Citation:
K. Jambhale et al., "Selection of Optimal Physiological Features for Accurate Detection of Stress," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022, pp. 2514-2517, doi: 10.1109/EMBC48229.2022.9871067.

DOI:
https://doi.org/10.1109/EMBC48229.2022.9871067

Selection of Optimal Physiological Features for Accurate Detection of Stress

Kiran Jambhale, Benjamin Rieland, Smridhi Mahajan, Prajakta Narsay, Nilanjan Banerjee, *Member, IEEE,* Abhijit Dutt, *Member, IEEE,* and Ramana Vinjamuri, *Senior Member, IEEE*

Abstract— **Stress is an established risk factor in the development of addiction and in reinstating drug seeking. Substance use disorder (SUD) is a dangerous epidemic that affects the brain and behavior. Despite this growing epidemic and its subsequent consequences, there are limited management and treatment options, pharmacotherapies and psychosocial treatments available. To this end, there is a need for new and improved personalized devices and treatments for the detection and management of SUD. Based on documented negative effects of stress in SUD, in this paper, our objective was to select a few significant physiological features from a set of 8 features collected by a chest-worn *RespiBAN Professional* in 15 individuals. We used three machine learning classifiers on these optimal physiological features to detect stress. Our results indicate that best accuracies were achieved when electrodermal activity (EDA), body temperature and chest-worn accelerometer were considered as features for the classification. Challenges, implications and applications were discussed. In the near future, the proposed methods will be replicated in individuals with SUD.**

I. Introduction

Substance use disorder (SUD) is a dangerous disease that affects an individual's brain and behavior. This leads to uncontrolled use of illicit drugs, alcohol, excessive use of legal drugs or other addictive behaviors. The prevalence (51.5 million adults with SUD and mental illnesses, Substance Abuse and Mental Health Services Administration (SAMHSA) 2019[1]) and the rate of increase (6% from 2018 to 2019) of SUD in the US deem it a rapidly growing epidemic [2]. Furthermore, during COVID19 pandemic, one of the serious challenges faced by the American Society of Addiction Medicine (ASAM) is the treatment of homeless individuals with SUD because of their compromised immune systems [3]. The National Institute on Drug Abuse (NIDA) estimates that the total expenditure of drug-related complications exceeds 500 billion dollars when healthcare costs and job losses are considered [4]. Despite this growing epidemic and its subsequent consequences, there are limited management and treatment options, pharmacotherapies, and psychosocial treatments available for SUD. To this end NIDA's mission and strategic plan emphasize the importance of the development of new and improved strategies and treatments for detection and management of SUD [4].

Decades of research has shown that stress increases risk of substance abuse [5][6][7] and could be a hindrance to effective treatment of substance abuse. Currently, measurement of stress is usually done through self-report [8] that has several practical limitations. As society becomes more receptive to the use of wearable biomedical sensors in the form of smart watches and other sensors [9], we envisage that the use of wearables to measure physiological signals such as electrodermal activity (EDA) will be ubiquitous.

EDA sensors could inform us about several vital features about the human body such as emotions, stress, etc. [10]. Our main goal in this research is to uncover a robust method for measurement of stress using EDA sensors. In order to be useful, such measurement must be accurate and fast. These requisite demands make it especially challenging because these sensors generate massive amounts of data and subsequent processing needs enormous computing power. Hence, it is necessary to identify and design efficient detection and estimation algorithms. EDA sensors provide continuous recording of a subject's stress and emotion that cannot be obtained in any other way. However, more studies are required to measure the accuracy and reliability of this technique. Previous research has shown that a strong causal relationship exists between stress and emotion [11]. As both emotion and stress could be measured independently using EDA sensors, it is possible that by adding emotions, we would be able to measure stress more accurately.

The negative effects of SUD manifest themselves in substance users in the form of stress, among other symptoms. When one is physically dependent on a substance to get them through their daily activities and experiences, stepping away from that substance can be psychologically stressful. This may lead to relapse or a prolonged experience of withdrawal symptoms, or post-acute withdrawal syndrome [3] that also induces stress. When under stress, the body's sympathetic nervous system is activated, and this results in psychological and physiological changes within the body [4]. Since psychological changes are not capable of being tracked well using devices and external methods of treatment, tracking physiological changes is the best approach when tracking SUD-related stress. These changes within the body's function include increased heart rate, greater electrodermal activity, and a higher body temperature among others. Therefore, the observation of changes within these parameters will allow for detection of the onset of SUD relapse.

