ICC 2022 - IEEE International Conference on Communications | 978-1-5386-8347-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICC45855.2022.9838970

Feasibility Study of Stress Detection with Machine Learning through EDA from Wearable Devices

Lili Zhu*, Pai Chet Ng*§, Yuanhao Yu**, Yang Wang**,
Petros Spachos*, Dimitrios Hatzinakos§, and Konstantinos N. Plataniotis§
*School of Engineering, University of Guelph, Guelph, ON, Canada
§Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada
**Huawei Technologies Canada, ON, Canada

***Abstract*—The recent pandemic has brought tremendous changes to everyone's life, causing stress about losing loved ones, losing jobs, and having changes in sleep or eating habits. This study investigates the feasibility of utilizing Electrodermal Activity (EDA) collected from wearable devices to detect people's stress. EDA can quantify the changes in sympathetic dynamics by measuring sweat produced by our sweat glands. Currently, the adoption of EDA sensors to commercially off-the-shelf smartwatches is still in the infancy stage, and only a few brands have the EDA sensors implemented into their smartwatch. To facilitate our feasibility study, we need the datasets that contain the EDA signals collected from wearable devices. This paper uses two publicly available datasets containing the EDA signals collected from research-grade wearable devices. We cast the stress detection problem as a binary classification problem and trained the classifiers with three popular machine learning methods: K-Nearest Neighbor, Logistic Regression, and Random Forests. According to experimental results, Random Forests achieves an accuracy of 85.7% to classify stress from non-stress status. The results verified that wearable devices with EDA sensors have the potential to predict stress status.**

***Index Terms*—Electrodermal activity, Stress Detection, Wearable Device, Smartwatches, Machine Learning, KNN, Logistic Regression, Random Forest.**

I. Introduction

All aspects of people's livings have undergone tremendous changes, since the global spreading of the COVID-19 pandemic in early 2020. According to a survey released by World Health Organization (WHO) in October 2020, the pandemic has caused the interruption or cessation of crucial mental health services in 93% of countries worldwide, and people's needs for mental health have increased significantly during the same period [1]. Some preventive measures, such as lockdowns and social distancing, aim to combat the spread of COVID-19. Unfortunately, many people may still face problems such as increased alcohol and drug use, insomnia, and anxiety due to mental stress. Research exposed that suicides over 65 years old increased by 30% in Hong Kong during the Severe Acute Respiratory Syndrome (SARS) epidemic in 2003 [2].

Mental health is the foundation of overall health and well-being. Therefore, researchers from all fields and societies are calling for increased investments in mental health. However, the attention and investment should not be limited to the intervention and treatment of existing apparent symptoms, but more importantly, it should also consider the not-so-obvious stress. Considering the pervasiveness of wearable devices, this paper studies the feasibility of using the Electrodermal Activity (EDA) data collected from smartwatches to detect human stress. An EDA sensor measures the change in skin conductance. Since skin conductance can reflect the human body's emotions and physiological response, EDA is often used as a physiological indicator to measure emotional changes. A few commercially available wearable devices (e.g., Fitbit sense) have integrated EDA sensors to deliver emerging applications, such as emotion monitoring to prevent excessive tension and anxiety [3].

Although EDA is a promising approach for stress and emotion monitoring, many wearable-based industries have yet to include it in their smartwatch designs. Unlike the medical-grade EDA sensors that need to be fixed at a specific position, wearable devices, especially smartwatches, are always worn on the human wrist with uncertain motion. Such dynamic motion by humans creates challenges to detect human stress with EDA sensors integrated in wearable devices. To verify the feasibility of EDA sensors for the above purposes, we explore two public datasets that provide EDA signals collected from wearable devices. The WESAD and VerBIO datasets contain various sensing modalities from wearable devices. In this work, we focus on the EDA signals. We applied three machine learning methods: K-nearest neighbors (KNN), Logistic Regression, and Random Forest, to compare their performances on classifying stress and non-stress status. In addition, the classification results of training with all features and selected features are examined. We verified that the EDA data collected by the wearable device can be used to detect stress status and can be utilized for further study.

