*Cognitive Computing and Internet of Things,* Vol. 43, 2022, 111–117
https://doi.org/10.54941/ahfe1001845

# Accurate Stress Detection from Novel Real-Time Electrodermal Activity Signals and Multi-Task Learning Models

**Stefan de Vries[1], Reon Smits[1], Michalina Tataj[1], Mieke Ronckers[1], Mayra van der Pol[1], Fransje van Oost[1], Esmee Adam[2], Hanneke Smaling[2], and Erwin Meinders[1]**

[1]Mentech, Eindhoven, The Netherlands
[2]Department of Public Health and Primary Care, Leiden University Medical Center, The Netherlands

## ABSTRACT

Stress often is associated with physical and mental health issues. To prevent these issues, an early detection of stress is essential. However, for people with an intellectual disability effectively expressing stress can be difficult and therefore, the necessary intervention can be delayed. An automatic stress detection system could help caregivers in early detection of stress development. This can be achieved using wearable sensors that continuously record physiology. The changes in physiological signals, like in skin conductance can be used to classify moments of stress. The devices recording these signals are however, not always suitable for long term measurements. The present study evaluates a newly developed sock integrated skin conductance sensor (SentiSock) that does not restrict movement and stays comfortable over time. To assess if the sensor can be used for stress detection a comparison was made with the Empatica E4, a commonly used wrist-based skin conductance sensor. Both sensors were worn by 28 participants (mean age 39.25 ± 17.04) in a lab study where stress was induced using mathematical exercises. The data was used to train a multitask learning neural network for each device, following an identical procedure. The models were validated using a 5-fold cross validation that resulted in an average balanced accuracy of 0.824 (SD = 0.018) for Empatica E4 and 0.834 (SD = 0.019) for SentiSock. This demonstrated that both sensors can be used to detect stress adequately in lab conditions. Given these results, SentiSock will be further investigated for long term measurements.

**Keywords:** Stress detection, Skin conductance, Multitask learning, Wearables

## INTRODUCTION

With a prevalence close to 30 percent in the general population in recent years, stress affects many people in society (Salari et al., 2020). Stress does not only affect one's mental health but also has implications for physical health (Cohen et al., 2007). It is therefore, highly important to detect stress in early stages to reduce the impact it may have on an individual. This may be especially true for people unable to express their stress effectively, for example people with a severe intellectual disability (Doodeman et al., 2022). Furthermore, stress in people with limited communicative abilities can express itself as challenging behavior (Scott and Havercamp, 2014), resulting in a reduced quality of life (Gur, 2016) and increased stress and burden in caregivers (Panicker and Ramesh, 2019). It is, therefore, important to detect the early built-up of stress automatically to ensure a timely response to changes in the patients' well-being. This may be achieved using technological solutions that allow for continuous monitoring of stress.

There are numerous ways to automatically detect stress (Gedam and Paul, 2021). While many methods have shown promising results, for a successful implementation of the technology in daily life, a method that does not limit persons' freedom of movement is required. A widely used solution is the use of wireless wearable devices. For stress detection, sensors that record physiological signals are most commonly used. This is due to the strong relation between physiology and psychological stress (Giannakakis et al., 2019), especially for electrodermal activity (EDA). EDA refers to the electric conductivity of the skin. During stressful episodes EDA amplitude increases together with the drastic increase in the number of peaks in the signal (Boucsein, 2012).

EDA is a signal that can be recorded at many locations on the body. However, not all body locations give the same measurement quality. The best locations to record EDA are at the fingertips, hand palm, forehead and foot (Dooren et al., 2012). The current available sensors generally measure EDA on the wrist, although this is found to be a suboptimal location (Dooren et al., 2012). The main motivation to record EDA at the wrist is that it causes minimal inconvenience and restrictions for the subject. The free movement may, however, introduce more movement artefacts, reducing the effectivity of these sensors in tasks that require movement of the upper body. In these cases measurement of EDA on the foot may result in a better performance (Liu and Du, 2018).

There is a clear relation between emotional state and EDA measured on the foot (Dooren et al., 2012; Frederiks et al., 2019; Liu and Du, 2018). However, using current approaches it is uncomfortable to record EDA on the foot. It has been suggested to integrate an EDA sensor into a sock to tackle this issue (Liu and Du, 2018). While there have been several studies that integrated an EDA sensor into a sock (Frederiks et al., 2019; Healey, 2011), no viable solution exists so far. To investigate the feasibility of a sock with integrated EDA sensor for stress detection, the present study evaluates a newly developed sensor integrated wearable called the SentiSock (Mentech, Eindhoven, the Netherlands). This sensor was designed for stress detection in people with an intellectual disability or dementia.

