BIOPHYSICAL SIGNAL PROCESSING FOR AUTOMATIC ANXIETY CLASSIFICATION IN A VIRTUAL REALITY EXPOSURE THERAPY SYSTEM

Nicolae JINGA[1], Cătălin-Dumitru PETRESCU[2], Oana MITRUȚ[3], Alin MOLDOVEANU[4], Florica MOLDOVEANU[5], Livia PETRESCU[6]

This paper presents a signal processing framework for automatic anxiety level classification in a virtual reality exposure therapy system. Two types of biophysical data (heart rate and electrodermal activity) were recorded, pre-processed and passed through a feature extraction procedure that provided input for the real-time anxiety level classification algorithm. We showcase the mathematical and engineering techniques behind the procedures and conclude with the challenges encountered in our research and future development ideas. The proposed method provides a good estimate of the level of anxiety while having at the same time a reduced level of complexity, allowing implementation on equipment with limited computing resources.

Keywords: anxiety; signal processing; feature extraction; classification; biophysical data

1. Introduction

PhoVR - Immersive Treatment of Phobias through Adaptive Virtual Reality and Biofeedback – is a virtual reality exposure therapy system (VRET) for phobias therapy (acrophobia, claustrophobia and fear of public speaking) that integrates biophysical data (heart rate - HR and electrodermal activity - EDA) acquisition and aims to automatically adapt the level of exposure or inform the psychotherapist about the patient's degree of anxiety so that he/she can personally adjust the exposure intensity. The advent of sensors technology and the

[1] PhD student, Dept. of Computer Science, National University of Science and Technology POLITEHNICA, Bucharest, Romania, e-mail: nicolae.jinga@stud.acs.upb.ro
[2] Associate Prof., Dept. of Automatic Control and Systems Engineering, National University of Science and Technology POLITEHNICA, Bucharest, Romania, e-mail: catalin.petrescu@upb.ro
[3] Associate Prof., Dept. of Computer Science, National University of Science and Technology POLITEHNICA, Bucharest, Romania, e-mail: oana.balan@upb.ro
[4] Prof., Dept. of Computer Science, National University of Science and Technology POLITEHNICA, Bucharest, Romania, e-mail: alin.moldoveanu@upb.ro
[5] Prof., Dept. of Computer Science, National University of Science and Technology POLITEHNICA, Bucharest, Romania, e-mail: florica.moldoveanu@upb.ro
[6] Lecturer, Faculty of Biology, University of Bucharest, Romania, e-mail: livia.petrescu@bio.unibuc.ro

development of advanced mathematical methods has expanded the field of computer-based emotion recognition systems in various domains, as shown in the literature [1-7].

The PhoVR Control Panel ensures communication with the hardware equipment (Virtual Reality headset and biophysical data acquisition equipment) and allows real-time visualization of the recorded physiological data. The processing of physiological signals involves the following steps:

- preprocessing of the acquired signal;
- feature extraction;
- anxiety level classification.

2. Related work

As the interest in the field of affective computing is constantly increasing, various methods for signal processing and feature extraction have been tested with promising results. Raw biophysical data is preprocessed in a pipeline that has the following steps: data filtering which removes unwanted components from the signal (noise smoothing, excluding high frequency intervals), downsampling, normalization and standardization.

Electrodermal Activity (EDA) signal is further decomposed into Skin Conductance Level (SCL) and Skin Conductance Response (SCR). This step is usually performed by two methods: Continuous Decomposition Analysis (CDA) and Discrete Decomposition Analysis (DDA). CDA is the recommended method for the deconvolution of skin conductance data. Prior to feature extraction, a segmentation stage of the acquired data is recommended, by extracting parts of the raw biophysical signal of different lengths, called time-windows.

