International Journal of Neural Systems, Vol. 0, No. 0 (2005) 1–17

STRESS DETECTION USING WEARABLE PHYSIOLOGICAL AND SOCIOMETRIC SENSORS

Oscar Martinez Mozos*
DETCP, Technical University of Cartagena,
Plaza del Hospital, n1, 30202, Cartagena, Spain
E-mail: oscar.mozos@upct.es
www.upct.es

Virginia Sandulescu
Faculty of Automatic Control and Computer Science, Politehnica University of Bucharest,
313 Splaiul Independentei, Bucharest, 060042, Romania
E-mail: s_virg@yahoo.com
www.upb.ro

Sally Andrews
Division of Psychology, Nottingham Trent University,
Burton Street, Nottingham, NG1 4BU, UK
E-mail: s.andrews@lincoln.ac.uk
www.ntu.ac.uk

David Ellis
Department of Psychology, Lancaster University,
Bailrigg, Lancaster, LA1 4YW, UK
E-mail: d.a.ellis@lancaster.ac.uk
www.lancaster.ac.uk

Nicola Bellotto
School of Computer Science, University of Lincoln,
Brayford Pool, Lincoln, LN67TS, UK
E-mail: nbellotto@lincoln.ac.uk
www.lincoln.ac.uk

Radu Dobrescu
Faculty of Automatic Control and Computer Science, Politehnica University of Bucharest,
313 Splaiul Independentei, Bucharest, 060042, Romania
E-mail: rd_dobrescu@yahoo.com
www.upb.ro

Jose Manuel Ferrandez
DETCP, Technical University of Cartagena,
Plaza del Hospital, n1, 30202, Cartagena, Spain
E-mail: jm.ferrandez@upct.es
www.upct.es

Stress remains a significant social problem for individuals in modern societies. This paper presents a machine learning approach for the automatic detection of stress of people in a social situation by combining two sensor systems that capture physiological and social responses. We compare the performance using different classifiers including support vector machine, AdaBoost, and k-nearest neighbour. Our experimental results show that by combining the measurements from both sensor systems, we could accurately discriminate between stressful and neutral situations during a controlled Trier social stress test (TSST). Moreover, this paper assesses the discriminative ability of each sensor modality individually and considers their suitability for real time stress detection. Finally, we present an study of the most discriminative features for stress detection.

Keywords: activity monitoring; assistive technologies; physiology; sensors; signal classification; sociometric badges; stress; stress detection; wearable technology

1. Introduction

While the world's population continues to rise, the ratio of caregivers to those who require care is rapidly decreasing [3,5]. New technologies, however, can offset this concern by automatically monitoring the physical and mental health of individuals across a variety of contexts. Specifically, stress remains a significant social problem for individuals and societies [2,26,67,82]. Moreover, stressful situations encountered in everyday life can lead to additional negative mental states including depression and anxiety. While early detection could improve an individual's quality of life, the automatic detection of stress may also reveal key mechanisms and trigger points that underlie related negative behaviours and cognitions.

Stress is a natural reaction of the human body to an outside perturbing factor. Typical physiological responses include variations in heart rate, pulse, skin temperature, pupil dilation, and electro-dermal activity amongst others [60,77,81]. While small levels of stress may have beneficial effects on the body, stress often negatively impacts attention, memory, and decision-making [63,71]. High levels of stress in the long-term also correlate with a variety of negative health outcomes including anxiety, depression and premature ageing [29,59].

Psychosocial stress remains a large problem for society as a whole and has a detrimental impact on health care systems and the economies. According to the Mental Health Foundation in the UK [2], around 12 million adults in the UK visit their general practitioner (GP) each year with mental health problems, many of which are related to or brought on by stress. As a consequence, 13.3 million working days are lost per year due to stress related illnesses. Moreover, the World Health Organization [5] recently calculated that stress costs around 8.4 million pound sterlings to UK enterprises. Therefore, treating negative mental states has become a priority in our societies, and numerous international organizations also consider stress to be a significant global problem, particularly within the work place [26,67,82]. Unfortunately, current treatments offered by the National Health Service (NHS) in the UK (e.g. Cognitive Behavioural Therapy (CBT) [13]) have an average patient waiting time of 3-6 months [58].

In this paper, we aim to detect stressful behaviours by analysing measurements provided by several, non-invasive wearable sensors. In particular, we use a wearable sensor that provides real time physiological responses such as electrodermal activity and photoplethysmogram [1]. In addition, we use a sociometric badge [4] to measure the social activity including body movement and voice. The measurements obtained from both sensors are processed and used as input to different classifiers that were trained to discriminate between stressful and neutral situations during a Trier Social Stress Test (TSST) [42]. In this work, we use a binary classification and do not differentiate between different levels of stress. Our results show that the combination of physiological and activity measurements is able to discriminate with a high level of confidence between stressful and neutral situations on people while engaging across several activities. The main novelty of our systems is the combination of wearable sensors that makes the system a realistic option for measuring stress in social situations. As far as we now this is the first time that these sensor modalities are combined to detect stress. Finally we present a study of the most discriminative

*Corresponding author

features obtained from the different sensors.

In this work, we train a personal classifier for each participant because an individuals response to a specific stressor is often highly variable. Therefore, we adopted a personalised systems approach.

The remainder of this paper is organised as follows. After considering some related work in Sect. 2, we introduce the physiological wearable sensor and the sociometric badge in Sect. 3. In Sect. 4 we describe our experimental setup. The process for data collection is explained in Sect. 5, and our classification methods are introduced in Sect. 6. Feature selection is explained Sect. 7. Experimental results are presented in Sect. 8. Finally, we conclude in Sect. 9.

2. Related Work

Research concerning the automatic monitoring of mental states has grown exponentially during the last decade. The resulting technologies aim to help patients monitor their own conditions while also supporting carer-providers. Examples include the detection and monitoring of stress [47, 62, 68, 74, 83], classification of different emotional states [18, 20, 40, 41], depression monitoring [7, 37, 73], obsessive compulsive disorder [12], behaviour classification [55], or cardiac states [48, 49, 65].

Various systems have been proposed for stress monitoring and/or detection using different physiological sensors. For example, the work reported in [68] uses a wearable sensor to record electrodermal activity (EDA). This work applies the Montreal Imaging Stress Task (MIST) [28] to induce states of stress on participants, and the results indicate that EDA measurements can be used to discriminate stress from cognitive loads with 82.8% accuracy. Although this work uses cross-validation for reporting results, it does not present results of individual classifiers for each participant. In contrast, we study the personalization of the detector for each participant.

Similarly, a mobile elecrocardiogram (ECG) sensor was used in [62] to monitor stress by analysing ultra short term heart rate variability features during a Stroop test [72]. The authors conclude that ultra short term features could be used for stress monitoring, but they do not apply any machine learning classification tool to support this claim. In comparison, we present results by applying classifiers to the collected data.

In [43], the authors analyzed EDA measurements together with skin temperature to discriminate between stressed and unstressed situations. Authors report correct classification rates of 96.6% in the test data. However, the paper does not detail the final applied classification procedure.

Finally, the work in [83] combines EDA signals, blood volume pulse, pupil diameter, and skin temperature to detect stress in participants while they took part in a modified Stroop test. The results report classifications of 90.1% accuracy. However, this system seems difficult to implement as a daily life wearable technology due to the measurement of pupil diameter. In contrast, our proposed system is easier to wear by a patient.

The above mentioned works use physiological measurements in isolation and do not include social information. In contrast, we present a multi-modal system for stress monitoring that includes a sociometric sensor that records social activity and thus increases final detection rates. In addition, we focus on personalized systems by analysing individual results for each participant. Moreover, our experimental setup is based on the TSST which, rather than generate a stressful situation using a computer-based task, provides neutral and stressful conditions in a social setting.

Voice is also used as indicator for stress. The *StressSense* system introduced in [47] is a voice based stress detection mobile app that uses microphones to record the voice of the participant and classifies the obtained features using Gaussian Mixture Models. This work reports classification accuracies of 81% and 76% for indoor and outdoor environments, respectively. In our research, we also use the voice as one of the modalities for stress detection, but we combine it with other physiological and activity measurements, thus increasing the individual classification accuracy of our results.

Multi-modal approaches for stress detection have also been developed where activity and voice signals are added to physiological measurements in order to improve stress detection and monitoring.

