrist skin. Particular attention must be also paid to the correct positioning of the EEG electrodes of the Emotiv headset. Anyhow, thanks to the fixed configuration and robustness of the hardware solutions, the setup time of the entire system can be achieved in less than 5–10 min per participant. All devices are connected to the same PC via Bluetooth.

2.1 Synchronization of the devices

The main issue in the combination of multiple *independent* devices is the lack of a hardware driven synchronization method. The data must then be managed and analyzed, devising software solutions to perform the temporal alignment of the various signals. That is critical since errors could occur in the generation of the clock of the electronics, thus potentially affecting the processing with temporal shifts in the recorded data. The resulting desynchronization must be avoided as it impairs the study of interactions between signals that underlies the concept of Network Physiology. Such issue was here solved by running a custom designed synchronization method that foresees the usage of the quantity that is available from all devices: the acceleration.

The process can be subdivided into the following steps:

(a) Used wearable devices connected to a rigid bar for time synchronization.
(b) A participant wearing the wearable devices.

Fig. 1 Wearable devices used for physiological signal acquisition: (A) Empatica E4; (B) Emotiv EPOC PLUS; (C) Smartex

1. the identification of the principal motion directions for each device;
2. the alignment and fastening of devices to a rigid support: the industrial Velcro achieved very good performances both in term of stability of the mount and removability capabilities, Fig. 1;
3. the motion of the rigid support (together with the sensors) in order to define a non uniform acceleration pattern: a sinusoidal path is suggested since it is periodic and easy to be performed;
4. the synchronization of the collected, low-pass filtered, acceleration signals with the one used as reference ($(a(t)_r)$).

The last two are performed both at the beginning and at the end of the recording sessions. This is fundamental to compensate any modifying factor that can cause dilatations of the time bases. The synchronization is performed as a linear warping of the time with respect to a reference signal (Fig. 2), in this case the one provided by the Smartex sensor. Equation (1) reports the formulation, where $t_{r,n}^{f,i}$ stands respectively for a time instant t collected from the *r*eference or *n*th series, aligned at the *i*nitial or *f*inal phase of the data record. The modified temporal instant $\tilde{t_n}$ can be computed as:

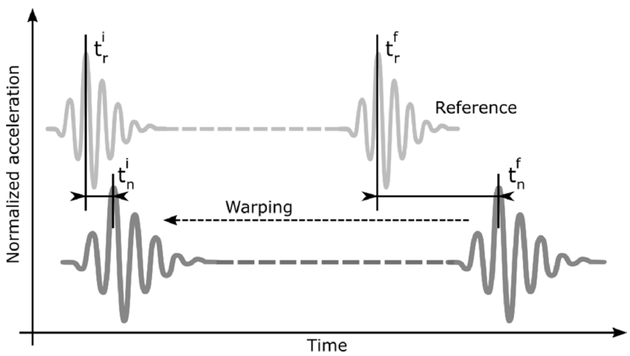
$$\tilde{t_n} = \frac{\left(t_r^f - t_r^i\right)}{t_n^f - t_n^i} \times \left(t_n - t_n^i\right) + t_r^i. \tag{1}$$

Alternative quantities than the acceleration can be considered, the method is general and can be adapted accordingly to the required hardware configuration with no major modifications.

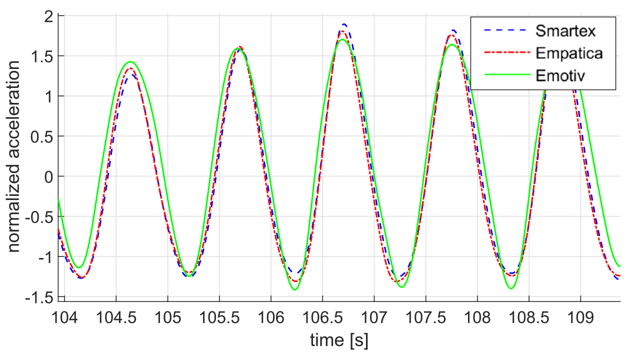
[1] http://www.smartex.it

[2] https://www.empatica.com

[3] https://www.emotiv.com



(a) Synchronization scheme.



(b) Example of synchronized data from the considered hardware architecture.

Fig. 2 Temporal synchronization of the wearable devices by mean of the acceleration signals

3 Experimental protocol

17 healthy participants, with and age ranging between 18 and 30, were monitored. The recording sessions were conducted between 10.30 and 12.00 a.m. to avoid possible differences due to the time of the day. The participants were seated in front of a PC in a comfortable room at constant illumination and were instructed to not speak and to limit their movements during the test.

