Towards Classification and Prediction of Stress Patterns using Multiple Physiological Signals

Clarissa Almeida Rodrigues[1], William da Rosa Fröhlich[1], Sandro José Rigo[1], Eliza Kern de Castro[1], Andreia Rodrigues[1], Rodrigo Marques Figueiredo[1], Ana Paula Mallmann[1]

[1] University of Vale do Rio dos Sinos
clarissa.ar@gmail.com, williamrfrohlich@gmail.com, rigo@unisinos.br, elisa.kerndecastro@gmail.com, andreiakrodrigues@yahoo.com.br, marquesf@unisinos.br, apmallmann@unisinos.br

Abstract. The stress is increasing in our society in the last years, due the large and tiring routines besides few time to rest. Keeping this in mind, this paper intends to determine patterns in stress' events using physiological signs, because these signals are a reliable source to identify stress states. The literature shows that the use of physiological signs as a source for stress patterns identification is a promising investigation subject and there are few studies evaluating the effect of combining several different signals. The objective of this article is to investigate the possible integration of data obtained from electrocardiographic (ECG), electrodermal activity (EDA) and electromyography (EMG) to detect stress patterns using wearable sensors to acquisition of biofeedback and propose algorithms to set some patterns. It was developed a dataset to made the pre-processing in all of data to evaluate the plausibility and develop an adequate database for the application of machine learning techniques establishing as a reference the obtained annotated data.

Keywords: Wearable sensors, Stress, Biofeedback.

1 Introduction

Stress can affect all aspects of our lives, including our emotions, behaviors, thinking ability, and physical health, making our society sick – both mentally and physically. Among the effects that the stress and anxiety can cause are heart diseases, such as coronary heart disease and heart failure [5]. Due this information, this research will present a proposal to help people handling stress using the benefit of technology development and to set patters of stress status as way to propose some intervention, once the first step to controlling stress is to know the symptoms of stress.

The stress symptoms are very board and can be confused with others diseases according The American Institute of Stress [15], for example the frequent headache, irritability, insomnia, nightmares, disturbing dreams, dry mouth, problems swallowing, increased or decreased appetite, or even cause other diseases such as frequent colds and infections. In view of the wide variety of symptoms caused by stress, this research intends to define, through physiological signals, the patterns generated by the body and obtained by wearable sensors and develop a standardized database to apply the machine learning.

2 Problem and Motivation

According to a research of The American Institute of Stress [15], 77% of people regularly experience physical symptoms caused by stress, 51% of them are related to fatigue. In other hand, advances in sensor technology, wearable devices and mobile growth would help to online stress identification based on physiological signals and delivery of psychological interventions. Currently with the advancement of technology and improvements in the wearable sensors area, made it possible to use these devices as a source of data to monitor the user's physiological state. The majority of the wearable devices consist of low-cost board that can be used to the acquisition of physiological signals [1, 10]. After the data are obtained it is necessary apply some filters to clear signal, without noise or distortions aiming to use some Machine Learning approaches to model and predict these stress states [2, 11].

The wide-spread use of mobile devices and microcomputers, as Raspberry Pi, and its capabilities presents a great possibility to collect, and process those signs with an elaborated application. These devices can collect the physiological signals and detect specific stress states to generate interventions following the predetermined diagnosis based on the standards already evaluated in the system [9, 6]. During the literature review it was evident the presence of few works dedicated to evaluating comprehensively the complete cycle of biofeedback, which comprises using the wearable devices, applying Machine Learning patterns detection algorithms, generate the psychologic intervention, besides monitoring its effects and recording the history of events [9, 3].

3 Background and Related Works

Stress is identified by professionals using human physiology, so wearables sensors could help on data acquisition and processing, through machine learning algorithms on biosignal data, suggesting psychological interventions. Some works [6, 14] are dedicated to define patterns as experiment for data acquisition simulation real situations. Jebelli, Khalili and Lee [6] showed a deep learning approach that was used to compare with a baseline feedforward artificial neural network.

Schmidt et al. [12] describes Wearable Stress and Affect Detection (WESAD), one public dataset used to set classifiers and identify stress patterns integrating several sensors signals with the emotion aspect with a precision of 93% in the experiments. The work of Gaglioli et al. [4] describe the main features and preliminary evaluation of a free mobile platform for the self-management of psychological stress.