Our group at UMBC in collaboration with psychiatrists and clinicians is poised to develop wearable devices that can help in the detection and management of SUD. In this paper, using WESAD [12], a publicly available dataset that has tracked

*Research supported by National Science Foundation (CAREER Award HCC-2053498) and START Award from UMBC.

K. Jambhale, B. Rieland, S. Mahajan, P. Narsay, N. Banerjee, A. Dutt, and R. Vinjamuri are with the Sensorimotor Control Laboratory (Vinjamuri Lab) and the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County (UMBC), Baltimore, 21220, USA (e-mail: rvinjam1@umbc.edu).

physiological features across multiple different emotional states, our objective was to identify the best physiological features for accurate detection of stress. The mental and social stress induced during the stress condition of the study made this dataset a good choice for studying the effects of SUD in which the same type of stress is induced.

II. Methods and Analysis

A. Experiment

To study the physiological responses to SUD-related stress, a publicly available multimodal dataset for wearable stress and affect detection (WESAD), was used. The WESAD dataset offers data collected by wearable devices from 12 male subjects and 3 female subjects with a mean age of 27.5 ± 2.4 years. The two wearable devices were *RespiBAN Professional* (chest-worn) and Empatica E4 (wrist-worn), which were equipped with sensors to track 3 axis accelerometers worn on the chest (ACC), respiration (RESP), electrocardiogram

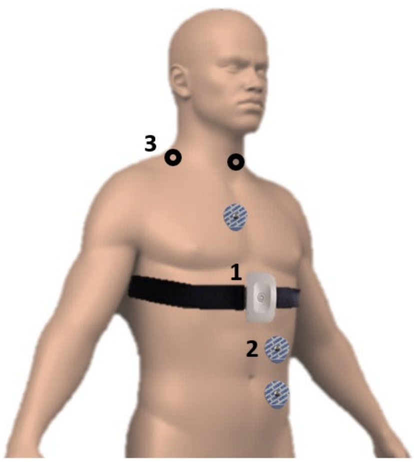

Fig. 1. *RespiBAN Professional*'s placement of electrodes. 1. *RespiBAN Professional* with temperature, EDA and control module. 2. Three ECG electrodes and 3. Two EMG electrodes on the back where the shoulder meets the neck.

(ECG), electrodermal activity (EDA), electromyograph (EMG), and body temperature (TEMP) as shown in Fig. 1. During the WESAD study, the participants were guided through various activities to simulate one of four emotional states - baseline, stress, amusement, and meditation as shown in Fig. 2. Here are brief descriptions of these states.

- Baseline: Subjects were sitting/standing at a table and reading neutral material
- Stress: Subjects were exposed to the Trier Social Stress Test in which they had to complete highly strenuous tasks - mental arithmetic (mental stress) and public speaking (social stress).
- Amusement: Subjects were shown funny video clips
- Mediation: Subjects were guided through meditation exercises

After each condition, the participants were asked to fill out a questionnaire which consisted of the Positive and Negative Affect Schedule, State-Trait Anxiety Inventory, and Self-Assessment Manikins tests. These tests offered prompts relating to different emotional states to which the participants

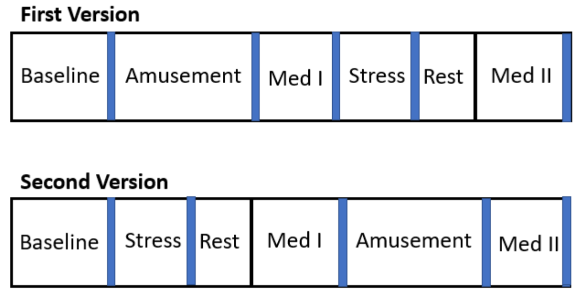

Fig. 2. The two protocols tested under this study. The blue bars indicate the times when the study participants filled out questionnaires for self-report. Adapted from [12].

had to assign a number rating. These ratings were used as a standard ground truth to evaluate the models for stress detection.

The data collected by *RespiBAN Professional* were used to conduct a computational analysis on which features would best predict the onset of stress in a repeatable, timely, and accurate manner. *RespiBAN Professional* was chosen due to its higher volume of data points (2+ million per subject) as compared to Empatica E4's 20-50,000 data points per subject.

B. Preprocessing and Analysis

This exploratory and predictive data analyses have been performed using MATLAB® Machine Learning Toolbox and Python. For Python, we have made use of various Python libraries like pandas, sklearn, matplotlib, and numpy. For the various algorithms that were implemented, we have used various tools provided by the sklearn library. The integrated development environment (IDE) used was Jupyter Lab in congruence with the Anaconda platform.