Three different levels of stress were induced to the participants. The first was a rest condition induced watching a relaxing video. The second was induced playing a serious game, which consisted in following a point moving on the screen using the mouse and trying to avoid some obstacles. The third was obtained through a mental arithmetic task using an online tool: participants had to perform sums and subtractions of 3-digit number and write the solution in a text-box using the keyboard. Each participant performed 2 recording sessions: one for the mental arithmetic task and one playing the serious game. Each recording session was structured in this manner:

- rest (12 min);
- mental arithmetic/serious game (7 min);
- rest (12 min).

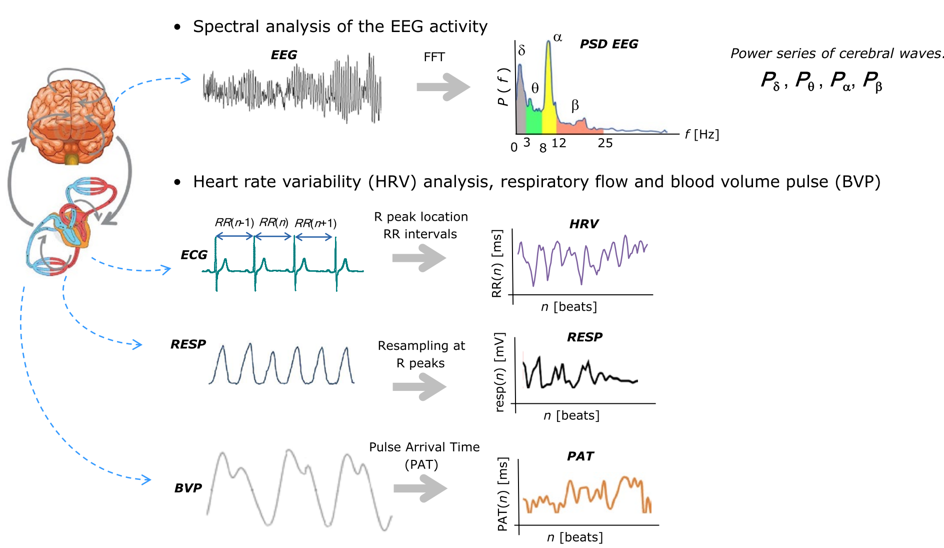
No pen and paper or other supports were allowed. Also finger counting was discouraged.

4 Data processing

The data was analyzed offline using MATLAB and following the procedure described in Zanetti et al. (2018). Figure 3 shows a schematic representation of the analysis performed on the acquired signals.

The R peaks in the ECG were detected using the template matching algorithm from Speranza et al. (1993), reconstructing in this way the R-R tachogram. The respiratory signal

Fig. 3 Time series extraction procedure from the acquired physiological signals



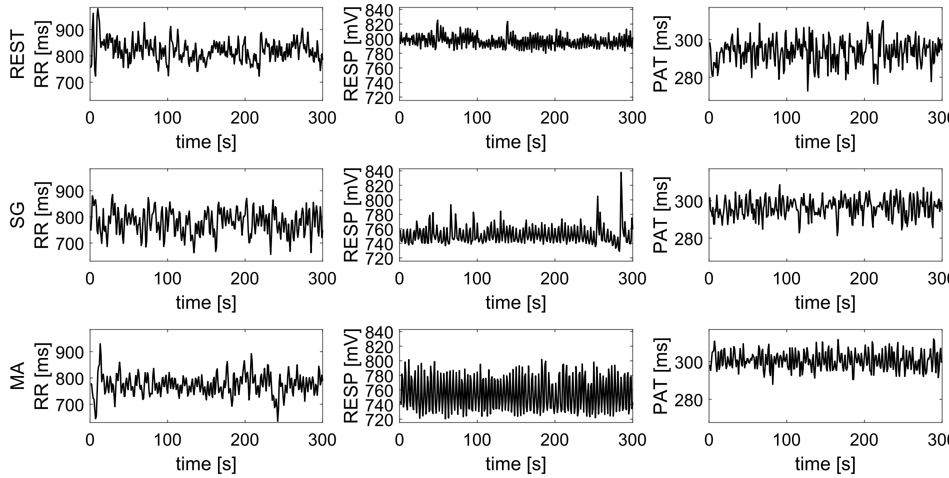


Fig. 4 RR, respiratory and PAT time series during rest (REST), serious game (SG) task, and mental arithmetic (MA)