Heart Rate (HR) can be obtained either from electrocardiogram (ECG) signal or from photoplethysmography (PPG) signal. The current trend is to mainly use the PPG signal because of its simple setup and large availability of devices that use it (fitness bracelets and smart watches). However, in some circumstances, accuracy of HR estimation based on PPG signal is lower than in case of ECG [8] especially due to its high sensitivity to motion [9] and breathing [10]. Consequently, HR extraction from the PPG signal is performed using complex algorithms [11][12][13]. Also, the reduction of motion-induced artifacts requires the use of additional acceleration sensors [14].

Among the feature extraction methods relevant in the literature, we mention the Fisher's Linear Discriminant [15], Minimum redundancy-Maximum relevance [16], Pearson Correlation Coefficient [16][17], Stepwise Linear Regression [18], Fractal dimension features (FD) and Statistical Higher Order Crossings (HOC) [19], Fast correlation based filter (FCBF) [20], Feature fusion [21], Principal Component Analysis, Covariate Shift Adaptation [22]. These algorithms have been applied on the DEAP (Database for Emotion Analysis using Physiological signals) [22] and MAHNOB-HCI [23] datasets which are public emotion databases used for research purposes.

In this article, we present the procedure for extraction of the features required by the anxiety level classification algorithm. Our approach is focused mainly on reducing the complexity of signal processing algorithms involved in EDA and HR signal analysis in order to allow their implementation on hardware devices with limited computational resources (BITalino).

3. Data acquisition system

The biophysical data acquisition system is used for continuous patient monitoring during the therapy session. For this purpose, we measured two parameters that are influenced by the patient's anxiety level [24]:

- heart rate variability - represented by the time interval between two successive waves (or inter-beat interval);
- electrodermal activity - assessed by measuring skin conductance.

Following a thorough analysis that included the technical characteristics, ease of use, ergonomics, and cost of the equipment, we decided to employ the BITalino device [25] which is a portable equipment capable of acquiring the two previously mentioned parameters. BITalino measures the electrical conductance of the skin using two electrodes placed on the medial phalanges of the index and middle fingers. Heart rate is assessed based on the measurement of blood volume variations in the capillaries using the optical plethysmography (PPG) method.

The device is composed of a central unit based on an ATMEGA328 8-bit microcontroller that integrates a processor, data, and program memory as well as a series of peripheral devices (timers, communication interfaces, analog-to-digital converter). Also, at the level of the central unit, a wireless communication module based on the Bluetooth Low-Energy technology and a power unit using a Li-Ion type battery are implemented. The device allows connecting a maximum of 6 analog and 2 digital sensors. In our application, 2 analog sensors are used:

- EDA sensor – composed of a signal conditioning module and a cable equipped with two self-adherent (replaceable) Ag-AgCl electrodes;
- PPG sensor – used to measure HR and built on the principle of optical plethysmography.

The data received via the Bluetooth interface is acquired by the biophysical data processing application at a sampling rate of 1000 samples per second. The data processing application runs on the computer in parallel with the Control Panel, communicating with it through a WebSocket connection. It has the role of extracting the value of the heart rate and the two components that describe the electrodermal activity (the tonic component – Skin Conductance Level (SCL) and the phasic component – Skin Conductance Response (SCR)).

4. Preprocessing of the physiological signals

The preprocessing of the acquired physiological signals has the role of bringing the data into a form usable by the following processing techniques.

Unlike other physiological signal acquisition equipment (such as Shimmers Sensing [26] or Biopac [27]), the BITalino system does not include specialized firmware for extracting the heart rate information from the acquired PPG signal. For this reason, it was necessary to develop and implement an algorithm for calculating heart rate based on the PPG signal (Fig. 1).