The dataset comprises various physiological signals and a label that specifies if the subject is stressed, amused, meditating or relaxed. As a first step in preprocessing the missing values in the dataset were addressed. The missing values have been dealt with by substitution methods such as mean or mode as per the distribution of the data. After this, the data was normalized to have optimal distributions. Normalization assigned equal weights and statistical importance to each variable so that no single variable drives the model performance in one direction and the prediction is not skewed.

Feature selection was performed to find the most contributing physiological features to accurate detection of stress using different methods like logistic regression, linear regression, and principal component analysis (PCA). We have also performed sequential forward feature selection using quadratic discriminant analysis to find the best features (see Table 1). This helped us in feature analysis for each subject.

C. Classification

Three types of classification were performed: (1) 2-way: stress vs amusement (2) 3-way: stress vs amusement vs meditation, and (3) 4-way: stress vs amusement vs meditation vs baseline. We have performed predictive analysis using 3 approaches - logistic regression, decision trees, and XGBoost (gradient boosted decision trees). We have also performed 5-fold cross-validation and the results of the predictive analysis

TABLE I. SIGNIFICANCE OF THE FEATURES

	ECG	EDA	EMG	Resp	Temp	X	Y	Z
S2	2	5	1	3	6	0	4	6
S3	2	5	3	1	6	0	4	6
S4	2	4	3	1	6	0	5	6
S5	2	6	3	1	4	0	5	6
S6	2	6	1	3	4	0	5	6
S7	1	5	2	3	6	0	4	6
S8	1	5	3	2	4	0	6	6
S9	2	6	1	3	5	0	4	6
S10	2	6	1	3	4	0	5	6
S11	0	4	1	3	5	2	6	6
S13	1	6	3	2	4	0	5	6
S14	2	4	1	3	5	0	6	6
S15	2	6	1	3	0	4	5	5
S16	1	6	2	3	4	0	5	6
S17	3	6	4	2	5	0	1	6

are measured using accuracy and area under the curve as metrics. While the accuracy is a standardized metric, it becomes essential in such problem statements to also carefully assess the accuracy that could not be achieved to minimize the false positives. Hence, for binary classification, we have also studied the area under the curve to check the degree of separation between true positives and false positives.

III. RESULTS

In this study, as a first step, in order to determine most significant combination of physiological features, we performed sequential forward feature selection using quadratic discriminant analysis. For each subject we looked at what will be the most significant 2, 3, 4, 5, and 6 features. For a total of these 6 cases, the results from this analysis are summarized in Table 1. All the features are mapped against the subject which corresponds to a value. The maximum value can be six and the minimum can be zero. The numerical value is the number of times the feature was selected in the combination with other features corresponding to the subject. If the value is six, it means that the feature was always present in all the combinations and is the most relevant feature for a particular subject. The value zero signifies the absence of that feature in the any combination carried out for that subject. Hence, it is the least relevant feature to be considered. A total of six experiments were conducted with various feature counts ranging from a combination of two features to combination of eight features. This selection was different for different subjects.

The most used feature is the accelerometer Z axis (Z) (perpendicular to the subject's chest). This feature was considered in every combination for all but one of the 15 subjects. The second most used feature was EDA, which was included all 6 cases with 8 subjects. The third feature most used was temperature (Temp), being included in all 6 cases for 4 subjects. Finally, the least significant feature in any combination was accelerometer X axis (X), since it was used in 0 combinations for 13 subjects. However, considering that accelerometer as a whole contributed to the classification accuracy, this variable cannot be ignored for a future analysis.

After conducting feature analysis with logistic regression and PCA on each individual subject, several features stood out as more important than the others. Specifically, these features had a much higher correlation to a subjects' mental and emotional states and were EDA, temperature, and the accelerometer (z-axis). Table 2 showed the results obtained from logistic regression with 2, 3 and 4-way multivariate classification. Overall, logistic regression had an average accuracy (ACC) of 0.969, and an average AUC-ROC of 0.985 across all subjects.