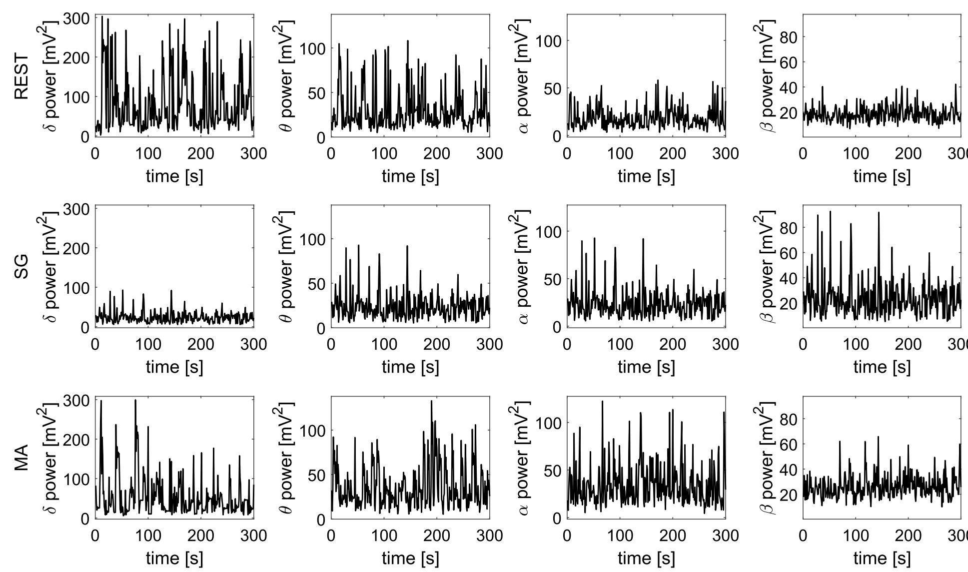


Fig. 5 EEG power series in the δ, θ, α, and β bands during rest (REST), serious game (SG) task, and mental arithmetic (MA) of a recording session for the AF3 electrode

was then resampled accordingly to the timing of the identified R peaks. The pulse arrival time (PAT) was obtained computing the time that elapses between the R peak in the EEG and the corresponding point of maximum derivative in the BVP signal (Orini et al. 2012). Figure 4 shows the RR, the respiratory, and the PAT time series for one subject during the three different mental stress levels. The time series were resampled at 1 Hz.

For what concerns the EEG, the power spectral density (PSD) in the δ (0.5–3 Hz), θ (3–8 Hz), α (8–12 Hz), β (12–25 Hz) bands was computed using the periodogram. A sliding window of 2 s and a 50% of overlap was used. The MATLAB function *bandpower()* was used to compute the PSD specifying the band of interest and the sampling frequency of the input signal. Figure 5 reports an example of EEG power series of a recording session for the AF3 electrode.

5 Feature extraction

The work follows the approach fostered by network physiology (Bashan et al. 2012), in which each organ system is seen as a node of a complex network of physiological interactions. To investigate these interactions, the proposed method exploits information-theoretic measures starting from the time series computed as reported in Sect. 4. For every signal and for every possible couple, this computes then the self-entropy S_y, the mutual information $I(X;\, Y)$, and the conditional mutual information $I(X;\, Y|Z)$ (Faes et al. 2016, 2017).

5.1 Information-theoretic measures

Given a dynamic process Y, its present sample Y_n and past states $\mathbf{V}_n^Y = [Y_{n-1}, Y_{n-2}, \ldots]$, the amount of information contained in Y_n, which can be predicted by its past, can be computed as follows:

$$S_Y = H(Y_n) - H(Y_n|\mathbf{V}_n^Y), \tag{2}$$

where $H(Y_n)$ is the Shannon entropy, defined as $H(Y_n) = -\sum p(Y_n) \ln p(Y_n)$, and $H(Y_n|\mathbf{V}_n^Y)$ the conditional entropy.

Considering instead two distinct dynamic processes X and Y, the mutual information $I(X_n;Y_n)$ measures the amount of information that can be obtained about the present value of a random variable observing another one, and it is defined as:

$$\begin{aligned} I(X_n;Y_n) &= H(X_n) - H(X_n|Y_n) \\ &= H(Y_n) - H(Y_n|X_n) \\ &= H(X_n) + H(Y_n) - H(X_n, Y_n), \end{aligned} \tag{3}$$

where $H(X_n)$ and $H(Y_n)$ are the marginal entropies, $H(X_n|Y_n)$ and $H(Y_n|X_n)$ are the conditional entropies, and $H(X_n, Y_n)$ the joint entropy.