A lab study with stress-inducing experiments was designed to evaluate the stress detection capability of the SentiSock in healthy participants. By comparing the SentiSock to the wrist-based EDA sensor Empatica E4 (Empatica Inc, Boston, United States of America), the study aims to demonstrate the accuracy and capability of the sock-integrated EDA sensor for stress detection. This will serve as a first step into the direction of stress detection for people with special needs.

## DATASET

A within-subject lab experiment was used to determine the relation between subjective and physiological stress. First, three baseline measurements were performed: sitting, standing, and walking while watching a neutral video for 5-minutes each. Then, participants experienced three counterbalanced conditions (sitting, standing, and walking), which consisted each of a 5-minute arithmetic task followed by a 5-minute neutral rest period. During the arithmetic task, participants had to solve equations and add the individual digits of that solution under time pressure. The task was designed to elicit stress by adding feedback sounds and a ticking digital clock that changed color. The video shown during the baseline and rest periods was the Windows "3D Pipes" screensaver. After completion of the three conditional tests, the participants were exposed to a cold pressor test (CPT) to induce physical stress. During the CPT participants immersed their dominant hand in cold water (2-5°C) for a maximum of 3 minutes. At the end of the complete experiment, the participants were asked to fill in a questionnaire containing items on demographics (gender, age, etc.) and health conditions (medication use, health problems, exercise, etc.). The total duration of the experiment was 75 minutes. The experiment was carried out in a controlled environment in the lab of Mentech in Eindhoven.

Real-time physiological responses were measured by multiple wearable devices: skin conductance on both feet with the SentiSock, and on the wrist using the Empatica E4. After each task or rest period, subjective emotional responses were measured using the 3-item SAM-scale (Bradley and Lang, 1994).

For this experiment, participants were recruited using convenience sampling; either through the researchers' personal network or volunteering networks such as 'NLvoorelkaar' and 'EindhovenDoet'. Given the changes in physiology when aging, the study explicitly aimed to include people from all ages. Participation was voluntary, and an informed consent was given prior to participation. Participants were given a small gift for their participation (i.e., water bottle). The data were collected between 01-11-2021 and 10-02-2022.

## MODEL DEVELOPMENT

### Procedure

To investigate the performance of the SentiSock for stress detection, two models were trained. One model used the SentiSock EDA data, while the other model used the Empatica E4 EDA data. The models were trained using the following general procedure, further explained in the following sections. First, a data cleaning step was conducted. Subsequently, EDA features were extracted for both datasets, followed by a feature selection procedure. The models were compiled using the same architecture and trained using an identical validation scheme. The procedure was kept identical to avoid any factors other than the sensors from affecting the performance.

### Data Cleaning

The first step was to exclude data that was not suitable for model training. Data was excluded based on two criteria. All periods with corrupt Empatica E4 or SentiSock EDA signal were excluded. Additionally, all participants who did not experience both stressful and non-stressful states, as indicated by their SAM scores, were omitted. The SAM scores were binarized, by labeling values larger than 5 were as moments of stress, and below 5 as rest. Following the feature extraction, the feature dataset samples with identical or invalid values (e.g. not a number, or infinity) were removed and a minimum 30 examples per class per person were required to include the features in the study.

### Feature Extraction

The features were calculated from the raw EDA signal. All features were calculated using a sliding window of 20 seconds with an overlap of 10 seconds. In total, 30 features were extracted. All features were based on the statistical and temporal sets in the TSFEL Python package (Barandas et al., 2020). These features were extracted for EDA signal from both the Empatica E4 and the SentiSock.

### Feature Selection

The features were selected using the variance influence factor (VIF) (Witten, 2013). This factor represents the multicollinearity of the features, indicating which features do not add new information and may be removed. Features with a VIF greater than 5 were excluded. This procedure was applied for both SentiSock and Empatica E4. For SentiSock a total of 9 features were included, while for Empatica E4 10 features were included.

### Model Architecture

The selected model used a multitask learning (MTL) neural network based on people-as-tasks (Jaques et al., 2016; Taylor et al., 2020). In this architecture, a shared layer represents the general physiological changes related to stress, while the task (or personal) layers represent the person specific physiological changes related to stress.