The main contributions of this paper are summarized below:

- Our work focuses on the feasibility of using EDA data collected from wearable devices to detect stress.
- The stress classification in this study is based on a data-driven approach and achieved an accuracy of 85.7% when Random Forest is used.
- The reliability of training with extracted statistical and EDA-related features from EDA signals was assessed.

The rest of this paper is organized as follows: Section II includes a literature review. Section III introduces the system

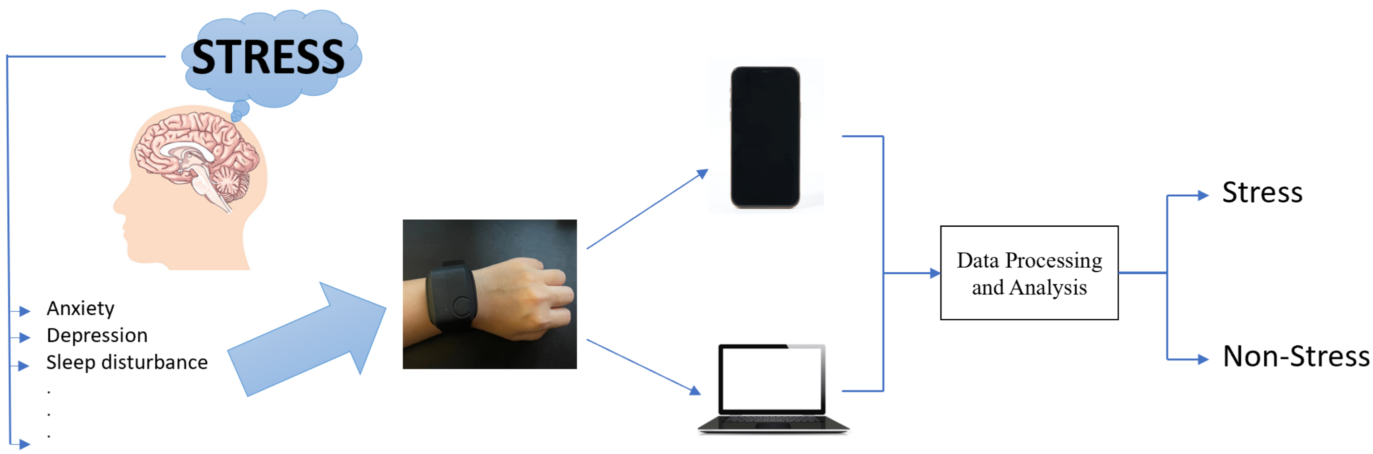

Fig. 1: A general stress detection pipeline of this study.

overview and related concepts. Section IV presents the methods we adopted for the experiment. Section V analyzes the performances of the methods and the observations from the results. Section VI summarizes this work.

II. Related Works

EDA has been applied to assess emotion and stress responses in several scenarios. In [4], they designed an experiment to explore if EDA is a helpful indicator of emotional responses when people are working alone, collaborating, and competing with others. The experiment results showed that higher EDA occurred when participants attended the collaborative task than the competitive task. Another experiment in [5] found a strong relation between EDA signals and the self-reported arousal scores given by the participants who were asked to read emotional content aloud. On the other hand, the feasibility of using EDA to detect the stress arousal of people who are in underwater situations has been examined in [6]. All these experiments have proved the feasibility of using EDA for stress monitoring. However, they have not examined the feasibility of using EDA data collected from wearable devices.