The conditional mutual information $I(X_n;Y_n|Z_n)$ is instead defined as:

$$\begin{aligned} I(X_n;Y_n|Z_n) &= I(X_n;Y_n, Z_n) - I(X_n;Z_n) \\ &= I(Y_n;X_n, Z_n) - I(Y_n;Z_n). \end{aligned} \tag{4}$$

where $I(X_n;Y_n|Z_n)$ is the expected value of the mutual information between X_n and Y_n, given the value of a third variable Z_n, measuring the fraction of the information shared between X_n and Y_n that is not shared with Z_n.

For the practical computation of the above quantities, under the hypothesis of Gaussian distribution of y, it is possible to apply the formulas described in Barnett et al. (2009) and Porta et al. (2015). For what concerns S_Y, it can be computed as:

$$S_Y = \frac{1}{2} \log \frac{\sigma_Y^2}{\sigma_\epsilon^2}, \tag{5}$$

where σ_Y^2 is the variance of Y and σ_ϵ^2 is the variance of the prediction error ϵ of an Auto Regressive model fitting Y:

$$Y_n = \sum_{i=1}^{p} a_i Y_{p-i} + \epsilon \tag{6}$$

where p is the model order, which is computed using the Akaike information criterion (Schwarz 1978).

Given the covariance Σ and precision Σ^{-1} matrices of X and Y:

$$\Sigma = \begin{bmatrix} \sigma_X^2 & \sigma_{XY}^2 \\ \sigma_{XY}^2 & \sigma_Y^2 \end{bmatrix} \tag{7}$$

$$\Sigma^{-1} = \begin{bmatrix} \gamma_X^2 & \gamma_{XY}^2 \\ \gamma_{XY}^2 & \gamma_Y^2 \end{bmatrix}, \tag{8}$$

$I(X_n;Y_n)$ and $I(X_n;Y_n|Z_n)$ can be computed as (Gelfand and IAglom 1959):

$$I(X_n;Y_n) = -\frac{1}{2} \log \left(1 - \frac{\sigma_{XY}^2}{\sigma_X^2 \sigma_Y^2}\right) \tag{9}$$

$$I(X_n;Y_n|Z_n) - \frac{1}{2} \log \left(1 - \frac{\gamma_{XY}^2}{\gamma_X^2 \gamma_Y^2}\right), \tag{10}$$

where Z_n contains all the variables except X_n and Y_n.

5.2 Application

The experimental testing protocol produced 3 time series from the cardio-respiratory part and 56 (14×4) from the EEG, for a total of 59. These were processed as described in Sect. 5.1 for every possible combination, obtaining 3481 features: 59 from the computation of the self entropy, 1711 from the mutual information and 1711 from the conditional mutual information. To compare the time series among different participants, all extracted features were initially normalized with respect to the baseline resting conditions. Given the three mental states, i.e. rest (REST), mental arithmetic (MA) and serious game (SG) and the feature f_i, for $i = 1, 2, 3, \dots, 3481$, the normalized feature $f_i^{j,*}$, where $j = \{REST, MA, SG\}$, was computed as follows:

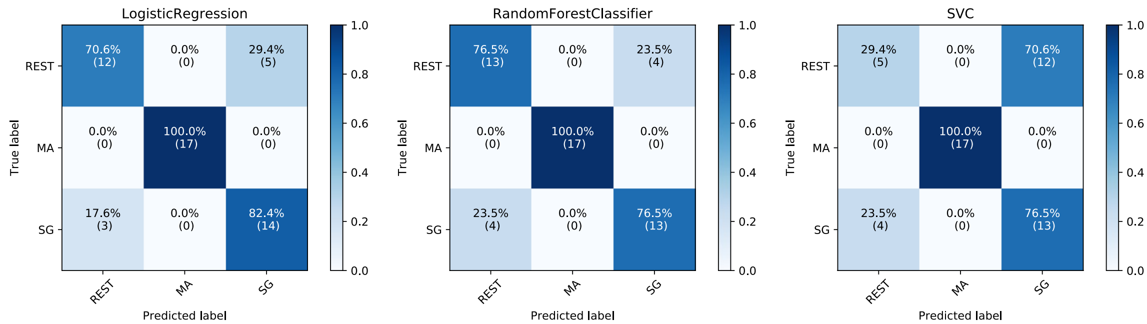
$$f_i^{j,*} = \frac{f_i^j}{f_i^{REST}}. \tag{11}$$

