A Low Cost EDA-based Stress Detection Using Machine Learning

Elahe Hosseini
Electrical and Computer Engineering
University of California, Davis
Davis, USA
ehosseini@ucdavis.edu

Ruijie Fang
Electrical and Computer Engineering
University of California, Davis
Davis, USA
rjfang@ucdavis.edu

Ruoyu Zhang
Electrical and Computer Engineering
University of California, Davis
Davis, USA
ryuzhang@ucdavis.edu

Anna Parenteau
Department of Psychology
University of California, Davis
Davis, USA

Sally Hang
Department of Psychology
University of California, Davis
Davis, USA

Setareh Rafatirad
Computer Science
University of California, Davis
Davis, USA

Camelia Hostinar
Department of Psychology
University of California, Davis
Davis, USA

Mahdi Orooji
Electrical and Computer Engineering
University of California, Davis
Davis, USA

Houman Homayoun
Electrical and Computer Engineering
University of California, Davis
Davis, USA

2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 978-1-6654-6819-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/BIBM55620.2022.9995093

***Abstract*—Stress is an inevitable part of our lives in modern society since in many situations people are exposed to various stressors daily. According to studies, long-term stress can cause mental and physical diseases such as depression, anxiety, high blood pressure, heart attacks, and stroke. Therefore, stress detection is one of the crucial areas of study to maintain a healthy life. Recently, by developing commercial wearable technologies, real-time and continuous data collection for personal stress monitoring becomes more feasible. Under stress conditions, there are notable changes in physiological signals such as heart rate, respiration, perspiration, and eye pupil dilation. Previous studies have shown that Electrodermal Activity (EDA), also known as Galvanic Skin Response (GSR), can identify stress. EDA measures changes in perspiration by detecting the changes in the electrical conductivity of the skin. This paper focuses on stress detection using only EDA wearable sensors and applied machine learning techniques. First, 87 different features are extracted from EDA signals. Then, the data are normalized per subject because of differences in individuals' physiological responses. Finally, five dominant features in stress detection are selected. We used a publicly available dataset, namely, the wearable stress and affect detection dataset (WESAD) in this study. The results show that the One-Leave-Out method is capable of detecting stress with 97.03% accuracy.**

***Index Terms*—Applied Machine Learning, Health care information system, Stress Detection, Physiological Signals**

I. Introduction

One in four people in modern society experience stress as a mental health problem [1]. There are many mental and physical consequences of stress including depression, anxiety, and suicide in severe cases, as well as high blood pressure, strokes, and heart attacks [2]. Studies have shown that stress may affect the immune system and enhance the possibility of cancer. In addition, there are some other stress effects on humans' relationships and work functioning that can decrease their quality of life. Recently, people are more aware of these effects and the demand for counseling to cope with stress is increasing. On the other hand, currently, the most common way for psychological and physiological specialists to determine stress is based on questionnaires [3]. This approach depends on the individuals' answers to the questionnaire which is subjective and may be time-consuming. Because some questionnaires have good reliability. Therefore, automatic stress detection can replace the traditionally questionnaire-based method to minimize uncertainty and improve society's well-being. Many studies have indicated that there is a strong correlation between the human body's physiological signals, such as Photoplethysmogram (PPG), Electrodermal Activity (EDA), respiration rate, pupil size, body temperature, and stress [4]–[7]. By exploiting these physiological signals, we are capable to detect stress automatically. With the emergence of wearable devices, continuous and real-time physiological data collection for health monitoring becomes more practical.

The Hypothalamicpituitary-Adrenal (HPA) and autonomic nervous system (ANS) are major systems for stress response. The nervous system is divided into the central and peripheral nervous systems (CNS and PNS). ANS is a part of PNS which is composed of sensory and motor neurons and operates between the central nervous system and internal organs, such as the heart, the lungs, and the sweat glands [8].

Studies show that among all the physiological signals, Electrodermal Activity (EDA), or Galvanic Skin Response (GSR), is one of the most reliable stress indicators [9]–[13]. It measures the change in sweat glands' activity which reflects changes in arousal and is driven unconsciously by ANS in order to meet behavioral requests. There are about three million sweat glands across the body that produce moisture when they are triggered. An EDA signal has two components: A tonic (skin conductance level, SCL) reflects a steadily increasing baseline conductivity, and a phasic response (skin conductance response, SCR) that reflects a rapidly changing reaction to a specific arousing stimulus, which may be evident as pulses or bursts [14]. GSR signals can be collected with devices like the *RespiBAN professional*[1], *Empatica E4*[2], and *Shimmer3 GSR+*[3].