The use of EDA was examined in different scenarios targeting different groups of users. For example, EDA was used to evaluate the construction workers' perceived risk when they were on site [7]. EDA was used to detect stress for patients who were going to have surgeries in [8]. In [9], they examined the stress responses of drivers via multiple electrocardiogram (ECG) and EDA data when they were driving in a simulator with different car setups. In [10], again with drivers, they used Fisher projection and linear discriminant analysis to detect drivers' stress levels based on EDA data collected under different driving conditions, and the methods had a recognition rate of 81.82%. In [11], they concluded that adolescents who have a major depressive disorder (MDD) showed noticeable low EDA during continuously recording periods, which indicates a dysfunctional regulation of the sympathetic part of the autonomic nervous system with the adolescent with MDD. Unfortunately, none of the above work targets general users of all ages undergoing a regular daily life routine. Our work focuses on EDA on wearable devices because these devices can be worn in all scenarios to detect human stress without intruding on human daily life activities.

When it comes to COVID-19, the psychological stress of medical workers during the COVID-19 pandemic are discussed in [12], [13], were calling for active intervention strategy to help medical workers relieve stress. A summary of the physiological metrics which can be utilized to monitor the physical health and mental well-being of COVID-19 positive individuals and the front-line workers is shown in [14], and the authors are calling for adopting wearable devices with physiological sensors to assist in alleviating the negative mental impacts brought by COVID-19.

To this end, our study verified the feasibility of filling the gap by using a wearable device and machine learning methods to detect stress based on EDA signals. It provides a direction for the industries and fastens their process in adopting EDA sensors to their smartwatches for real-world use cases.

III. System Overview

In this study, we use the EDA data collected by wearable devices and utilized the EDA data to test the feasibility of detecting stress status. The proposed framework is illustrated in Fig. 1. It consists of the EDA signals collected from wearable devices and will be exported to a computing device for further analysis. This section presents the characteristics of EDA signals and the wearable device, along with a brief description of the two public datasets we used in this paper.

A. EDA signal characteristics

Since EDA refers to changes in skin conductance in response to sweat secretion (usually a small amount), it can be measured with an EDA device, which records the electrical signal by electrodes applied to the skin. EDA signal is composed of two main components: the general tonic level and the phasic level, shown in Fig. 2.

The general tonic level EDA is correlated to the signal's slow-acting components and background characteristics. The

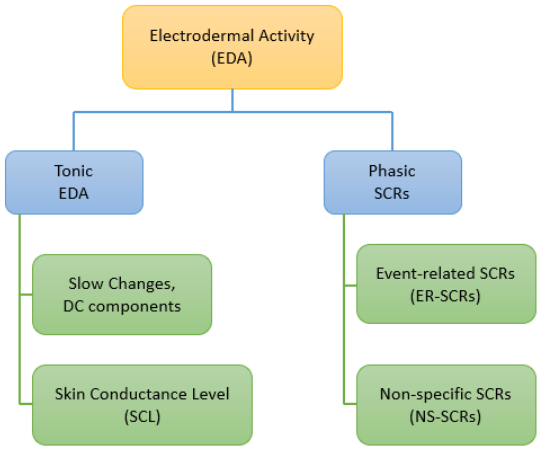

Fig. 2: Components of EDA signals.

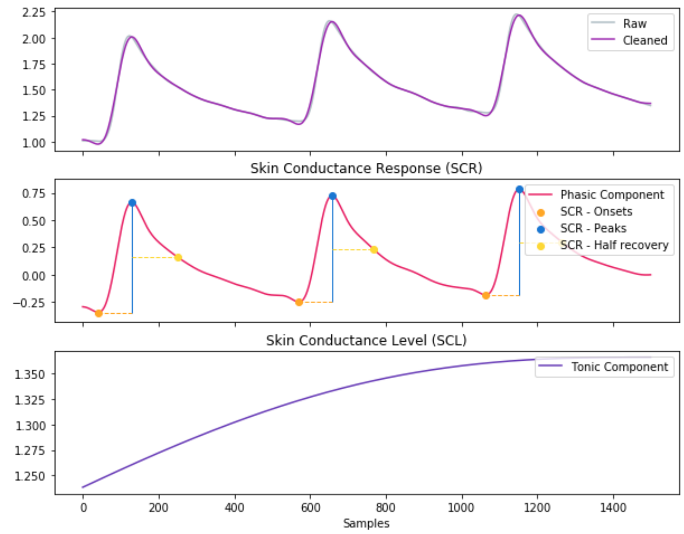

Fig. 3: An example of the features of EDA signals.