The shared layer of the model consisted of 30 neurons and used Swish activation. Each personal layer had 10 neurons and also used Swish activation. The output of each personal layer used sigmoid activation. The model was compiled using an Adam optimizer with a learning rate of 0.0001 and used the binary cross-entropy loss function.

### Model Validation

The models were validated by calculating a set of performance metrics from the output associated with the person. Since each person has their own personal layer, it was investigated how the model performs on new data from a person on their own personal layer. The calculated metrics included the f1-score, sensitivity, specificity, and balanced accuracy. To use all available data a 5-fold cross-validation was used. To ensure that each person was

**Table 1.** 5-fold cross validation results (mean ± SD) for SentiSock and Empatica E4 models.

| Model | f1-score | Sensitivity | Specificity | Balanced Accuracy |
|---|---|---|---|---|
| SentiSock | 0.843 ± 0.017 | 0.813 ± 0.025 | 0.855 ± 0.022 | 0.834 ± 0.019 |
| Empatica E4 | 0.832 ± 0.022 | 0.807 ± 0.022 | 0.840 ± 0.033 | 0.824 ± 0.018 |

well represented in each fold, the folds were created separately for each person. Afterwards the mean of the different folds was taken to evaluate the performance of the model.

## RESULTS

In total, 51 participants completed the whole experiment. After the data cleaning, a sample of 28 participants remained for the analysis (20 males and 8 females; mean age = 39.25, *SD* = 17.04, range = 22–69). In total 12 participants were excluded because they did not experience both stress and restful periods as indicated by their SAM scores. Four participants were excluded because of missing Empatica E4 recordings, that were either corrupt or not recorded. Additionally, five participants were removed because of invalid features from the Empatica E4 recording. The remaining excluded participants did not have at least 30 examples for both rest and stressful moments.

The results of the model validation using a 5-fold cross-validation on the Empatica E4 and SentiSock EDA are shown in Table 1.

## DISCUSSION

The present study evaluated the performance of the SentiSock, a sock-integrated EDA sensor for stress detection, by comparing its stress detection capability with that of the Empatica E4, a widely used wrist-based sensor. The Empatica E4 and the SentiSock model both performed high on f1-score, sensitivity, specificity and balanced accuracy. The performance of the SentiSock model was in line with previous research on foot-based sensors (Liu and Du, 2018). The present study illustrated a negligible difference in performance of the stress detection models between the two sensors. Interestingly, the suboptimal location of the wrist for EDA measurement (Dooren et al., 2012), did not lead to lower performance. The choice to use a foot-based or wrist-based EDA sensor may therefore, be more driven by the application or use-case. For example, daily use in a care setting for real-time stress detection requires a comfortable sensor that can be integrated in the daily care process (Poh et al., 2010). The wrist-based sensors use metal studs to capture the skin conductance. These studs press into the skin and eventually may lead to bruises and skin damage (Jackson et al., 2019). On the contrary, a sock-integrated sensor consists of printed electrodes and electrode pads that feel comfortable and seamlessly integrate with the garment of the sock. Future studies should examine the user-friendliness, acceptability and applicability of the sock-integrated sensor in clinical practice with people with cognitive and communication impairments.

While the study demonstrated that both Empactica E4 and SentiSock can be used to detect stress, it should be noted that the models were trained and validated with lab experiments. The lab experiment did not induce stress in all volunteers , which largely explains the number of excluded participants. Furthermore, additional samples were omitted due to missing or corrupted data from the Empatica E4, which may have been caused by the streaming platform that was used. Due to the validation through a lab experiment, it was not possible to assess the performance under daily-life conditions, different types of stress, and how the system would perform over long periods of time. To conclude a sock-integrated EDA sensor is viable for stress detection. Future studies should validate the sock-integrated sensor for stress detection in long term care.

## CONCLUSION

The current study demonstrated the capability of stress detection with a sock-integrated EDA sensor. The performance of the sensor was evaluated with a neural network model, trained with labeled physiological responses to emotional data of 28 test persons. The stress metrics obtained from the foot-sensor, including the f1-score, balanced accuracy, sensitivity, and specificity, were comparable with the stress metrics obtained from the wrist-sensor. The high balanced accuracy of 0.834 demonstrates the capability of accurate stress detection. Furthermore, a sock-integrated EDA sensor may be more comfortable to wear than the wristband. This will be especially relevant in cases where the sensor will be worn for long periods of time, like in continuous stress detection. Additionally, the added comfort may help in target groups that may not accept noticeable, uncomfortable or restrictive wearables.