The work in [74] presents a activity-aware mental stress detection scheme that combines Electrocardiogram (ECG), galvanic skin response (GSR), and accelerometer measurements from participants across three activities, i.e. sitting, standing, and walking, while users were subjected to mental stressors. This work applies the Stroop test [72] to induce stress. In our work, however, we are interested in detecting stress in social situations, therefore we apply the TSST to our participants. In addition, we use the sociometric badge to increase the modalities for stress detection by including voice and body movement.

Authors in [45] introduce a multi-modal approach in which a rich set of activity and physiological measurements are used to monitor stress. In addition, they use a combination of facial expressions, eye movements, head movements, heart rate, skin temperature, GSR, mouse finger pressure, behavioural data from user interaction activities with the computer, and performance measures. The obtained results show that the presented activity measurements strongly correlate with the task workload and ensure reliability for further classification approaches. However, the proposed system is thought to be implemented in a desktop-based workplace environment, which restricts the situations in which it can be tested. Our system, however, is based on wearable sensors that allow the participants to freely move during their activities.

In addition to stress, the work presented in [75] aimed to detect deception while playing a poker game by including measurements like voice variation, skin conductance, and heart rate. The experimental results show that it is possible to develop simple linear models with high accuracy that can be used to identify stress and bluffing for real money nolimit Hold'em tournaments. However, this system has been tested only in poker games where spikes in stress levels are artificially high. In contrast, we aimed to develop stress detectors in different social activities.

The *LiveNet* system presented in [76] is a complete wearable hardware and software system for long term monitoring which, according to the authors, has potential applications for monitoring soldiers and Parkinson's patients. It could also detetct epilepsy seizures. However, these applications remain under development. The LiveNet system can also include a sociometric badge similar to the one we use in our research, but no classification results are presented using that sensor.

Finally, in [64], we applied the same wearable physiological sensor to detect stress in people. However, that work only reports on 5 participants. In contrast, in this work we combine the physiological signals with the activity signals from the sociometric sensor and thus we improve classification results. In addition we use 18 participants in our experiments thus providing better reliability in the results.

One of the main contributions of our work is the use of the sociometric badge for stress detection. This badge provides a way of capturing unconscious *social signals* [56] that have recently provided researchers with a novel and indirect way of capturing an individual's thoughts and cognitive states. Methods for assessing social signals also offer new tools for measuring levels of stress, particularly where verbal or written reports concerning underlying cognitive states may be incomplete or inaccurate. Therefore, we think the sociometric badge can be a powerful tool to monitor stress.

3. Sensor Modalities

In this section we describe the two sensor devices used during our experiments and the features extracted from them.

To measure the physiological signals we use a wireless sensor [1] that is worn as a wristband (Fig. 1 left) on the non-dominant hand of a subject and it is equipped with a set of electrodes situated on the fingers. This sensor connects wireless to a computer through a communication station. This setup allows the participant wearing the sensor to move freely during the experiments while signals are sent wirelessly to a computer.

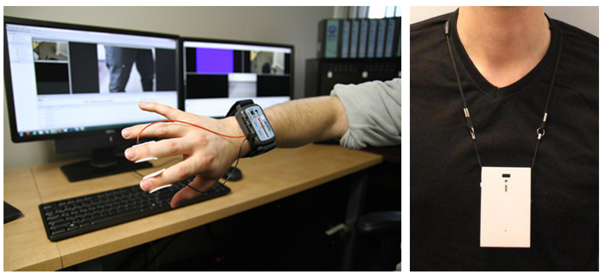

Fig. 1. The left picture shows the wireless sensor worn as a wristband. The right image depicts the sociometric sensor worn as a conference badge.

From the electrodes situated on the fingers we obtain the three different measurements: the electrodermal activity (EDA), the photoplethysmogram (PPG), and heart rate variability (HRV). The EDA signal, also called, skin conductance activity or galvanic skin response, is an indication of perspiration.

A transient increase on the EDA signal is proportional to sweat secretion [27] and it is related to stress [17]. The PPG signal, also termed blood volume pulse, is obtained using a pulse oxiometer which illuminates the skin and measures the differences in light absorption. The amount of light that returns to the PPG sensor is proportional to the volume of blood in the tissue [57]. Finally, the HRV signal represents the beat-to-beat variability over a given period of time and is computed by calculating the standard deviation of the average of normal-to-normal heartbeats [57].

During our experiments, the EDA and PPG physiological signals were acquired at a 1000 Hz sampling frequency. Following acquisition, the signals were down-sampled to 10 Hz. A filtering and artefact removal approach was also applied by using the routines included in the AcqKnowledge software [1]. Specifically, the PPG signal was filtered using a band pass filter, with a low frequency of 0.5 Hz and a high frequency of 35 Hz. The EDA signal was filtered using a low pass filter with the cut-off frequency of 0.5 Hz. We obtain templates for the PPG signal by applying autocorrelation. The final list of recorded physiological features and their corresponding description are provided in Table 1.

Table 1. List of physiological features from the Biopac wearable sensor.

Feature	Description
eda	raw electro-dermal activity (EDA).
eda_f	filtered EDA.
ppg	raw pulse plethysmograph (PPG).
ppg_t	PPG template using autocorrelation.
ppg_r	PPG rate variability.
hrv	heart rate variability (HRV).

The second sensor we used in our system is the sociometric sensor [53], a small device that is worn around the neck like a conference badge (Fig. 1 right). The sociometric badge is equipped with a microphone to record speech, an accelerometer to measure degree and direction of people's movement, a Bluetooth transmitter to measure the proximity of other sensors, and an infrared transmitter to measure when two sensor wearers are facing one another. In this work we only used the measurements provided by the microphone and the accelerometer. The data recorded from the sociometric badge can be translated into useful measures of social behaviour [56].

In our experiments we recorded the signals from the sociometric badge at 10Hz. The samples were stored on the badge's internal memory and downloaded to the computer in an off-line process. The list of features that we used in our system based on the measurements provided by the sociometric badge is as follows:

(1) Body movement (bm): the normalized acceleration magnitude over the 3 axis of movement.
(2) Body movement activity (bm_{act}): absolute value of the first derivative of the accelerometer's energy.
(3) Body movement rate (bm_r): second derivative of energy. It indicates the direction of change in someone's activity level.
(4) Posture activity (pos_{act}): absolute angular velocity.
(5) Posture rate (pos_r): angular acceleration.
(6) Posture left right (pos_{lr}): orientation angle of the badge in left-right plane.
(7) Posture front back (pos_{fb}): orientation angle of the badge in front-back plane.
(8) Speak ($voiced$): takes values 0 when the person is not speaking, and 1 when the person is speaking.
(9) Silence ($unvoiced$): takes values 1 when the person is speaking, and 0 when the person is speaking.
(10) Speech volume front (vol_f): average absolute value of amplitude of the front microphone.
(11) Speech volume back (vol_b): average absolute value of amplitude of the back microphone.
(12) Volume consistency front ($volc_f$): measurement of change in speech volume.
(13) Volume consistency back ($volc_b$): measurement of change in speech volume.
(14) Frequency and amplitude front ($hz0_f$, $amp0_f$), ($hz1_f$, $amp1_f$), ($hz2_f$, $amp2_f$), ($hz3_f$, $amp3_f$): pairs of (frequency, amplitude) for the 1st / 2nd / 3rd / 4th strongest peak in the frequency spectrum.
(15) Frequency and amplitude back ($hz0_b$, $amp0_b$), ($hz1_b$, $amp1_b$), ($hz2_b$, $amp2_b$), ($hz3_b$, $amp3_b$): pairs of (frequency, amplitude) for the 1st / 2nd / 3rd / 4th strongest peak in the frequency spectrum.
(16) Front pitch ($pitch_f$): pitch of the voice from the

front mic correlated with the fundamental frequency of the voice signal.

(17) Back pitch ($pitch_b$): pitch of the voice from the front mic correlated with the fundamental frequency of the voice signal:

(18) Front pitch volume ($frontp_v$) and back pitch volume ($backp_v$) as defined in [52–54].

More detailed descriptions of the previous features can be found in [52–54].