slow-acting components feature slow climbing and slow declining over time according to the individual's response, skin dryness, or autonomous adjustment ability. The measurement of the tonic component is the skin conductance level (SCL), and changes in SCL are considered the general changes in autonomic arousal.

The phasic component refers to the fast changes in the EDA signal called the skin conductance response (SCR) [15]. The phasic response is above the tonic level, with more significant changes and faster speeds, displayed as a burst or peak. SCR is sensitive to specific emotional stimulation events. Event-related skin conductance responses (ER-SCRs) will burst within 1 to 5 seconds after the emotional stimulation occurs; non-specific skin conductance responses (NS-SCRs) spontaneously occur in the human body and do not relate to any stimuli. The event-related SCRs are the main focus when analyzing EDA data since they can reflect the arousal and engagement of the subjects. The four indicators that can be used to characterize a typical SCR are shown in Fig. 3. These four indicators can be summarized as the following four EDA features:

- Skin conductance level (SCL): Tonic level of skin's electrical conductivity.
- Skin conductance response (SCR): Phasic change in skin's electrical conductivity.
- Non-specific SCR (NS-SCRs): Rate of NS-SCRs that occur in the absence of identifiable stimuli over time.
- Event-related SCR (ER-SCR): SCRs that can be attributed to a specific eliciting stimulus.

B. Wearable device

Wearable devices use software and hardware components to achieve powerful functions through data and cloud interaction. In particular, these devices combine technologies such as multimedia, sensors, and wireless communication with daily wear to realize hardware terminals with functions such as user interaction, entertainment, and physiological monitoring.

Medical treatment will be one of the main directions of wearable devices' development among these functions. Since the advancement of science and technology, people have paid more attention to their health and hope to get their health information anytime and anywhere. EDA is one of the physiological signals that need to be monitored. Table I shows the wearable devices that currently have EDA sensors.

C. Data collection

In order to evaluate the individual EDA characteristics of respondents, their EDA activity should be collected during a neutral baseline period [16]. In this condition, no stimuli are presented. Respondents sit comfortably in a relaxed position. The recorded EDA activity reflects the spontaneous variability of the signal, consisting only of the tonic level and non-specific skin conductance responses (NS-SCR) only. After the baseline period, the respondents will receive variable stimuli. As EDA peaks occur within 1 to 5 seconds after the onset of a stimulus, researchers certainly want to present any material long enough for respondents to process its content. Furthermore, placing cool-off stimuli of appropriate duration between the stimuli of interest might be helpful to allow the EDA signal to return to baseline. Then the respondents' EDA signal can start from the baseline level for the next stimulus. Important characteristics of the signal are the following:

- Latency. The duration from stimulus onset to the onset of the phasic burst.
- Peak amplitude. The amplitude difference between onset and peak.
- Rise time. The duration from onset to peak.
- Recovery time. The duration from peak to 100% recovery.

D. Public available EDA datasets

Unlike other physiological indicators, such as Electroencephalography (EEG), ECG, and Photoplethysmogram (PPG), open EDA datasets have always been scarce. The EDA data in some open datasets, such as PPG-DaLiA [19] and CLAS [20], were acquired from the wrist by the wearable device, while the

Product	Form	Sensors	Functions
Empatica E4	Wristband	EDA, PPG, ST	Collect physiological data
MyFeel	Wristband	HR, EDA, ST	Monitor stress level
MOXO Sensor	Wristband	EDA	Measure emotional reactions
HEALBE GoBe2	Wristband	Impedance, HR, EDA	Track calorie intake, body hydration, sleep quality, heart rate, and stress levels
U-Check-It	Wristband	EDA	Monitor stress hormones
Fitbit Sense	Smartwatch	ECG, EDA, ST, PPG, Red and infrared sensors for SpO2	Stress management,monitor HR&HRV, SpO2, Breathing Rate, Sleep quality, ST, and ECG

Table I: Wearable devices with EDA sensor available on the market.