## REFERENCES

Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P.J., Liu, H., Schultz, T., Gamboa, H., 2020. TSFEL: Time Series Feature Extraction Library. SoftwareX. https://doi.org/10.1016/j.softx.2020.100456

Boucsein, W., 2012. Principles of Electrodermal Phenomena, in: Boucsein, W. (Ed.), Electrodermal Activity. Springer US, Boston, MA, pp. 1–86. https://doi.org/10.1007/978-1-4614-1126-0_1

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: The self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry 25, 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Cohen, S., Janicki-Deverts, D., Miller, G.E., 2007. Psychological stress and disease. JAMA 298, 1685–1687. https://doi.org/10.1001/jama.298.14.1685

Doodeman, T.W.M., Schuengel, C., Sterkenburg, P.S., 2022. Expressions of stress of people with severe intellectual disabilities and sensitive caregiving to regulate stress: A qualitative study. J. Intellect. Dev. Disabil. 1–10. https://doi.org/10.3109/13668250.2021.2011171

Dooren, M.V., Vries, J., Janssen, J.H., 2012. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. Physiol. Behav. https://doi.org/10.1016/j.physbeh.2012.01.020

Frederiks, K., Sterkenburg, P., Barakova, E., Feijs, L., 2019. The effects of a bioresponse system on the joint attention behaviour of adults with visual and severe or profound intellectual disabilities and their affective mutuality with their caregivers. J. Appl. Res. Intellect. Disabil. JARID 32, 890–900. https://doi.org/10.1111/jar.12581

Gedam, S., Paul, S., 2021. A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3085502

Giannakakis, G., Grigoriadis, D., Giannakaki, K., Simantiraki, O., Roniotis, A., Tsiknakis, M., 2019. Review on psychological stress detection using biosignals. https://doi.org/10.1109/TAFFC.2019.2927337

Gur, A., 2016. Challenging behavior, functioning difficulties, and quality of life of adults with intellectual disabilities. Int. J. Dev. Disabil. 64, 45–52. https://doi.org/10.1080/20473869.2016.1221233

Healey, J., 2011. GSR Sock: A New e-Textile Sensor Prototype, in: 2011 15th Annual International Symposium on Wearable Computers. Presented at the 2011 15th Annual International Symposium on Wearable Computers, pp. 113–114. https://doi.org/10.1109/ISWC.2011.36

Jackson, D., Sarki, A.M., Betteridge, R., Brooke, J., 2019. Medical device-related pressure ulcers: A systematic review and meta-analysis. Int. J. Nurs. Stud. https://doi.org/10.1016/j.ijnurstu.2019.02.006

Jaques, N., Taylor, S., Nosakhare, E., Sano, A., Picard, R.W., 2016. Multi-task Learning for Predicting Health , Stress , and Happiness.

Liu, Y., Du, S., 2018. Psychological stress level detection based on electrodermal activity. Behav. Brain Res. https://doi.org/10.1016/j.bbr.2017.12.021

Panicker, A.S., Ramesh, S., 2019. Psychological status and coping styles of caregivers of individuals with intellectual disability and psychiatric illness. J. Appl. Res. Intellect. Disabil. 32, 1–14. https://doi.org/10.1111/jar.12496

Poh, M.-Z., Swenson, N.C., Picard, R.W., 2010. A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity. IEEE Trans. Biomed. Eng. 57, 1243–1252. https://doi.org/10.1109/TBME.2009.2038487

Salari, N., Hosseinian-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulpoor, Shna, Mohammadi, M., Rasoulpoor, Shabnam, Khaledi-Paveh, B., 2020. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. Glob. Health 16, 57. https://doi.org/10.1186/s12992-020-00589-w

Scott, H., Havercamp, S., 2014. Mental health for people with intellectual disability: the impact of stress and social support. Am. J. Intellect. Dev. Disabil. https://doi.org/10.1352/1944-7558-119.6.552

Taylor, S., Jaques, N., Nosakhare, E., Sano, A., Picard, R.W., 2020. Personalized Multitask Learning for Predicting Tomorrow's Mood, Stress, and Health. IEEE Trans. Affect. Comput. https://doi.org/10.1109/TAFFC.2017.2784832

Witten, D., 2013. An Introduction to Statistical Learning: with Applications in R. undefined.