4. Experimental Protocol

To check the validity of our stress detector system we prepared an experimental setup where participants experienced different stressful situations. The work in [29] reviews more than 200 stress experiments, and concludes that the most effective tasks for inducing stress include public speaking and cognitive tasks. Therefore, our final design is based on the Trier Social Stress Test (TSST) [42], which includes both public speaking and cognitive tasks that place participants under high cognitive load.

The TSST is a very popular controlled experimental set-up and it has been used in more than 4000 sessions during the last few decades [29]. This test consists of a neutral task followed by a public speaking task, a cognitive task, and a final neutral task. Each neutral task consists of 2 minutes of predefined standard questions including: "How do you find the weather today?" and "How did you get here?". The public speaking segment is a 5 minute interview for a desired job. After this, a cognitive task involves the participant counting back in steps of 13, starting from 1022. All the previous tasks are performed in front of a live audience and video camera. The camera, however, is only used to increase stress levels [29] and recordings are not stored. The neutral tasks are considered as *non-stressful* situations, while the speaking and cognitive tasks are considered *stressful* conditions. The TSST experiment is used in this work as a valid proof of concept to show the capabilities our system to detect stress in social situations.

In addition we ask the participants to fill in a State Trait Anxiety Inventory (STAI) [70] to measure their levels of anxiety. The STAI questionnaire is divided into two sections. The state section contains 20 items that measure state anxiety, that is, how an individual feels right now; and the trait section contains a further 20 items that measure trait anxiety, that is, how an individual feels generally. Items are rated on a 4-point scale running from (1=Almost never to 4=Almost always). The state section of the STAI was completed before and after the TSST session to ensure our protocol was having the desired effect on stress levels.

Our protocol for the TSST was as follows. When a participant entered the room, she was given verbal and written information about the procedures involved in the experiment. The participant was then asked to fill in a consent form and to confirm that she did not suffer from any cardiovascular or anxiety disorder that might be affected by experiencing stress or that might affect the results. After being briefed, the participant was asked to fill in the STAI form. This provided an estimate current and general levels of stress. After fitted with both sensors, the participant was asked predefined neutral questions for 2 minutes in order to determine each individual's baseline measurements, which we used as a neutral state. Afterwards, the participant was asked to sit at a desk and prepare a presentation for a mock job interview for 3 minutes. They were provided with pen and paper. When 3 minutes of time expired, she was asked to hand back the sheet of paper, stand up in a predefined location inside the room, and begin her presentation. The participant was encouraged to speak continuously during 5 minutes. If the participant stopped during the presentation, at the first pause, she was told about the remaining time and asked to continue. At the next pause, she was asked a set of predefined typical interview questions including: "What are your strengths/weaknesses?", "Where do you see yourself in 5 years?" and so on. Following the presentation, the participant was asked to complete a cognitive task by counting backwards in steps of 13 from 1022 and a 5 minutes timer was started. If the participants made a mistake, she was asked to start the countdown again from the beginning (from 1022). At the end of this cognitive task, the participant was given a short time to relax while being debriefed. Another two minutes of neutral questions were then recorded. Finally, the participant was asked to re-complete the STAI questionnaire. A flowchart indicating the steps of our protocol is presented in Fig. 2.

The times for each tasks during the TSST session are tentative and for most tasks, the duration of each differs by a few seconds between participants.

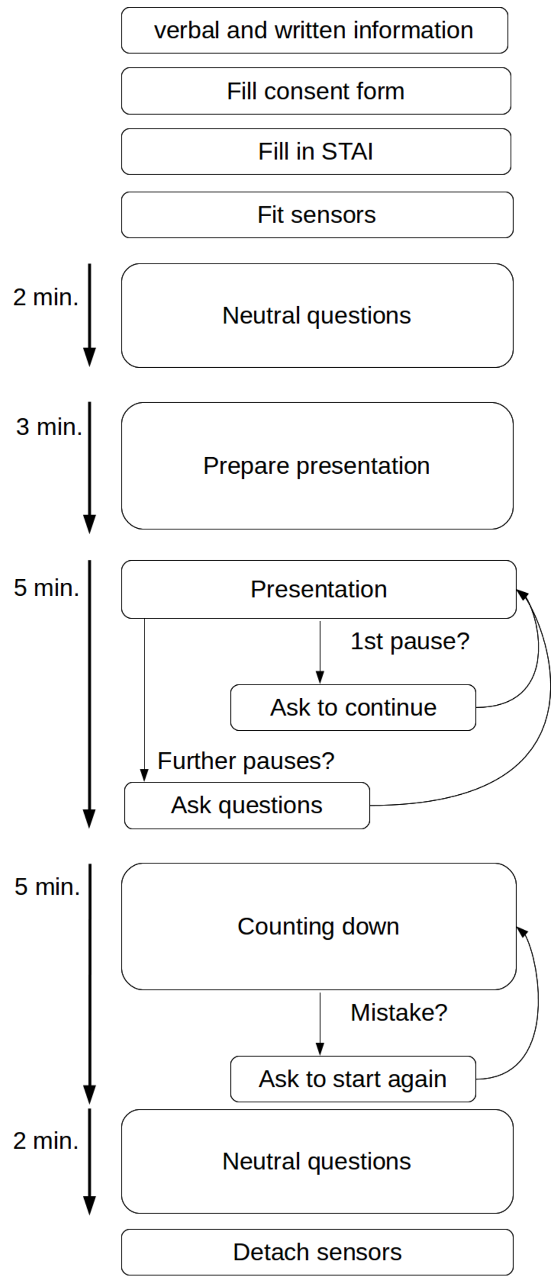

Fig. 2. Flow chart of our TSST session.

5. Data Collection

Eighteen participants $P_1 - P_{18}$, who completed the TSST, were volunteering students from the School of Psychology at the University of Lincoln. The participants were aged 18 to 39, and included males and females. All participants signed a consent form before taking part. In addition, ethical approval for the experiment was obtained from the School of Psychology Research Ethics Committee at the University of Lincoln, following the British Psychological Society (BPS) ethical guidelines.

Each participant took part individually in one TSST session as introduced in Sec. 4. The participants wore both sensors at the same time during the TSST session, and the signals from both sensors were recorded and synchronised using the closest timestap. The final sample frequency for both sensor modalities was 10Hz. This value is below the maximum recommended for EDA measurements [14]. The complete sets of signals for each sensor are listed in Sect. 3. The final set of sampled signals for each time step t is thus defined as:

$$x^t = \{x^t_{phy}, x^t_{badge}\}, \tag{1}$$

where x^t_{phy} is the vector containing the physiological signals as:

$$x^t_{phy} = \{eda^t, eda^t_t, ppg^t, ppg^t_t, hrv^t\}, \tag{2}$$

and x^t_{badge} is the vector containing the activity signals from the sociometric as:

$$\begin{aligned} x^t_{badge} = \{ & bm^t, bm^t_{act}, bm^t_r, pos^t lr, pos^t_{fb}, \\ & voiced^t, unvoiced^t, vol^t_f, vol^t_b, volc^t_f, volc^t_b, \\ & hz0^t_f, amp0^t_f, hz1^t_f, amp1^t_f, hz2^t_f, amp2^t_f, \\ & hz3^t_f, amp3^t_f, hz0^t_b, amp0^t_b, hz1^t_b, amp1^t_b, \\ & hz2^t_b, amp2^t_b, hz3^t_b, amp3^t_b, pitch^t_f, pitch^t_b \\ & frontp^t_v, backp^t_v\}. \end{aligned} \tag{3}$$

The synchronised signals recorded from each participant P_k were stored in the corresponding dataset D_k. The total number of synchronized samples for each participant P_k is shown in Table 2. As explained in Sect. 4, our TSST sessions are composed of stressful and neutral scenarios. Therefore, each entry x_t in the dataset D_k is labeled as *stressed*, or *neutral* depending on the corresponding task (c.f. Sect. 6), i.e. we assume the participant is not stressed during the neutral activities as discussed in Sect. 4. Thus, the recorded dataset D_k for each person was composed of the measurements obtained at each time interval x_t together with their corresponding label as $D_k = \{(x_t, l_t)\}$, with $l_t \in L = \{stress, neutral\}$. The corresponding number of stressful and neutral samples in each dataset are also shown in Table 2. The average sample size for each participant is 13745 which is much higher than the dimension of the feature vector.