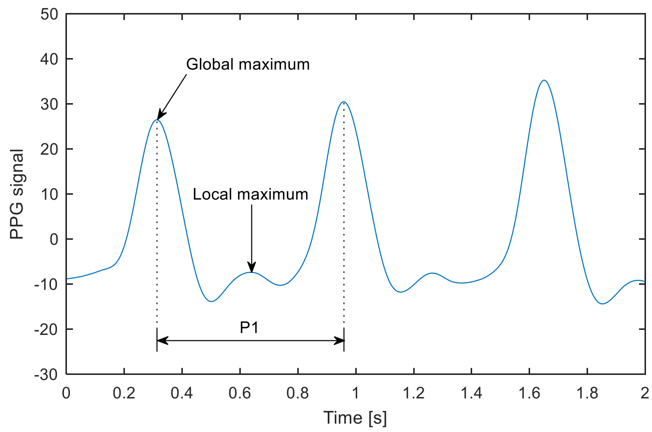

Fig. 1. Algorithm for HR extraction from the PPG signal

The pulse duration (P1) is the interval between two successive global maxima of the signal. The heart rate computing algorithm can detect the maximum points in the signal and differentiate between the local and global ones. The detection of the maximum points is performed by calculating the derivative of the signal and locating the moments when it changes its sign from positive values (increasing signal) to negative values (decreasing signal).

To clear the noises and artefacts that accompany the PPG signal, we filtered it using a Butterworth-type bandpass filter of order II with a bandpass between 0.5 and 10 Hz.

We calculated the derivative by evaluating the difference between two successive samples of the signal. For detecting the maximum points, we implemented a simple finite-state machine, that identifies the minimum and maximum points based on the derivative's change of sign.

- State 1 – corresponds to signal sections with negative derivative (falling signal). In this state, the current value of the derivative is tested for sign change detection and the algorithm transitions to state 1. Also, on sign change detection (which signals the identification of a minimum point), the current value of the signal is stored in a variable V_{min}.

- State 2 – corresponds to signal sections with positive derivative (rising signal). In this state, the current value of the derivative is tested for sign change detection and transitions back to state 1. Once a sign change is detected (corresponding to the identification of a maximum point), the difference between the current value of the signal (Vmax) and the previous minimum value present in the variable Vmin is calculated. This difference represents the amplitude of the rising signal segment that precedes the identified maximum point – ΔV.

The differentiation between a local maximum and a global one is performed by comparing the amplitude of the rising segment of the signal ΔV preceding the detected maximum with the difference between the highest and the lowest value of the signal recorded in the last 1.25 seconds noted with AMP. The duration of the interval was chosen so that it certainly contains a pulse.

A maximum is considered global and is taken as a benchmark in determining the pulse duration if $\Delta V \geq$ AMP. The following figure shows an example in which two maximum points are classified based on the previously stated criterion (Fig. 2).

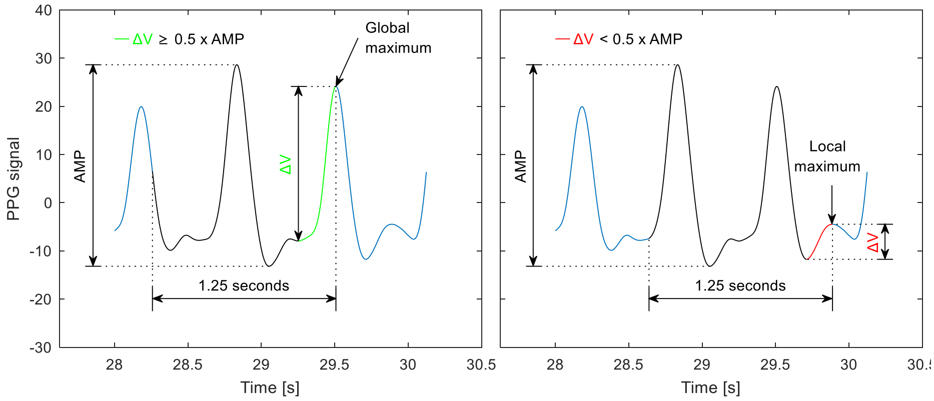

Fig. 2. Classification of two maximum points

The data acquisition equipment (BITalino) cannot be configured to use different sampling rates on each channel. For this reason, the sampling frequency must be chosen in such a way as to effectively allow the acquisition of the signal with the widest frequency band.