In addition to logistic regression, we repeated similar classification using decision tree. The decision tree not only favored the same features as logistic regression (EDA, temperature, and accelerometer Z Axis), it also performed slightly better than Logistic Regression, with an average AUC-ROC of 0.998 and an average accuracy of 0.968. Finally, we used another classification algorithm XGBoost. This classifier outperformed both logistic regression and decision tree, while

TABLE II. CLASSIFICATION ACCURACY

		AUC-ROC	ACC	Top 3 Features							
				ECG	EDA	EMG	Resp	Temp	X	Y	Z
S2	2Way	1	1		X			X		X	
	3Way	1	.999		X			X			X
	4Way	.999	.997	X	X						X
S3	2Way	.89	.82	X	X			X			
	3Way	.945	.949		X			X			X
	4Way	.965	.886					X		X	X
S4	2Way	1	1		X					X	X
	3Way	.999	.996		X			X			X
	4Way	.999	.994		X			X	X		
S5	2Way	1	1		X					X	X
	3Way	.999	.996		X			X			X
	4Way	.999	.994		X			X	X		
S6	2Way	.997	.988		X			X		X	
	3Way	.982	.944					X		X	X
	4Way	.948	.873		X			X			X
S7	2Way	1	1	X	X	X					
	3Way	.999	.995		X			X		X	
	4Way	.947	.803		X			X			X
S8	2Way	.999	.998	X						X	X
	3Way	.999	.996					X		X	X
	4Way	.999	.999					X		X	X
S9	2Way	.993	.982		X			X			X
	3Way	.995	.981		X			X	X		
	4Way	.982	.971		X			X	X		
S10	2Way	1	1	X	X		X				
	3Way	.999	.999		X			X			X
	4Way	.999	.997		X			X			X
S11	2Way	1	1					X	X	X	
	3Way	.997	.971		X			X			X
	4Way	.988	.929		X			X			X
S13	2Way	1	1	X	X	X					
	3Way	.998	.981		X			X		X	
	4Way	.998	.981		X			X		X	
S14	2Way	.919	.903		X			X	X		
	3Way	.873	.819					X		X	X
	4Way	.932	.886		X			X			X
S15	2Way	1	1		X				X	X	
	3Way	.999	.998		X			X	X		
	4Way	.998	.992		X			X	X		
S16	2Way	1	1		X	X	X				
	3Way	.999	.998		X			X			X
	4Way	.999	.998		X			X			X
S17	2Way	1	1		X			X	X		
	3Way	.999	.995		X			X			X
	4Way	.999	.997		X			X			X

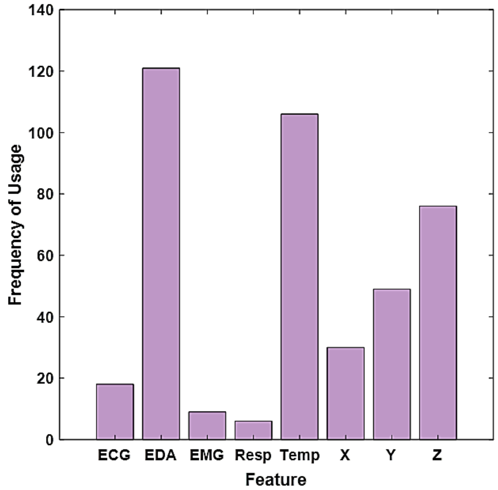

Fig. 3. Significance of the features based on how frequently they were used in three different classifiers for best accuracies. EDA, temperature and accelerometer Z stand out as the important features.

also favoring the EDA, temperature, and accelerometer Z axis features. Its average AUC-ROC was 0.998 and average accuracy 0.995.

IV. Discussion

The objective of this paper was to select optimal physiological features that can predict stress accurately. For 15 subjects, we employed 3 classifiers to perform 3 types of classification. In summary, there were a total of 45 tests runs with each of the 3 classifiers, for a total of 135 total individual tests. Of the 135, EDA was the most relevant feature, chosen as one of the top three, 121 times. Temperature was the second most relevant feature, being chosen 106 times as a top feature. Finally, the third most relevant feature was the accelerometer Z Axis, making the top 3 features 76 times. These along with other features were shown in Fig. 3. The task of classification of whether the subject is stressed or not stressed depends upon various factors and so considering only the most important features don't give relevant results. Hence finding out relevant combinations of different features was a vital task.

Similar to other studies [13][14][15], that have reported previously we found that EDA, temperature, and accelerometer are important features in stress detection. Our objective was not to find one optimal feature, but multiple features that can help in precise detection. Fusion of multimodal features improves the detection of accuracy[15]. As shown in Table 2, optimal features changed from subject to subject. This enables us to build personalized models that are unique to individuals.