The main contributions of this paper can be summarized as follows:

- We propose a low-cost machine learning-based stress detection system using only the EDA sensor. In the preprocessing phase, natural and noisy EDA is clean, and then in the feature extraction phase, different features of phasic and tonic (SCR and SCL) EDA are extracted.
- We introduce a 'normalization per subject' method that can eliminate the variation in individuals' body physiological signals.
- To build an effective and efficient stress detection system, we selected the most informative features as the input of the ML model. Based on our experiments, Mean of Tonic (SCL), Max of Tonic (SCL), Max of Phasic (SCR) Peaks, Standard Deviation of Phasic (SCR) Rise Time, and Number of Peaks in Phasic (SCR) are the most discriminative statistical features for stress detection.
- We conducted experiments on the public Wearable Stress and Affect Detection (WESAD) dataset [8] to evaluate our proposed stress detection system. Results show that the proposed method detects stress with 97.03% accuracy.

The rest of the paper is organized as follows. Section II describes an overview of the proposed methodology for stress detection. In Section III, we describe the data preprocessing stage to be ready for feature extraction. Section IV discusses feature extraction, normalization and feature selection. Section V presents the experimental results. Finally, conclusions are discussed in Section VI .

II. Methodology

Fig. 1 shows our proposed EDA-based stress detection system. There are three main stages in this pipeline: 1) the preprocessing stage, 2) feature extraction and feature selection, and 3) machine learning-based classification. According to Fig. 1, the preprocessing stage consists of up-sampling and denoising signal using a low-pass filter, then segmenting the EDA signal into 60 seconds windows to make data suitable for feature extraction. In the next stage, we use two different algorithms to extract the features. In the final stage, the five most dominant features are fed into the AdaBoost classifier to predict the stress state. In the following sections of the paper, we explain each stage of the pipeline in detail.

[1]https://www.pluxbiosignals.com/collections/wearables
[2]https://www.empatica.com/research/e4/
[3]https://shimmersensing.com/product/shimmer3-gsr-unit/

III. Data Preprocessing

In this stage, we first up-sample and filter the "wrist" EDA raw signal from WESAD. Then clean data are segmented into windows for further analysis and feature extraction. Further details of each step are discussed below.

A. Up-sampling and Denoising

The data considered in this paper belongs to baseline (20 minutes) and TSST test to stimulus stress (10 minutes) affective states from *Empatica E4* wristband with a 4 Hz sample rate. We upsample the EDA signal to 8 Hz, because in this study one of the toolkits that we utilize to extract features is NeuroKit2 [15], and this toolkit is capable to process EDA with a higher than 8 Hz sample rate. The natural raw EDA signals contain lots of noise due to power line interferences, movements, and faulty connections. Therefore, a low-pass Butterworth filter is applied to the EDA signal with a cutoff frequency of 5 Hz divided by sampling. Fig. 2 shows an example of the raw EDA data from subject 9 in the WESAD dataset collected with a 4 Hz sampling rate from *Empatica E4* wristband and the data after upsampling to 8 Hz and filtering to denoise the signal.

B. Segmentation

The 30 minutes experiment is segmented into 60 seconds sliding windows with a 0.25-second shift. In [6], the effect of different window sizes on detection rate is investigated and this study shows that the classifier obtains higher accuracy on longer window sizes. Out of all segments, 65% belong to the no-stress class and 35% represent the stress condition.

IV. Feature Extraction

In this stage of our proposed stress detection system, we analyzed EDA signals. In order to properly analyze the EDA signal, it is necessary to correctly extract the phasic component from the original signal. Then we need to find the maximum amplitude value in each onset-offset window of the EDA signal to calculate the number of peaks, amplitude values, rise time, amplitude, and recovery time. Essentially, onsets and offsets occur when the phasic component of the signal crosses above $0.01\mu S$ and below $0\mu S$, respectively. Therefore, we divided EDA into high variation phasic (SCR) and smooth tonic (SCL) components. We utilize NeuroKit2 [15] and cvxEDA [16] algorithm to detect peaks and onsets. Fig. 3 presents the characteristics of EDA. Finally, after finding all existing EDA characteristics, 87 different statistical features are calculated.