The PPG signal used to determine the heart rate has a base frequency in the range of 0.75 - 3 Hz, significantly higher than the components of the electrodermal response. Moreover, for the correct determination of the interval between two successive pulses, we need to analyze the segments of the signal that show very rapid variations in time, which requires a high sampling frequency. For this reason, it is the PPG signal that determines the frequency at which both signals will be acquired, namely 1000 samples per second.

The EDA signal sampled 1000 times per second will be difficult to process because it is a slow signal, and its analysis is performed over long time intervals that would contain a very large number of samples. For this reason, it is necessary to reduce the sampling frequency to a value adapted to its dynamics, namely 20 samples/second.

The resampling of the EDA signal is done by a decimation process where we keep a single sample at every 50 acquired samples. To avoid the aliasing phenomenon introduced by reduction of the sampling frequency, a preliminary filtering of the signal is necessary. Filtering is performed using a 4th order Butterworth low-pass filter with a pass frequency of 10 Hz (Nyquist frequency of the resampled signal).

5. Feature extraction

During this stage, we performed the real-time implementation of the algorithms for calculating the characteristic values of the EDA and HR signals (extracted from the PPG signal).

The separation of tonic and phasic components from the resampled EDA signal is performed using a set of two filters, a low-pass filter for estimating the tonic component and a high-pass filter for the phasic component.

Both filters have a pass frequency of 0.05 Hz and are implemented as finite impulse response filters designed using a Hamming window with a size of 401 samples. The high-pass filter was designed as a complementary filter to the first one, having the transfer function:

$$H_{high-pass}(q^{-1}) = q^{-200} - H_{low-pass}(q^{-1}) \tag{1}$$

The frequency characteristics of the filters are presented in the following figure (Fig. 3):

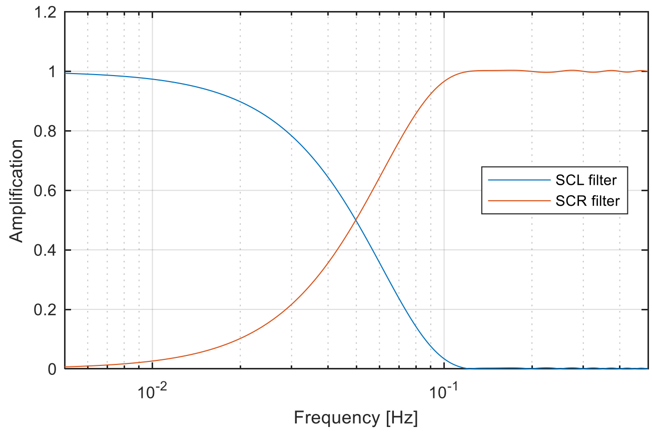

Fig. 3. Frequency characteristics of the filters

Both EDA preprocessing (filtering, resampling, and component separation) and PPG signal preprocessing for heart rate extraction are performed in real-time as the data packets are received from the BITalino acquisition system.

6. Anxiety level classification

The biophysical signal characteristics computed in real time are received by a classifier whose role is to estimate the level of anxiety perceived by the subject during the virtual reality therapy session.

Most implementations of anxiety level classifiers are based on neural networks that are trained to find an estimate of anxiety level based on the biophysical features presented at the input.

Due to the high computational resources required by neural networks implementation, we chose an alternative approach based on a regression model:

$$y = f\left(x, c\right) + e \tag{2}$$

where:

y – represents the dependent variable (anxiety level);
x – the vector containing independent variables (biophysical features);
c – the vector of model parameters;
e – represents an additive error term;