6 Results

Different classification algorithms were tested for the classification of the stress status: (1) support vector classification (SVC), (2) random forest (RF), and (3) logistic regression (LR). The hyper-parameters for each classifier were optimized by a grid search: C, γ, and kernel for SVC, depth and the number of estimators for RF, and C and penalty for LR (Buitinck et al. 2013). A leave-one-person-out cross validation was applied to test the accuracy of the considered classification algorithms.

LR and RF achieved the best classification accuracy, equal to 84.3% and 84.3% respectively, Fig. 6 reports the confusion matrices.

The most remarkable result concerns the classification of the mental arithmetics status: all classifiers correctly classified this task for 100% of the cases, and at the same time other tasks were never misclassified with it. It follows that the feature values for the mental arithmetic strongly characterize the task, making it well distinguishable from others. The outcome is that a heavy mental stress status can be reliably be recognized by the proposed method.



(a) Logistic Regression (C: 1.0, penalty: L2) **(b)** RandomForest (depth: 4, estimators: 8) **(c)** SVC (C: 1.0, γ : 0.0001, kernel: RBF)

Fig. 6 Classification results for different classifiers. The best result was obtained by logistic regression and random forest classifiers, with an accuracy of 84.3%

As for the remaining classes, these present some misclassified results, proof that the considered feature base presents some similarities in these two stress states. However, since the logistic classifier and random forest classifier have correctly recognized about 80% rests and serious-games, this represents a sub-optimal but anyhow sufficiently accurate classification outcome for the applicability of the method.

The SVC with the low classification accuracy wrongly recognized many rests as the serious-game. The soft-margin SVM with RBF kernel allows some examples placed on the wrong side to be ignored based on C parameter, on fitting. Since the classifier with a low C parameter like this classifier ignores many examples placed on the wrong side, it is considered that the classifier was built so that many rest states placed on the serious-game side were ignored, in this result. However, the overall classification accuracy decreased by using a higher C parameter; therefore, it is said that the SVM algorithm is not suitable for this dataset.

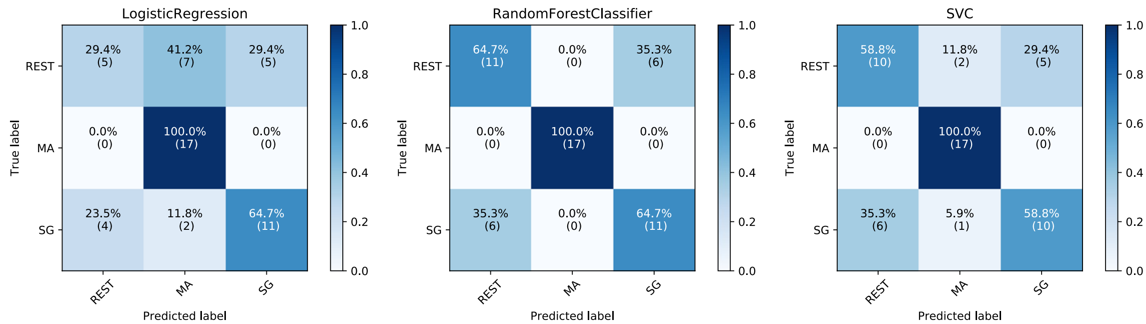
Random Forest algorithm builds a set of decision trees based on feature importance. This can be exploited to investigate what feature is important for classification. Table 1 reports the values, as normalized percentage, of the ten most important features identified by the RF model. Such features count for the 59.3% of the overall feature importance score. The most important features are the ones relative to EEG signals. In particular, 4 features out of 10 are relative to the mutual information shared between pairs of electrodes in which one is positioned in the frontal part and the other in the occipital part of the head. Among the most important features there are also the self-entropies of the electrodes FC6 and T7.

Since the Emotive EPOC is a quite invasive device for a real-life scenario, we tested the accuracy of the classification algorithm using the features provided only from the cardio-respiratory series. Since in this case we would have only 9 features, we added to them more traditional features for stress measurement, i.e. the LF/HF ratio, the mean and the standard deviation of the RR series, the respiratory frequency and the mean of the phasic component (Greco et al. 2016) of the EDA signal, which is provided by the Empatica E4 wristband. In this case the best obtained accuracy was of 76.5% using the RF classifier (Fig. 7).