In terms of the wearables, some studies [13, 14] evaluate the usability of devices to monitory the signals and the patient's well-being. Pavic et al. [13] showed a research performed to monitor cancer patients remotely and as the majority of the patients have a lot of symptoms but cannot stay at hospital during all treatment. The authors emphasize that was obtained good results and that this system is viable, as long as the patient is not a critical case, as it does not replace medical equipment or the emergency care present in the hospital.

Henriques et al. [5] focus was to evaluated the effects of biofeedback in a group of students to reduce anxiety, in this paper was monitored the heart rate variability with two experiments with duration of four weeks each. The work of Wijman [8] describes the use of EMG signals to identify stress, this experiment was conducted with 22 participants, evaluating both the wearables signals and questionnaires.

4 Approach and Uniqueness

In this section will be described the uniqueness of this research and the devices that was used. This solution is being proposed by several literature study about stress patterns and physiological aspects but with few results, for this reason, our project will address topics like experimental study protocol on signals acquisition from patients/participants with wearables to data acquisition and processing, in sequence will be applied machine learning modeling and prediction on biosignal data regarding stress (Fig. 1).

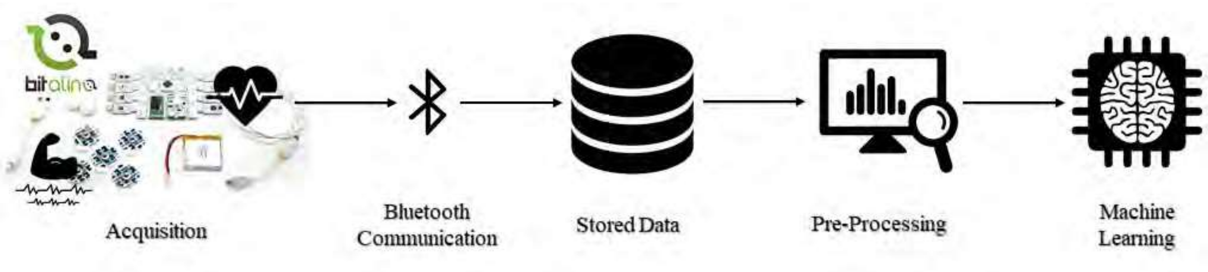

Fig. 1. The acquisition system

The protocol followed to the acquisition of signals during all different status is the Trier Social Stress Test (TSST) [7], recognized as the gold standard protocol for stress experiments. The estimated total protocol time, involving pre-tests and post-tests, is 116 minutes with a total of thirteen steps, but applied experiment was adapted and it was established with ten stages: Initial Evaluation: The participant arrives, with the scheduled time, and answer the questionnaires; Habituation: It will take a rest time of twenty minutes before the pre-test to avoid the influence of events and to establish a safe baseline of that organism; Pre-Test: The sensors will be allocated (Fig. 2), collected saliva sample and applied the psychological instruments.

Fig. 2. Sensors Allocated

The next step is Explanation of procedure and preparation: The participant reads the instructions and the researcher ensures that he understands the job specifications, in sequence, he is sent to the room with the jurors (Fig. 3), composed of two collaborators of the research, were trained to remain neutral during the experiment, not giving positive verbal or non-verbal feedback; Free Speech: After three minutes of preparation, the participant is requested to start his speech, being informed that he cannot use the notes.

Fig. 3. Participant and the jurors

This will follow the Arithmetic Task: the jurors request an arithmetic task in which the participant must subtract mentally, sometimes, the jurors interrupt and warn that the participant has made a mistake; Post-Test Evaluation: The experimenter receives the subject outside the room for the post-test evaluations; Feedback and Clarification: The investigator and jurors talk to the subject and clarify what the task was about; Relaxation technique: A recording will be used with the guidelines on how to perform a relaxation technique, using only the breathing; Final Post-Test: Some of the psychological instruments will be reapplied, saliva samples will be collected, and the sensors will still be picking up the physiological signals.

