Interacting with Computers Advance Access published April 8, 2015

doi:10.1093/iwc/iwv010

Neuro-Fuzzy Physiological Computing to Assess Stress Levels in Virtual Reality Therapy

G. Tartarisco[1,*], N. Carbonaro[3], A. Tonacci[2], G.M. Bernava[1], A. Arnao[1], G. Crifaci[1], P. Cipresso[4], G. Riva[4], A. Gaggioli[4], D. De Rossi[3], A. Tognetti[3] and G. Pioggia[1]

[1]*Institute of Clinical Physiology (IFC), National Research Council of Italy (CNR), Messina, Italy*
[2]*Institute of Clinical Physiology (IFC), National Research Council of Italy (CNR), Pisa, Italy*
[3]*Interdepartmental Research Center "E. Piaggio", University of Pisa, Pisa, Italy*
[4]*Istituto Auxologico Italiano, Milan, Italy*
**Corresponding author: gennaro.tartarisco@ifc.cnr.it*

This paper reports the design and assessment of a neuro-fuzzy model to support clinicians during virtual reality therapy. The implemented model is able to automatically recognize the perceived stress levels of the patients by analyzing physiological and behavioral data during treatment. The model, consisting of a self-organizing map and a fuzzy-rule-based module, was trained unobtrusively recording electrocardiogram, breath rate and activity during stress inoculation provided by the exposure to virtual environments. Twenty nurses were exposed to sessions simulating typical stressful situations experienced at their workplace. Four levels of stress severity were evaluated for each subject by gold standard clinical scales administered by trained personnel. The model's performances were discussed and compared with the main machine learning algorithms. The neuro-fuzzy model shows better performances in terms of stress level classification with 83% of mean recognition rate.

RESEARCH HIGHLIGHTS

- Stress levels were predicted on the basis of physiological computing using a neuro-fuzzy model during virtual reality therapy.
- Features were extracted from ECG and respiration obtaining high accuracy and optimization of computational costs.
- The neuro-fuzzy model shows better performance than the more frequently adopted classifiers.
- This approach may enhance the use of physiological computing for stress treatment in clinical practice.

Keywords: human computer interaction; empirical studies in HCI; virtual reality; stress recognition; artificial intelligence; physiological computing

Received 27 June 2014; Revised 13 February 2015; Accepted 10 March 2015

1. INTRODUCTION

Psychological stress contributes to many chronic diseases suffered by citizens in today's society. Exposure to prolonged stress is known to increase the risk of physical and mental health problems, including depression and disabling anxiety conditions. Furthermore, chronic stress can lead to immunodepression and dysregulation of the immune response, thus significantly enhancing the risk of contracting a disease or negatively altering its course (Gao *et al.*, 2013). According to Cohen *et al*. (1995), stress is a biopsychosocial phenomenon in which 'environmental demands tax or exceed the adaptive capacity of an organism, resulting in psychological and biological changes that may place a person at risk for disease'. This definition emphasizes that in dealing with stress, it is not only necessary to consider the environmental demands, but also the appraisals of such demands, as well as the physiological systems that come into play. Recently, there has been growing interest toward the use of virtual reality (VR) as a new technology-based strategy for assessment and management of psychological stress. VR is employed together with cognitive-behavioral therapy to gradually and repeatedly expose patients to events that have been previously identified as potential stressors. The key-idea underlying this approach is that virtually 'inoculating' the stressor in combination with the acquisition of effective coping skills could prepare the patient to face similar situations in daily life (Serino *et al.*, 2014; Wood *et al.*, 2013). However, an open question in the application of VR for stress assessment and management, is how to accurately evaluate the stress response during the exposure (Darwish and Hasseinein, 2011). There are many biomarkers that can be used as good indicator of stress levels such as the secretion of cortisol (Miller *et al.*, 2007) or the salivary enzyme alpha-amylase (Takai *et al.*, 2004), but even though these biomarkers are more reactive to acute stress their analysis is not feasible in real time. The use of wearable biosensors for monitoring physiological and behavioral correlates of stress is a potential alternative strategy (Gaggioli *et al.*, 2014; Popovic *et al.*, 2009; Riva *et al.*, 2010). In particular, the ECG signal, acquired by smart, minimally obtrusive sensors, is considered one of the most reliable physiological parameters that can be extracted from human body, with a wide range of features, in both time and frequency domains, that could be extremely useful indicators of a subject's stress state (Gomes *et al.*, 2013; Okada *et al.*, 2013; Sweeney *et al.*, 2013). However, the analysis and interpretation of ECG measures is extremely time-consuming and requires specific technical skills, which are often unavailable to psychotherapists. To address this need, a neuro-fuzzy physiological computing model was designed and validated to assist the therapist/researcher in the assessment of stress levels is proposed during VR sessions. The system was tested in a field trial, in which a sample of nurses was presented with virtual stressful situations experienced at their workplace.

2. RELATED STUDIES

One of the most investigated cognitive-behavioral techniques for the management of stress is the Stress Inoculation Training (SIT; Meichenbaum and Novaco, 1985). SIT is implemented through gradual and repeated exposure to events, which have been previously identified as potential stressors. A key objective of this technique is to 'inoculate' the stressor in combination with the acquisition of effective coping skills in order to prepare the patient to face similar situations in daily life. In recent years, there has been growing interest toward the use of VR to support SIT (for a review, see Serino *et al.*, 2014). By exposing the patient to realistic simulations of typical stressful situations, VR is thought to further enhance the efficacy of inoculation (Chittaro, 2013). However, a key issue in the application of VR in the SIT approach is how to accurately evaluate the stress response during the exposure to the simulated stressor. The use of biosensors for monitoring physiological and behavioral correlates of stress has been proposed as a potential solution to this need (Gaggioli *et al.*, 2013a,b; Pallavicini *et al.*, 2013; Tartarisco *et al.*, 2012). Such systems integrate sensors together with on-body signal conditioning and pre-elaboration, as well as management of energy consumption and wireless communication protocol (Anliker *et al.*, 2004; Vuorela *et al.*, 2010). ECG, despite monitoring of other physiological signals such as galvanic skin response, surface electromyography and pupil diameter, is probably the most significant and reliable stress-related information that can be collected during VR exposure. Actually, there is substantial evidence that heart rate variability (HRV) indices can be used to estimate activity of the autonomic nervous system in relation to affective and cognitive states, including mental stress (Berntson and Cacioppo, 2004; Kimhy *et al.*, 2009; Melillo *et al.*, 2011; Salahuddin *et al.*, 2007; Sloan *et al.*, 1994). For example, HRV features have been used to discriminate between subjects reporting high and low levels of stress during the day, with an overall accuracy of 66.1%. (Kim *et al.*, 2008). Another study compared short-term HRV measures using short-term ECG recording in students undergoing university examinations (Melillo *et al.*, 2011). By applying linear discriminant analysis on non-linear features of HRV for automatic stress detection, these authors were able to obtain a total classification accuracy of 90%. The relationship between stress and cardiac autonomic regulation was also investigated in a sample of psychotic patients, using experience sampling in combination with cardiac monitoring (Kimhy *et al.*, 2009). They found that momentary increases in stress were significantly associated with increased sympathovagal balance and parasympathetic withdrawal. A fundamental issue in the measurement of stress is that this response is idiosyncratic, because it depends on an individual's perception of challenges and the skills which she/he can use to face those challenges. Furthermore, it is important for the system to be tailored to the individual characteristics. One possible approach to developing adaptive systems for stress recognition was identifying a subject's baseline and stress threshold in a laboratory setting by eliciting sympathetic and parasympathetic responses, and then using this information to differentiate between stress and no-stress in daily life (Cinaz *et al.*, 2013; Morris *et al.*, 2010). It is complex to create models to assist the therapist in the assessment of stress levels during VR sessions. Traditional approaches are reported to be extremely time-consuming and the data analysis is not easy for clinicians without any particular technical skills. For this reason, the employment of new methods that could help the therapist by simplifying his work and reducing the time required for data analysis is growing exponentially in the stress monitoring field (Campos *et al.*, 2013; Villarejo *et al.*,