All classifiers could correctly recognize mental arithmetics even without Emotive EPOC features. However, the logistic regression classifier and the SVC have recognized some rest and serious-game as mental arithmetic. Especially, the LR classifier wrongly recognized about 70% of rest as others and the classification accuracy decreased more 20% than the logistic regression classifier with Emotive EPOC features. Conversely, although the SVC also have recognized several rests and serious-games as mental arithmetic; however, it has correctly recognized rests than the SVC with Emotive EPOC features and the classification accuracy was also improved 5%. Therefore, it is thought that the dataset without Emotive EPOC is suitable for SVC, and is not for logistic regression. The random forest classifier has increased a few numbers of incorrect classification between rest and mental arithmetic; therefore, the classification accuracy has also decreased about 8% than the classifier with Emotive EPOC features. However, the random forest classifier has no any rest and serious-game recognized as mental arithmetic, unlike other classifiers, and has kept enough high classification accuracy even without Emotive EPOC

Table 1 Top 10 feature importance of Random Forest classifier

	Feature	Importance (%)		Feature	Importance (%)
1	S-δ_{FC6}	7.024	6	MI-α_{F8}-β_{AF4}	6.068
2	MI-δ_{F3}-δ_{O1}	7.015	7	MI-β_{AF3}-β_{P8}	5.872
3	S-θ_{T7}	6.937	8	MI-β_{F3}-β_{P8}	5.672
4	MI-α_{FC5}-α_{P8}	6.648	9	MI-δ_{F4}-α_{F8}	4.188
5	MI-α_{P7}-α_{O2}	6.189	10	MI-β_{F7}-θ_{T8}	3.677



(a) Logistic Regression (C: 10.0, penalty: L2)

(b) RandomForest (depth: 6, estimators: 32)

(c) SVC (C: 10.0, γ: 0.1, kernel: RBF)

Fig. 7 Classification results for different classifiers. The best result was obtained by random forest classifiers, with an accuracy of 76.5%

Table 2 Feature importance of Random Forest classifier without Emotive EPOC features

	Feature	Importance (%)		Feature	Importance (%)
1	Mean-RR	24.033	6	Mean-phasicEDA	5.918
2	std-RR	16.624	7	S-RESP	5.487
3	S-RR	8.998	8	MI-RR-RESP	5.123
4	S-PAT	7.369	9	RESPfreq	4.378
5	CMI-RESP-RR	6.971	10	CMI-PAT-RESP	3.547

Table 3 Comparison of our work with respect to similar works in the literature

Study	Acquired signals	Mental states	Accuracy (%)
Proposed	ECG, Resp., BVP, EEG	3	84.6
Masood and AlGhamdi (2019)	EEG, Heart Rate Monitor, EDA	2	90.0
Vuppalapati et al. (2018)	EEG	4	83.43
Huysmans et al. (2018)	ECG, EDA	2	79.0
Subhani et al. (2017)	EEG	4	83.4
Hou et al. (2015)	EEG	4	67.1
Smets et al. (2015)	ECG, Resp., EDA, Temp.	2	84.6
Sandulescu et al. (2015)	EDA, PPG	2	83.08
Mohino-Herranz et al. (2015)	ECG, TEB	2	67.7

features. In both datasets, Random Forest has been the best classifier; in conclusion, it is clear that it is suitable for our recognition.

Table 2 shows the feature importance of the random forest classifier. In this case, the most important features are relative to the ECG signal; i.e. the mean and the standard deviation of the RR series and its self-entropy.

Table 3 shows the comparison of our work with respect to others found in the literature and analyzed in Sect. 1. The framework proposed in this paper falls among the best results. For such a reason, it is possible to claim that the Network Physiology paradigm can be a good framework to detect stressful situations, even among different subjects.

7 Conclusion

The simultaneous recording of ECG, BVP, respiration and EEG signals, provided by wearable devices, was exploited to distinguish between 3 different mental stress states, i.e. rest, sustained attention, and stress, elicited in 17 participants. An approach based on the new field of Network Physiology was used. Information theoretic measures, such as self entropy, mutual information and conditional mutual information, were used to train different classifier algorithms. The best results were obtained by LR and RF classifiers with an accuracy of 84.6%. An accuracy of 76.5% was instead obtained by RF using only features provided by the cardio-respiratory signals. These results are comparable with the ones found in the literature (Table 3).