Since the physiological signals response across different individuals is different, the normalization technique on the extracted features is useful for machine learning classification. Equation 1 illustrate min-max normalization:

$$x_{norm} = \frac{x - min(x)}{max(x) - min(x)} \tag{1}$$

Therefore, we normalize data per subject in this paper. The next step in this stage is feature selection. Section V discussed the effectiveness of normalization per subject on classification

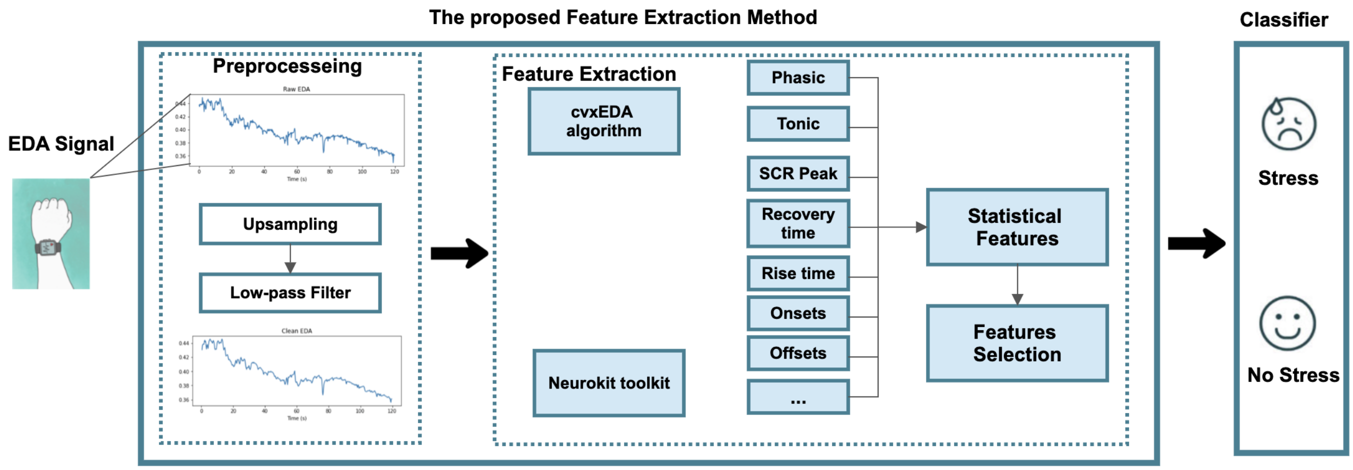

Fig. 1. The proposed low cost EDA-based stress detection system

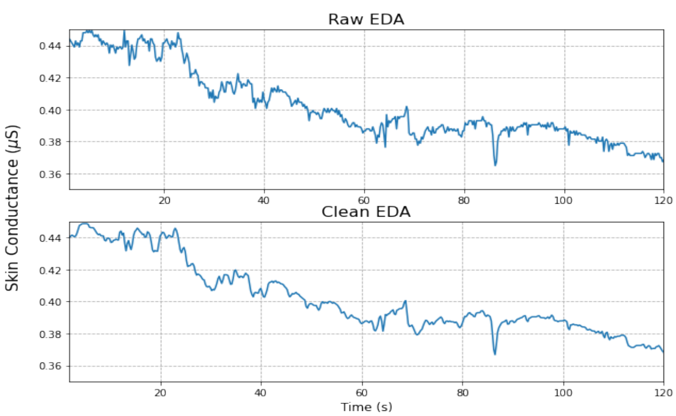

Fig. 2. Raw EDA data and EDA data after preprocessing

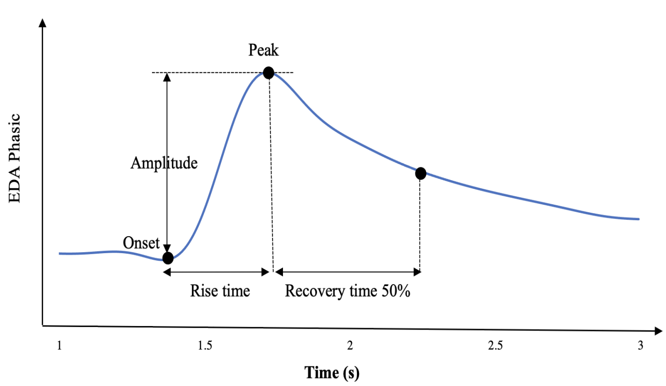

Fig. 3. EDA phasic characteristics

accuracy. We select the most informative and discriminative features for machine learning-based classification by using extensive experiments to test and evaluate classification on each feature.