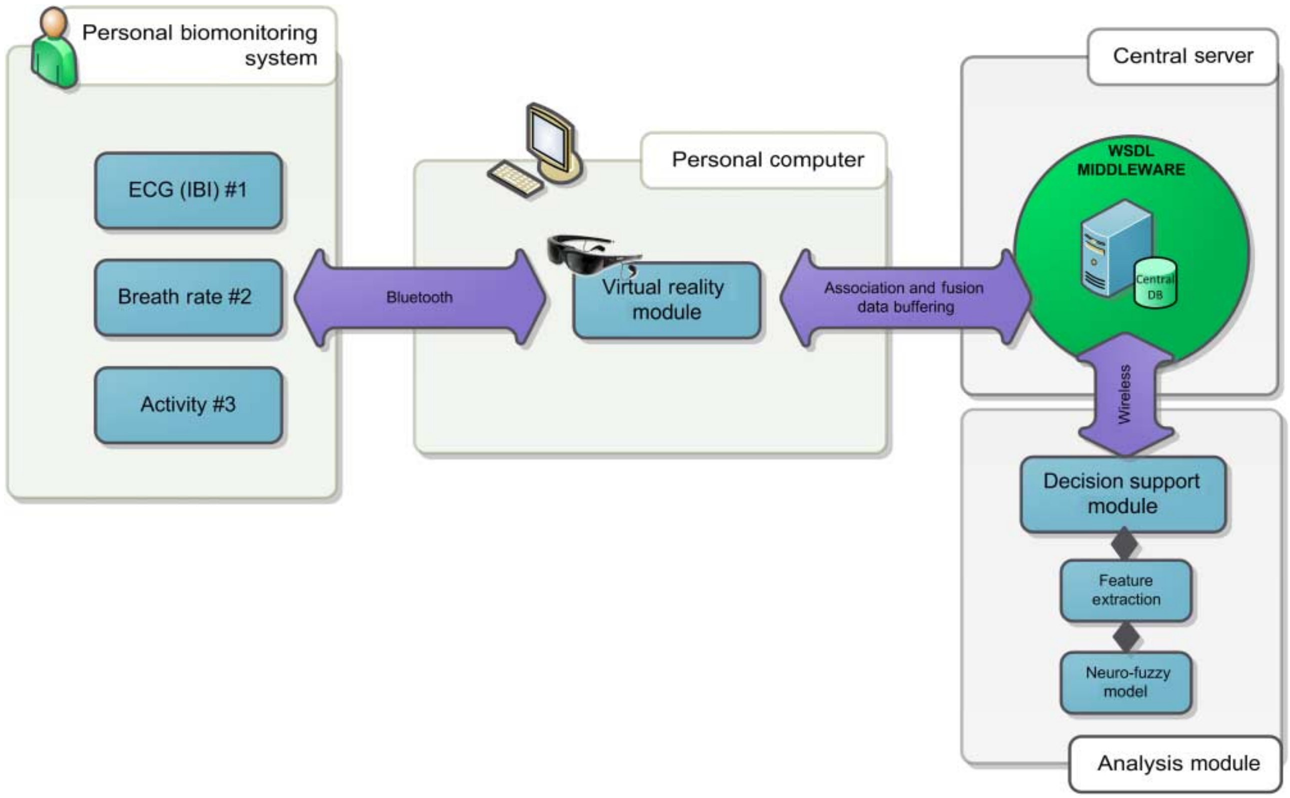

Figure 1. System architecture overview.

2013). Among the several different models that could be employed, the mixture of artificial neural networks (ANNs) (Chakroborty, 2013; Liu *et al.*, 2013; Sansone *et al.*, 2013; Sun and Cheng, 2012) and fuzzy rule-based algorithms (Gacek, 2013; Haseena *et al.*, 2011a,b; Javadi, 2013; Rajendra Acharya *et al.*, 2013; Selvaraj *et al.*, 2013) have proved to be extremely useful in classification in general. For this reason, this approach could be a convenient alternative for classifying stress in different categories depending on the parameters extracted from an ECG signal. A study (Haseena *et al.*, 2011a,b) demonstrated the high reliability of this mixed approach in classifying arrhythmias from ECG signal with a percentage up to 99.58%, a considerably high value that revealed once more the effectiveness of merging ANNs with fuzzy-rule-based algorithms.

3. MATERIALS AND METHODS

The neuro-fuzzy model is part of a general data acquisition and processing architecture (Fig. 1). The general architecture is composed of a personal biomonitoring system (PBS), a personal computer (PC), a central server (CS) and the analysis module (AM). The PBS is dedicated to tracking physiological and behavioral parameters of the user. The PC is devoted to hosting the VR, as well as to collecting and sending data to the CS. The AM consists of a feature extraction and selection modules and the neuro-fuzzy model.

3.1. Personal biomonitoring system

The PBS is an ergonomic chest band, easily worn during daily activities. It integrates a tri-axial accelerometer, an ECG module and electrodes in a comfortable platform as shown in Fig. 2. The chest band collects, fuses and analyzes the meaningful physiological parameters to study stress effects, i.e. inter-beat interval (IBI) and breathing rate (BR). IBI is defined as the time in milliseconds between two normal R to R waves of the ECG. Moreover, the level of user activity is analyzed using the signal magnitude area (SMA) index (Luinge and Veltink, 2004). The SMA index will be useful to automatically exclude from our analysis the motion effects that affect the physiological stress related response. The SMA is extracted from a tri-axial accelerometer integrated in the PBS (Bouten *et al.*, 1997). The ECG module consists of a three-lead sensor and an electronic 256 Hz sampling front-end based on the INA321 instrumentation amplifier. INA321 rejects the common-mode amplifying the input differential ECG signal. The core of the system is the low power microcontroller

Figure 2. The chest band and its electronic board.

Figure 3. PBS data packet format.

(MSP family made by Texas Instruments, MSP430FG439) able to pre-process ECG and respiration signals.

A Kalman filter with a predictor stage allows a QRS complex to be accurately detected (Carbonaro *et al.*, 2011) and the IBI signal to be extracted. In order to collect the BR, a polyvinylidene piezoelectric cable is integrated in the chest band. The cable transduces the mechanical forces produced during the chest movements into an electrical signal, from which analysis allows the BR to be extracted, as previously demonstrated (Carbonaro *et al.*, 2013). Each measurement is stored in the CS using a data packet format containing the timestamp and the variables (Fig. 3).

3.2. VR module

The VR consists of a head-mounted display (Vuzix VR Bundle with twin high-resolution 640×480 LCD displays, 920 000 pixels, iWear® 3D compliant) able to show a 3D view of the virtual scene. The virtual scenes were realized by means of NeuroVR-2 (http://www.neurovr2.org) (Riva *et al.*, 2011). The virtual environment was rendered using a PC (iMac with CPU Intel® Core™i5 and graphic processor Nvidia GeForce GT 540M). A joystick (Xbox Controller) enables the user to interact with the environment. The VR is used for stress inoculation, reproducing a number of virtual stressful scenarios realized on the basis of storyboards tailored to selected stressful work-related situations.