Based on literature [14] and wearable devices available the signals that was selected to analysis is the ECG, EDA and EMG for an initial experiment. This experimental study protocol on data acquisition started with 71 participants, where data annotation each step was done manually, from protocol experiment, preprocessing data based on features selection. In the Machine Learning step, it was evaluated the metrics of different algorithms as Decision Tree, Random Forest, AdaBoost, KNN, K-Means, SVM.

The experiment was made using the BITalino Kit - PLUX Wireless Biosignals S.A. (Fig. 4) composed by ECG sensor, which will provide data on heart rate and heart rate variability; EDA sensor that will allow measure the electrical dermal activity of the sweat glands; EMG sensor that allows the data collect the activity of the muscle signals.

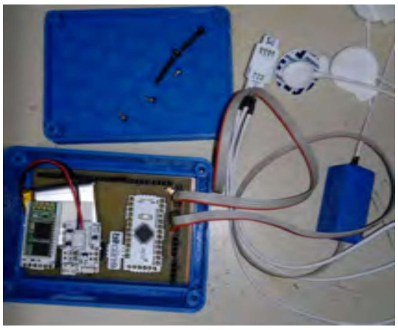

Fig. 4. BeWell prototype

5 Results

This section will describe the results in the pre-processing step and how it was made, listing all parts regarded to categorization and filtering data, evaluating the signal to know if it has plausibility and create a standardized database. The developed code is written in Python due to the wide variety of libraries available, in this step was used the libraries NumPy and Pandas, both used to data manipulation and analysis.

In the first step it is necessary read the files with the raw data and the timestamp, during this process the used channels are renamed to the name of the signal, because the BITalino store the data with the channel number as name of each signals. In sequence, the data timestamp is converted to a useful format, with goal to compare with the annotations, after time changed to the right format all channels unused are discarded to avoid unnecessary processing. The next step is to read the annotations taken manually in the experiment, as said before, to compare the time and classify each part of the experiment with its respective signal.

After all signals are classified with its respective process of the TSST, each part of the experiment is grouped in six categories, which will be analyzed later. The first category is the "*baseline*", with just two parts of the experiment, representing the beginning of the experiment, when the participants had just arrived. The second is called of "*tsst*" comprises the period in which the participant spoke, the third category is the "*arithmetic*" with the data in acquired in the arithmetic test.

The others two relevant categories are the "*post_test_sensors_1*" and "*post_test_sensors_2*", with its respective signals in the parts called with the same name. Every other part of the experiment was categorized as "*no_category*", in sequence, this category is discarded in function of it will not be necessary in the machine learning stage. After the dataframe is right with all signals properly classified, the columns with the participants number and the timestamp are removed of the dataframe. The next step is evaluated the signal, to verify if the signal is really useful in the process of machine learning. For this, it is analyzed the signals using the BioSPPy library, which performs the data filtering process and makes it possible to view the data.

Finally, the script checks the volume of data present in each classification and returns the value of the smallest category. This is done because it was found that the categories have different volumes of data, which would become a problem in the machine learning stage, by offering more data from a determinate category than from the others. Due this fact, the code analyzes the others categories and reduce its size until all categories stay with the same number of rows in each category (); after this the dataframe is exported in a CSV file, to be read in the machine learning stage.

```
In [46]: minimum_size = comparator(new_data)

Baseline: 139230
TSST: 73931
Arithmetic: 346907
Post Test Sensors I: 584527
Post Test Sensors II: 741600
73931
```

Call the **"data_generator"** Method to Generate the File with Standardized Data

```
In [47]: data_save = data_generator(new_data, minimum_size)
```

Call the **"comparator"** Verify the DataFrame

```
In [48]: comparator(data_save)

Baseline: 73931
TSST: 73931
Arithmetic: 73931
Post Test Sensors I: 73931
Post Test Sensors II: 73931
73931
```

Fig. 5. Standardization of data

6 Conclusion

The purpose of this article is to describe some stages of the development of a system for the acquisition and analysis of physiological signals to determine patterns in these signals that would detect stress states. During the development of the project was verified that there are data gaps in the dataframe in the middle of the experiment in some participants; A hypothesis about the motivation of this had happened is the sampling of the acquisition of BITalino regarding communication issues in some specifics sampling rates.