Dataset	Subjects	Device	Activities	Modalities	Year	Ref.
VerBIO	55	Empatica E4, Actiwave	Public speaking	EDA, ECG, BVP	2020	[17]
WESAD	15	RespiBAN, Empatica E4	Reading, watching videos, public speaking, meditation, a mental arithmetic task	EDA, ECG, BVP, RESP	2018	[18]

Table II: Information about public available EDA datasets used in this work.

EDA data could also be recorded from fingertips or palm in other datasets such as ITMDER [21]. The datasets' objectives mainly focused on detecting emotions, stress, and attention distraction.

In this work, we chose two available public datasets, which are VerBIO [17] and WESAD [18], as summarized in Table II. These datasets were chosen for two main reasons. First, the EDA data in the two datasets were all recorded by the wearable device Empatica E4. Empatica E4 is a wristband of which the electrodes are built to detect the EDA signals on the wrist. Second, the experiments in the two datasets have similar activities. No physical-intense activity was included in VerBIO and WESAD, and public speaking was required to perform in both of them. In this way, there can be more possibilities to cross-compare the data. The objectives of creating these two datasets were both for emotions/stress detection. Furthermore, these datasets have relatively detailed demographic descriptions, including the age and gender of the subjects, so that it provides more possibilities to conduct experiments with different purposes for future study. The followings are the overviews of these two datasets.

1) VerBIO: The objective of creating the VerBIO dataset and analyzing the data is to understand if the stress caused by public speaking impacts physiological signal changes. EDA, ECG, and Blood Volume Pulse (BVP) signals were recorded during 344 public speaking sessions given by 55 subjects. The subjects had to deliver their speeches to real or virtual audiences in different sessions. A virtual reality device was adopted to stimulate the public speaking scenario with virtual audiences. Empatica E4 recorded the EDA data at a frequency of 4 Hz. This dataset uses self-report as ground truth labels.

2) WESAD: WESAD was built for emotional states recognition. It filled the gap between the authors' previous laboratory research on stress and emotion by including three different emotional states: neutral, stress, and entertained. This dataset includes EDA, ECG, EMG, respiration, body temperature, and triaxial acceleration data recorded from 15 subjects. Watching amusing videos, performing public speaking and mental arithmetic tasks, and meditation were designed to elicit the subjects' different emotional statuses. Again, Empatica E4 was used to obtain the EDA data in WESAD, and the frequency is 4 Hz. In addition, the dataset also includes self-reports obtained from subjects through multiple surveys.

IV. Methodology

A stress detection system should be able to maps the signal inputs (i.e., EDA in this paper) to an output indicating a person's stress level. We formulated stress detection as a binary classification problem. Let $E = [e_1, e_2, \ldots, e_t, \ldots, e_n]$ be the input EDA data, where e_t denote the EDA measurement in Siemens at discrete time t, the stress detection problem can be defined as follows:

$$s = C(E_w, \theta) \tag{1}$$

where $s = \{0, 1\}$ is the stress state, with $s = 1$ indicates stress and $s = 0$ non-stress, $E_w \in E$ is a subset of EDA signals defined according to the segmentation window $w < n$, C is the classification model, and θ is the corresponding coefficients. These classification models and their corresponding coefficients can be learned with any supervised machine learning methods.

A. Data pre-processing

During data pre-processing, four main steps should be performed.