3.3. Feature extraction module

In order to reduce the number of artifacts and increase the robustness of the signal, an auto-regressive (AR) model implemented by a predictive adaptive filter was applied to pre-process the IBI signal. The filter coefficients were estimated on the initial RR interval and updated beat to beat by the least mean square algorithm. The AR model predicts the RR duration, i.e. the expected position of the next QRS peak, even when the signal is affected by artifacts (Varanini *et al.*, 2014). Following the processing of the RR time series, relevant features were extracted using a rectangular non-overlapping window of 5 min (epochs). For each epoch, five time-domain linear features were extracted according to the International Guidelines of HRV (Malik *et al.*, 1996a,b): RR mean (mRR) and standard deviation (σRR); root mean square of successive differences of intervals (RMSSD); the number of successive differences of intervals which differ by more than 50 ms (pNN50% expressed as a percentage of the

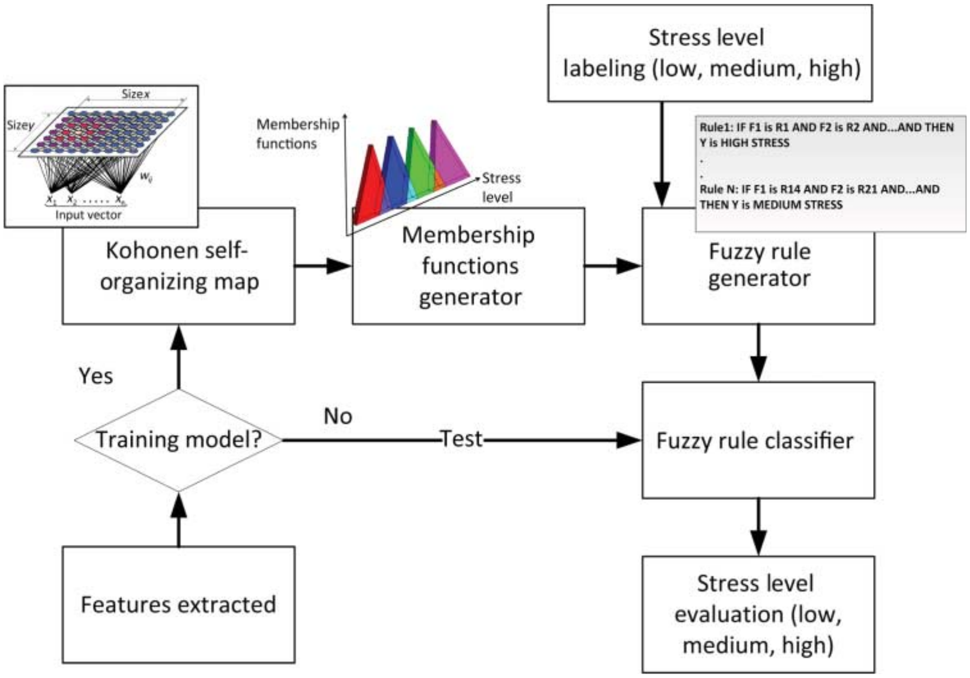

Figure 4. The neuro-fuzzy model for automatic stress level classification.

total number of heartbeats analyzed); SDANN, the standard deviation of the average NN intervals calculated over short periods, usually 5 min. Moreover, in the frequency domain, in order to evaluate the global sympathetic-parasympathetic balance, the low frequency (LF)/ high frequency (HF) ratio between the LF (0.03–0.15 Hz) and HF (0.15–0.40 Hz) powers was extracted. This feature was assessed using an estimation of the power spectral density analysis according to the Burg spectral estimation (Burg, 1967), where the optimal order p was estimated according to the Akaike information criterion (Akaike, 1969). For each epoch, three non-linear features were extracted: the two standard deviations (SDs) of the Poincaré plot and the sample entropy (SmEn). The Poincaré plot is a useful method for investigating and combining the differences in the cardiac rhythms during the performed tasks. It is a graphical representation created by plotting the RR time series, RR(n), on the x-axis versus the shifted RR time series, RR($n+1$), on the y-axis (Brennan *et al.*, 2001). The SmEn is an index extracted to evaluate the complexity and irregularity of RR time series (Richman and Moorman, 2000). Finally, the BR was collected. The SMA was used to remove portions of physiological signals related to movements over a specified empirical threshold which affects HRV parameters, because the hypothesis was to recognize overall stress levels mainly related to mental workload elicited by VR exposure. The epochs affected by movements were detected and removed by applying a simple threshold-based algorithm to the SMA signal (Curone *et al.*, 2010; Mathie *et al.*, 2003).

3.4. Feature selection module

Once features were extracted, we selected a subset of original features to improve interpretability and performances and reduce computational costs of the neuro-fuzzy model. We decided to perform feature selection in two steps. First, we adopted the ReliefF algorithm (Robnik-Šikonja and Kononenko, 2003) to rank features in function of their relevance. The ReliefF is an efficient and robust solution against noisy data. The algorithm iterates for every instance of the dataset searching for the k nearest neighbors belonging to the same class (nearest hits H) and the k nearest neighbors belonging to other classes (nearest misses M). The second step of the feature selection procedure consisted in the use of the Davies–Bouldin (DB) cluster evaluation index (Bezdek and Pal, 1998), a criterion to select the best N ranked features. The DB index is defined as the ratio of the sum of within-class scatter to between-class separation. The ranked features are incrementally added in a multi-dimensional feature space, and if classes of datasets are well-separated we obtain a smaller DB value and vice versa.

3.5. The neuro-fuzzy model

The neuro-fuzzy model is a combination of a self-organizing map (SOM) and a rule-based fuzzy model (Fig. 4). The input variables of the model are the extracted features of HRV and BR, normalized in the range of [0, 1]. An SOM

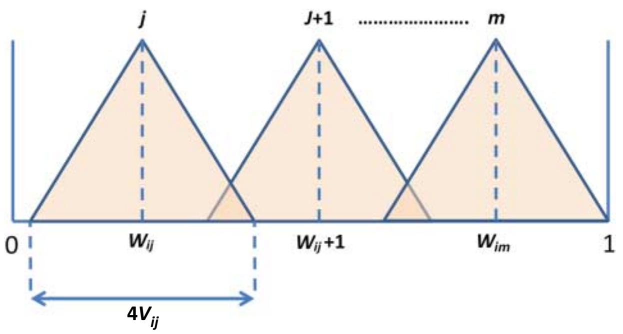

Figure 5. Generation of the fuzzy membership function for the *i*th input.

is a type of ANN that provides a topological mapping using unsupervised learning (Kohonen, 1995). It produces a discretized representation of the input space (a map). An SOM consists of components called neurons. For each neuron a weight vector of the same dimension as the input data vectors is associated. A model based on SOMs is identified by means of a training phase, where a competitive process analyzes the input examples. The training phase starts randomly initializing the weights. The test phase is used to automatically classify an unknown input vector.

The structure of the trained SOM was used to generate the membership functions (Chi *et al.*, 1995). The number of triangular fuzzy regions is equal to the number of artificial neurons of the SOM. The weight of one single artificial neuron affects one single fuzzy region. Assuming the *i*th input variable at the end of the training phase, each triangular fuzzy region is centered in the mean value of the i weights w_{ij} of the corresponding artificial neuron j, while the variance v_{ij} corresponds to the width (Fig. 5). Therefore, considering m as the number of artificial neurons, the centers of the triangular membership functions are $w_{i1}, w_{i2} \ldots w_{im}$. The corresponding regions were set to $[w_{i1} - 2v_{i1}, w_{i1} + 2v_{i1}]$, $[w_{i2} - 2v_{i2}, w_{i2} + 2v_{i2}], \ldots, [w_{im} - 2v_{im}, w_{im} + 2v_{im}]$, as shown in Fig. 5.

In order to improve the accuracy of the system (Chi *et al.*, 1995), the neighboring fuzzy regions whose distance is below a pre-specified threshold *thr* are combined in a single fuzzy region, obtaining a trapezoidal shape as reported in Fig. 6. The neighboring fuzzy regions satisfying the following equation were merged:

$$\frac{lh_j + rh_j}{2} - \frac{lh_{j-1} + rh_{j-1}}{2} \leq thr, \tag{1}$$

where ll_j, lh_j, rh_j and rl_j reported in Fig. 6, represent the four corners of the new merged region.