Because there are unlimited possibilities to create the model function, we chose to use a second order approximation of the non-linear multivariable function f using Taylor series:

$$y = f(a) + (x-a)^T \nabla f(xa) + (x-a)^T H(x_0)(x-a) \tag{3}$$

where $\nabla f(a)$ and $H(a)$ are the gradient vector, respectively Hessian matrix of the estimation function $f(x)$ evaluated in the reference point a. If the reference point is chosen $a=0$, due to the symmetry of the Hessian matrix, the regression model can be rewritten as the following quadratic function:

$$y = \sum_{\substack{i=1..P \\ j=1..P \\ i \le j}} \alpha_{i,j} x_i x_j + \sum_{i=1}^{P} \beta_i x_j + \gamma \tag{4}$$

where x_k are the characteristic values of the biophysical signals and $\alpha_{i,j}$, β_i and γ are the coefficients of the estimator determined by an experimental identification procedure described in [28].

The value calculated by this estimation function is a number between 0 and 10, where 0 corresponds to a situation where the patient is completely relaxed and 10 indicates an extreme level of anxiety. Anxiety is classified into three levels (Low, Medium, or High) based on the value of the estimation function, as follows:

- Low Level –in the range [0,3]
- Medium Level –in the range (3,7)
- High Level - in the range (7,10]

To avoid the instability of the computed anxiety rate, the classifier analyzes the estimated anxiety level for each window in the signal for the previous 10 seconds and selects the level that has the highest frequency of occurrence.

We performed a set of experiments in which the subjects were exposed to phobic stimuli of varying degrees of intensity. During the experiments, the biophysical parameters (EDA and HR) were recorded and after each exposure to stimuli, the subjects self-assessed their level of anxiety.

The parametric identification procedure aimed to determine the set of numerical values that lead to the closest estimate of the anxiety level relative to the self-assessed one during the experiments.

The quadratic function has 7 characteristics of the acquired biophysical signals as input parameters. The 7 features selected following the preliminary tests (carried out in the first stage of the project) are summarized in the following table (Table 1):

Table 1

Characteristics of the biophysical signals

Parameter	**Name**	**Description**
x_1	HR_mav	Average duration of pulses
x_2	HR_std	The standard deviation of the pulse duration
x_3	SCR_mav	The average of the absolute

		values of the phasic component – SCR
x_4	SCL_mar	The mean of the tonic component values relative to the baseline value
x_5	SCR_mar	The mean of the phasic component values relative to the baseline value
x_6	GSR_wamp	Willison amplitude of the EDA signal (not decomposed into components)
x_7	SCR_wamp	Willison amplitude of the phase component

The value of the quadratic function corresponding to the classifier is evaluated in real time, as the parameters of the biophysical signals become available.

After performing some tests, we identified two deficiencies of the biophysical data acquisition system:

The first problem is related to the quality of the signal provided by the optical sensor used to measure heart rate. This problem is generated by the sensor's fixing system on the finger which does not ensure stable contact especially when the subject moves the hand on which the sensor is connected.

Another problem that arose during the experiments was the impossibility of measuring the electrodermal response in the case of certain subjects who presented a very high conductance level. BITalino device offers a conductance measurement range between 0 and 25 µS that is suitable for most people, but other alternatives that allow the dynamic adaptation of the minimum/maximum values are being sought.

Since the number of subjects participating in the initial experiments was small (the dataset used for parameters identification consisted of 105 items), it is possible that the level of generalization of the classifier may not be high enough.

For this reason, we aim to improve its performance by adjusting the parameter values as more information is acquired while using the system. Thus, the current version of the system records the values of the physiological parameters throughout the therapy sessions. Biophysical parameters are acquired and recorded at a rate of 1000 samples/second to allow additional features to be computed later (if this can help increase classification accuracy).

Apart from the physiological parameters, the system records the moments of time when the patients perform different tasks in the virtual reality exposure scenarios. These moments correspond to a peak exposure to the phobic stimuli and recording the corresponding timestamps allows the localization of the segments in the biophysical signals that reflect the patients' response to stimuli.

Also, after performing each task, a self-assessment of the degree to which the patients were able to control their emotions is recorded.

All this recorded information (evolution of the biophysical parameters, timestamps for performing tasks and self-assessments of the degree of emotion control) will be used later to improve the accuracy and degree of generalization of the anxiety level classifier.