V. Experimental Results and Evaluations

We conduct three sets of experiments in this section:

- ML-based classification on all extracted feature
- Investigation of the effectiveness of normalizing extracted features per subject on classification and illustrate its importance in physiological signals analysis
- Selecting the most informative features reduces the ML training and testing time alongside increases the performance of machine learning algorithms.

We use three different machine learning algorithms: 1) AdaBoost, 2) Random Forest, and 3) SVM to evaluate our extracted features for stress detection.

A. WESAD Dataset

This study is conducted based on a publicly available multimodal dataset called WESAD [8] for wearable stress and affect detection. This dataset contains physiological and motion data recorded from the *Empatica E4* (on wrist) and *RespiBAN professional* (on chest) devices. All 15 (12 male, 3 female) participants in this completed the Trier Social Stress Test (TSST) [17] in order to induce moderate psychological stress in a laboratory setting. The goal of this dataset is to elicit three affective states: "baseline" (sitting or neutral reading), "amusement" (watching eleven funny video clips), and "stress" (being exposed to the TSST). This dataset contains electrocardiogram (ECG), blood volume pulse (BVP), electromyogram (EMG), respiration (RESP), skin temperature (TEMP), 3-axis accelerometer (ACC), Heart Rate (HR), and EDA. There are two different versions of the study protocol in this data set. In our study, we consider baseline states as a non-stress class to compare with stress class in a binary classification task. Our model is created based on the EDA data collected from the *Empatica E4* wristband with a 4 Hz sample rate.

B. Classification Results

In this paper, three sets of experiments have been conducted on EDA extracted features to predict stress conditions.

The first goal includes the performance improvement with normalization per subject. The second goal is to select the most dominant and discriminative features to increase the accuracy of the classification task and reduce the execution time. The third goal is to investigate three different machine learning algorithms in the proposed stress detection system and comparing with the state-of-the-art methods. We employ leave-one-subject-out as our cross-validation scheme, which means the classification models are trained on all data except one unseen subject which is used in testing. The classification performance is evaluated in terms of average accuracy, precision, and F1-score. All experiments are summarized in Table. I and Fig. 4.

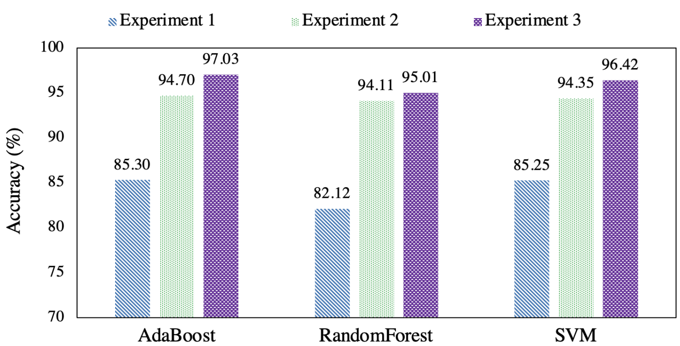

Fig. 4. Leave-One-Out average accuracy on three different experiments for stress detection using different machine learning algorithms

TABLE I
THE AVERAGE PERFORMANCE OF THE PROPOSED METHOD FOR DIFFERENT EXPERIMENTS

Experiments	Model	Accuracy	Precision	F1-Score
1	AdaBoost	85.30	84.80	83.70
	RandomForest	83.21	83.56	81.63
	SVM	85.25	86.60	84.40
2	AdaBoost	94.70	94.96	94.55
	RandomForest	94.11	94.26	93.87
	SVM	94.35	94.35	94.14
3	AdaBoost	**97.03**	**97.36**	**97.03**
	RandomForest	95.01	95.62	94.86
	SVM	95.95	96.42	95.76

1) Experiment 1: Classification on All Extracted Features: In this experiment, we trained a classifier on all extracted features. Results show 85.30% accuracy on average with the AdaBoost classifier, which means even without considering differences in an individual's physiological response to a stress condition, EDA extracted features can discriminate stress vs. no stress conditions. Table. I displayed the evaluation results of this experiment as *Experiment 1*. Also, Fig. 4 visualized the results as the blue columns.