The fuzzy rules were generated according to the following method (Wang and Mendel, 1992): i.e. the label of the fuzzy region corresponding to the maximum membership value is associated to each input. According to such a strategy, the number of fuzzy rules is the same of the elements of the training samples. A general example of the rule is below reported:

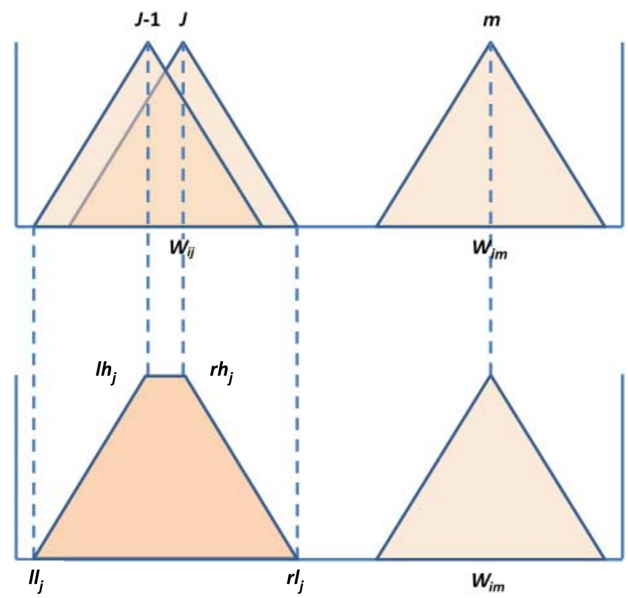

Figure 6. Trapezoidal function obtained for neighboring triangular regions.

IF
feature1 is R1 AND feature2 is RN AND feature3 is R2 AND feature4 is R3 AND
feature5 is R6, AND feature6 is R8 .... AND feature M is R3
THEN it is
No/Low/Medium/High Stress level

The output of the model for each input pattern is evaluated according to the centroid defuzzification formula.

$$Z = \frac{\sum_{i=1}^{k} D_p^i O^i}{\sum_{i=1}^{k} D_p^i} \tag{2}$$

where Z is the output of the model, k is the number of rules, O^i is the class generated by the rule i and D_p^i measures how the input vector fits the ith rule. D_p^i is the membership of the output of the ith rule. The output Z is within [0, 3] in the case of four stress level classifications (0, no stress; 1, low stress; 2, medium stress; 3, high stress) and is adapted taking the nearest smaller integer value. Finally, the SOM of the neuro-fuzzy model was a 5 × 5 map of artificial neurons. The length of the training phase was 600 epochs. The fuzzy membership functions were set using a *thr* equal to 0.1. In order to avoid repetitions or conflicts, only 633 rules were selected, i.e. the rules supported by more than five examples.

3.6. Other classifiers

In order to test the classification performances, the neuro-fuzzy model was compared with the most common machine learning models, i.e. instance-based learning (IBL) (Aha *et al.*, 1991), Naïve Bayes (NB) (Duda *et al.*, 1999), multilayer perceptron (MLP) (Duda *et al.*, 1999), J48 decision tree (Quinlan, 1993), random forest (RF) decision tree (Breiman, 2001), single SOM (Kohonen, 1995) and support vector machine (SVM) (Vapnik, 1995). All the classifiers were implemented using the Java libraries provided by Weka 3 data mining software (Witten and Frank, 2005). To guarantee the optimal performance of IBL, NB, J48 and RF, the default parameters were used as suggested by literature (Amancio *et al.*, 2014). As regards the Weka SVM library, a radial basis function kernel combined with the other default parameters was selected. The MLP parameters were the following: one hidden layer, learning rate of 0.9 with a decay of 0.1, momentum of 0.6 with a decay of 0.1 and 800 training epochs. Finally, the single SOM was set as a 5×5 map trained for 600 epochs. For all the neural network, the training was run until a minimum average square error (MSE) of <0.001 or an increasing MSE was found in the training set.

3.7. Model validation metric

The performances of the neuro-fuzzy model where assessed by using the confusion matrix. The generic elements i,j of the confusion matrix indicate how many times a pattern belonging to the class i was classified as belonging to the class j. In particular, if $i=j$, the generic element represents the correct classification rate (CR_i), while if $i \neq j$ the generic element represents the misclassification rate (ER_{ij}) as reported in Table 1.

In order to check the generalization capability of the model, the data set of input features was divided into training set and testing set, normalized and the tenfold cross-validation and the leave-one-subject-out cross-validation (Breiman *et al.*, 1984) methods were respectively carried out. In the former, the original dataset was randomly partitioned into 10 equal size subsamples, where a single subsample was retained as the validation data for testing the model, and the remaining 9 subsamples were used as training data. The cross-validation process was then repeated 10 times, with each of the 10 subsamples used exactly once as the validation data. In the latter, the number of folds corresponds to the number of subjects ($n=18$), where each subsample consists of data belonging to a single subject. In this case, the cross-validation process was then repeated 18 times, with each of the 18 subsamples used exactly once as the validation data. In both cases, the obtained confusion matrices were averaged producing an estimated confusion matrix whose elements represents the sensitivity of the model to predict the specific clinical class and are expressed in terms of mean percentage ± SD. The overall accuracy of the model is given by the mean value of sensitivity of each recognized class.

Table 1. Generic confusion matrix for a single cross-validation process.

	Predicted class			
	Class A	Class B	Class C	Class D
Clinical class				
Class A	$\mathbf{CR_A}$	ER_{AB}	ER_{AC}	ER_{AD}
Class B	ER_{BA}	$\mathbf{CR_B}$	ER_{BC}	ER_{BD}
Class C	ER_{CA}	ER_{CB}	$\mathbf{CR_C}$	ER_{CD}
Class D	ER_{DA}	ER_{DB}	ER_{DC}	$\mathbf{CR_D}$

3.8. Participants

The experimental sample included 20 nurses between 25 and 60 years old recruited from the personnel of the University Hospital 'G. Martino' of Messina, Italy. Each participant completed a screening interview involving both males and females. Subjects undergoing pharmacotherapy or with neurological diseases, psychosis, alcohol, drug dependence, migraine, headache or vestibular abnormalities were excluded. Participants were instructed to avoid caffeine, tobacco and strenuous exercise at least 4 h before the beginning of the session. Written consent was given by the subjects before the study conducted in laboratory settings. The trials were approved by the Ethical Committee of the University Hospital 'G. Martino', Messina, on 21 January 2013.

3.9. Experimental setup design

The experimental study was conducted in laboratory setting, reproducing appropriate stress stimuli with VR in order to develop the model for automatic stress level detection by using only physiological data. Each participant was exposed to VR scenarios simulating typical situations experienced by nurses at the workplace (Fig. 7), i.e. being reproached by colleagues, managing an emergency and coping with a patient's criticism, as well as simulating relaxing scenarios.

The storyboard design guidelines were written asking a representative sample of nurses to participate in focus groups and in-depth interviews, while the stressful scenarios were played by real actors and included in the virtual environments (Riva *et al.*, 2011) after a video post-production. For this study, 10 stressful scenarios were selected by clinicians:

(i) Managing patients' relatives,
(ii) Managing patients' complaints,
(iii) Managing a medical emergency situation,
(iv) Relationship with colleagues,
(v) Managing relatives'/caregivers' anxiety,
(vi) Distribution of work tasks,
(vii) Patient–doctor communication,
(viii) Managing patient's anxiety,
(ix) Unsuccessful collaboration/communication with colleagues,
(x) Arguments with medical doctors.

Figure 7. Setup of the VR system used in the clinical setting. Top left: personal bio-monitoring system. Bottom right: example of a stressful scenario used in the training.