1) Data segmentation: Since EDA data are usually extracted from different activities, and the duration of the activities usually lasts for several minutes to hours, researchers need to split the EDA data into segments with the same lengths so that the data format will be consistent and the computation cost will be affordable. This study segmented all data and labels by a 30-second non-overlapping sliding window for next-step processing. In WESAD, the data came from more than one activity and had four ground truth labels.

Dataset	Subjects	Selected Subjects	Segments	
			Stress	Non-stress
VerBIO	55	18	284	322
WESAD	15	15	317	572

Table III: The segmentation overview of VerBIO and WESAD.

Dataset	All	Training	Validation	Testing
VerBIO	18	15	1	2
WESAD	15	12	1	2

Table IV: The number of subjects for training, validation, and testing set.

For the binary classification objective in this study, we merged the labels to stress and non-stress categories for each dataset according to the stress status indicated on the labels. After applying the sliding window, there are 606 attributes from VerBIO and 889 attributes from WESAD. Table III shows the segmentation results of VerBIO and WESAD.

2) Components seperation: Further data pre-processing is still necessary for EDA data has two components that make the data seem not intuitive. Meanwhile, the electrodes of the EDA sensor will slightly move due to skin moisture and arm or body movements, resulting in artifacts being generated. Therefore, the artifact removal methods need to be applied to the raw data, but the SCR and SCL components must be extracted. The Python library NeuroKit, which is developed specifically for processing and analyzing physiological signals, including EDA data, can apply filters to EDA data and extract the informative SCR components. In our research, the NeuroKit is used to pre-process the EDA data.

3) Data splitting: When the pre-processing is finished, the data are split into training and testing sets. The splitting is based on subjects for subject-independent experiments are more beneficial than subject-dependent ones so that the research can explore more insights of the relations between the emotional changes and the EDA data based on individuals. We took 10% of the subjects as testing set (rounded up), and the rest are used for training. The leave-one-subject-out method is used to validate the training, which means one subject is left for validation in every training round. Table IV shows the data splitting results.

4) Feature extraction: Training with all data features would be time-consuming and increase the computation cost. Statistical features and additional SCR features were computed and extracted to form a feature vector used to train the data. In the end, seven features are chosen to establish the feature vector. The feature vector can be expressed as:

$$\begin{aligned} FeatureVector =&[mean_{EDA}, min_{EDA}, \\ &max_{EDA}, std_{EDA}, \\ &mean_{SCRonsets}, mean_{SCRamp}, \\ &mean_{SCRrecovery}] \end{aligned} \tag{2}$$

B. Classification algorithms

Three machine learning methods are applied to perform the classification task.

1) K-nearest neighbor: KNN classifies by measuring the distance between different feature values, that is, given a training dataset, find the K instances in the training dataset closest to the input instance. The input instance will be classified according to most of the K instances belonging to which specific class.

2) Logistic Regression: Logistic Regression is a linear classification algorithm that investigates a sample's probability belonging to a certain category. Logistic regression calculates the best decision boundary that can distinguish the categories to the most.

3) Random Forest: Random Forest is randomly composed of many decision trees in the forest, and no correlation exists between each decision tree. When there is a new sample, each decision tree in the forest decides to which category the sample belongs. Then the decision tress vote to determine the final classification result based on which category is more selected.

C. Evaluation metrics

The accuracy, recall, precision, and F1 score, all of which are common evaluation methods in statistics, are used to evaluate the classification performances, as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$F1score = 2 * \frac{precision * recall}{precision + recall} \tag{6}$$

$TP = True\ positive, TN = True\ negative,$
$FP = False\ positive, FN = False\ negative.$