2) Experiment 2: Classification on All Normalized Extracted Features: Individuals may suffer from anxiety or other mental illness. Even in healthy individuals, drinking caffeine can affect the nervous system which is responsible for physiological signals responses in each condition. These individual differences in humans motivate us to find new approaches to analyzing EDA features for each person based on her/his specific body response. Therefore, we normalize each feature for each person. The experiments prove the above state about the difference in individual's body signals and show more than 10 percent improvement in terms of accuracy after normalization. Table. I and Fig. 4 illustrate this experiment as *Experiment 2*.

3) Experiment 3: Classification on Selected Normalized Features: There are 87 features extracted from EDA components, but they are not all useful for the machine learning training process and they can degrade the classifier's performance. Therefore, we select the 5 best features for the ML classifier. First, we conducted extensive experiments and trained ML classifiers on each extracted feature using the Leave-One-Subject-Out method. Then we selected the best feature that achieved the highest accuracy. In the next iteration, we add another feature to the best feature and train and test the model. We repeated this procedure until the best performance is obtained. Fig. 5 depicted our approach to selecting features by adding one by one feature the previous list of features. The final selected features are: 1) Mean of Tonic (SCL), 2) Max of Tonic (SCL), 3) Max of Phasic (SCR) Peaks, 4) Standard Deviation of Phasic (SCR) Rise Time, and 5) Number of Peaks in Phasic (SCR).

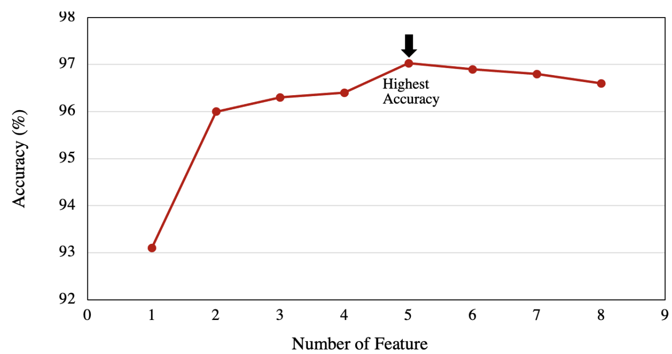

Fig. 5. Feature selection

C. Comparison with Other Works

In this paper, we only used EDA data collected from *Empatica E4* a wearable device that may not have very high accuracy compared to the professional and medical-grade devices like *RespiBAN*. We compare our proposed with the existing state-of-the-art methods in this section on the WESAD dataset. The comparison results are shown in Table. II. The proposed framework shows promising performance and effectiveness compared to the state-of-the-art methods and outperforms the existing methods in terms of accuracy. Lai *et al.* achieved 97.75% accuracy, but they utilized 7 physiological signals where some of them (such as ECG and EMG) can not be collected by commercial wearable devices. Therefore, we are capable to predict stress using only EDA sensor from a wearable device with high accuracy. Although Aqajari *et al.* [11] proposed an EDA-based stress detection, which utilized the deep learning features, extracting these features needs time and computational power which is not proper for embedded and wearable devices with limited resources.

TABLE II
COMPARISON OF THE PROPOSED FRAMEWORK RESULTS WITH THE STATE-OF-THE-ART RESULTS IN WESAD

	Signals	Device	Accuracy
Schmidt *et al.* [8]	ACC, TEMP, EDA, EMG, RESP, ECG, BVP	*Empatica E4* *RespiBAN*	92.28
Bobade *et al.* [4]	ACC, TEMP, EDA, EMG, RESP, ECG, BVP	*Empatica E4* *RespiBAN*	95.21
Lai *et al.* [7]	ACC, TEMP, EDA, EMG, RESP, ECG, BVP	*Empatica E4* *RespiBAN*	97.75
Siirtola [6]	TEMP, BVP, HR	*Empatica E4*	87.40
Hsieh *et al.* [10]	EDA	*Empatica E4* *RespiBAN*	92.38
Aqajari *et al.* [11]	EDA	*RespiBAN*	91.60
Ours	**EDA**	***Empatica E4***	**97.03**