Table 2. Sample Size for Each Participant

Participant	Total $\|D_k\|$	Stress	Neutral
P_1	12800	8200	4600
P_2	11620	8400	3220
P_3	13550	8350	5200
P_4	13450	8460	4990
P_5	13740	8760	4980
P_6	13000	8310	4690
P_7	15940	8600	7340
P_8	13610	7810	5800
P_9	12900	8420	4480
P_{10}	14200	8900	5300
P_{11}	15680	8660	7020
P_{12}	13530	8660	4870
P_{13}	14120	8560	5560
P_{14}	13900	8810	5090
P_{15}	13350	8350	5000
P_{16}	14500	8800	5700
P_{17}	13480	8760	4720
P_{18}	14040	8820	5220
Average $\pm$ std	13745 $\pm$ 685	8535 $\pm$ 220	5210 $\pm$ 608

Some examples of the collected features during the TSST session are given in Figure 3. In particular, we show the features PPG template (ppg_t), EDA (eda_f), HRV (hrv), and body movement rate (bm_r) for participants P_6 and P_{16}. The signals are shown with the corresponding task during the TSST session.

6. Classification

In order to predict the state of each person at each point in time we compared different classification approaches. In particular, we use SVM [15, 25], AdaBoost [31], and K-nearest neighbour (KNN) [10]. In this work we use a binary classification and do not differentiate between different levels of stress.

SVM-based predictors are popular in different areas [9, 19, 23, 24, 32, 38, 39] including the classification mental states using physiological signals [6,8,12, 30, 37, 40, 55, 74, 83]. SMVs were used in our experiments because the size of a trained SVM model is often much smaller than the volume of training data required in order to be successful. They can also be developed in real time. In this work we compare two different kernel types: radial basis function (RBF) and linear.

AdaBoost is a popular boosting classifier applied in several areas [11,51,80] including biomedical, cognitive signals, and diagnosis [16, 21, 34, 50, 61, 78, 79]. Adaboost is a meta classifier that improves the classification capabilities of a weak classifier. Following [51], we created a single-feature weak classifier from each feature described in Section 3:

$$h_j(x) = \begin{cases} 1 \text{ if } p_j f_j(x) < p_j \theta_j \\ 0 \text{ otherwise.} \end{cases} \quad (4)$$

where θ_j is a threshold and p_j is either -1 or 1 and thus representing the direction of the inequality. The AdaBoost algorithm determines for each weak classifier $h_j(x)$ the optimal values for θ_j and p_j, such that the number of misclassified training examples is minimized. The output of the AdaBoost algorithm is an ensemble of weak classifiers weighted by their discriminative power. The final ensemble may contain a subset of weak classifiers, i.e. a subset of the original features.

KNN classifiers are instance-based clasifiers that are used in several medical applications [22,44,66,69]. The KNN algorithm assigns the label of the k training instances that are closer in the feature space using a majority vote.

In this paper, we trained one individual classifiers C_k for each participant P_k using the corresponding dataset D_k. As introduced in the previous Sect. 5, each dataset D_k contains the synchronized physiological and activity signals for the corresponding participant P_k. For each individual C_k, we used 75% of the corresponding dataset D_k for training and the remaining 25% for testing. To create these sets we used a random stratified selection to ensure the same class distribution in both the training and test sets. Afterwards, the attribute values of the training and test sets were scaled to the range $[-1, 1]$.

The classification results for each personalized classifier C_k are evaluated using *accuracy*, *precision*, and *recall*, which are defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

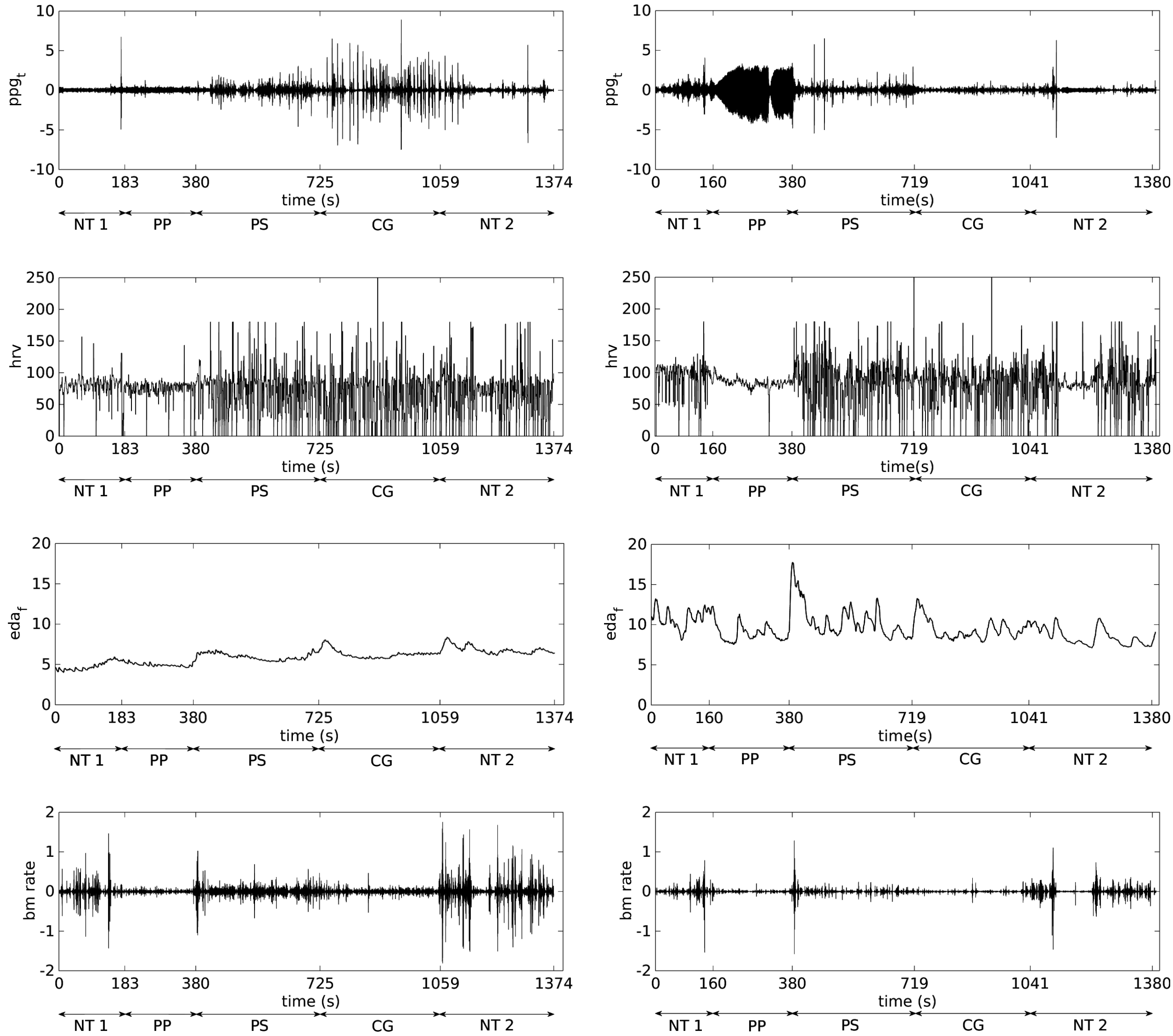

Fig. 3. Example signals recorded during the TSST sessions for participant P_6 (left column), and participant P_{16} (right column). The x-axis additionally presents the task the person was carrying out at each instant: NT 1 (First Neutral Task), PP (Presentation Preparation), PS (Public Speaking), Cognitive Task (CG), NT 2 (Second Neutral Task).

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

and finally *recall* is given by:

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

where TP indicates true positives, TN true negatives, and FN false negatives. In our case we consider *positive* the examples labeled as *stress*. The values for *accuracy*, *precision*, and *recall* lie in the range $[0, 1]$ with values closer to 1 indicating better results.

7. Feature Selection

The classification results may improve by selecting a subset of the features that provides best discrimination capabilities. In this paper, we apply a correlation-based feature selection (CFS), which is a filter method that selects a subset of features that are highly correlated with the class and uncorrelated with each other [35]. In our experiments, we apply CFS to each individual dataset D_k to select the best features for each participant. The resulting reduced datasets are then used to train the corresponding individual classifiers C_k. We used this method as preprocessing for the SVM and KNN classifiers.