According to previous studies, all of the scenes were effective in eliciting stress, a strong emotional response and feeling of presence that was similar (Gorini *et al.*, 2010a,b) or even greater (Villani *et al.*, 2012) than the real scene. The relaxing scenarios were presented together with relaxing audio narratives based on Guided Imagery procedures and developed according to ‘emotive engineering’ principles (Rossman, 2010). The relaxation environments were created on the basis of similar virtual relaxing environments that were used and validated in previous studies (Ferrer-García *et al.*, 2009; Gorini *et al.*, 2010a,b; Manzoni *et al.*, 2008; Pallavicini *et al.*, 2009) selecting five relaxing scenarios: a beach, a lake, a campfire, a mountain summit and a desert. Each of them were associated with different pre-recorded audio narratives. The protocol was based on 5 weeks of VR experience (two sessions per week). During each session, participants were asked to wear the PBS in order to collect IBI and BR values during exposure, the head-mounted display for VR and sitting in a quiet room on a comfortable chair. Participants also had a joystick, which allowed them to explore and to interact with the VR environment. Each VR session lasted ∼1 h and included an initial discussion with a psychologist. Following the introduction, we recorded physiological baseline for 15 min and then the participants were exposed to 30–40 min of a stressful or relaxing scenario. For each VR session, one stressful or relaxing scenario was randomly selected and administered. Features were extracted from epochs of 5 min giving subjects the time necessary to feel comfortable seated, to adapt to the change in environment and the signal to stabilize from the previous task. The starting point of the feature extracted was detected with the integrated accelerometer on the chest. In particular, epochs below the empirical threshold of SMA index were analyzed. This criterion was applied to prevent HRV feature changes from being affected by physical movements. Before and after each exposure to VR scenarios, a psychometric assessment based on questionnaires (Pallavicini *et al.*, 2013) was administered to each participant and used to label each instance of dataset and validate the perception of no-stress, low, medium and high stress. For this purpose, the psychological questionnaire of State Anxiety Inventory Y1 (STAI-Y1, self-rated) (Spielberg *et al.*, 1983) was used immediately before and after each VR session. The STAI-Y1 comprises a questionnaire of 20 items that ask how much participants agree with sentences about their current state. Based on the answers, the STAI assigns a score from 20 to 80 to the participant’s anxiety state. These data were transformed to four ordinal groups for classifiers corresponding to no, low, medium and high stress. More specifically, we identified the following ranges: no stress <30, low <40, medium 40–55, high >55 in agreement with other studies available in literature (Orbach-Zinger *et al.*, 2012).

4. RESULTS

The inputs of the neuro-fuzzy model were: mRR, σRR, RMSSD, pNN50%, SDANN, LF/HF, SD1, SD2, SmEn, BR. The IBI recordings of <5 min were removed since according to literature (Malik *et al.*, 1996a,b), reliable features cannot be extracted. Two subjects out of 20 were excluded from the analysis due to missing data and/or artifacts. A total of

156 h of data recording were collected. Of the acquired data, 11% was removed because it was over the SMA threshold. A total of 139 h of valid data corresponding to 1668 valid epochs of 5 min (673 classified as no-stress, 322 as low stress, 343 as medium stress and 330 as high stress) were taken into account. The sample sizes were equalized before the training of the classifiers by selecting a random subsample of 350 no-stress sessions, obtaining a total balanced dataset of 1345 instances. To confirm that the stress induction was successful, we performed a psychometric assessment. In particular, we extracted the change in anxiety making the difference between anxiety state scores measured after and before the VE experience. This parameter was compared with the time of duration of each VR scenario using an *a priori* statistical evaluation (Spearman's correlation test). The two measures resulted uncorrelated, with a correlation coefficient of 0.203 and a statistical significance of $P = 0.56$. We also evaluated the Spearman correlation between each physiological feature and time on task each 5 min (T5, T10, T15, T20, T25, T30), obtaining a mean correlation coefficient of 0.311 and no statistical significance with time ($P > 0.05$) respectively for all features. The mean STAY scores before and after each VR session compared with a paired *t*-test showed a significant difference ($P < 0.01$). Regarding stress VR sessions, the mean score shown before ($M = 34.5$, SD = 6.4) was lower than that assessed after ($M = 55.5$, SD = 7.1). A mirror result was obtained for relaxed VR sessions. The mean score shown before ($M = 33.2$, SD = 7.1) was higher than that assessed after the 'relaxed' exposure ($M = 25.1$, $SD = 4.2$). After this preliminary analysis, we reported in Table 1 the performances of all the classifiers in terms of percentage of classification error, applying both the randomized 10-fold cross-validation and the leave-one-subject-out methods. In the case of 10-fold cross-validation, all the classifiers achieved a mean error rate of ~14% with no significant differences between methods as confirmed by a one-way ANOVA ($P > 0.3$). In the case of the leave-one-subject out instead, ANOVA, the effect was significant, $F(2, 105) = 6.03$, $P < 0.05$, $\eta^2 = 0.06$. Post hoc analysis with Bonferroni test showed that the difference between the neuro-fuzzy model and each other classifier was significant ($P < 0.05$). The mean error of classification of ~15% was higher than the other classifier as reported in Table 2.

In order to analyze in more detail the ability of the model to discriminate the four different stress levels, in Table 3 the mean confusion matrix of the neuro-fuzzy model for the leave-one-subject-out validation is reported. We observed that the no-stress and high-stress classes were classified with a high degree of sensitivity (more than 90%), while low and medium stress were detected with a lower sensitivity (~80%). A second analysis using features selected from ReliefF algorithm combined with DB index was performed. Once the features were ranked we calculated the DB index for 2–10 features at increments of one feature as reported in Fig. 8.

Table 2. Percentage of classification error (mean ± SD) using all 10 features.

Classifiers	10 × *cross validation*	*Leave-one-subject-out*
IBL	13.24 (±0.44)	23.21 (±8.93)
NB	14.69 (±2.78)	26.58 (±11.17)
J48	13.81 (±2.11)	27.25 (±9.52)
MLP	14.78 (±0.52)	21.16 (±10.66)
RF	14.76 (±0.57)	25.61 (±6.31)
SVM	13.62 (±0.47)	23.12 (±6.32)
SOM	13.78 (±0.22)	20.09 (±8.52)
Neuro-fuzzy	13.51 (±0.61)	14.82 (±6.79)

Aware that the minimum value of the extracted DB index indicates the optimal cutoff point for the discrimination of the four stress levels, we selected the following first five features: RMSSD, SmEn, LF/HF, BR and pNN50. All the classifiers were re-tested using the five selected features. A paired *t*-test for each classifier revealed that the accuracy of classification between all and five selected features was not statistically significant ($P < 0.05$). This result is also reported in Fig. 9. Finally, an ANOVA analysis at the 95% CI revealed a significant difference also among classifiers based on five selected input features $F(2, 101) = 4.02$, $P < 0.05$, $\eta^2 = 0.03$. The post hoc Bonferroni test demonstrated that the neuro-fuzzy model with five input features was statistically equivalent to MLP and SOM and higher than that of other classifiers. The mean value of accuracy of SOM + FUZZY was equal to 81.58 (SD = 6.42), for MLP 77.87 (SD = 11.13) and 77.58 (SD = 9.42) for SOM.

5. DISCUSSION

The main objective of this study was to test the applicability of the neuro-fuzzy model and physiological sensing for determining the user's overall stress level during VR therapy. The performances of the implemented model were compared with the main machine learning algorithms showing that the four stress levels could be recognized with an overall accuracy of ~83%. More specifically, the ANOVA post hoc Bonferroni analysis showed a significance of neuro fuzzy with the leave-one-subject-out approach also evidencing that the model is more suitable for detecting the perceived stress regardless of the subject. This outcome is important because it highlights that the system is able with a low error to manage the non-linearity of physiological reactions and at the same time the different effects and components of the individual's physiology. A statistical analysis performed to better characterize the collected data demonstrated that the use of VR scenarios is able to affect a change in anxiety ($P < 0.01$) and that this change is not affected by the duration of VR scenarios as demonstrated by the Spearman correlation index of 0.203 ($P = 0.56$). Moreover, the Spearman analysis performed at each 5-min interval

Table 3. Mean confusion matrix for stress classification using the neuro-fuzzy model with leave-one-subject-out cross-validation repeated 18 times.