With respect to the current state of the art, the novelty of our approach consists in using the new approach of Network Physiology on signal acquired from "low-invasive" and consumer wearable devices to distinguish among different levels of mental stress. These results are quite promising and suggest that an inter-subject model using the parameters provided by Network Physiology is feasible. Future development will foresee the improvement of the classification accuracy, using only the devices related to the cardio-respiratory signals for their lower invasiveness. Indeed these devices are more suitable for applications in real-life scenarios.

Acknowledgements This research was developed with the support of the IEEE Smart Cities Initiative-Student Grant Program and AUSILIA project financed by Provincia Autonoma di Trento (2015–2018).

References

Barnett L, Barrett AB, Seth AK (2009) Granger causality and transfer entropy are equivalent for gaussian variables. Phys Rev Lett 103(23):238701

Bashan A, Bartsch RP, Kantelhardt JW, Havlin S, Ivanov PC (2012) Network physiology reveals relations between network topology and physiological function. Nat Commun 3:702

Buitinck L, Louppe G, Blondel M, Pedregosa F, Mueller A, Grisel O, Niculae V, Prettenhofer P, Gramfort A, Grobler J et al (2013) API design for machine learning software: experiences from the scikit-learn project. arXiv preprint arXiv:13090238

Calvaresi D, Cesarini D, Sernani P, Marinoni M, Dragoni AF, Sturm A (2017) Exploring the ambient assisted living domain: a systematic review. J Ambient Intell Human Comput 8(2):239–257

Cohen S, Janicki-Deverts D, Miller GE (2007) Psychological stress and disease. JAMA 298(14):1685–1687

Curtis BM, O'Keefe JH Jr (2002) Autonomic tone as a cardiovascular risk factor: the dangers of chronic fight or flight. In: Mayo Clinic proceedings, Elsevier, vol 77, pp 45–54

Eda N, Ito H, Shimizu K, Suzuki S, Lee E, Akama T (2017) Yoga stretching for improving salivary immune function and mental stress in middle-aged and older adults. J Women Aging 30(3):227–241

Esler M (2017) Mental stress and human cardiovascular disease. Neurosci Biobehav Rev 74:269–276

Faes L, Porta A, Nollo G, Javorka M (2016) Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks. Entropy 19(1):5

Faes L, Porta A, Javorka M, Nollo G (2017) Efficient computation of multiscale entropy over short biomedical time series based on linear state-space models. Complexity 2017:1768264

Gauche R, Lima RM, Myers J, Gadelha AB, Neri SG, Forjaz CL, Vianna LC (2017) Blood pressure reactivity to mental stress is attenuated following resistance exercise in older hypertensive women. Clin Interv Aging 12:793

Gelfand IM, Yaglom A (1959) Calculation of the amount of information about a random function contained in another such function. American Mathematical Society, Providence

Greco A, Valenza G, Lanata A, Scilingo EP, Citi L (2016) CVXEDA: a convex optimization approach to electrodermal activity processing. IEEE Trans Biomed Eng 63(4):797–804

Hassard J, Teoh KR, Visockaite G, Dewe P, Cox T (2017) The cost of work-related stress to society: a systematic review. J Occup Health Psychol 23(1):1–17

Hou X, Liu Y, Sourina O, Tan YRE, Wang L, Mueller-Wittig W (2015) EEG based stress monitoring. In: Systems, man, and cybernetics (SMC), 2015 IEEE international conference on, IEEE, pp 3110–3115

Huysmans D, Smets E, De Raedt W, Van Hoof C, Bogaerts K, Van Diest I, Helic D (2018) Unsupervised learning for mental stress detection-exploration of self-organizing maps. In: 11th International joint conference on biomedical engineering systems and technologies, BIOSTEC 2018, SciTePress

Jaques N, Taylor S, Sano A, Picard R, et al. (2017) Predicting tomorrow's mood, health, and stress level using personalized multitask learning and domain adaptation. In: IJCAI 2017 Workshop on artificial intelligence in affective computing, pp 17–33

Kemeny ME (2003) The psychobiology of stress. Curr Dir Psychol Sci 12(4):124–129

Kikhia B, Stavropoulos TG, Meditskos G, Kompatsiaris I, Hallberg J, Sävenstedt S, Melander C (2018) Utilizing ambient and wearable sensors to monitor sleep and stress for people with BPSD in nursing homes. J Ambient Intell Human Comput 9(2):261–273