In the case of Adaboost we do not need to apply any filtering method since the algorithm already applies a weighted selection of weak classifiers to create the final ensemble. In our case, the best T features can be obtained by selecting the T weak classifiers with highest weights in the ensemble [11, 51, 80].

8. Experimental Results

To ensure our approach for stress detection was valid, we ran a series of experiments that involved analysing the recorded physiological and activity signals of people under stress. We run one TSST for each participant, and each participant wore both sensors during the session.

We first present stress detection results when applying different classifiers to the synchronised physiological and activity signals. In addition, we analyse the capabilities of each individual sensor modality, and compare the classification results with those obtained in combination. Moreover, we analyse the discrimination capabilities of each feature. Finally, we discuss the results of the STAI questionnaires.

8.1. *Stress Detection using Physiological and Activity Sensors*

In a first experiment we analysed the results of applying our proposed system using a combination of physiological and sociometric sensor.

We trained an individual classifier C_k for each participant P_k using the data D_k corresponding to her TSST session. Each participant wore the sociometric badge and the wearable sensor during the corresponding TSST session and features from both sensors were synchronised using the closest timestamp (cf. Section 5).

We then divided each dataset D_k into disjoint training and test sets. The training set contains 75% of the samples in D_k and the test set contains the remainder 25%. To create these sets we used a random stratified selection to ensure the same class distribution in both the training and test sets. Afterwards, the attribute values of the training and test sets were scaled to the range $[-1, 1]$.

We first present classification results using the full set of features. We compare four classifiers: SVM with a RBF kernel, SVM with linear kernel, AdaBoost, and KNN. For the RBF-based SVM we selected the best C and γ for each participant P_k using grid-search and cross-validation in the corresponding training data [36]. The parameters $C = 10$ for the linear SVM, $T = 300$ for AdaBoost, and $k = 3$ for KNN were empirically selected.

Table 3 shows the classification results for the four methods averaging over the 18 participants. As we can see from the table, in our experiments the AdaBoost classifier and the RBF-based SVM provide similar classification results, although AdaBoost provides slightly higher correct accuracy and precision rates. In both cases, the results are higher than in the linear SVM and KNN classifiers.

Table 3. Classifications for the combined sensors.

Method	Accuracy	Precision	Recall
AdaBoost (T=300)	0.94±0.03	0.94±0.03	0.96±0.02
RBF kernel SVM	0.93±0.03	0.93±0.03	0.96±0.01
Linear kernel SVM	0.85±0.04	0.84±0.04	0.94±0.02
KNN	0.87±0.03	0.87±0.03	0.94±0.02

In Table 4 we present in more detail the classification results of the AdaBoost classifier for each participant. Values for accuracy, precision and recall are similar for all the participant indicating that our personalised classification approach is suitable for different individuals.

Table 4. Classification Results using Combined Sensors and AdaBoost

Participant	Accuracy	Precision	Recall
P_1	0.95	0.95	0.97
P_2	0.91	0.92	0.96
P_3	0.87	0.89	0.92
P_4	0.96	0.96	0.98
P_5	0.99	0.99	0.99
P_6	0.89	0.90	0.93
P_7	0.96	0.97	0.97
P_8	0.93	0.94	0.94
P_9	0.93	0.94	0.95
P_{10}	0.92	0.92	0.95
P_{11}	0.95	0.95	0.96
P_{12}	0.96	0.96	0.97
P_{13}	0.91	0.92	0.94
P_{14}	0.97	0.96	0.98
P_{15}	0.91	0.90	0.95
P_{16}	0.95	0.96	0.96
P_{17}	0.98	0.98	0.99
P_{18}	0.96	0.96	0.97

In addition, we include the average confusion matrix when using the Adaboost classifier in Table 5. The rows indicate the original label of the examples in the test set and the columns indicate the predicted label by the classifier. We observe high classification results in the main diagonal (true classifications) of the confusion matrix, which indicates that the classifier correctly labels the test examples with high accuracy.

Table 5. Average Confusion Matrix for Combined Sensors using AdaBoost

		Predicted Label	
	%	*Stress*	*Neutral*
Original Label	*Stress*	**96.05**	3.95
	Neutral	9.00	**91.00**

To exemplify the behaviour of our classifier we show its prediction for participants P_6 and P_{16} during a complete TSST session in Fig. 4. The continuous line indicates the right label for each time instant (1 for stress and -1 for no-stress), while red diamonds indicate the predicted values by the classifier. The plots show that most of the time the classifier predicts the correct state of the person during the TSST session. We can see in the plot several false alarms. This is likely due to the slow rate of change for some physiological properties between relaxed and stressed situations. Future work will consider these transitions in more detail.

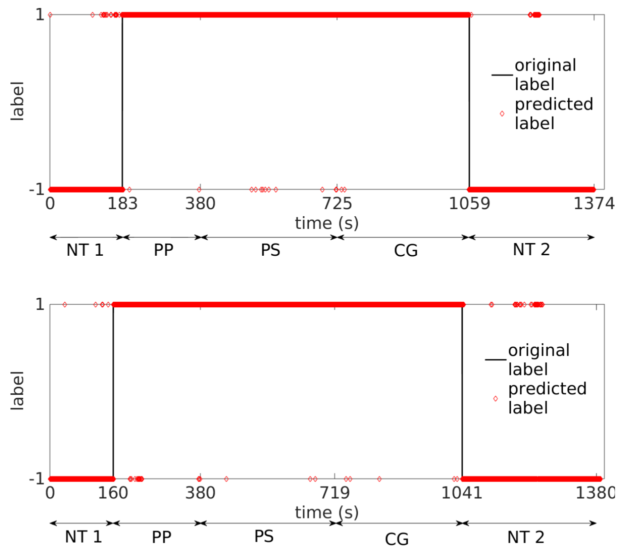

Fig. 4. Classifier behaviour during a complete TSST session for participants P_6 (top) and P_{16} (bottom) using both sensors. The y-axis represents the output of the classifier which is 1 for stressful tasks, and -1 for non-stressful tasks. The x-axis indicates each instant time in seconds during the TSST and the corresponding task: NT 1 (First Neutral Task), PP (Presentation Preparation), PS (Public Speaking), Cognitive Task (CG), NT 2 (Second Neutral Task). The continuous line indicates the right label for each time instant, and red diamonds indicate the predicted values from the classifier.

In addition we analysed the capability of each single sensor to detect stressful situations in the participants. In this way we can compare the detector capabilities of each sensor with respect the combination of both. We believe this information can be useful when deciding he sensor device to use in different situations. We include results using Adaboost since we showed in the previous section that this classifier provides the best classification results in our TSST sessions. For the physiological results we use as input for the classifier the feature vectors x_{phy}, whereas for the activity results we classify the vector x_{badge} (see Sec. 5). Table 6 compares single sensor modalities and the combination of both. In all cases we have applied our AdaBoost classifier. As we can see in the table the combined modality provides better classification results. We also observe that the sociometric sensor provides better prediction results when used as single device. This may be due to the fact that the TSST session requires the participant to speak continuously and participants are also free to move during the experiments. A more detailed analysis of the features is provided in the following section.

Table 6. Comparison of single and combined modalities

Method	Accuracy	Precision	Recall
Physiological	0.79±0.08	0.79±0.09	0.86±0.06
Sociometric	0.89±0.03	0.90±0.03	0.92±0.03
Physiological + Sociometric	0.94±0.03	0.94±0.03	0.96±0.02

8.2. *Feature Analysis*

In this section we apply a selection of features and analyse its impact in the final classification results. As introduced in Section 6, we applied a CFS filtering to each individual dataset D_k. The resulting subsets of selected features are shown in Table 7.