V. Results and Discussion

We performed the binary stress and non-stress classification tasks with both the full features and the feature vector with selected features. From Table V, we can observe that Logistic Regression and Random Forest have better performance on VerBIO and WESAD, respectively. Logistic Regression has an 85.3% accuracy for VerBIO, while Random Forest achieved an accuracy of 85.7% on WESAD. However, the performance of KNN is not as satisfactory as Logistic Regression and Random Forest on both datasets. Although VerBIO is a balanced dataset while WESAD is unbalanced, the results indicate that the EDA data extracted from the wrist by Empatica E4 in both datasets can detect the stress status. However, due to the lacking of open EDA datasets, the training samples are insufficient for the growing demand for EDA research. The similarity of VerBIO and WESAD offers an opportunity that EDA datasets with similar activities and labels can be considered to merge and train together to expand the training samples and improve the classification accuracy.

Model	Data	Metrics	VerBIO	WESAD
KNN	Full	F1 score	0.788	0.569
		Accuracy	0.794	0.605
	Features	F1 score	0.588	0.412
		Accuracy	0.588	0.664
Logistic Regression	Full	F1 score	0.828	0.407
		Accuracy	0.853	0.731
	Features	F1 score	0.615	0.743
		Accuracy	0.706	0.849
Random Forest	Full	F1 score	0.824	0.566
		Accuracy	0.824	0.588
	Features	F1 score	0.800	0.785
		Accuracy	0.765	0.857

Table V: Evaluation of the classifiers on the classification task with full features and selected features. (Full: All features; Features: Selected features.)

In addition, all the three classifiers offer better accuracy when training with full features than with selected features of VerBIO. However, an opposite conclusion can be drawn from the results of WESAD. This indicates that it is unreliable to determine whether training with all or selected features is safe for different datasets. Since the wearable device on the human body will be slightly displaced due to people's activities and the EDA signal is derived from direct contact with the skin, artifacts will inevitably be generated. The peaks caused by the artifacts will affect the data quality and classification results accuracy. Consequently, the extracted SCR features may be affected, resulting in comparing training with all features and different combinations of exacted features, and advanced artifact removal algorithms are necessary.

The proposed framework and the experiment results show the feasibility of adopting wearable devices with EDA sensors to detect stress. Empatica E4, a research-grade wearable device, is widely used in academia for its convenience and raw data provision. Verifying the feasibility of using EDA to detect stress can provide confidence to the industry that integrating EDA sensors into smartwatches is the mainstream to develop future smartwatch products.

VI. Conclusion

This study evaluated the feasibility of classifying stress and non-stress status with EDA signals by applying three machine learning models, KNN, Logistic Regression, and Random Forest, on the datasets VerBIO and WESAD. We trained the models with all features and selected features separately, and there has no consistency that training with selected features could be better than with all features, or vice versa. According to the results, Random Forest offered the best binary classification performances on WESAD, with an accuracy of 85.7%, making EDA data collected from wearable devices a promising approach for stress detection systems.

References

[1] WHO, "COVID-19 disrupting mental health services in most countries, WHO survey," 2020, Available at: https://www.who.int/news/item/05-10-2020-covid-19-disrupting-mental-health-services-in-most-countries-who-survey, Accessed: 2022-02-15.

[2] S. M. S. Chan, F. K. H. Chiu, C. W. L. Lam, P. Y. V. Leung, and Y. Conwell, "Elderly suicide and the 2003 sars epidemic in hong kong," *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, vol. 21, no. 2, pp. 113–118, 2006.

[3] Fitbit debuts sense, its most advanced health smartwatch; world's first with eda sensor for stress management, plus ecg app, spo2 and skin temperature sensors. [Online]. Available: https://www.businesswire.com/news/home/20200825005373/en/

[4] P. Sariñana-González, Á. Romero-Martínez, and L. Moya-Albiol, "Cooperation induces an increase in emotional response, as measured by electrodermal activity and mood," *Current Psychology*, vol. 36, no. 2, pp. 366–375, 2017.