	Predicted stress			
% correct (mean ± SD)	No stress	Low stress	Medium stress	High stress
Clinical stress				
No stress	**90.32 (±5.78)**	5.85 (±2.22)	1.93 (±1.02)	1.90 (±0.75)
Low stress	1.97 (±1.21)	**79.36 (± 10.92)**	16 (±9.67)	2.67 (±1.99)
Medium stress	2.37 (±0.66)	13.41 (±10.49)	**80.84 (±12.45)**	3.38 (±2.38)
High stress	1.45 (±0.61)	1.56 (±1.49)	6.78 (± 5.34)	**90.21 (±3.56)**

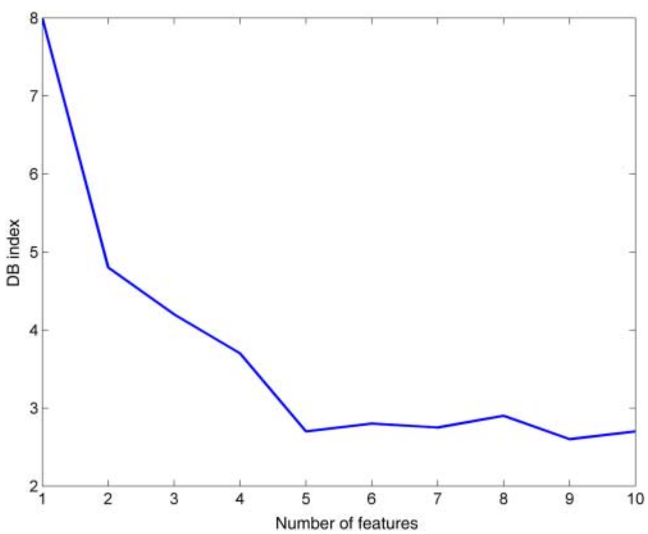

Figure 8. DB index extracted ranking features with ReliefF algorithm and computed feature by feature.

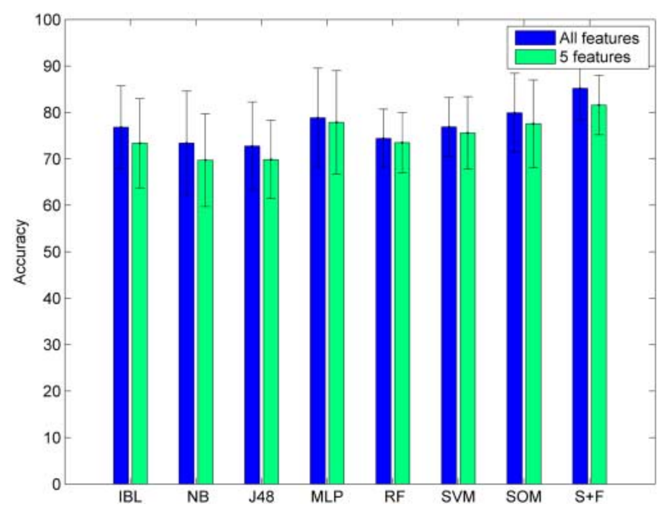

Figure 9. Bar plot of all classifiers' accuracy, respectively, for all features and the five features selected.

of physiological data also showed with a correlation of 0.311 ($P = 0.45$) that the physiological signals do not change within the time on task. In this analysis, it is important to highlight that the effects of VR scenarios were excluded because they were selected randomly. A further analysis in Table 2 shows the confusion matrix with the mean sensitivity of each class of stress, respectively, of 90.32, 79.36, 80.84, and 90.21% obtained after leave-one-subject-out validation. We observed that ~16% of the low-stress instances were classified as medium stress and ~13% of medium-stress instances were classified as low-stress. Considering the high similarity between the low and medium classes, these results are encouraging. The minor misclassification error may between low and medium stress may be due to the stimuli provided in these instances which did not induce the required stress state, resulting in psychophysiological values of different magnitudes, or may be due to the subjects' difficulty identifying such close stress levels clearly. A second analysis based on comparison between classification performances with all the available features and the selected subsets identified with the ReliefF algorithm combined with the DB method has revealed that five features (RMSSD, SmEn, LF/HF, BR and pNN50) seem to provide more information to the classification process than the others. Although the mean accuracy of the neuro-fuzzy model was lower with selected features (~81%), the paired *t*-test demonstrated that there was no significant difference ($P < 0.05$) with performances using all features. Moreover, ANOVA analysis with the Bonferroni post hoc test confirmed the significance of the neuro-fuzzy model. This result highlights how by selecting a subset of features we can reduce the performance costs of the mode, obtaining a good compromise of accuracy to classify the stress level. This approach could be very promising in the near future for developing a mobile application. An important consideration is about the SMA index extracted from the integrated accelerometer. It was used to remove epochs of signals with movements of the subject over a certain threshold, because as demonstrated by literature, they affect the changes of HRV features. Of course with this approach, one limitation is that we cannot monitor the stress level when the subject is moving (i.e. walking or running), but in our specific setting this aspect had a low impact, because the patient was comfortably seated. In the near future, we will perform experiments to use this model to infer stress level during daily life with a mobile platform and the SMA index will be used and further investigated to contextualize changes of HRV features.

6. CONCLUSION

The paper reports the design and validation of a neuro-fuzzy model and physiological sensing for stress monitoring. The platform assists the therapist in analyzing and interpreting a patient's stress levels during VR exposure. The original contribution of this work concerns the development of an SOM combined with a fuzzy rule-based algorithm dedicated to the classification of four stress levels. In particular, the model was trained unobtrusively recording electrocardiogram, breath rate and activity, during stress inoculation provided by the exposure to virtual environments. The platform also provides a framework, where fuzzy rules can be continuously updated and integrated with those learned from the training data for a more comprehensive definition of the stress response. This strategy allows a closed-loop approach that is lacking in current strategies for the evaluation and treatment of psychological stress. The assessment is conducted throughout the virtual experiences and enables tracking of the individual's psycho-physiological status in the context of stress inoculation using realistic stressful scenarios. The performances of the neuro-fuzzy model were compared with the most common machine learning algorithms. The leave-one-subject-out cross-validation analysis shows how the neuro-fuzzy model achieves the best performances, with an accuracy of above 83%. A feature-selecting method allows obtaining a reduced feature set able to achieve similar discrimination of stress level results. Future insights may contribute to enhance the use of physiological computing for stress treatment in clinical practices.

ACKNOWELEDGEMENTS

A special thanks to the University Hospital 'G. Martino', Messina, Italy for support and assistance. Authors are grateful to the colleagues who provided a valuable contribution to the trial.

FUNDING

This study was funded by European Union INTERSTRESS project (ICT Grant Number: FP7-247685).

REFERENCES

Aha, D.W., Kibler, D. and Albert, M.K. (1991) Instance-based learning algorithms. Mach. Learn., 6, 37–66.

Akaike, H. (1969) Fitting autoregressive models for prediction. Ann. Inst. Statist. Math., 21, 243–247.

Amancio, D.R., Comin, C.H., Casanova, D., Travieso, G., Bruno, O.M., Rodrigues, F.A. and da Fontoura Costa, L. (2014) A systematic comparison of supervised classifiers. PloS One, 9, e94137.

Anliker, U., Ward, J.A., Lukowicz, P., Troster, G., Dolveck, F., Baer, M., and Vuskovic, M. (2004) AMON: a wearable multiparameter medical monitoring and alert system. IEEE Trans. Inf. Technol. Biomed., 8, 415–427.

Berntson, G.G. and Cacioppo, J.T. (2004) Heart rate variability: stress and psychiatric conditions. Dyn. Electrocardiography, 7, 57–64.