VI. CONCLUSION

This paper introduces a framework for stress detection from EDA signal using machine learning algorithms. First, we extract 87 different features of EDA from raw signal, its phasic and tonic components. Second, we utilize normalization per person method because of difference between individuals' body responses and physiological signals. Then, we select the most dominant and informative features for machine learning classification. The analysis of feature importance shows that with 5 features, the classifier performance increases. In addition, the execution time and computational cost decrease because the input size of the classifier decreases after the feature selection stage. According to our studies, we can acquire 97.03% accuracy and 97.36% F1-Score. Moreover, in comparison of the proposed framework with state-of-the-art research works, the results indicate a promising improvement.

REFERENCES

[1] *Communications, N.. World health report*, 2001. http://www.who.int/whr/2001/media$_centre/press_release/en/$.

[2] B. S. McEwen, "Brain on stress: how the social environment gets under the skin," *Proceedings of the National Academy of Sciences*, vol. 109, no. Supplement 2, pp. 17180–17185, 2012.

[3] E. Andreou, E. C. Alexopoulos, C. Lionis, L. Varvogli, C. Gnardellis, G. P. Chrousos, and C. Darviri, "Perceived stress scale: reliability and validity study in greece," *International journal of environmental research and public health*, vol. 8, no. 8, pp. 3287–3298, 2011.

[4] P. Bobade and M. Vani, "Stress detection with machine learning and deep learning using multimodal physiological data," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 51–57, IEEE, 2020.

[5] H. J. Han, S. Labbaf, J. L. Borelli, N. Dutt, and A. M. Rahmani, "Objective stress monitoring based on wearable sensors in everyday settings," *Journal of Medical Engineering & Technology*, vol. 44, no. 4, pp. 177–189, 2020.

[6] P. Siirtola, "Continuous stress detection using the sensors of commercial smartwatch," in *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, pp. 1198–1201, 2019.

[7] K. Lai, S. N. Yanushkevich, and V. P. Shmerko, "Intelligent stress monitoring assistant for first responders," *IEEE Access*, vol. 9, pp. 25314–25329, 2021.

[8] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM international conference on multimodal interaction*, pp. 400–408, 2018.

[9] F. Seoane, I. Mohino-Herranz, J. Ferreira, L. Alvarez, R. Buendia, D. Ayllón, C. Llerena, and R. Gil-Pita, "Wearable biomedical measurement systems for assessment of mental stress of combatants in real time," *Sensors*, vol. 14, no. 4, pp. 7120–7141, 2014.

[10] C.-P. Hsieh, Y.-T. Chen, W.-K. Beh, and A.-Y. A. Wu, "Feature selection framework for xgboost based on electrodermal activity in stress detection," in *2019 IEEE International Workshop on Signal Processing Systems (SiPS)*, pp. 330–335, IEEE, 2019.

[11] S. A. H. Aqajari, E. K. Naeini, M. A. Mehrabadi, S. Labbaf, A. M. Rahmani, and N. Dutt, "Gsr analysis for stress: Development and validation of an open source tool for noisy naturalistic gsr data," *arXiv preprint arXiv:2005.01834*, 2020.

[12] S. A. H. Aqajari, E. K. Naeini, M. A. Mehrabadi, S. Labbaf, N. Dutt, and A. M. Rahmani, "pyeda: An open-source python toolkit for preprocessing and feature extraction of electrodermal activity," *Procedia Computer Science*, vol. 184, pp. 99–106, 2021.

[13] L. Shu, J. Xie, M. Yang, Z. Li, Z. Li, D. Liao, X. Xu, and X. Yang, "A review of emotion recognition using physiological signals," *Sensors*, vol. 18, no. 7, p. 2074, 2018.

[14] IMotions, "Galvanic skin response—the complete pocket guide," 2017.

[15] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. H. A. Chen, "NeuroKit2: A python toolbox for neurophysiological signal processing," *Behavior Research Methods*, vol. 53, pp. 1689–1696, feb 2021.

[16] A. Greco, G. Valenza, A. Lanata, E. P. Scilingo, and L. Citi, "cvxeda: A convex optimization approach to electrodermal activity processing," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 797–804, 2015.

[17] 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*, vol. 28, no. 1-2, pp. 76–81, 1993.