Although our objective was to detect stress as a trigger in the context of substance abuse, stress leads to several other complications, thus the developed methods have multiple applications in different domains. Stress can either be a trigger or can aggravate many pathological conditions. Stress affects cognitive functions, weakens memory, increases blood pressure, causes cardiac disorders, diabetes, to list a few. In each of these disease conditions, stress manifests differently. Thus, multimodal fusion of features helps in improving the detection accuracy across domains.

V. Conclusion

Motivated by the detrimental effects of stress in SUD we set out to find what are the optimal physiological signals easily measured by wearables that can help in accurate detection of stress. We found that EDA, body temperature and chest-worn accelerometer contribute significantly to accurate detection and classification of stress and other emotional states. In the near future, we will separately analyze mental and social stressors. This will assist in the detection of situational stressors in individuals with SUD. We will also look at unique optimal physiological features in individuals to customize detection models. In this paper, we focused on stress and other emotional states, but in our future works we will also focus on detecting different levels within stress itself. Soon, the proposed methods will be tested in individuals with SUD.

Acknowledgment

We would like to thank Dr. David Gorelick, Professor of Psychiatry, University of Maryland School of Medicine for his valuable inputs and suggestions in data analysis and preparation of this manuscript.

References

[1] "SAMHSA." [Online]. Available: https://www.samhsa.gov/.
[2] "Drug Abuse Statistics." [Online]. Available: https://drugabuse.com/statistics-data/.
[3] "Treating Unhoused People with Addiction During Covid-19." [Online]. Available: https://www.asam.org/Quality-Science/covid-19-coronavirus/treating-unhoused-people-with-addiction-during-covid-19.
[4] "NIDA." [Online]. Available: https://www.drugabuse.gov/about-nida.
[5] M. Lijffijt, K. Hu, and A. C. Swann, "Stress Modulates Illness-Course of Substance Use Disorders: A Translational Review," *Front. Psychiatry*, vol. 5, Jul. 2014.
[6] R. Sinha, "Chronic Stress, Drug Use, and Vulnerability to Addiction," *Ann. N. Y. Acad. Sci.*, vol. 1141, p. 105, 2008.
[7] N. E. Goeders, "The impact of stress on addiction," *Eur. Neuropsychopharmacol.*, vol. 13, no. 6, pp. 435–441, 2003.
[8] S. Carreiro, D. Smelson, M. Ranney, K. J. Horvath, R. W. Picard, E. D. Boudreaux, R. Hayes, and E. W. Boyer, "Real-Time Mobile Detection of Drug Use with Wearable Biosensors: A Pilot Study," *J. Med. Toxicol.*, vol. 11, no. 1, pp. 73–79, Mar. 2015.
[9] C. Goldfine, J. T. Lai, E. Lucey, M. Newcomb, and S. Carreiro, "Wearable and Wireless mHealth Technologies for Substance Use Disorder," *Current Addiction Reports*, vol. 7, no. 3. Springer, pp. 291–300, 01-Sep-2020.
[10] J. W. Ahn, Y. Ku, and H. C. Kim, "A novel wearable EEG and ECG recording system for stress assessment," *Sensors (Switzerland)*, vol. 19, no. 9, May 2019.
[11] J. Du, J. Huang, Y. An, and W. Xu, "The Relationship between stress and negative emotion: The Mediating role of rumination," *Clin. Res. Trials*, vol. 4, no. 1, 2018.
[12] P. Schmidt, A. Reiss, R. Duerichen, and K. Van Laerhoven, "Introducing WeSAD, a multimodal dataset for wearable stress and affect detection," *ICMI 2018 - Proc. 2018 Int. Conf. Multimodal Interact.*, pp. 400–408, Oct. 2018.
[13] T. Iqbal, P. Redon-Lurbe, A. J. Simpkin, A. Elahi, S. Ganly, W. Wijns, and A. Shahzad, "A Sensitivity Analysis of Biophysiological Responses of Stress for Wearable Sensors in Connected Health," *IEEE Access*, vol. 9, pp. 93567–93579, 2021.
[14] S. Elzeiny and M. Qaraqe, "Automatic and Intelligent Stressor Identification Based on Photoplethysmography Analysis," *IEEE Access*, vol. 9, pp. 68498–68510, 2021.
[15] M. D. Hssayeni and B. Ghoraani, "Multi-Modal Physiological Data Fusion for Affect Estimation Using Deep Learning," *IEEE Access*, vol. 9, pp. 21642–21652, 2021.