The procedure for improving the classifier involves re-identifying the coefficients of the anxiety level estimation function using an extended dataset consisting of all the information acquired during the use of the system by as many patients as possible. Using data that describes the physiological response of many different individuals has the effect of increasing the generalizability of the classifier.

We also aim to implement the classifier in an alternative version based on a neural network. This will be possible only when the amount of acquired information contains at least several thousand elements because the training of neural networks can only be done correctly under these conditions.

7. Conclusions and future work

Considering the desire to market the PhoVR system to individual users with a mild-to-moderate phobia, our aim is to reduce the purchase cost of the system together with its accessories. One of the substantial costs of acquisition is the biophysical data acquisition hardware. The proposed solution represents an adaptation of the anxiety level estimation method presented in [28] to allow real time implementation on a low-cost signal acquisition device (BITalino). The main contributions of the proposed solution are the algorithm for HR estimation and replacing the off-line Discrete Decomposition Analysis with a real time filtering for extracting Tonic and Phasic components of the EDA signal. Feature extraction and anxiety level classification were not modified because are simple to be implemented in real time with low computational effort.

In the experiments performed, the HR signal was acquired in an experimental protocol which involved positioning the left hand (used for the placement of the sensors) on a stable surface (a table) and avoiding movements during the experiment. To eliminate these constraints that can influence the therapy sessions, we will test various sensors that have a more stable fixing mechanism (as those integrated in medical equipment for monitoring heart rate and/or oxygen saturation). We will also investigate the use of sensors containing multiple optical detectors that have high immunity to contact pressure variations. The use of this type of sensor will require adapting the signal processing algorithm so that it can integrate the information provided by several detectors.

Acknowledgements

This work was supported by project TRAVEE (Virtual Therapist with Augmented Feedback for Neuromotor Recovery) through a grant of the Ministry of Research, Innovation and Digitization, CCCDI - UEFISCDI, project number PN-III-P2-2.1-PTE-2021-0634, within PNCDI III.

R E F E R E N C E S

[1] *M. Malik*, “Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the European Society of Cardiology and the North American Society for Pacing and Electrophysiology,” Ann. Noninvasive Electrocardiol., **vol**. 1, no. 2, 1996, pp. 151–181.

[2] *J. Machajdik and A. Hanbury*, “Affective image classification using features inspired by psychology and art theory,” in Proceedings of the 18th ACM international conference on Multimedia, 2010

[3] *R. Francese, M. Risi, and G. Tortora*, “A user-centered approach for detecting emotions with low-cost sensors,” Multimed. Tools Appl., **vol.** 79, no. 47–48, 2020, pp. 35885–35907

[4] *S. Vijayakumar, R. Flynn, and N. Murray*, “A comparative study of machine learning techniques for emotion recognition from peripheral physiological signals,” in 2020 31st Irish Signals and Systems Conference (ISSC), 2020

[5] *V. Doma and M. Pirouz*, “A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals,” J. Big Data, **vol**. 7, no. 1, 2020

[6] *M. Benedek and C. Kaernbach*,“ Decomposition of skin conductance data by means of nonnegative deconvolution,” Psychophysiology, **vol**. 47, no. 4, pp. 647–658, 2010

[7] *Mitruț, O., Moldoveanu, A., Petrescu, L., Petrescu, C., Moldoveanu, F.*, “A Review of Virtual Therapists in Anxiety and Phobias Alleviating Applications,“ International Conference on Human-Computer Interaction, HCII 2021: Virtual, Augmented and Mixed Reality, (J.Y.C. Chen, Fragomeni, G. - editors), Lecture Notes in Computer Science, **vol.** 12770, pp. 71-79. Springer, Cham. https://doi.org/10.1007/978-3-030-77599-5_6