Table 7. Selected subset of features according to CFS

D_k	size	best features
D_1	7	$amp3_f$, $volc_b$, $hz0_f$, bm, vol_b, pos_{lr}, pos_{fb}
D_2	5	$amp3_f$, $volc_b$, pos_r, bm_r, $frontp_v$
D_3	7	vol_f, $amp3_f$, $volc_b$, $volc_f$, $backp_v$, bm_{act}, ppg
D_4	6	$amp3_f$, $volc_b$, $hz0_f$, $pitch_b$, bm_{act}, ppg_r
D_5	5	$amp3_f$, $volc_b$, $amp2_b$, pos_r, bm_r
D_6	5	$amp3_f$, pos_r, $frontp_v$, bm_{act}, ppg
D_7	4	$amp3_f$, $volc_b$, $frontp_v$, ppg_t
D_8	4	$amp3_f$, $volc_b$, $volc_f$, pos_{fb}
D_9	7	$amp3_f$, $volc_b$, $hz1_b$, vol_b, $poslr$, ppg, ppg_r
D_{10}	7	$amp3_f$, $volc_b$, $pitch_f$, pos_{act}, bm_{act}, $poslr$, ppg_r
D_{11}	4	$amp3_f$, $volc_b$, $poslr$, ppg
D_{12}	6	$amp3_f$, $volc_b$, pos_r, $pitch_f$, pos_{fb}, ppg_r
D_{13}	6	$amp3_f$, $volc_b$, $hz1_b$, pos_r, vol_b, bm_{act}
D_{14}	4	vol_f, $amp3_f$, $volc_b$, pos_{fb}
D_{15}	16	vol_f, $voiced$, $hz2_f$, $amp3_f$, $volc_b$, $hz0_f$, pos_r, $frontp_v$, $backp_v$, pos_{act}, $poslr$, ppg_t, ppg_r, $hz2_b$, $hz1_b$, eda
D_{16}	5	$amp3_f$, $volc_b$, $pitch_b$, pos_{act}, bm_{act}
D_{17}	7	$amp3_f$, $volc_b$, bm_r, pos_{act}, bm_{act}, ppg_t, ppg_r
D_{18}	4	$amp3_f$ $volc_b$, $frontp_v$, pos_{act}

We applied the classifiers to the subset of selected features shown in Table 7. The new classification results, which are provided in Table 8, do not show any significant improvement. This may be due to the fact that the original feature vector has a dimension 38, which is low. However, it is interesting to see that we can get high classification results, as for example for SVM with RBF kernel, using a very small subset of features. Moreover, for some participants the selected features by CFS correspond only to one sensor. For example, for participant P_1 we can detect the stress using only the features of the sociometric badge as indicated in Table 7.

Table 8. Classification on selected features

Method	Correct classification rate
RBF kernel SVM	0.92±0.03
Linear kernel SVM	0.80±0.03
KNN	0.62±0.03

As explained in Section 7, the AdaBoost algorithm creates a ranked ensemble of weak classifiers weighted by their discriminative capabilities. When each weak classifiers is constructed using a single feature, then the ranked ensemble is representative of the best features for the classification [11, 51, 80]. The features can appear several times with different weights. To compare the feature selection by AdaBoost with the CFs filter method, in Table 9 we indicate the first 5 most discriminative features, in order of discrimination capabilities, that are selected by AdaBoost for each personalised classifier C_k. The most frequent features that appear in all the classifiers are eda (16% appearance), pos_{act} (15%), $hz3_f$ (14%), bm_{act} (13%), and $amp3_f$ (13%). The rest of features do not appear on the top five or they do with less than 10%.

Table 9. Best 5 features from AdaBoost

Classifier	1st	2nd	3rd	4th	5th
C_1	$hz3_f$	eda	ppg_t	$amp0_f$	bm_{act}
C_2	$hz3_f$	$pitch_f$	eda	pos_{act}	bm_{act}
C_3	pos_{act}	$pitch_b$	$volc_b$	$hz3_f$	$volc_b$
C_4	$hz3_f$	eda	$pitch_f$	pos_{act}	pos_{act}
C_5	$amp3_f$	eda	pos_{lr}	$hz3_f$	eda
C_6	$amp0_b$	$hz3_f$	$hz0_f$	bm_{act}	eda
C_7	$hz3_f$	eda	$pitch_f$	bm_{act}	pos_{act}
C_8	$hz3_f$	ppg	$amp3_f$	eda	$volc_b$
C_9	$amp3_f$	bm_{act}	bm_{act}	$amp1_b$	ppg
C_{10}	bm_{act}	eda	pos_{act}	$hz3_f$	bm_{act}
C_{11}	$hz3_f$	bm_{act}	ppg	ppg_t	$amp3_f$
C_{12}	$amp3_f$	pos_{act}	$hz3_f$	$pitch_f$	ppg
C_{13}	$amp3_f$	pos_{act}	eda	$hz3_f$	ppg
C_{14}	$amp3_f$	$hz3_b$	bm_{act}	eda	pos_{act}
C_{15}	vol_b	$backp_v$	eda	$amp3_f$	bm_{act}
C_{16}	pos_{act}	eda	$amp3_f$	bm_{act}	pos_{act}
C_{17}	pos_{act}	$amp3_f$	eda	ppg_t	pos_{act}
C_{18}	$hz3_f$	$pitch_f$	pos_{act}	$amp3_f$	$amp3_f$

8.3. *Generalization*

In order to check how well our method generalizes we trained the AdaBoost classifier using the first 9 participants and then tested the resulting model in the remainder 9 participants. The classification results are shown in Table 10. The results show a decrease in the accuracy of the classifier when using several participants for training. These results are to be expected given the large between-subject variance to stress [33].

Table 10. General AdaBoost Classifier

Accuracy	Precision	Recall
0.67	0.68	0.76

8.4. *STAI Questionnaire Analysis*

We recorded participants self-reported anxiety levels using the STAI [70] before and immediately after the TSST session (all Cronbach alphas > 0.9 [46]). A paired sample t-test demonstrated that participants felt more stressed following the TSST [$t(34) = 3.54$, $p > 0.01$] suggesting that our procedure provided a stressful environment. While changes in state-levels of anxiety indicate that participants felt more stressed following a TSST session, it does not confirm how levels of stress varied across each part of the protocol independently. During the TSST, we assume that participants became more relaxed during the neutral sections, but this is difficult to validate with 100% accuracy because any additional measures of self-report would both disrupt the flow of the experiment and reduce the ecological validity of the procedure by reminding participants they were taking part in an experiment. Future research aims to explore these changes beyond a laboratory setting in conjunction with real-time self-report.

9. Conclusion

In this paper we have presented a novel approach for stress detection using a combination of wearable physiological and sociometric sensors. The experiments were carried out under controlled conditions during different TSST sessions. Our wearable system allows the state of a participant to be determined at any instant by providing an accurate decision regarding his/her stress state at any time. Our classification results demonstrate that our method and analysis provides a useful tool for real-time stress detection. In the future this may allow other researchers to consider for example, the effect of real-time feedback and even reveal specific triggers that lead to high and unhealthy levels of stress.

Although the TSST is a controlled setup and does not completely represent general everyday activities, it is useful for our work because it provides a reliable and safe protocol to generate a stressful environment. A controlled situation like this is the logical first step towards validating any new system. Moreover, a personal interview is an example of a social activity that many people will have to face. However, future work will include longer experiments in order to use these technologies in daily life activities, although artificially manipulating stress levels for longer periods will be difficult to achieve experimentally. Alternative methods will be required to analyse the transitions between relaxed and stressed states in more detail.

The combined sensor solution presented in this paper may not be a perfect general solution for detecting stress on a day-to-day basis because the sensors can be uncomfortable in the long term, in particular the wireless physiological sensor, although the sociometric sensor can be worn as a simple badge during longer periods of time. However, one of the main contributions of this paper is to show the individual capabilities of each sensor modality to detect stress, and to present results that will aid in the future selection of appropriate sensors in different situations. The sociometric badge can be worn as part of an anyone's daily activities (e.g. in an office environment) however, the physiological sensor may be more suitable within controlled environments (e.g. patients in hospital). In addition, new physiological and activity sensors continue to be developed which are smaller and more ergonomic, and the results presented in this paper can easily transfer across.

Classification results using only the sociometric badge indicate that easy to wear activity recording sensors are likely to be suitable for monitoring everyday stress. While the classification rates were lower when compared to a combination of both physiological and sociometric sensors, they still remained high. In addition, the sociometric badge allows for the recording of social interactions between participants. Thus, studying stress levels while interacting in a variety of social activities is another future area of research that we would like to explore in the future. Combined signals from social interactions could be combined with individual variation to increase the modalities when detecting stress in individuals and groups.

Finally, similar sensors to those contained within the sociometric badge can be readily found in other devices including intelligent bracelets, smart phones or smart watches. Therefore, a further series of studies is likely to explore the usefulness of these devices alongside self-report when monitoring everyday levels of stress over longer periods of time that go beyond the psychological laboratory.