[5] C. Marzi, A. Greco, E. P. Scilingo, and N. Vanello, "Towards a model of arousal change after affective word pronunciation based on electrodermal activity and speech analysis," *Biomedical Signal Processing and Control*, vol. 67, p. 102517, 2021.

[6] H. F. Posada-Quintero, J. P. Florian, A. D. Orjuela-Cañón, and K. H. Chon, "Electrodermal activity is sensitive to cognitive stress under water," *Frontiers in physiology*, vol. 8, p. 1128, 2018.

[7] B. Choi, H. Jebelli, and S. Lee, "Feasibility analysis of electrodermal activity (eda) acquired from wearable sensors to assess construction workers' perceived risk," *Safety science*, vol. 115, pp. 110–120, 2019.

[8] A. Anusha, P. Sukumaran, V. Sarveswaran, A. Shyam, T. J. Akl, S. Preejith, M. Sivaprakasam *et al.*, "Electrodermal activity based pre-surgery stress detection using a wrist wearable," *IEEE journal of biomedical and health informatics*, vol. 24, no. 1, pp. 92–100, 2019.

[9] P. Zontone, A. Affanni, R. Bernardini, L. Del Linz, A. Piras, and R. Rinaldo, "Emotional response analysis using electrodermal activity, electrocardiogram and eye tracking signals in drivers with various car setups," in *2020 28th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 1160–1164.

[10] Y. Liu and S. Du, "Psychological stress level detection based on electrodermal activity," *Behavioural brain research*, vol. 341, pp. 50–53, 2018.

[11] A. Mestanikova, I. Ondrejka, M. Mestanik, I. Hrtanek, E. Snircova, and I. Tonhajzerova, "Electrodermal activity in adolescent depression," in *Pulmonary Infection and Inflammation*. Springer, 2016, pp. 83–88.

[12] X. Shen, X. Zou, X. Zhong, J. Yan, and L. Li, "Psychological stress of icu nurses in the time of covid-19," *Critical Care*, vol. 24, 2020.

[13] W. Wu, Y. Zhang, P. Wang, L. Zhang, G. Wang, G. Lei, Q. Xiao, X. Cao, Y. Bian, S. Xie *et al.*, "Psychological stress of medical staffs during outbreak of covid-19 and adjustment strategy," *Journal of medical virology*, vol. 92, no. 10, pp. 1962–1970, 2020.

[14] D. R. Seshadri, E. V. Davies, E. R. Harlow, J. J. Hsu, S. C. Knighton, T. A. Walker, J. E. Voos, and C. K. Drummond, "Wearable sensors for covid-19: a call to action to harness our digital infrastructure for remote patient monitoring and virtual assessments," *Frontiers in Digital Health*, vol. 2, p. 8, 2020.

[15] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe, "A guide for analysing electrodermal activity (eda) & skin conductance responses (scrs) for psychological experiments," *Psychophysiology*, vol. 49, no. 1, pp. 1017–1034, 2013.

[16] A. Greco, G. Valenza, and E. P. Scilingo, *Advances in Electrodermal activity processing with applications for mental health*. Springer, 2016.

[17] M. Yadav, M. N. Sakib, E. H. Nirjhar, K. Feng, A. Behzadan, and T. Chaspari, "Exploring individual differences of public speaking anxiety in real-life and virtual presentations," *IEEE Transactions on Affective Computing*, 2020.

[18] 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*, 2018, pp. 400–408.

[19] A. Reiss, I. Indlekofer, P. Schmidt, and K. Van Laerhoven, "Deep ppg: Large-scale heart rate estimation with convolutional neural networks," *Sensors*, vol. 19, no. 14, p. 3079, 2019.

[20] V. Markova, T. Ganchev, and K. Kalinkov, "Clas: a database for cognitive load, affect and stress recognition," in *2019 International Conference on Biomedical Innovations and Applications (BIA)*. IEEE, 2019, pp. 1–4.

[21] J. Pinto, "Exploring physiological multimodality for emotional assessment," 2019.