[8] *Esgalhado F.; Batista A.; Vassilenko V.; Russo S.; Ortigueira M.*, “Peak Detection and HRV Feature Evaluation on ECG and PPG Signals.” Symmetry, **vol.** 14, no. 6, 2020, p. 1139. https://doi.org/10.3390/sym14061139

[9] *Lu G., Yang F.*, “Limitations of Oximetry to Measure Heart Rate Variability Measures.” Cardiovascular Engineering, **vol.** 9, issue 3, 2009, pp. 119–125, https://doi.org/10.1007/s10558-009-9082-3

[10] *Jan HY., Chen MF., Fu TC. et al.* “Evaluation of Coherence Between ECG and PPG Derived Parameters on Heart Rate Variability and Respiration in Healthy Volunteers With/Without Controlled Breathing.” Journal of Medical and Biological Engineering, **vol.** 39, issue 5, 2019, pp. 783–795, https://doi.org/10.1007/s40846-019-00468-9

[11] *Alqaraawi A., Alwosheel A., Alasaad A.*, “Heart rate variability estimation in photoplethysmography signals using Bayesian learning approach”, Healthcare Technology Letters, 2016 Jun 13; **vol.** 3(2), pp. 136-142. https://doi.org/10.1049/htl.2016.0006

[12] *P. Fan, H. Peiyu, L. Shangwen and D. Wenfeng,* “Feature extraction of photoplethysmography signal using wavelet approach,” 2015 IEEE International Conference on Digital Signal Processing (DSP), Singapore, 2015, pp. 283-286, doi: 10.1109/ICDSP.2015.7251876

[13] *Kazemi K., Laitala J., Azimi I., Liljeberg P., Rahmani AM.,* "Robust PPG Peak Detection Using Dilated Convolutional Neural Networks, " Sensors (Basel) 2022; **vol.** 22, no. 16, 2022, p. 6054, https://doi.org/10.3390/s22166054
[14] *Pollreisz, D., TaheriNejad, N.,* "Detection and Removal of Motion Artifacts in PPG Signals" Mobile Networks and Applications **vol.** 27, pp. 728–738 (2022). https://doi.org/10.1007/s11036-019-01323-6
[15] *J. Atkinson and D. Campos*, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," Expert Syst. Appl., **vol**. 47, pp. 35–41, 2016
[16] *H. J. Yoon and S. Y. Chung*, "EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm," Comput. Biol. Med., **vol**. 43, no. 12, pp. 2230–2237, 2013
[17] *I. Daly et al*., "Identifying music-induced emotions from EEG for use in brain-computer music interfacing," in 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), 2015.
[18] *Y. Liu and O. Sourina*, "EEG databases for emotion recognition," in 2013 International Conference on Cyberworlds, 2013
[19] *S. Koelstra et al*., "Single trial classification of EEG and peripheral physiological signals for recognition of emotions induced by music videos," in Brain Informatics, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 89–100, 2010
[20] *M. Ben and Z. Lachiri*, "Emotion classification in arousal valence model using MAHNOB-HCI database," Int. J. Adv. Comput. Sci. Appl., **vol**. 8, no. 3, 2017
[21] *S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena*, "EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation," Scientific World Journal, vol. 2014, p. 627892, 2014
[22] DEAP Dataset. Available online: http://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html
[23] *M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic*, "A multimodal database for affect recognition and implicit tagging," IEEE Trans. Affect. Comput., **vol**. 3, no. 1, pp. 42–55, 2012
[24] *L. Petrescu et al*., "Machine learning methods for fear classification based on physiological features," Sensors (Basel), **vol.** 21, no. 13, 2021, pp. 4519
[25] BITalino. Available: https://www.pluxbiosignals.com/collections/bitalino
[26] Shimmer Sensing. Available: https://shimmersensing.com/
[27] Biopac. Available: https://www.biopac.com/
[28] *L. Petrescu et al*., "Integrating biosignals measurement in virtual reality environments for anxiety detection," Sensors (Basel), **vol**. 20, no. 24, 2020, p. 7088