It evaluate the results obtained when reducing this acquisition rate, however, it is necessary to carefully evaluate the extent to which the reduction in the sampling rate will interfere with the results. During the evaluation of the plausibility of the signals, it was verified that there are evident differences between the signals patterns in the different stages of the process, thus validating the protocol followed in the acquisition of the standards. The next step in this project is implement the machine learning stage, applying different algorithms as SVM, Decision Tree, Random Forest, AdaBoost, KNN and K-Means; besides to evaluate the results using metrics like Accuracy, Precision, Recall and F1.

The next steps of this research will support the confirmation of the hypothesis raised about being able to define patterns of physiological signals to detect stress states. From the definition of the patterns, a system can be applied that identifies the acquisition of the signals and, in real time, performs the analysis of these data based on the machine learning results. Therefore we can detect the state of the person and that the psychologist can indicate a proposal intervention and monitor whether the decrease is occurring.

References

1. Attaran, N., Brooks, J. and Mohsenin, T. (2016) “A Low-Power Multi Physiological Monitoring”. Processor for Stress Detection. IEEE SENSORS. 30 Oct. - 3 Nov. 2016. DOI: 10.1109/ICSENS.2016.7808776
2. Choi, J. and Gutierrez-Osuna, R. (2009) "Using Heart Rate Monitors to Detect Mental Stress". Sixth International Workshop on Wearable and Implantable Body Sensor Networks, Berkeley, CA, 2009, pp. 219-223
3. Gaggioli, A. et al. (2014). “Positive technology: A free mobile platform for the self-management of psychological stress”. Annual Review of Cyber Therapy and Telemedicine, 12(May), 25–29
4. Henriques, G. et al. (2011) “Exploring the effectiveness of a computer-based heart rate variability biofeedback program in reducing anxiety in college students”. Applied Psychophysiology Biofeedback, 36(2), 101–112
5. Herrmann-Lingen, C. and Wachter, R. (2020) “Psychological stress and incidence of atrial fibrillation”. European Journal of Preventive Cardiology. The European Society of Cardiology 2020. DOI: 10.1177/2047487319898022
6. Jebelli, H., Khalili M. M. and Lee S. A. (2018) “Continuously Updated, Computationally Efficient Stress Recognition Framework Using Electroencephalogram (EEG) by Applying Online MultiTask Learning Algorithms (OMTL)”. IEEE J Biomed Health Inform
7. Kudielka, B. M., Hellhammer, H. and Kirschbaum, C. (2007) “Ten years of research with the trier social stress test. Social Neuroscience, (October 2015), 56–83.
8. Wijsman, J. et al. (2010) “Trapezius muscle EMG as predictor of mental stress”. In Wireless Health 2010 (WH '10). ACM, New York, NY, USA, 155-163

9. Paredes, P. et al. (2014) “PopTherapy: coping with stress through pop-culture“. Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare, May 20-23, 2014
10. Saha, S. S. et al. (2018) “DU-MD: An Open-Source Human Action Dataset for Ubiquitous Wearable Sensors”. Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV). Kitakyushu, Japan
11. Sano, A. and Picard, R. W. (2013) “Stress Recognition using Wearable Sensors and Mobile Phones”. Humaine Association Conference on Affective Computing and Intelligent Interaction. DOI 10.1109/ACII.2013.117
12. Schmidt, P. et al. (2018) “Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection”. In Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI '18). ACM, New York, NY, USA, 400-408
13. Pavic, M. et al (2019) “Feasibility and Usability Aspects of Continuous Remote Monitoring of Health Status in Palliative Cancer Patients Using Wearables”. Oncology and Informatics – Research Article. 98:386–395. DOI: 10.1159/000501433
14. Ponciano, V. et al. (2020) “Detection of diseases based on Electrocardiography and Electroencephalography signals embedded in different devices: An exploratory study”, Brazilian Journal of Development, Curitiba, v. 6, n.5, p.27212-27231 May. 2020. ISSN 2525-8761. DOI: 10.34117/bjdv6n5-247
15. The American Institute of Stress (2020) “Stress Effects”. The American Institute of Stress. 10 Jul. 2020. www.stress.org/stress-effects