Bezdek, J.C. and Pal, N.R. (1998) Some new indexes of cluster validity. IEEE Trans. Syst. Man Cybern. B, Cybern., 28, 301–315.

Bouten, C.V., Koekkoek, K.T., Verduin, M., Kodde, R. and Janssen, J.D. (1997) A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. IEEE Trans. Biomed. Eng., 44, 136–147.

Breiman, L. (2001) Random Forests. Machine Learning, Vol. 45, pp. 5–32. Springer.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984) Classification and Regression Trees. Belmont, CA: Wadsworth International Group.

Brennan, M., Palaniswami, M. and Kamen, P. (2001) Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability? IEEE Trans. Biomed. Eng., 48, 1342–1347.

Burg, J.P. (1967) Maximum Entropy Spectral Analysis. 37th Annual Int. Meeting. Society of Exploration Geophysics.

Campos, L.A., Pereira Jr, V.L., Muralikrishna, A., Albarwani, S., Brás, S. and Gouveia, S. (2013) Mathematical biomarkers for the autonomic regulation of cardiovascular system. Front. Physiol., 4, 279.

Carbonaro, N., Anania, G., Mura, G.D., Tesconi, M., Tognetti, A., Zupone, G. and De Rossi, D. (2011) Wearable Biomonitoring System for Stress Management: A Preliminary Study on Robust ECG Signal Processing. IEEE Int. Symp. World of Wireless, Mobile and Multimedia Networks (WoWMoM), pp. 1–6. IEEE.

Carbonaro, N., Tognetti, A., Anania, G., De Rossi, D., Cipresso, P., Gaggioli, A. and Riva, G. (2013) A Mobile Biosensor to Detect Cardiorespiratory Activity for Stress Tracking. Proc. 7th Int. Conf. Pervasive Computing Technologies for Healthcare, pp. 440–445. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Chakroborty, S. (2013) Accurate Arrhythmia Classification Using Auto-Associative Neural Network. 35th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4247–4250. IEEE.

Chi, Z., Wu, J. and Yan, H. (1995) Handwritten numeral recognition using self-organizing maps and fuzzy rules. Pattern Recognit., 28, 59–66.

Chittaro, L. (2013) Anxiety induction in virtual environments: an experimental comparison of three general techniques. Interact. Comput., iwt049.

Cinaz, B., Arnrich, B., La Marca, R. and Tröster, G. (2013) Monitoring of mental workload levels during an everyday life office-work scenario. Pers. Ubiquitous Comput., 17, 229–239.

Cohen, S., Kessler, R. and Gordon, L. (1995) Strategies for Measuring Stress in Studies of Psychiatric and Physical Disorders. In Cohen, S., Kessler, R. and Gordon, L. (eds). Measuring Stress, pp. 3–28. New York: Oxford University Press.

Curone, D., Tognetti, A., Secco, E.L., Anania, G., Carbonaro, N., De Rossi, D. and Magenes, G. (2010) Heart rate and accelerometer data fusion for activity assessment of rescuers during emergency interventions. IEEE Trans. Inf. Technol. Biomed., 14, 702–710.

Darwish, A. and Hassanien, A.E. (2011) Wearable and implantable wireless sensor network solutions for healthcare monitoring. Sensors, 11, 5561–5595.

Duda, R.O., Hart, P.E. and Stork, D.G. (1999) Pattern Classification. John Wiley and Sons.

Ferrer-García, M., Gutiérrez-Maldonado, J., Caqueo-Urízar, A. and Moreno, E. (2009) The validity of virtual environments for eliciting emotional responses in patients with eating disorders and in controls. Behav. Modif., 33, 830–854.

Gacek, A. (2013) Data structure-guided development of electrocardiographic signal characterization and classification. Artif. Intell. Med., 59, 197–204.

Gaggioli, A., Cipresso, P., Serino, S., Campanaro, D.M., Pallavicini, F., Wiederhold, B.K. and Riva, G. (2013a) Positive technology: a free mobile platform for the self-management of psychological stress. Stud. Health Technol. Inf., 199, 25–29.

Gaggioli, A., Pioggia, G., Tartarisco, G., Baldus, G., Corda, D., Cipresso, P. and Riva, G. (2013b) A mobile data collection platform for mental health research. Pers. Ubiquitous Comput., 17, 241–251.

Gaggioli, A. *et al.* (2014) Experiential virtual scenarios with real-time monitoring (interreality) for the management of psychological stress: a block randomized controlled trial. J. Med. Internet Res.,16, e167.

Gao, K., Boyd, C., Wiederhold, M.D. and Wiederhold, B.K. (2013) VR mobile solutions for chronic stress reduction in young adults. Stud. Health Technol. Inform., 199, 88–93.

Gomes, P., Kaiseler, M., Lopes, B., Faria, S., Queiros, C. and Coimbra, M. (2013) Are Standard Heart Rate Variability Measures Associated with the Self-Perception of Stress of Firefighters in Action? 35th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2571–2574. IEEE.

Gorini, A., Griez, E., Petrova, A. and Riva, G. (2010a) Assessment of the emotional responses produced by exposure to real food, virtual food and photographs of food in patients affected by eating disorders. Ann. Gen. Psychiatry, 9, 30.

Gorini, A., Pallavicini, F., Algeri, D., Repetto, C., Gaggioli, A. and Riva, G. (2010b) Virtual reality in the treatment of generalized anxiety disorders. Stud. Health Technol. Inform., 154, 39–43.

Haseena, H.H., Joseph, P.K. and Mathew, A.T. (2011a) Classification of arrhythmia using hybrid networks. J. Med. Syst., 35, 1617–1630.

Haseena, H.H., Mathew, A.T. and Paul, J.K. (2011b) Fuzzy clustered probabilistic and multi layered feed forward neural networks for electrocardiogram arrhythmia classification. J. Med. Syst., 35, 179–188.

Javadi, M. (2013) Combining neural networks and ANFIS classifiers for supervised examining of electrocardiogram beats. J. Med. Eng. Technol., 37, 484–497.

Kim, D., Seo, Y., Cho, J. and Cho, C.H. (2008) Detection of Subjects with Higher Self-Reporting Stress Scores Using Heart Rate Variability Patterns During the Day. 30th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society, EMBS 2008, pp. 682–685. IEEE.

Kimhy, D., Delespaul, P., Ahn, H., Cai, S., Shikhman, M., Lieberman, J.A. and Sloan, R.P. (2009) Concurrent measurement of 'real-world' stress and arousal in individuals with psychosis: assessing the feasibility and validity of a novel methodology. Schizophr Bull, 36, 1131–1139.

Kohonen T. (1995) Self-Organizing Maps. Series in Information Sciences, 30. Springer, Heidelberg.

Liu, S.H., Cheng, D.C. and Lin, C.M. (2013) Arrhythmia identification with two-lead electrocardiograms using artificial neural networks and support vector machines for a portable ECG monitor system. Sensors, 13, 813–828.

Luinge, H.J. and Veltink, P.H. (2004) Inclination measurement of human movement using a 3-D accelerometer with autocalibration. IEEE Trans. Neural Syst. Rehabil. Eng., 12, 112–121.

Malik, M., Bigger, J.T., Camm, A.J., Kleiger, R.E., Malliani, A., Moss, A.J. and Schwartz, P.J. (1996a) Heart rate variability standards of measurement, physiological interpretation, and clinical use. Eur. Heart J., 17, 354–381.

Malik, M., Bigger, J.T., Camm, A.J., Kleiger, R.E., Malliani, A., Moss, A.J. and Schwartz, P.J. (1996b) Heart rate variability standards of measurement, physiological interpretation, and clinical use. Eur. Heart J., 17, 354–381.