10. Acknowledgements

This work was partially funded by a Sectoral Operational Programme Human Resources Development 2007-2013 from the Ministry of European Funds (POS- DRU/159/1.5/S/132397), and by a Research Investment Grant from the University of Lincoln (RIF2014-31).

References

1. Biopac `http://www.biopac.com`.
2. Mental Health Foundation in UK `http://www.mentalhealth.org.uk`.
3. The Organisation for Economic Co-operation and Development (OECD) `http://www.oecd.org`.
4. Sociometric badge `http://www.sociometricsolutions.com`.
5. World Health Organization (WHO) `http://www.who.org`.
6. U. R. Acharya, S. V. Sree, S. Chattopadhyay and J. S. Suri, Automated diagnosis of normal and alcoholic EEG signals, *International Journal of Neural Systems* **22**(3) (2012) p. 1250011.
7. U. R. Acharya, V. K. Sudarshan, H. Adeli, J. Santhosh, J. E. W. Koh and A. Adeli, Computer-Aided Diagnosis of Depression Using EEG Signals, *Journal on Multimodal User Interfaces* **73**(5-6) (2015) 329–336.
8. U. R. Acharya, R. Yanti, J. W. Zheng, M. M. R. Krishnan, J. H. Tan, R. J. Martis and C. M. Lim, Automated diagnosis of epilepsy using CWT, HOS and texture parameters, *International Journal of Neural Systems* **23**(3) (2013) p. 1350009.
9. H. Adeli and A. Karim, Fuzzy-wavelet rbfnn model for freeway incident detection, *Journal of Transportation Engineering* **126**(6) (2000) 464–471.
10. D. W. Aha, D. Kibler and M. K. Albert, Instance-based learning algorithms, *Machine learning* **6**(1) (1991) 37–66.
11. K. O. Arras, O. M. Mozos and W. Burgard, Using boosted features for the detection of people in 2D range data, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, (Roma, Italy, 2007), pp. 3402–3407.
12. S. Aydin, N. Arica, E. Ergul and O. Tan, Classification of obsessive compulsive disorder by EEG complexity and hemispheric dependency measurements, *International Journal of Neural Systems* **25** (2015) p. 1550010.
13. A. Beck, *Cognitive therapy and the emotional disorders* (International Universities Press, Inc., Madison, CT, 1975).
14. M. Benedek and C. Kaernbach, A continuous measure of phasic electrodermal activity, *Journal of Neuroscience Methods* **190**(1) (2010) 80–91.
15. C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
16. R. Boostani and M. H. Moradi, A new approach in the bci research based on fractal dimension as feature and adaboost as classifier, *Journal of Neural Engineering* **1**(4) (2004) p. 212.
17. H. Boucsein, *Electrodermal Activity* (Plenum, New York, USA, 1992).
18. C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann and S. Narayanan, Analysis of emotion recognition using facial expressions, speech and multimodal information, *Proceedings of the 6th International Conference on Multimodal Interfaces*, (New York, NY, USA, 2004), pp. 205–211.
19. E. Castillo, D. Peteiro-Barral, B. Guijarro Berdinas and O. Fontenla-Romero, Distributed one-class support vector machine, *International Journal of Neural Systems* **25**(7) (2015) p. 1550029.
20. G. Chanel, C. Rebetez, M. Btrancourt and T. Pun, Emotion assessment from physiological signals for adaptation of game difficulty, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* **41**(6) (2011) 1052–1063.
21. B. Cheng, *Applied Informatics and Communication: International Conference, ICAIC 2011,Xi'an, China, August 20-21, 2011, Proceedings, Part I* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011), Berlin, Heidelberg, ch. Emotion Recognition from Physiological Signals Using AdaBoost, pp. 412–417.
22. M. A. Chikh, M. Saidi and N. Settouti, Diagnosis of diabetes diseases using an artificial immune recognition system2 (airs2) with fuzzy k-nearest neighbor, *Journal of Medical Systems* **36**(5) (2012) 2721–2729.
23. J. S. Chou and A. D. Pham, Multifractal analysis and relevance vector machine-based automatic seizure detection in intracranial, *International Journal of Neural Systems* **25**(6) (2015) p. 1550020.
24. J. S. Chou and A. D. Pham, Smart artificial firefly colony-based support vector regression for enhanced forecasting in civil engineering, *Computer-Aided Civil and Infrastructure Engineering* **30**(9) (2015) 715–732.
25. C. Cortes and V. Vapnik, Support-vector network, *Machine Learning* **20** (1995) 273–297.
26. T. Cox, A. Griffiths and E. Rial-González, *Research on work-related stress* (European Agency for Safety and Health at Work, Luxembourg, 2000).
27. C. Darrow, The rationale for treating the change in galvanic skin response as a change in conductance, *Psychophysiology* **1** (1964) 31–38.
28. K. Dedovic, R. Renwick, N. K. Mahani, V. Engert, S. J. Lupien and J. C. Pruessner, The montreal imaging stress task: using functional imaging to investigate the effects of perceiving and processing psy-

chosocial stress in the human brain, *Journal of Psychiatry and Neuroscience* **30**(5) (2005) 319–325.
29. S. S. Dickerson and M. E. Kemeny, Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research, *Psychological Bulletin* **130**(3) (2004) 355–391.
30. O. Faust, U. Acharya, L. Min and B. Sputh, Automatic identification of epileptic and background EEG signals using frequency domain parameters, *International Journal of Neural Systems* **20**(2) (2010) 159–176.
31. Y. Freund and R. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, *Computational Learning Theory (Eurocolt)*, (Barcelona, Spain, 1995).
32. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection, *IEEE Transactions on Biomedical Engineering* **55**(2) (2002) 512–518.
33. M. J. Goldstein, Individual differences in response to stress, *American Journal of Community Psychology* **1**(2) (1973) 113–137.
34. G. Gosztolya, R. Busa-Fekete and L. Tóth, Detecting autism, emotions and social signals using adaboost, *INTERSPEECH Annual Conference of the International Speech Communication Association*, (Lyon, France, 2013), pp. 220–224.
35. M. A. Hall, Correlation-based feature selection for machine learning, PhD thesis, The University of Waikato (1999).
36. C.-W. Hsu, C.-C. Chang and C.-J. Lin, A practical guide to support vector classification `http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf`, (2010).
37. J. Joshi, R. Goecke, S. Alghowinem, A. Dhall, M. Wagner, J. Epps, G. Parker and M. Breakspear, Multimodal assistive technologies for depression diagnosis and monitoring, *Journal on Multimodal User Interfaces* **7**(3) (2013) p. 217228.
38. A. Karim and H. Adeli, Comparison of the fuzzy wavelet rbfnn freeway incident detection model with the california algorithm, *Journal of Transportation Engineering* **128**(1) (2002) 21–30.
39. A. Karim and H. Adeli, Radial basis function neural network for work zone capacity and queue estimation, *Journal of Transportation Engineering* **129**(5) (2002) 494–503.
40. C. D. Katsis, G. Ganiatsas and D. I. Fotiadis, An integrated telemedicine platform for the assessment of affective physiological states, *Diagnostic Pathology* **1** (2006) p. 16.
41. K. Kim, S. Bang and S. Kim, Emotion recognition system using short-term monitoring of physiological signals, *Medical and Biological Engineering and Computing* **42**(3) (2004) 419–427.
42. C. Kirschbaum, K. M. Pirke and D. H. Hellhammer, The Trier Social Stress Test – A tool for investigating psychobiological stress responses in a laboratory setting, *Neuropsychobiology* (1993) 76–81.
43. M. H. Lee, G. Yang, H. K. Lee, and S. Bang, Development stress monitoring system based on personal digital assistant (PDA), *Proceedings of the IEEE International Conference of the Engineering in Medicine and Biology Society (EMBC)*, (San Francisco, CA, USA, 2004), pp. 2364–2367.
44. J. Li, H. Liu, S.-K. Ng and L. Wong, Discovery of significant rules for classifying cancer diagnosis data, *Bioinformatics* **19**(suppl 2) (2003) ii93–ii102.
45. W. Liao, W. Zhang, Z. Zhu and Q. Ji, A real-time human stress monitoring system using dynamic bayesian network, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (Washington, USA, 2005).
46. C. L.J., Coefficient alpha and the internal structure of tests, *Psychometrika* **16**(3) (1951) p. 297334.
47. H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez and T. Choudhury, Stresssense: Detecting stress in unconstrained acoustic environments using smartphones, *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, (ACM, New York, NY, USA, 2012), pp. 351–360.
48. R. Martis, U. Acharya and H. Adeli, Current methods in electrocardiogram characterization, *Computers in Biology and Medicine* **48** (2014) 133–149.
49. R. Martis, U. Rajendra Acharya, H. Adeli, J. Tan, L. Tong, C. Chua, T. Loon and S. Jie, Computer aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation, *Biomedical Signal Processing and Control* **13** (2014) 295–305.
50. M. Miyaji, H. Kawanaka and K. Oguri, Driver's cognitive distraction detection using physiological features by the adaboost, *International IEEE Conference on Intelligent Transportation Systems*, (St Louis, USA, 2009), pp. 1–6.
51. O. M. Mozos, C. Stachniss and W. Burgard, Supervised learning of places from range data using AdaBoost, *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, (Barcelona, Spain, 2005), pp. 1742–1747.
52. D. O. Olguin, B. N. Waber, T. Kim, A. Mohan, K. Ara and A. Pentland, Sensible organizations: Technology and methodology for automatically measuring organizational behavior, *IEEE Transactions on Systems, Man, and Cybernetics Societ* **39**(1) (2009) 43–55.
53. D. Olguín-Olguín and A. Pentland, Social sensors for automatic data collection, *Proceedings of the 14th Americas Conference on Information Systems*, (Toronto, ON, Canada, 2008), pp. 1–10.
54. D. Olguin-Olguin and A. Pentland, Sensor-based organisational design and engineering, *Int. J. of Organisational Design and Engineering* **1**(1/2) (2010) 69–97.