Kopyt A, Zuzewicz K, Bartuzi P (2017) Experimental identification of a mathematical model of human operator working under mental stress. Acta Bioeng Biomech 19(3):177–185

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436

Masood K, AlGhamdi M (2019) Modeling mental stress using a deep learning framework. IEEE Access 7:68446–68454

Mohino-Herranz I, Gil-Pita R, Ferreira J, Rosa-Zurera M, Seoane F (2015) Assessment of mental, emotional and physical stress through analysis of physiological signals using smartphones. Sensors 15(10):25607–25627

Orini M, Citi L, Barbieri R (2012) Bivariate point process modeling and joint non-stationary analysis of pulse transit time and heart period. In: Engineering in medicine and biology society (EMBC), 2012 Annual international conference of the IEEE, IEEE, pp 2831–2834

Panicker SS, Gayathri P (2019) A survey of machine learning techniques in physiology based mental stress detection systems. Biocybern Biomed Eng 39:444–469

Pisoni T, Conci N, De Natale FG, De Cecco M, Nollo G, Frattari A, Guandalini GM (2016) AUSILIA: assisted unit for simulating independent living activities. In: Smart cities conference (ISC2), 2016 IEEE international, IEEE, pp 1–4

Porta A, Faes L, Nollo G, Bari V, Marchi A, De Maria B, Takahashi AC, Catai AM (2015) Conditional self-entropy and conditional joint transfer entropy in heart period variability during graded postural challenge. PLoS One 10(7):e0132851

Sandulescu V, Andrews S, Ellis D, Bellotto N, Mozos OM (2015) Stress detection using wearable physiological sensors. In: International work-conference on the interplay between natural and artificial computation, Springer, pp 526–532

Schwarz G et al (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464

Selye H (1974) Stress without distress. J.B. Lippincott Company, Philadelphia, p 171

Seoane F, Mohino-Herranz I, Ferreira J, Alvarez L, Buendia R, Ayllón D, Llerena C, Gil-Pita R (2014) Wearable biomedical measurement systems for assessment of mental stress of combatants in real time. Sensors 14(4):7120–7141

Sharma N, Gedeon T (2012) Objective measures, sensors and computational techniques for stress recognition and classification: a survey. Comput Methods Progr Biomed 108(3):1287–1301

Smets E, Casale P, Großekathöfer U, Lamichhane B, De Raedt W, Bogaerts K, Van Diest I, Van Hoof C (2015) Comparison of machine learning techniques for psychophysiological stress detection. In: International symposium on pervasive computing paradigms for mental health, Springer, pp 13–22

Speranza G, Nollo G, Ravelli F, Antolini R (1993) Beat-to-beat measurement and analysis of the R-T interval in 24 h ECG Holter recordings. Med Biol Eng Comput 31(5):487–494

Steinheuser V, Ackermann K, Schönfeld P, Schwabe L (2014) Stress and the city: impact of urban upbringing on the (re)activity of the hypothalamus–pituitary–adrenal axis. Psychosom Med 76(9):678–685

Subhani AR, Mumtaz W, Saad MNBM, Kamel N, Malik AS (2017) Machine learning framework for the detection of mental stress at multiple levels. IEEE Access 5:13545–13556

Vaccarino V, Sullivan S, Hammadah M, Wilmot K, Al Mheid I, Ramadan R, Elon L, Pimple PM, Garcia EV, Nye J et al (2018) Mental stress-induced-myocardial ischemia in young patients with recent myocardial infarction: sex differences and mechanisms. Circulation 137(8):794–805

Vanitha V, Krishnan P (2016) Real time stress detection system based on EEG signals. Biomed Res 27:271–275

Vuppalapati C, Raghu N, Veluru P, Khursheed S et al (2018) A system to detect mental stress using machine learning and mobile development. In: 2018 International conference on machine learning and cybernetics (ICMLC), IEEE, vol 1, pp 161–166

Zanetti M, Faes L, De Cecco M, Fornaser A, Valente M, Guandalini G, Nollo G (2018) Assessment of mental stress through the analysis of physiological signals acquired from wearable devices. In: Italian forum of ambient assisted living, Springer, pp 243–256

Zheng R, Yamabe S, Nakano K, Suda Y (2015) Biosignal analysis to assess mental stress in automatic driving of trucks: palmar perspiration and masseter electromyography. Sensors 15(3):5136–5150