Manzoni, G.M., Gorini, A., Preziosa, A., Pagnini, F., Castelnuovo, G., Molinari, E. and Riva, G. (2008) New Technologies and relaxation: an exploratory study on obese patients with emotional eating. The role of presence. J. Cybertherapy Rehabil., 1, 182–192.

Mathie, M.J., Coster, A.C.F., Lovell, N.H. and Celler, B.G. (2003) Detection of daily physical activities using a triaxial accelerometer. Med. Biol. Eng. Comput., 41, 296–301.

Meichenbaum, D. and Novaco, R. (1985) Stress inoculation: a preventative approach. Issues Ment. Health Nurs., 7, 419–435.

Melillo, P., Bracale, M. and Pecchia, L. (2011) Nonlinear heart rate variability features for real-life stress detection. Case study: students under stress due to university examination. Biomed. Eng. Online, 10, 96.

Miller, G.E., Chen, E. and Zhou, E.S. (2007) If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. Psychol. Bull.,133, 25–45.

Morris, M.E., Kathawala, Q., Leen, T.K., Gorenstein, E.E., Guilak, F., Labhard, M. and Deleeuw, W. (2010) Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. J. Med. Internet Res., 12, p. e10.

Okada, Y., Yoto, T.Y., Suzuki, T., Sakuragawa, S. and Sugiura, T. (2013) Wearable ECG Recorder with Acceleration Sensors for Monitoring Daily Stress: Office Work Simulation Study. 35th

Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC), pp. 4718–4721. IEEE.

Orbach-Zinger, S., Ginosar, Y., Elliston, J., Fadon, C., Abu-Lil, M., Raz, A. and Eidelman, L.A. (2012) Influence of preoperative anxiety on hypotension after spinal anaesthesia in women undergoing Caesarean delivery. Br. J. Anaesth., 109, 943–949.

Pallavicini, F., Algeri, D., Repetto, C., Gorini, A. and Riva, G. (2009) Biofeedback, virtual reality and mobile phones in the treatment of Generalized Anxiety Disorder (GAD): a phase-2 controlled clinical trial. J. Cybertherapy Rehabil., 2, 315–327.

Pallavicini, F., Gaggioli, A., Raspelli, S., Cipresso, P., Serino, S., Vigna, C. and Riva, G. (2013) Interreality for the management and training of psychological stress: study protocol for a randomized controlled trial. Trials, 14, 191.

Popovic, S., Horvat, M., Kukolja, D., Dropuljic, B. and Cosic, K. (2009) Stress inoculation training supported by physiology-driven adaptive virtual reality stimulation. Stud. Health Technol. Inf., 144, 50–54.

Quinlan, J.R. (1993) C4. 5: Programs for Machine Learning, Vol. 1. Morgan Kaufmann.

Rajendra Acharya, U., Faust, O., Adib Kadri, N., Suri, J.S. and Yu, W. (2013) Automated identification of normal and diabetes heart rate signals using nonlinear measures. Comput. Biol. Med., 43, 1523–1529.

Richman, J.S. and Moorman, J.R. (2000) Physiological time-series analysis using approximate entropy and sample entropy. Am. J. Physiol.-Heart Circ. Physiol., 278, H2039–H2049.

Riva, G., Raspelli, S., Algeri, D., Pallavicini, F., Gorini, A., Wiederhold, B.K. and Gaggioli, A. (2010) Interreality in practice: bridging virtual and real worlds in the treatment of posttraumatic stress disorders. Cyberpsychol. Behav. Soc. Network., 13, 55–65.

Riva, G., Gaggioli, A., Grassi, A., Raspelli, S., Cipresso, P., Pallavicini, F. and Donvito, G. (2011) NeuroVR 2—a free virtual reality platform for the assessment and treatment in behavioral health care. Stud. Health Technol. Inf., 163, 493–495.

Robnik-Šikonja, M. and Kononenko, I. (2003) Theoretical and empirical analysis of ReliefF and RReliefF. Mach. Learn., 53, 23–69.

Rossman, M.L. (2010) Guided imagery for self-healing: an essential resource for anyone seeking wellness. HJ Kramer.

Salahuddin, L., Cho, J., Jeong, M.G. and Kim, D. (2007) Ultra Short Term Analysis of Heart Rate Variability for Monitoring Mental Stress in Mobile Settings. 29th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society, EMBS 2007, pp. 4656–4659. IEEE.

Sansone, M., Fusco, R., Pepino, A. and Sansone, C. (2013) Electrocardiogram pattern recognition and analysis based on artificial neural networks and support vector machines: a review. J. Healthcare Eng., 4, 465–504.

Selvaraj, J., Murugappan, M., Wan, K. and Yaacob, S. (2013) Classification of emotional states from electrocardiogram signals: a non-linear approach based on hurst. Biomed. Eng. Online, 12, 44.

Serino, S., Triberti, S., Villani, D., Cipresso, P., Gaggioli, A. and Riva, G. (2014) Toward a validation of cyber-interventions for stress disorders based on stress inoculation training: a systematic review. Virtual Real., 18, 73–87.

Sloan, R.P., Shapiro, P.A., Bagiella, E., Boni, S.M., Paik, M., Bigger Jr, J.T. and Gorman, J.M. (1994) Effect of mental stress throughout the day on cardiac autonomic control. Biol. Psychol., 37, 89–99.

Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R. and Jacobs, G.A. (1983) Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.

Sun, Y. and Cheng, A.C. (2012) Machine learning on-a-chip: a high-performance low-power reusable neuron architecture for artificial neural networks in ECG classifications. Comput. Biol. Med., 42, 751–757.

Sweeney, K.T., Kearney, D., Ward, T.E., Coyle, S. and Diamond, D. (2013) Employing Ensemble Empirical Mode Decomposition for Artifact Removal: Extracting Accurate Respiration Rates from ECG Data During Ambulatory Activity. 35th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC), pp. 977–980. IEEE.

Takai, N., Yamaguchi, M., Aragaki, T., Eto, K., Uchihashi, K. and Nishikawa, Y. (2004) Effect of psychological stress on the salivary cortisol and amylase levels in healthy young adults. Arch. Oral Biol., 49, 963–968.

Tartarisco, G., Baldus, G., Corda, D., Raso, R., Arnao, A., Ferro, M. and Pioggia, G. (2012) Personal health system architecture for stress monitoring and support to clinical decisions. Comput. Commun., 35, 1296–1305.

Vapnik V. (1995) The Nature of Statistical Learning Theory. Springer.

Varanini, M., Tartarisco, G., Billeci, L., Macerata, A., Pioggia, G. and Balocchi, R. (2014) An efficient unsupervised fetal QRS complex detection from abdominal maternal ECG. Physiol. Meas., 35, 1607.

Villani, D., Repetto, C., Cipresso, P. and Riva, G. (2012) May I experience more presence in doing the same thing in virtual reality than in reality? An answer from a simulated job interview. Interact. Comput., 24, 265–272.

Villarejo, M.V., Zapirain, B.G. and Zorrilla, A.M. (2013) Algorithms based on CWT and classifiers to control cardiac alterations and stress using an ECG and a SCR. Sensors, 13, 6141–6170.

Vuorela, T., Seppa, V.P., Vanhala, J. and Hyttinen, J. (2010) Design and implementation of a portable long-term physiological signal recorder. IEEE Trans. Inf. Technol. Biomed., 14, 718–725.

Wang, L.X. and Mendel, J.M. (1992) Generating fuzzy rules by learning from examples. IEEE Trans. Syst. Man Cybern., 22, 1414–1427.

Witten, I.H. and Frank, E. (2005) Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann.

Wood, D.P., Mclay, R.L., Webb-Murphy, J., Wiederhold, M.D., Spira, J.L., Pyne, J.M. and Wiederhold, B.K. (2013) Virtual reality graded exposure therapy with arousal control for the treatment of combat related posttraumatic stress disorder: a follow up case series. Stud. Health Technol. Inf., 199, 141–145.