55. P. B. Patel and T. Marwala, Caller behaviour classification using computational intelligence methods, *International Journal of Neural Systems* **20**(1) (2010) 87–93.
56. A. Pentland, *Honest signals: How they shape our world* (MIT Press, Cambridge, MA, 2008).
57. E. Peper, R. Harvey, L. I-Mei, H. Tylova and D. Moss, Is there more to blood volume pulse than heart rate variability, respiratory sinus arrhythmia, and cardiorespiratory synchrony?, *Biofeedback* **35**(2) (2007) 54–61.
58. A. Perkins, Saving money by reducing stress, *Harvard Business Review* **72**(12) (1994).
59. D. Rai, K. Kosidou, M. Lundberg, R. Araya, G. Lewis and C. Magnusson, Psychological distress and risk of long-term disability: population-based longitudinal study, *Journal of Epidemiology and Community Health* **66**(7) (2011) 586–592.
60. P. Renaud and J.-P. Blondin, The stress of stroop performance: physiological and emotional responses to color-word interference, task pacing, and pacing speed, *International Journal of Psychophysiology* **27** (1997) 87–97.
61. M. Sabeti, S. Katebi and R. Boostani, Entropy and complexity measures for {EEG} signal classification of schizophrenic and control participants, *Artificial Intelligence in Medicine* **47**(3) (2009) 263 – 274.
62. L. Salahuddin, J. Cho, M. Jeong and D. Kim, Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings, *IEEE International Conference of the Engineering in Medicine and Biology Society, EMBC*, (Lyon, France, 2007), pp. 4656–4659.
63. C. Sandi, Stress and cognition, *Wiley Interdisciplinary Reviews:Cognitive Science* **4** (2013) 245–261.
64. V. Sandulescu, S. Andrews, D. Ellis, N. Bellotto and O. M. Mozos, Stress detection using wearable physiological sensors, *International Work-Conference on the Interplay between Natural and Artificial Computation (IWINAC)*, (Elche, Spain, 2015).
65. Z. Sankari and H. Adeli, Heartsaver: A mobile cardiac monitoring system for auto-detection of atrial fibrillation, myocardial infarction and atrio-ventricular block, *Computers in Biology and Medicine* **41**(4) (2011) 211–220.
66. M. Sarkar and T.-Y. Leong, Application of k-nearest neighbors algorithm on breast cancer diagnosis problem, *Proceedings of the American Medical Informatics Association (AMIA) Symposium*, (Los Angeles, CA, 2000), p. 759.
67. S. Sauter, L. Murphy, M. Colligan, N. Swanson, J. Joseph Hurrell, J. Frederick Scharf, R. S. P. Grubb, L. Goldenhar, T. Alterman, J. Johnston, A. Hamilton and J. Tisdale, Stress...at work `http://www.cdc.gov/niosh/docs/99-101.`, (1999).
68. C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Troster and U. Ehlert, Discriminating stress from cognitive load using a wearable EDA device, *IEEE Transactions on Information Technology in Biomedicine* **14**(2) (2010) 410–417.
69. M. Shouman, T. Turner and R. Stocker, Applying k-nearest neighbour in diagnosing heart disease patients, *International Journal of Information and Education Technology* **2**(3) (2012) p. 220.
70. C. D. Spielberger, R. L. Gorsuch, R. E. Lushene, P. R. Vagg and G. A. Jacobs, *State-Trait Anxiety Inventory* (Consulting Psychologists Press, Palo Alto, CA, USA, 1983).
71. R. S. Stawski, M. J. Sliwinski and J. M. Smyth, Stress-related cognitive interference predicts cognitive function in old age, *Psychology and Aging* **21**(3) (2006) 535–544.
72. J. Stroop, Studies of interference in serial verbal reactions, *Journal of Experimental Psychology* **18** (1935) p. 643661.
73. D. E. Sturim, P. A. Torres-Carrasquillo, T. F. Quatieri, N. Malyska and A. McCree, Automatic detection of depression in speech using gaussian mixture modeling with factor analysis, *INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association, Florence, Italy, August 27-31, 2011*, (Florence, Italy, 2011), pp. 2981–2984.
74. F. T. Sun, C. Kuo, H. T. Cheng, B. Senaka, P. Collins and M. Griss, Activity-aware mental stress detection using physiological sensors, *MOBICASE 2010, LNICST*, **76** (Springer Berlin Heidelberg, 2012), pp. 211–230.
75. M. Sung and A. Pentland, Pokermetrics: Stress and lie detection through non-invasive physiological sensing, PhD thesis, MIT Media Laboratory (2005).
76. M. Sung, C. Marci and A. Pentland, Wearable feedback systems for rehabilitation, *Journal of NeuroEngineering and Rehabilitation* **2**(17) (2005) Open Access, doi:10.1186/1743-0003-2-17.
77. J. Taelman, S. Vandeput, A. Spaepen and S. Van Huffel, Influence of mental stress on heart rate and heart rate variability, *4th European Conference of the International Federation for Medical and Biological Engineering*, eds. J. Vander Sloten, P. Verdonck, M. Nyssen and J. Haueisen, *IFMBE Proceedings* **22** (Springer Berlin Heidelberg, 2009), pp. 1366–1369.
78. J. R. Villar, S. Gonzalez, J. Sedano, C. Chira and J. M. Trejo-Gabriel-Galan, Improving human activity recognition and its application in early stroke diagnosis, *International Journal of Neural Systems* **25**(4) (2015) p. 1450036.
79. R. Vincent, J. Pineau, P. de Guzman and M. Avoli, Recurrent boosting for classification of natural and synthetic time-series data, *Proceedings of the Canadian Conference on Artificial Intelligence*, (Canada, 2007), pp. 192–293.
80. P. Viola and M. Jones, Robust real-time object detection, *International Journal of Computer Vision* **4** (2001) 137–154.

81. M. Wikgren, M. Maripuu, T. Karlsson, K. Nordfjäll, J. Bergdahl, J. Hultdin, J. Del-Favero, G. Roos, L. Nilsson, R. Adolfsson and K. Norrback, Short telomeres in depression and the general population are associated with a hypocortisolemic state, *Biological Psychiatry* **71**(4) (2012) 294–30.
82. World Health Organization, *Guidelines for the management of conditions specifically related to stress* (WHO Press, Geneva, 2013).
83. J. Zhai and A. Barreto, Stress detection in computer users based on digital signal processing of noninvasive physiological variables, *IEEE International Conference of the Engineering in Medicine and Biology Society (EMBC)*, (New York, USA, 2006), pp. 1355–1358.